Fast, Simple, Accurate Multi-atlas Segmentation of the Brain

Sean Murphy[1], Brian Mohr[1], Yasutaka Fushimi[2],
Hitoshi Yamagata[3], and Ian Poole[1]

[1] Toshiba Medical Visualization Systems Europe,
2 Anderson Pl, Edinburgh, EH6 5NP, UK
`bmohr@tmvse.com`
[2] Department of Diagnostic Imaging and Nuclear Medicine, Kyoto University
Graduate School of Medicine, 54 Shogoin Kawaharacho,
Sakyoku, Kyoto, 606-8507, Japan
[3] Toshiba Medical Systems Corporation,
1385 Shimoishigami, Otawara, 324-8550, Japan

Abstract. We are concerned with the segmentation of structures within the brain particularly the gyri of the cerebral cortex, but also subcortical structures from volumetric T1-weighted MRI images. A fully automatic multi-atlas registration based segmentation approach is used to label novel data. We use a standard affine registration method combined with a small deformation (non-diffeomorphic), non-linear registration method which optimises mutual information, with a cascading set of regularisation parameters. We consistently segment 138 structures in the brain, 98 in the cortex and 40 in the sub-cortex. An overall Dice score of 0.704 and a mean surface distance of 1.106 mm is achieved in leave-one-out cross validation using all available atlases. The algorithm has been evaluated on a number of different cohorts which includes patients of different ages and scanner manufacturers, and has been shown to be robust. It is shown to have comparable accuracy to other state of the art methods using the MICCAI 2012 multi-atlas challenge benchmark, but the runtime is substantially less.

1 Introduction

The aim of this work is to provide fully automated, accurate segmentation of the gyrus regions and substructures in T1 weighted images. Fully automatic segmentation has applications in visualisation, localisation of pathology, navigation, neurosurgery planning, radio-therapy planning and in understanding the morphometry and longitudinal changes of the brain, particularly with respect to neurological conditions like Alzheimer's, schizophrenia and Parkinson's [1,2,3].

The notion of a medical image atlas is introduced in [4]. The atlas can be decorated with a variety of extra information such as points, curves, structures and probabilistic maps. Most schemes for segmenting the brain are based on volumetric registration although a deformable surface model approach is used in [5,6]. Methods based on registration are dependent on the accuracy of that registration.

S. Ourselin and M. Modat (Eds.): WBIR 2014, LNCS 8545, pp. 1–10, 2014.

An extensive evaluation [7], compares the accuracy of 14 non-linear registration algorithms in the context of brain parcellation. The paper concluded that there is a modest correlation between the degrees of freedom of a registration method and its accuracy. It also found that the relative accuracy of methods appeared to be little affected by subject population, labelling protocol and accuracy measurement. This suggests (at least for the evaluated algorithms) that they will generalize well to unseen populations and protocols. The top entrants were SPM (Ashburner, University College London), ART (NITRC, Mass), IRTK (Rueckert, Imperial College London) and SyN (Avants, University of Pennsylvania).

For an example on how to combine segmentation and registration in one complete framework see [8]. This approach shows how to use the expectation maximization algorithm to determine the variance and mean of the MR signals in tissue types/compartments which incorporates priors from probabilistic masks.

In [9], Fischl constructs a probabilistic model of the position of each compartment in the brain. He also constructs a model of the MRI-signal which is assumed to be drawn from a Gaussian distribution where the parameters are free to vary from point to point. These models are constructed by data-mining many samples. The probabilistic model also includes a Markov Random Field component and is used to drive non-linear registration as in [8].

