NeuroImage 143 (2016) 235–249

Contents lists available at ScienceDirect

NeuroImage

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

Fast and sequence-adaptive whole-brain segmentation using parametric Bayesian modeling

Oula Puonti [a,*], Juan Eugenio Iglesias [b,d], Koen Van Leemput [a,c]

[a] Department of Applied Mathematics and Computer Science, Technical University of Denmark, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, Denmark
[b] Basque Center on Cognition, Brain and Language (BCBL), Paseo Mikeletegi, 20009 San Sebastian - Donostia, Gipuzkoa, Spain
[c] Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, 149 13th St, Charlestown, MA 02129, USA
[d] Department of Medical Physics and Biomedical Engineering, University College London, Gower St, London WC1E 6BT, United Kingdom

ARTICLE INFO

Article history:
Received 10 June 2016
Accepted 5 September 2016
Available online 7 September 2016

Keywords:
MRI
Segmentation
Atlases
Parametric models
Bayesian modeling

ABSTRACT

Quantitative analysis of magnetic resonance imaging (MRI) scans of the brain requires accurate automated segmentation of anatomical structures. A desirable feature for such segmentation methods is to be robust against changes in acquisition platform and imaging protocol. In this paper we validate the performance of a segmentation algorithm designed to meet these requirements, building upon generative parametric models previously used in tissue classification. The method is tested on four different datasets acquired with different scanners, field strengths and pulse sequences, demonstrating comparable accuracy to state-of-the-art methods on T1-weighted scans while being one to two orders of magnitude faster. The proposed algorithm is also shown to be robust against small training datasets, and readily handles images with different MRI contrast as well as multi-contrast data.

1. Introduction

So-called *whole-brain segmentation* techniques aim to automatically label a multitude of cortical and subcortical regions from brain MRI scans. Recent years have seen tremendous advances in this field, enabling, for the first time, fine-grained comparisons of regional brain morphometry between large groups of subjects. Current state-of-the-art whole-brain segmentation algorithms are typically based on supervised models of image appearance in T1-weighted scans, in which the relationship between intensities and neuroanatomical labels is learned from a set of manually annotated training images.

This approach suffers from two fundamental limitations. First, segmentation performance often degrades when the algorithms are applied to T1-weighted data acquired on different scanner platforms or using different imaging sequences, due to subtle changes in the obtained image contrast (Han and Fischl, 2007; Roy et al., 2013). And second, the exclusive focus on only T1-weighted images hinders the ultimate translation of whole-brain segmentation techniques into clinical practice, where they hold great potential to support personalized treatment of patients suffering from brain diseases. This is because clinical imaging uses additional MRI contrast mechanisms to show clinically relevant information, including T2-weighted or fluid attenuated inversion recovery (FLAIR) images that are much more sensitive to certain pathologies than T1-weighted scans (e.g., white matter lesions or brain tumors). Although incorporating models of lesions into whole-brain segmentation techniques is an open problem in itself, a first necessary step towards bringing these techniques into clinical practice is to make them capable of handling the multi-contrast images that are acquired in standard clinical routine.

In this article, we present and validate the performance of a fast, sequence-independent whole-brain segmentation algorithm. The method, which is based on a mesh-based computational atlas combined with a Gaussian appearance model, yields segmentation accuracies comparable to the state of the art; automatically adapts to different MRI contrasts (even if multimodal); requires only a small amount of training data; and achieves computational times comparable to those of the fastest algorithms in the field (Zikic et al., 2014; Ta et al., 2014).

1.1. Current state of the art in whole-brain segmentation

Early methods for the segmentation of brain structures often relied on *parametric* models, in which the available training data were summarized in relevant statistics that were subsequently used to inform the segmentation of previously unseen subjects.

* Corresponding author.
E-mail address: oupu@dtu.dk (O. Puonti).

http://dx.doi.org/10.1016/j.neuroimage.2016.09.011

Because many distinct brain structures have similar intensity characteristics in MRI, these methods were typically built around detailed probabilistic models of the expected shape and relative positioning of different brain regions, using surface-based (Kelemen et al., 1998; Pizer et al., 2003; Patenaude et al., 2011; Cootes et al., 1998) or volumetric (Fischl et al., 2002; Pohl et al., 2006b) models. These anatomical models were then combined with supervised models of appearance to encode the typical intensity characteristics of the relevant structures in the training data, often using Gaussian models for either the intensity of individual voxels (Fischl et al., 2002; Pohl et al., 2006b) or for entire regional intensity profiles (Kelemen et al., 1998; Pizer et al., 2003; Patenaude et al., 2011; Cootes et al., 1998). The segmentation problem was then formulated in a Bayesian setting, in which segmentations were sought that satisfy both the shape and appearance constraints.

