2010 International Conference on Pattern Recognition

Segmentation of Anatomical Structures in Brain MR Images Using Atlases in FSL - A Quantitative Approach *

Octavian Soldea[1], Ahmet Ekin[1], Diana F. Soldea[2], Devrim Unay[3], Müjdat Çetin[2], Aytül Erçil[2], Mustafa Gökhan Uzunbaş[4], Zeynep Firat[5], Mutlu Cihangiroglu[5], and The Alzheimers Disease Neuroimaging Initiative

[1] *Video Processing and Analysis Group, Philips Research, Eindhoven, The Netherlands*
[2] *Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul, Turkey*
[3] *Electrical and Electronics Engineering, Bahcesehir University, Istanbul, Turkey*
[4] *The Computer Science Department at Rutgers State University, Piscataway, NJ, 08854, USA*
[5] *The Radiology Department of the Yeditepe University Hospital, Istanbul, Turkey*
E-mails: {Octavian.Soldea, Ahmet.Ekin}@philips.com, fdiana@su.sabanciuniv.edu, devrim.unay@bahcesehir.edu.tr, {mcetin@, aytulercil@}sabanciuniv.edu, uzunbas@cs.rutgers.edu, {zfirat, mcihangiroglu}@yeditepe.edu.tr

Abstract

Segmentation of brain structures from MR images is crucial in understanding the disease progress, diagnosis, and treatment monitoring. Atlases, showing the expected locations of the structures, are commonly used to start and guide the segmentation process. In many cases, the quality of the atlas may have a significant effect in the final result. In the literature, commonly used atlases may be obtained from one subject's data, only from the healthy, or depict only certain structures that limit their accuracy. Anatomical variations, pathologies, imaging artifacts all could aggravate the problems related to application of atlases. In this paper, we propose to use multiple atlases that are sufficiently different from each other as much as possible to handle such problems. To this effect, we have built a library of atlases and computed their similarity values to each other. Our study showed that the existing atlases have varying levels of similarity for different structures.

1 Introduction

The large availability of MR acquisition devices raises new challenges for the scientific community. While nowadays, MR devices induce strong magnetic fields, automatic segmentation is still a major conundrum. Although the quality of MR images continues to improve in accuracy and resolution, segmentation of structures is still a difficult task, especially in the presence of pathologies causing large deformations and large anatomical variations between people.

In this context, we analyze an atlas-based segmentation scheme that is advocated by widely accessible tools. In atlas-based segmentation, prior information are obtained from probabilistic maps encoded in the electronic atlases, such as MNI [4], Thalamus, Harvard-Oxford [1], Jülich [2], and Talairach [3]. This prior information is utilized for collecting tissue statistics and to start active contours based schemes [6]. The accuracy of this approach is very much dependent on the quality of the atlas. Some atlases are obtained from only one subject's brain, some model only healthy, and some are specialized to only certain structures. We envision a database of atlases from which multiple atlases depending on the structure of interest and pathology is selected to be used for segmentation. As the first step of this scheme, this paper analyzes the quality and the correlation of different atlases. The multiple atlas information can be integrated in expectation maximization schemes such as [8], [7].

We performed our comparison by using the atlases in FMRIB Software Library (FSL) [5]. We rigidly registered them by using FLIRT of FSL and quantified the correlation and accuracy values in the context of segmentation of cortical and sub-cortical regions, which are critical in evaluation of brain disorders such as Alzheimer's disease.

*This work was partially supported by the European Commission under Grant MTKI-CT-2006-042717

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.635

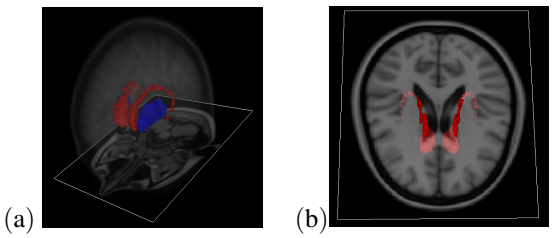


Figure 1. Talairach binary labeling, which is registered to MNI 152 space. Figures (a) and (b) show superpositions of structures and sectioning planes through the volumetric information. (a) - in Talairach, the structures are clearly separated. While Nucleus Caudate is depicted in red, Putamen is illustrated in blue. (b) - in Talairach, the Nucleus Caudate includes (an optional) tail.