Multi-atlas methods involve performing multiple independent registrations and fusing the results together, typically using per pixel majority voting as discussed in [10]. Many variations on this theme exist. The STAPLE algorithm [11] shows how to solve the segmentation problem while simultaneously evaluating the accuracy of the different raters (atlases). For a comparison between state of the art multi-atlas segmentation techniques see the MICCAI 2012 challenge on multi-atlas labelling [12]. Techniques used included variations of the expectation maximisation algorithm, variations of the STAPLE algorithm, patch based label fusion, trained classification methods, different registration approaches including log domain diffeomorphic demons, spline based methods and dense deformation field methods. The winning entry from the University of Pennsylvania performed affine registration using FLIRT, followed by non-rigid registrtion using AIR of the novel image to all atlases, followed by a label fusion algorithm and a corrective learning approach using AdaBoost classifier with a mixture of spatial, appearance and contextual features within a 5 x 5 x 5 window [13]. In section 3, we numerically compare our technique to those in the challenge by evaluating on the same training and test sets.

In the following sections we describe our method in chronological order, followed by a results section which is broken into three subsections, followed by a conclusions section.

2 Method

We also propose a multi-atlas based solution. The algorithm is outlined in figure 1, with details given in the following sub-sections. We first affinely register all available atlases (27) by maximising mutual information (MI) over the set of rigid body transformations using simulated annealing at a reduced scale of 8, and then over a 9 parameter search space which includes axis aligned scales using Powell's method [14] at a scale of 4. A specified number of these atlases (possibly all atlases) are selected to proceed to the non-linear registration phase by performing hierarchical clustering on the resultant transformations and selecting the most consistent subset. The selection trades runtime for accuracy. The multi-scale non-linear registration phase optimises MI over the set of dense displacement fields by gradient descent, with a semi-numerical expression for the gradient. Finally the results from these atlases are used to construct a per compartment probability map which are used as priors in the expectation maximisation (EM) fitting of a per compartment Gaussian intensity model. The final segmentations are arrived at by maximum a posteriori (MAP) classification.

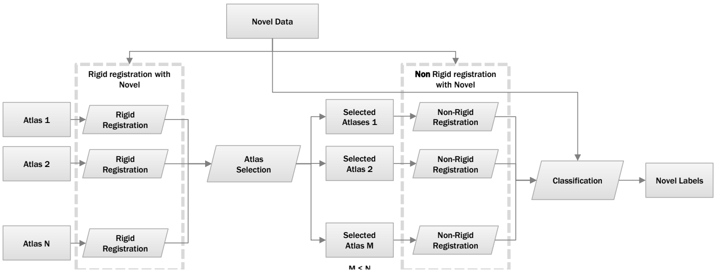


Fig. 1. Proposed workflow

2.1 Affine Registration

Affine registration between each atlas and the novel dataset is achieved by initially optimizing MI computed from a joint histogram using simulated annealing [15] over the class of rigid body transformations. The rotation is parameterised in terms of roll, pitch and yaw Euler angles around the center of the atlas volume. The novel and atlas volumes are both downscaled with anti-aliased smoothing by a factor of 8 prior to registration. The joint histogram is computed by sparsely sampling from the volumes. The transformation which aligns the center of the two volumes with no relative scaling or rotation is used to initialise the search. The simulated annealing optimizer has the effect of restarting in several locations, which avoids local minima and provides robustness. This is then followed by a search over the set of 9-parameter transformations which includes axis aligned scaling, using Powell's method at a reduced scale of 4.

2.2 Atlas Clustering

Clustering of the atlases with respect to the computed affine transformation is used to reduce the runtime of the algorithm by reducing the number of non-linear registrations to atlas volumes. Our method requires that a number of anatomical landmarks are manually marked on the *atlas* datasets. The novel volume is affinely registered to all atlas volumes. The distance metric $d(A, B)$ for atlases A and B is defined in 1, as the sum of the distances between corresponding landmarks.

$$d(A,B) = \sum_i \|T_A(\mathbf{x}_{A,i}) - T_B(\mathbf{x}_{B,i}))\| \quad (1)$$

$\|\cdot\|$ is the Euclidean norm, T_A and T_B are the affine transformations to the novel image from atlas A and B respectively, and $\mathbf{x}_{A,i}$ and $\mathbf{x}_{B,i}$ are the positions of predetermined anatomical landmarks, indexed by i, in atlases A and B respectively. In this way the pre-defined landmark positions $\mathbf{x}_i$, which have been identified offline by a clinician, are transformed to the novel coordinate system using the affine transformation T for each atlas.