More recently, *non-parametric* methods[1] have gained increasing attention in the field of whole-brain segmentation, mostly in the form of multi-atlas label fusion (Rohfling et al., 2004a; Heckemann et al., 2006; Isgum et al., 2009; Artaechevarria et al., 2009; Sabuncu et al., 2010; Rohfling et al., 2004b; Wang et al., 2013; Manjón et al., 2011; Rousseau et al., 2011; Tong and Wolz, 2013; Wu et al., 2014; Asman and Landman, 2013; Zikic et al., 2014; Iglesias and Sabuncu, 2015). In these methods, each of the manually annotated training scans is first deformed onto the target image using an image registration algorithm. Then, the resulting deformation fields are used to warp the manual annotations, which are subsequently fused into a final consensus segmentation. Although early methods used a simple majority voting rule (Rohfling et al., 2004a; Heckemann et al., 2006), recent developments have concentrated on exploiting local intensity information to guide the atlas fusion process. This is particularly helpful in cortical areas, for which accurate inter-subject registration is challenging (Sabuncu et al., 2010; Ledig et al., 2012). Label fusion methods have been shown to yield very accurate whole-brain segmentations (Landman and Warfield, 2012), but their accuracy comes at the expense of a high computational cost as a result of the multiple non-linear registrations that are required. Efforts to alleviate this issue include a local search using entire image patches, such that much faster *linear* registrations can be used (Manjón et al., 2011; Ta et al., 2014), as well as using rich contextual features so that only a single non-linear warp is needed (Zikic et al., 2014).

1.2. Existing methods that handle changes in MRI contrast

With the exception of simple majority voting (Rohfling et al., 2004a; Heckemann et al., 2006), all the methods reviewed above use *supervised* intensity models, in the sense that they explicitly exploit the specific image contrast properties of the dataset used for training. This poses limitations on their ability to segment images that were acquired with different scanners or imaging sequences than the training scans.

A generic way of making such methods work across imaging platforms is histogram matching (also known as intensity normalization), in which the intensity profiles of new images are altered so as to resemble those of the images used for training (Nyúl et al., 2000; Roy et al., 2013). However, histogram matching can only be used when the training and target data have been acquired with the same type of MRI sequence (e.g., T1-weighted), and it does not completely cancel the negative effects that intensity mismatches have on segmentation accuracy (Roy et al., 2013).

Another approach is to have the training dataset include images that are representative of all the scanners and protocols that are expected to be encountered in practice. However, this approach quickly becomes impractical due to the large number of possible combinations of MRI hardware and acquisition parameters. The situation is exacerbated for clinical data, due to the lack of standardized protocols to acquire multi-contrast MRI data across clinical imaging centers.

In contrast synthesis (Roy et al., 2013), the original scan is not directly segmented, but rather used to generate a new scan with the desired intensity profile, which is then segmented instead. The premise of this technique is that a database of scans acquired with both the source and target contrast is available, so that the relationship between the two can be learned (Iglesias et al., 2013a; Roy et al., 2013). This approach makes it unnecessary to manually annotate additional training data for each new set-up that is considered – a considerable advantage given that a manual whole-brain segmentation often takes several days per scan (Fischl et al., 2002). However, it still requires that additional example subjects are scanned with both the source and target scanner and protocol, which is not always practical.

Finally, a more fundamental way to address the problem is to perform whole-brain segmentation in the space of intrinsic MRI tissue parameters (Fischl et al., 2004b). However, this requires the usage of specific MRI sequences for which a physical forward model is available, which are not widely implemented on MRI scanning platforms, and particularly not on clinical systems.

1.3. Contribution: validation of a fast, sequence-adaptive whole-brain segmentation algorithm

In contrast to the aforementioned approaches to whole-brain segmentation, which rely on supervised models of the specific intensity profiles seen in the training data, in this paper we validate an *unsupervised* approach that automatically learns appropriate intensity models from the images being analyzed. At the core of the method is an intensity clustering algorithm (a Gaussian mixture model) that derives its independence of specific image contrast properties by simply grouping together voxels with similar intensities. This approach is well-established for the purpose of tissue classification (aimed at extracting the white matter, gray matter and cerebrospinal fluid) where it is typically augmented with models of MRI imaging artifacts (Wells et al., 1996a; Van Leemput et al., 1999a; Ashburner and Friston, 2005) and spatial models such as probabilistic atlases (Ashburner and Friston, 1997; Van Leemput et al., 1999a; Ashburner and Friston, 2005) or Markov random fields (Van Leemput et al., 1999b; Zhang et al., 2001).

Here we validate a method for whole-brain segmentation that is rooted in this type of approach, building on prior work from our group including a proof-of-concept demonstration in whole-brain segmentation (Van Leemput, 2009), as well as the automated segmentation methods for hippocampal subfields (Iglesias et al., 2015a) and subregions of the brainstem (Iglesias et al., 2015b) that are distributed with the FreeSurfer software package (Fischl et al., 2002). The method we validate here uses a mesh-based probabilistic atlas to provide whole-brain segmentation accuracy at the level of the state of the art, both within and across scanner platforms and pulse sequences. Unlike many other techniques, the method does not need any preprocessing such as skull stripping, bias field correction or intensity normalization. Furthermore, because the method is parametric, only a single non-linear registration (of the atlas to the target image) is required, yielding a very fast overall computational footprint.

An early version of this work, with a preliminary validation, was presented in Puonti et al. (2013). The current article adds a

[1] Note that the distinction between parametric vs. non-parametric methods here only refers to the overall segmentation approach that is taken – the pair-wise registrations in non-parametric segmentation methods can still be either parametric (e.g., B-splines, Rueckert et al. (1999)) or non-parametric (e.g., Demons, Thirion (1998)).