NeuroImage 49 (2010) 2352–2365

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Fast and robust multi-atlas segmentation of brain magnetic resonance images

Jyrki MP. Lötjönen [a,*], Robin Wolz [b], Juha R. Koikkalainen [a], Lennart Thurfjell [c], Gunhild Waldemar [d], Hilkka Soininen [e], Daniel Rueckert [b] and The Alzheimer's Disease Neuroimaging Initiative [1]

[a] Knowledge Intensive Services, VTT Technical Research Centre of Finland, P.O. Box 1300 (street address Tekniikankatu 1), FIN-33101 Tampere, Finland
[b] Department of Computing, Imperial College London, London, UK
[c] Medical Diagnostics R and D, GE Healthcare, Uppsala, Sweden
[d] Memory Disorders Research Group, Department of Neurology, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark
[e] Department of Neurology, University of Kuopio, Kuopio, Finland

ARTICLE INFO

Article history:
Received 12 July 2009
Revised 9 October 2009
Accepted 10 October 2009
Available online 24 October 2009

Keywords:
MRI
Segmentation
Atlases
Registration
Hippocampus

ABSTRACT

We introduce an optimised pipeline for multi-atlas brain MRI segmentation. Both accuracy and speed of segmentation are considered. We study different similarity measures used in non-rigid registration. We show that intensity differences for intensity normalised images can be used instead of standard normalised mutual information in registration without compromising the accuracy but leading to threefold decrease in the computation time. We study and validate also different methods for atlas selection. Finally, we propose two new approaches for combining multi-atlas segmentation and intensity modelling based on segmentation using expectation maximisation (EM) and optimisation via graph cuts. The segmentation pipeline is evaluated with two data cohorts: IBSR data ($N=18$, six subcortial structures: thalamus, caudate, putamen, pallidum, hippocampus, amygdala) and ADNI data ($N=60$, hippocampus). The average similarity index between automatically and manually generated volumes was 0.849 (IBSR, six subcortical structures) and 0.880 (ADNI, hippocampus). The correlation coefficient for hippocampal volumes was 0.95 with the ADNI data. The computation time using a standard multicore PC computer was about 3–4 min. Our results compare favourably with other recently published results.

Introduction

Brain MR imaging is playing an important role in neuroscience. Neurodegenerative brain diseases mark the brain with morphological signatures; detection of these signs may be useful to improve diagnosis, particularly in diseases for which there are few other diagnostic tools. For example, early and significant hippocampal atrophy in people who have memory complaints points to a diagnosis of Alzheimer's disease. Quantitative analysis and objective interpretation of images usually require segmentation of various structures from images. Reliable and accurate segmentation is a prerequisite for comprehensive analysis of images. Current state-of-the-art brain segmentation algorithms can be classified into algorithms that label voxels (a) into brain/non-brain (Ségonne et al., 2004; Smith, 2002); (b) into different tissue types such as white matter (WM), grey matter (GM), or cerebral spinal fluid (CSF) (Ashburner and Friston, 2005; Bazin and Pham, 2007; Pham and Prince, 1999; Scherrer et al., 2008; van Leemput et al., 1999; Zhang et al., 2001); or (c) algorithms that identify anatomical areas, e.g., hippocampus, thalamus, putamen, caudate, amygdala, and corpus callosum (Bazin and Pham, 2007; Chupin et al., 2009; Corso et al., 2007; Desikan et al., 2006; Fischl et al., 2002; Heckemann et al., 2006; Klein et al., 2005; Morra et al., 2008; Scherrer et al., 2008).

Atlas-based segmentation is a commonly used technique to segment image data. In atlas-based segmentation, an intensity template is registered non-rigidly to a target image and the resulting transformation is used to propagate the tissue class or anatomical structure labels of the template into the space of the target image. Many different approaches have been published using registration-based segmentation, for example, for segmenting subcortical structures (Avants et al., 2008; Bhattacharjee et al., 2008; Han and Fischl, 2007; Pohl et al., 2006). A comparison of different atlas-based segmentation algorithms was recently published by Klein et al. (2009). A review of registration techniques is presented in Gholipour et al. (2007).

The segmentation accuracy can be improved considerably by combining basic atlas-based segmentation with techniques from machine learning, e.g., classifier fusion (Heckemann et al., 2006;

* Corresponding author. Fax: +358 20 722 3499.
E-mail address: jyrki.lotjonen@vtt.fi (J.M.P. Lötjönen).

[1] Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu\ADNI). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. ADNI investigators include (complete listing available at http://www.loni.ucla.edu/ADNI/Collaboration/ADNI_Authorship_list.pdf.

1053-8119/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2009.10.026

Klein et al., 2005; Rohlfing et al., 2004; Warfield et al., 2004). In this approach, several atlases from different subjects are registered to target data. The label that the majority of all warped labels predict for each voxel is used for the final segmentation of the target image. Babalola et al. (2008) compared in a recent study different algorithms for the segmentation of subcortical structures. They found that multi-atlas segmentation produced the best accuracy from the algorithms tested. However, the major drawback of multi-atlas segmentation is that it is computationally expensive, limiting its every day use in clinical practice.