The distance metric is used to populate a square matrix with dimensions of the number of atlases (in this case 28). Hierarchical clustering, [16], is used to identify the atlases which perform consistent registrations. The most consistent cluster of the appropriate size is selected. In some cases the cluster may be too large, in which case a random subset of this set is taken. The selected atlases will continue to the non-linear registration, voting, and EM steps outlined below. Thus overall, atlas clustering identifies outliers from the affine registration stage, presumes these to be errors, and removes them from the following steps.

2.3 Non-linear Registration

$$W = \arg\max_{W'} \mathrm{MI}(R, W'[T]) \quad (2)$$

This stage finds a dense deformation field W (not necessarily diffeomorphic), parameterised as vectors on a Cartesian grid, which maximises (locally) the MI between the floating volume T and the reference R. The deformation-field is initialised from the affine phase and evolved to convergence, using a gradient ascent optimiser, with the deformation-field at iteration i given by:

$$W_{i+1} = E * (W_i + kV * F_i) \quad (3) \qquad F = \frac{\partial \mathrm{MI}(R, W[T])}{\partial W} \quad (4)$$

Convolution kernels V and E enforce the viscous fluid and elastic constraints, respectively; both modelled by Gaussian filters ($*$ is the convolution operator). F_i is the current force-field, calculated as the multivariate derivative of the MI with respect to the components of the deformation at each point as described in [17]. k is a free constant which controls the rate of ascent. It is beneficial to use a decreasing cascade of k's and of fluid regularisation parameters. Both these parameters are reduced geometrically when temporary convergence of MI is detected, with a lower bound on the fluid regularisation.

The non-linear registration algorithm is multi-scale where results from lower scales are used as the input to finer scales. The down scale factors used here are 4, 2 and 1. For this application, no elastic constraints are used. The registration is restricted to the domain of the brain in the atlas dataset dilated by 5 voxels.

2.4 Expectation-Maximization (EM)

The T1 weighted signal of each voxel x, in a given compartment, c, is assumed to be drawn from a Gaussian distribution: $P(x|c) = \mathcal{N}(x|\mu_c, \sigma_c^2)$. The parameters of each Gaussian distribution, the mean μ_c and variance σ_c^2, are unknown.

The compartment/class, c, of each voxel is also unknown and is modelled as a per voxel distribution $\rho_{c,i} = p(c|i, x)$ where i denotes the voxel. The Gaussian parameters and voxel classifications can be estimated by maximizing the joint probability density function of the entire image, using the EM algorithm [18] with sufficient independency assumptions. This iterative algorithm uses gradient descent to converge to a locally maximal configuration. The votes from each of the atlases can be used to inform this process as in [19,20], by constructing a probability mask for each compartment which takes the form of a prior $p(c|i)$. The variables in the algorithm at iteration t are denoted as $\rho_{c,i}^{(t)}$, $\mu_c^{(t)}$ and $\sigma_c^{2(t)}$. In this scheme background is also included as a class.

$$\rho_{c,i}^{(0)} = p(c|i) \tag{5}$$

$$\rho_{c,i}^{(t+1)} = \frac{\mathcal{N}(x_i|\mu_c^{(t)}, \sigma_c^{2(t)})p(c|i)}{\sum_c \mathcal{N}(x_i|\mu_c^{(t)}, \sigma_c^{2(t)})p(c|i)} \tag{6}$$

$$\mu_c^{(t)} = \frac{\sum_i x_i \rho_{c,i}^{(t)}}{\sum_i \rho_{c,i}^{(t)}} \tag{7}$$

$$\sigma_c^{2(t)} = \frac{\sum_i (x_i - \mu_c^{(t)})^2 \rho_{c,i}^{(t)}}{\sum_i \rho_{c,i}^{(t)}} \tag{8}$$