2 An Overview of FSL Atlases

In FSL, all the atlases are registered to the MNI 152 model. The FSL distribution includes the MNI 152 T1 weighted data and we choose one millimeter resolution. All the atlases provide binary labeled regions, while the atlases MNI, Thalamus (Thlm), Harvard-Oxford (HO), Jülich (Jü), and JHU also include probability maps, where the probability of pixel membership to a certain structure is expressed in percentages. We regard Tailarach labeling as a binary probability map. The structures in Tailarach (Tal) are clearly disjoint (see Figure 1 (a) and (b)). Note that JHU includes structure tracts therefore, we do not consider it in our analysis.

3 Methodology

Consider an arbitrary fixed atlas. In this context, segmentation begins with registering the MNI 152 space to the subject at analysis. The registration stage produces a matrix that defines a transformation. We apply this transformation matrix to the probability map structures in the atlas. Each transformed structure defines in the subject's volumetric space a segmentation result. Often segmentation techniques strongly rely on registration and vice versa [6]. In this context, we propose an accuracy analysis based on correlations.

3.1 Image Registration

Spatial alignment of the atlases to real subject scans is achieved by the FSL FLIRT affine registration tool (see Figure 2). Optimum parameter setting is performed by a trial-and-error approach, which resulted in the following: six, seven, nine, and twelve degrees-of-freedom (DOF), trilinear neighbor interpolation, and search over angles in the range of $\left[-\frac{\pi}{2}..\frac{\pi}{2}\right]$, around each of $x, y,$ and z axes.

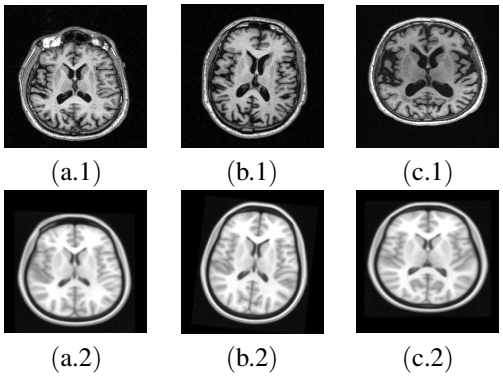


Figure 2. Snapshots from registrations of MNI 152 T1-weighted atlas to the volumetric scans of three subjects: (a) healthy control, (b) mild cognitive impaired, (c) Alzheimer's. For each pair, subject image is displayed at the top (e.g. a.1), while atlas at the bottom (e.g. a.2).

For each one of the different DOF we recommend the same best atlas and pairs of atlases. However, due to brain anatomy related issues, we believe that the best choice is seven, i.e. rotation, translation, and global scale. While six DOF might be too restrictive, nine and twelve DOF are less natural as compared to seven.

3.2 Correlation Analysis

In order to estimate the degree of match between two volumetric images we employ the Pearson's product moment correlation coefficient, which is defined by

$$\rho_{X,Y} = \frac{\Sigma(XY) - \Sigma(X)\Sigma(Y)}{\sqrt{\Sigma(X^2) - \Sigma^2(X)}\sqrt{\Sigma(Y^2) - \Sigma^2(Y)}}$$

where X and Y are the two probabilistic maps and $\Sigma()$ is the expected value operator. The resulting correlation value spans the range of [-1,+1] and indicates the degree of linear dependence or similarity between the two images. The closer the value to -1 or +1 is, the stronger the correlation.

4 Experimental Results and Discussion

We have chosen nine elderly subjects, grouped into three healthy controls (HC), three mild cognitive impaired (MCI, a transition stage between healthy controls and Alzheimer's) and three Alzheimer's patients. For each subject, a T1-weighted MR scan is collected using a 1.5 T scanner.

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu\ADNI). The ADNI was launched in 2003 as a $60 million dollar, 5-year public-private partnership overseen by the

National Institute on Aging. ADNI is the result of efforts of many co-investigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the US and Canada. The investigators within the ADNI did not participate in analysis or writing of this report, however.

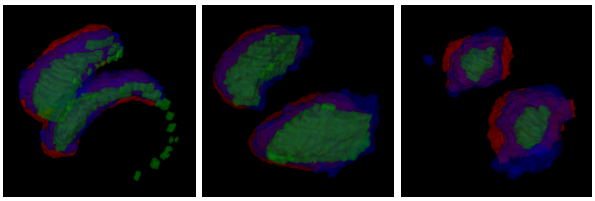
Aiming at the highest possible granularity of similarity analysis, we computed correlations between same structure probability maps among all the atlases. For example, we computed all the correlations between the Putamen probability maps, as indicated in MNI, Harvard-Oxford, and Talairach. In addition, we have computed correlations between same structure probability maps of all the subjects when segmentation was performed using all the atlases. For example, consider an ordering between subjects. We computed correlations between the Putamen segmentations achieved when using $Harvard - Oxford\ and\ the\ first\ subject$ versus $Talairach\ and\ the\ second\ subject$. Note that we computed correlations on MNI 152 projections.