Several factors affect the segmentation accuracy and computation time in multi-atlas segmentation (Fig. 1). First, all atlases are non-rigidly registered to the target (patient) image. During the non-rigid registration, an atlas is deformed in such a way that a similarity measure between the atlas and the target data is maximised. The selection of the similarity measure and the deformation model are central components in optimising the performance of non-rigid registration. A prolific number of solutions are available for similarity measures and for ways to deform the atlas. In this work, we study similarity measures although the deformation model also plays an important role. Second, when the majority voting is applied after non-rigid registration, the objective is to keep the number of atlases as low as possible because the computation time increases correspondingly. As shown in Heckemann et al. (2006), segmentation accuracy increases in a logarithmic way when new atlases are included, i.e., first rapidly and finally very slowly when the number of atlases is high. For these reasons, a compromise must be made when selecting the number of atlases. On the other hand, not only the number of atlases matters but also their quality. If an atlas is very similar to the target data, the inclusion of this atlas probably increases the segmentation accuracy more than less similar atlases. Appropriately implemented atlas selection improves the accuracy of multi-atlas segmentation (Aljabar et al., 2009). Third, the standard multi-atlas segmentation does not model and utilise the statistical distributions of intensities in different structures although this information could be highly valuable in improving the segmentation accuracy. Combining multi-atlas segmentation and intensity modelling as a post-processing step improves the segmentation accuracy (van der Lijn et al., 2008). This work investigates these three factors in more detail.

The ultimate objective of this study is to develop a segmentation method for the clinical practice. This means that we aim (1) to search methods to further improve the segmentation accuracy and (2) to speed up processing without compromising segmentation accuracy, in the context of multi-atlas segmentation. To be clinically feasible, the automatic segmentation algorithm should produce accuracy comparable with manual segmentation made by an expert, and require only a few minutes computation time in a stand-alone PC workstation. The major contribution of this work is the optimisation of the whole multi-atlas segmentation pipeline. We develop and compare different (1) similarity measures in non-rigid registration, (2) atlas-selection methods, and (3) methods to combine multi-atlas segmentation and intensity modelling.

In this article, methods for non-rigid registration, atlas-selection, and combination of multi-atlas segmentation and intensity modelling are first described. This is followed by describing the experiments to assess the multi-atlas segmentation pipeline. Finally, results for two data cohorts are shown and discussed. Part of the research presented in this work appeared previously in conference articles (Lötjönen et al., 2009; Wolz et al., 2009).

Materials and methods

In this section, the whole pipeline for multi-atlas segmentation is described: pre-processing, non-rigid registration, atlas selection, and combination of multi-atlas segmentation and intensity modelling as a post-processing step.

Pre-processing

Intensity normalisation of atlases

The intensity values of CSF, GM, and WM in the atlases were first normalised; the mean intensity values of CSF, GM, and WM were computed and mapped to pre-defined intensity values (see details in Intensity difference as a similarity measure section).

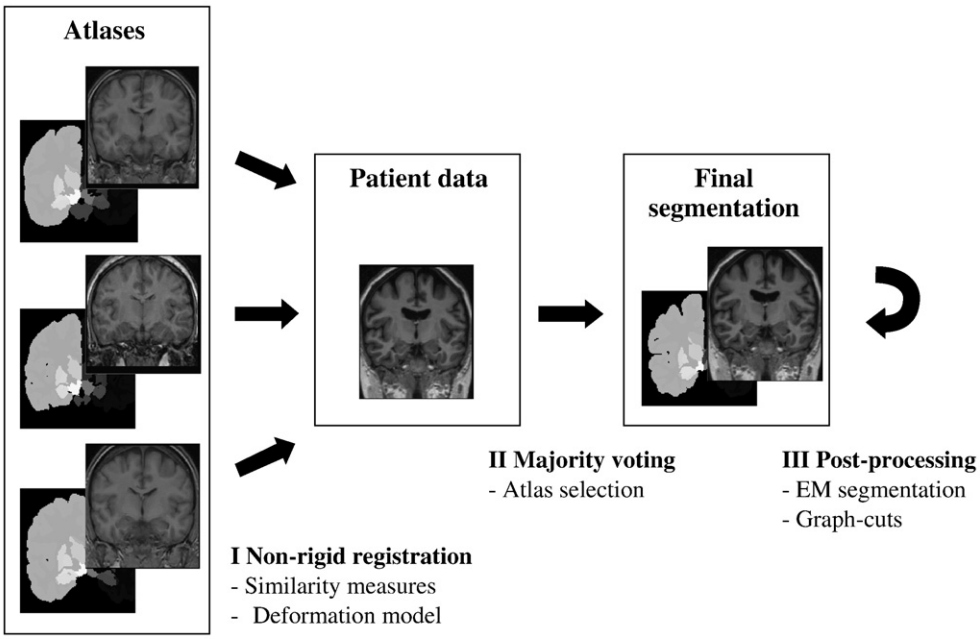

Fig. 1. Steps of multi-atlas segmentation: (I) non-rigid registration used to register all atlases to patient data, (II) classifier fusion using majority voting for producing class labels for all voxels, and (III) post-processing of multi-atlas segmentation result by various algorithms taking into account intensity distributions of different structures.