3 Results

The method has been evaluated in three ways. Firstly, on a unaltered subset of the OASIS database as discussed in section 3.1, using a leave one out strategy. Secondly, on data derived from the MICCAI multi-atlas labelling challenge, using a classic training/test split in section 3.2. Evaluating on this group allows for a like for like comparison to other entrants in the challenge. Thirdly, the results have been qualitatively evaluated on a broad range of private T1 weighted images from different cohorts, scanner manufacturers and geographic locations as discussed in 3.3.

Table 1. Top: 98 detected cortical structures. Middle: 40 detected subcortical structure. Bottom: 12 annotated anatomical landmarks. All identifiers have a left and right variant with the exception of those marked with *.

inferior temporal gyrus	cuneus	postcentral gyrus
superior parietal lobule	precuneus	frontal operculum
superior occipital gyrus	gyrus rectus	central operculum
occipital fusiform gyrus	frontal pole	parietal operculum
inferior occipital gyrus	temporal pole	supramarginal gyrus
anterior cingulate gyrus	planum polare	medial orbital gyrus
transverse temporal gyrus	lingual gyrus	middle frontal gyrus
posterior cingulate gyrus	angular gyrus	parahippocampal gyrus
supplementary motor cortex	occipital pole	middle temporal gyrus
precentral gyrus medial segment	fusiform gyrus	medial frontal cortex
postcentral gyrus medial segment	entorhinal area	lateral orbital gyrus
superior frontal gyrus medial segment	anterior insula	superior frontal gyrus
orbital part of the inferior frontal gyrus	subcallosal area	middle occipital gyrus
opercular part of the inferior frontal gyrus	planum temporale	middle cingulate gyrus
triangular part of the inferior frontal gyrus	precentral gyrus	anterior orbital gyrus
	posterior insula	superior temporal gyrus
	calcarine cortex	posterior orbital gyrus

Lateral Ventricle	CSF*	Hippocampus
Cerebral Exterior	Vessel	Optic Chiasm*
Cerebellum Exterior	Putamen	Inf Lat Vent
Cerebral White Matter	Caudate	4th Ventricle*
Cerebellum White Matter	Pallidum	3rd Ventricle*
Cerebellar Vermal Lobules I-V*	Amygdala	Accumbens Area
Cerebellar Vermal Lobules VIII-X*	Ventral DC	Basal Forebrain
Cerebellar Vermal Lobules VI-VII*	Brain Stem*	Thalamus Proper

Frontal Horn of the Lateral Ventricle	Floor of the Maxillary Sinus	Pineal Gland*
Optic Nerve Attachment Point	Pituitary Gland (Base)*	
Superior Aspect of Eye Globe	Centre of Eye Globe	

3.1 OASIS

28 datasets were manually segmented from the OASIS database [21] by Neuromorphometrics (NMM)[1]. The patients range in age from 30 to 96, with approximately four patients from each decade: 12 male and 16 female. The patients are all right-handed. NMM clinicians have annotated 138 structures in all. A complete list of the cortical and subcortical regions defined can be found in table 1. Each dataset was additionally annotated with up to 12 key anatomical landmarks as specified in 1. These are not detected in the novel dataset, but are used during the clustering stage. The OASIS volumes have been anonymised digitally by removing the face, and the signal to noise ratio has been improved by averaging together several repeated scans. Not every structure has been labelled in every dataset because it may not have been present.