On top of these granular correlation tables, we infer recommendations for using atlases as individuals and pairs. Moreover, we also differentiate between recommendations per structure as well as multi-structure segmentation approaches.

In this work, the volumetric and 3D visualization images are captured using our visualization software, which is implemented over VTK and ITK. In each image, each structure is visualized via an iso-surface. We have chosen a common iso-constant over all the brain structures. Establishing different constants as depending on the volumetric information of each atlas or analyzed subject is not in the scope of this work.

4.1 Comparing Correlations of Atlases Pairs

Nucleus Caudate is labeled in MNI, Harvard-Oxford, and Talairach. Putamen is labeled in MNI, Harvard-Oxford, and Talairach. Amygdala is labeled in Harvard-Oxford, Jülich, and Talairach. Thalamus is labeled in MNI, Thalamus, Harvard-Oxford, and Talairach. Hippocampus is labeled in Harvard-Oxford, Jülich, and Talairach. For each one of the structures Nucleus Caudate, Putamen, Amygdala, Thalamus, and Hippocampus, we illustrate the relative positions of the labels in different labeling atlases Figures 3, 4, and 5, respectively, and provide measurements of correlations in Table 1. In Table 1, we conclude that most correlated pairs are (MNI, Harvard-Oxford) for Nucleus Caudate, Putamen, and Thalamus, and (Harvard-Oxford, Jülich) for the Amygdala and Hippocampus.



(a) (b) (c)

Figure 3. A comparison of structures. (a) Caudate Nucleus.(b) Putamen. (a) and (b) MNI, Harvard-Oxford, and Talairach atlases are colored in red, blue, and green, respectively. (c) Amygdala. Harvard-Oxford, Jülich, and Talairach atlases, which are colored in red, blue, and green, respectively.

	Nucleus Caudate	Putamen	Amygdala	Thalamus	Hippocampus
MNI vs. HO	0.885	0.888	-	0.919	-
MNI vs. Tal	0.687	0.718	-	0.854	-
MNI vs. Thlm	-	-	-	0.84	-
HO vs. Tal	0.662	0.812	0.624	0.82	0.407
HO vs. Jü	-	-	0.714	-	0.681
HO vs. Thlm	-	-	-	0.881	-
Jü vs. Tal	-	-	0.529	-	0.271
Tal vs. Thlm	-	-	-	0.82	-
Best Pair	MNI vs. HO	MNI vs. HO	HO vs. Jü	MNI vs. HO	HO vs. Jü

Table 1. Correlations among all FSL Atlases per brain structure.

4.2 A Comparison of Atlas-Based Segmentations

In this section, we provide a recommendation on a unique atlas. We have chosen MNI, Harvard-Oxford, and Talairach for comparison over Nucleus Caudate, Putamen, and Thalamus. We have computed a leave-one out like test and show the results in Table 2. For example, when we want to decide which one of the MNI and Harvard-Oxford is better, we compare the correlations of MNI and Harvard-Oxford to Talairach. In Table 2, we conclude that the best atlas is Harvard-Oxford. While Harvard-Oxford and Talairach include all structures we analyze, a natural question is correlation to ground truths provided with registration and segmentations with subjects. The registrations to subjects do not change the correlations between the structures computed directly from atlases. We have computed aver-

	Nucleus Caudate	Putamen	Thalamus	Conclusion
HO vs. MNI / Tal	0.662 vs. 0.687	0.812 vs. 0.718	0.820 vs. 0.854	HO > MNI
HO vs. Tal / MNI	0.885 vs. 0.687	0.888 vs. 0.718	0.919 vs. 0.854	HO > Tal
MNI vs. Tal / HO	0.885 vs. 0.662	0.888 vs. 0.812	0.919 vs. 0.820	MNI > Tal

Table 2. Comparison of correlations between atlases. - Recommendation for one atlas. Each cell compares correlations of two atlases (e.g. HO vs. MNI) given a third one as reference (e.g. Tal), while the last column summarizes the conclusion of atlas comparisons. These conclusions are inferred summing correlations on lines.