We performed a leave-one-out cross validation and evaluated two accuracy metrics: Dice coefficient, see 9, and mean surface distance in millimetres, see 10.

$$\frac{2|X \cap Y|}{|X| + |Y|} \quad (9) \qquad \frac{1}{|S_X|} \sum_{i \in S_X} \operatorname*{arg\,min}_{j \in S_Y} \|i - j\| \quad (10)$$

$\|\cdot\|$ is the Euclidean norm. X and Y are the ground truth and generated segmentations, respectively, and S_X and S_Y the corresponding set of points on the surface of these. Table 2 shows the mean over all compartments and all volumes, for both of these statistics. Structures which are not present in the ground truth in all datasets do not contribute to the average Dice or surface distance as in [12]. Figure 2 shows the effect of the number of atlases used on the accuracy.

Table 2. Mean Dice and mean surface distance for 28 oasis datasets for all structures, for the cortex and for the non-cortex

	Overall	**Cortical**	**Non-Cortical**
Mean Dice	0.704	0.689	0.745
Surface Distance (mm)	1.094	1.119	1.026

3.2 MICCAI Multi-atlas Labelling Challenge

A separate set of 35 datasets were manually segmented from the OASIS database [21] by NMM for the purpose of the challenge. The patients ranged in age from 18 to 90. 15 datasets were used for training and the remaining 20 were used for testing. These volumes were pre-processed using bias-field correction and aligned manually to the AC-PC axis using translation and rotation. Because of this, the initial affine registration problem is somewhat easier than might be expected on truly novel data. A more realistic representation of the expected score can be seen on the pure OASIS dataset in section 3.1. Nonetheless, this set is interesting because it was used in the MICCAI multi-atlas labelling challenge [12]. Held in 2012, this challenge made the training set available and invited entrants to submit their methods and results to be independently evaluated on the test

---

[1] `http://www.neuromorphometrics.com/`

Table 3. The results of the MICCAI multi-atlas segmentation challenge [12]. Shown is the mean overall Dice, the mean Dice in the cortical areas, the mean Dice in the non-cortical areas and also estimated runtimes. The score and ranking of our method (TMVSE_BGM) was retrospectively calculated and is shown in the table in bold.

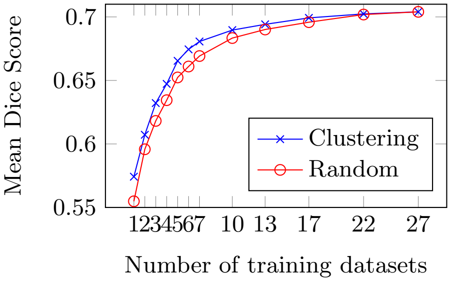
Overall Rank	Team Name	Reg Method	Mean Dice Overall	Mean Dice Cortical	Mean Dice Non-Cortical	Runtime Estimate (mins)
1	PICSL_BC	*ANTS*	0.7654	0.7388	0.8377	> 1200
2	NonLocalSTAPLE	*ANTS*	0.7581	0.7318	0.8296	> 1200
3	MALP_EM	*Nifti-Seg*	0.7576	0.7328	0.8252	> 50 (5 on GPU)
4	PICSL_Joint	*ANTS*	0.7499	0.7216	0.8271	> 1200
5	MAPER	*Nifti-Seg*	0.7413	0.7144	0.8144	> 50 (5 on GPU)
6	STEPS	*Nifti-Seg*	0.7372	0.7107	0.8095	> 50 (5 on GPU)
7	SpatialSTAPLE	*ANTS*	0.7372	0.7093	0.8130	> 1200
8	CIS_JHU	*LDDMM*	0.7357	0.7131	0.7971	10 - 450
8.5	**TMVSE_BGM**		**0.7346**	**0.7183**	**0.7807**	**> 5**
9	CRL_Weighted_STAPLE_ ANTS+Baloo	*ANTS*	0.7344	0.7122	0.7950	> 1200
10	CRL_Weighted_STAPLE_ ANTS	*ANTS*	0.7308	0.7066	0.7966	> 1200
11	CRL_STAPLE_ANTS+Baloo	*ANTS*	0.7290	0.7064	0.7919	> 1200
12	CRL_STAPLE_ANTS	*ANTS*	0.7280	0.7033	0.7951	> 1200
13	CRL_Probabilistic_STAPLE_ANTS+Baloo	*ANTS*	0.7251	0.7009	0.7911	> 1200
14	CRL_MV_ANTS+Baloo	*ANTS*	0.7247	0.6966	0.8012	> 1200
15	CRL_MV_ANTS	*ANTS*	0.7243	0.6951	0.8035	> 1200
16	DISPATCH	*Nifti-Seg*	0.7243	0.6965	0.8000	> 50 (5 on GPU)
17	CRL_Probabilistic_STAPLE_ANTS	*ANTS*	0.7223	0.6972	0.7907	> 1200
18	SBIA_SimRank+NormMS+WtROI	*DRAMMS*	0.7212	0.6940	0.7953	> 240
19	SBIA_BrainROIMaps_MV_IntCorr	*DRAMMS*	0.7193	0.6933	0.7904	> 240
20	SBIA_BrainROIMaps_JaccDet_IntCorr	*DRAMMS*	0.7186	0.6913	0.7927	> 240
21	BIC-IPL-HR	*ANIMAL*	0.7173	0.6888	0.7948	> 168
22	SBIA_SimMSVoting	*DRAMMS*	0.7172	0.6898	0.7918	> 240
23	UNC-NIRAL	*ANTS*	0.7171	0.6869	0.7992	> 1200
24	SBIA_SimRank+NormMS	*DRAMMS*	0.7162	0.6884	0.7919	> 240
25	BIC-IPL	*ANIMAL*	0.7107	0.6829	0.7864	> 168