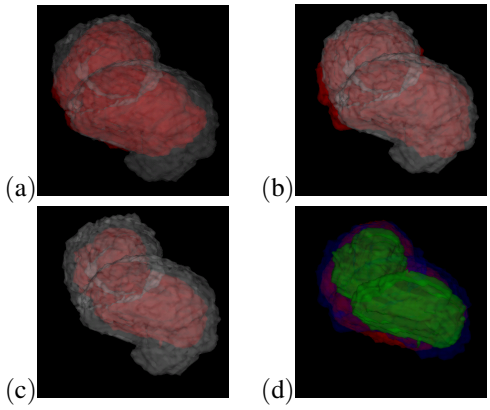


Figure 4. A comparison of Thalamus in MNI, Thalamus, Harvard-Oxford, and Talairach atlases. (a) , (b) , and (c) represent superpositions of MNI, Harvard-Oxford, and Talairach with the Thalamus (atlas) respectively. While MNI, Harvard-Oxford, and Talairach are colored in red, the Thalamus is colored in white. (d) illustrates a superposition of MNI, Harvard-Oxford, and Talairach atlases, which are colored in red, blue, and green, respectively.

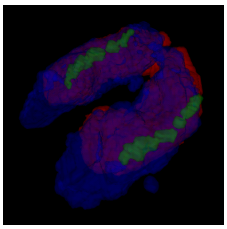


Figure 5. A comparison of Hippocampus in Harvard-Oxford, Jülich, and Talairach atlases. which are colored in red, blue, and green, respectively.

ages of correlations over all nine subjects as well as on the three separate groups, i.e. AD, MCI, and HC. We show the average values for Putamen in Table 3. Regardless of anatomical structure and disease grouping, we report on correlations that follow atlases based quantitative analysis, as presented in Table 1.

4.3 Segmentation Time Requirement

We have run FLIRT, i.e. the registration stage, in Linux, on top of a virtual machine. The underlying machine is a Core2Duo, with 3 GHz processors and 2 GB main memory. Usually, it takes less than half a minute for a registration between a study case and the MNI 152 space. The application of the registration transformations on the atlas probability maps takes less than ten seconds, in the same framework.

Putamen	Entire test set	Alzheimer's	MCI	Healthy controls
MNI vs. HO	0.9 ± 0.00279	0.9 ± 0.00007	0.89 ± 0.00314	0.90 ± 0.00011
MNI vs. Tal	0.77 ± 0.0139	0.78 ± 0.00021	0.76 ± 0.0158	0.78 ± 0.00019
HO vs. Tal	0.87 ± 0.014	0.88 ± 0.00023	0.85 ± 0.0156	0.88 ± 0.00025

Table 3. Correlations among segmentations for Putamen over all the subjects. We consider all subjects together as well as different disease groups.

5 Conclusion

We have analyzed the accuracy of an automatic volumetric segmentation scheme that employs atlases and registration. We provide recommendations for atlases (individually) and pairs of atlases, aiming fast initialization of active contours based segmentation schemes as well as improving the accuracy in expectation maximization methods. While we recommend Harvard-Oxford to be the best individual atlas, this one does not necessarily assure separation of the included structures. When the target is to separate the structures, the best atlas is Talairach. These two atlases include all the structures we have analyzed. When pairs of atlases are of interest, the best correlations are achieved among MNI and Harvard-Oxford for Nucleus Caudate, Putamen, and Thalamus structures. Harvard-Oxford and Jülich present the best correlations for Amygdala and Hippocampus. Future work will deal with extending the experiments on larger data sets.

References

[1] In *Harvard-Oxford Cortical and Subcortical Structural Atlases*. http://www.cma.mgh.harvard.edu/.

[2] S. Eickhoff, T. Paus, S. Caspers, M.-H. Grosbras, A. Evans, K. Zilles, and K. Amunts. Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *Neuroimage*, 36(3):511–521, July 2007.

[3] J. L. Lancaster, D. Tordesillas-Gutierrez, M. Martinez, F. Salinas, A. Evans, K. Zilles, J. C. Mazziotta, and P. T. Fox. Bias between mni and talairach coordinates analyzed using the icbm-152 brain template. *Human Brain Mapping*, 28(11):1194–1205, November 2007.

[4] J. Mazziota and et all. A probabilistic atlas and reference system for the human brain: International consortium for brain mapping (icbm). *Philosophical Transactions of the Royal Society*, 28(11):1194–1205, November 2007.

[5] S. Smith and et all. Advances in functional and structural mr image analysis and implementation as fsl. *NeuroImage*, 23 (Supplement 1):208–219, September 2004.

[6] G. Unal and G. Slabaugh. Estimation of vector fields in unconstrained and inequality constrained variational problems for segmentation and registration. *Journal of Mathematical Imaging and Vision*, 31(1):57 – 72, 2008.

[7] K. van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. A unifying framework for partial volume segmentation of brain MR images. *IEEE Trans on Medical Imaging*, 22(1):105–119, January 2003.

[8] K. S. Warfield, H. K. Zou, and M. W. Wells. Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentationn. *IEEE Trans. on Medical Imaging*, 23(7):903 – 921, 2004.