set. Since all methods were trained and evaluated on the same set, it provided a rare opportunity for the quantitative comparison between various methods. We have retrospectively trained and tested on this data in the same way. The results in table 3 show that are results are comparable to the state of the art in the field. The average Dice was calculated as described in section 3.1. The runtimes per registration are estimated from the specified registration methods cited in [12]. The methods used were "Nifty-Seg", see [22] which reported approximately 3.3 minutes on the CPU, or 20 seconds on a GPU, "ANTS-SyN" [23] which is reported to have runtimes in the order of 80 minutes[7], "DRAMMS" which reports to have a runtime of approximately 20% of ANTS (16 minutes) in [24], "LDDMM" which is reported to have a runtime in the range 40 secs to 30 minutes, on a high end server, depending on required accuracy [25] and "ANIMAL" with a reported time of 11.2 minutes. TMVSE registration (CPU based) has a runtime of 20 seconds on architecture comparable to that specified in [7]. Since all registrations are independent, multiplying by 15 gives the estimated runtime.

3.3 Subjective Evaluation

The results on 90 other T1 MR brain datasets have also informally examined. These datasets had no pre-processing applied to them and are representative of what might be expected when the method is applied to novel data. The

Fig. 2. The average dice achieved as a function of the number of atlases used, for both the random and clustering strategy. For the random strategy, the requested number of atlases are chosen at random. For the clustering strategy the most consistent cluster of atlases of the requested size are used.



datasets covered a range of different cohorts, acquisition parameters, acquisition directions, slice spacing, venders (Siemens, GE, Philips and Toshiba), resolutions, ages, pathologies and geographic locations. No major failure cases were reported, which suggests a robustness of the affine registration phase. In one case, a volume was segmented correctly even though only the left side of the patients head was present in the volume. Subjectively, the method appears to perform better on high resolution data. Examples of these results can be seen in figure 3.

It has been reported that the location of structures are mostly correct although there are often multiple errors per patient, which is typical of other competing approaches. Both under and over segmentation is observed. Segmentations are often noisy, although this can be corrected by post-processing. The multi-atlas approach boosts overall classification/Dice accuracy but it can sometimes come at the cost of unrealistic resultant segmentation shapes. Structures that are topologically inconsistent (not simply connected) with the ground truth have been observed in some cases. The method makes no explicit provision for pathological tissue types. When presented with pathology the method will label it as one (or more) of the known nearby compartments. This has been observed in at least one patient which a large tumour.

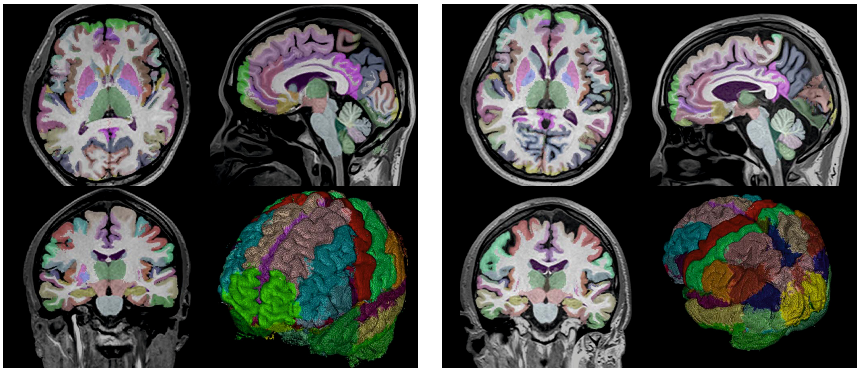


Fig. 3. 2 fully automatic results. 3D T1 Weighted images. Acquired: August 2012 on Toshiba's Titan 3T scanner, Kyoto. Resolution: 512 × 512. Voxel size: $0.49mm \times 0.49mm \times 0.6mm$.

4 Conclusions

A fast and simple method for segmenting structures in the brain is presented. The speed comes from its algorithmic simplicity. The cost of this simplicity is a loss in some advantageous properties that are afforded by the more complex methods like LDDMM and ANTS such as **a)** guaranteed, efficient invertibility, **b)** symmetry: invariance to the order of source and target or **c)** the natural emergence of a mathematical metric and associated space. However it seems that this simplicity does not imply low segmentation accuracy in a multi-atlas setting as shown by the comparable performance to other state-of the art techniques. The simplicity translates to favourable runtimes, with the next fastest CPU based non-rigid registration method in the MICCAI challenge taking an estimated 10 times longer which is useful for time critical applications such as image-guided intervention or for processing large numbers of datasets. We acknowledge that these are rough runtime estimates since they were not directly measured and had to be inferred, but they serve as reasonable order of magnitude estimates. We also acknowledge that our results are retrospective, having had time to optimize performance with sight of other methods and results from the challenge.

Acknowledgements. The authors would like to acknowledge Neuromorphometrics contribution to the project in the form of fully annotated datasets. We would also like to thank the commitee and entrants of the MICCAI 2012 grand challenge on multi-atlas labeling.

References

1. Petrella, J.R., Coleman, R.E., Doraiswamy, P.M.: Neuroimaging and early diagnosis of Alzheimer disease: a look to the future. Radiology 226, 315–336 (2003)
2. Andreasen, N.C., Olsen, S.A., Dennert, J.W., Smith, M.R.: Ventricular enlargement in schizophrenia: relationship to positive and negative symptoms. The American Journal of Psychiatry 139, 297–302 (1982)
3. Hutchinson, M., Raff, U.: Structural Changes of the Substantia Nigra in Parkinson's Disease as Revealed by MR Imaging. Imaging 21, 697–701 (2000)
4. Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J.: A probabilistic atlas of the human brain: theory and rationale for its development the international consortium for brain mapping (ICBM). Neuroimage 2(2PA), 89–101 (1995)
5. MacDonald, D., Kabani, N., Avis, D., Evans, A.C.: Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI. NeuroImage 12(3), 340–356 (2000)
6. Patenaude, B., Smith, S.M., Kennedy, D.N., Jenkinson, M.: A Bayesian model of shape and appearance for subcortical brain segmentation. NeuroImage 56(3), 907–922 (2011)
7. Klein, A., Andersson, J., Ardekani, B., Ashburner, J., et al.: Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. NeuroImage 46(3), 786–802 (2009)
8. Ashburner, J., Friston, K.: Multimodal image coregistration and partitioning–a unified framework. NeuroImage 6(3), 209–217 (1997)

9. Fischl, B.: Automatically Parcellating the Human Cerebral Cortex. Cerebral Cortex 14(1), 11–22 (2004)
10. Rohlfing, T., Brandt, R., Menzel, R., Maurer, C.R.: Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. NeuroImage 21, 1428–1442 (2004)
11. Warfield, S.K., Zou, K.H., Wells, W.M.: Simultaneous Truth and Performance Level Estimation (STAPLE): An Algorithm for the Validation of Image Segmentation 23(7), 903–921 (2004)
12. Landman, B., Warfield, S.: MICCAI 2012 workshop on multi-atlas labeling.. Challenge and Workshop on Multi-Atlas Labeling . . . (2012)
13. Wang, H., Das, S.R., Suh, J.W., Altinay, M., Pluta, J., Craige, C., Avants, B., Yushkevich, P.A.: A learning-based wrapper method to correct systematic errors in automatic image segmentation: consistently improved performance in hippocampus, cortex and brain segmentation. NeuroImage 55(3), 968–985 (2011)
14. Powell, M.: An efficient method for finding the minimum of a function of several variables without calculating derivatives. The Computer Journal 7(2), 155 (1964)
15. Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Optimization by simulated annealing. Science 220, 671–680 (1983)
16. Ward Jr., J.H.: Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association (1963)
17. Crum, W.R., Hill, D.L.G., Hawkes, D.J.: Information theoretic similarity measures in non-rigid registration. In: Taylor, C.J., Noble, J.A. (eds.) IPMI 2003. LNCS, vol. 2732, pp. 378–387. Springer, Heidelberg (2003)
18. Moon, T.: The expectation-maximization algorithm. IEEE Signal Processing Magazine 13 (1996)
19. Ashburner, J., Friston, K.J.: Unified segmentation. NeuroImage 26, 839–851 (2005)
20. Murgasova, M., Rueckert, D., Edwards, D., Hajnal, J.: Robust segmentation of brain structures in MRI. In: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (1), pp. 17–20 (June 2009)
21. Marcus, D.S., Wang, T.H., Parker, J., Csernansky, J.G., Morris, J.C., Buckner, R.L.: Open Access Series of Imaging Studies (OASIS): cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. Journal of Cognitive Neuroscience 19(9), 1498–1507 (2007)
22. Modat, M., Ridgway, G.R., Taylor, Z.A., Lehmann, M., Barnes, J., Hawkes, D.J., Fox, N.C., Ourselin, S.: Fast free-form deformation using graphics processing units. Computer Methods and Programs in Biomedicine 98(3), 278–284 (2010)
23. Avants, B., Tustison, N., Song, G., Cook, P.: A reproducible evaluation of ANTs similarity metric performance in brain image registration. Neuroimage 54(3), 2033–2044 (2011)
24. Ou, Y., Ye, D.H., Pohl, K.M., Davatzikos, C.: Validation of DRAMMS among 12 popular methods in cross-subject cardiac MRI registration. In: Dawant, B.M., Christensen, G.E., Fitzpatrick, J.M., Rueckert, D. (eds.) WBIR 2012. LNCS, vol. 7359, pp. 209–219. Springer, Heidelberg (2012)
25. Ceritoglu, C., Tang, X., Chow, M., Hadjiabadi, D., Shah, D., Brown, T., Burhanullah, M.H., Trinh, H., Hsu, J.T., Ament, K.A., Crocetti, D., Mori, S., Mostofsky, S.H., Yantis, S., Miller, M.I., Ratnanather, J.T.: Computational analysis of LDDMM for brain mapping. Frontiers in Neuroscience 7, 151 (2013)