NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 32 (2006) 1591 – 1607

Comparative assessment of statistical brain MR image segmentation algorithms and their impact on partial volume correction in PET

Habib Zaidi,* Torsten Ruest, Frederic Schoenahl, and Marie-Louise Montandon

Division of Nuclear Medicine, Geneva University Hospital, CH-1211 Geneva 4, Switzerland

Received 27 October 2005; revised 28 April 2006; accepted 10 May 2006
Available online 7 July 2006

Magnetic resonance imaging (MRI)-guided partial volume effect correction (PVC) in brain positron emission tomography (PET) is now a well-established approach to compensate the large bias in the estimate of regional radioactivity concentration, especially for small structures. The accuracy of the algorithms developed so far is, however, largely dependent on the performance of segmentation methods partitioning MRI brain data into its main classes, namely gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). A comparative evaluation of three brain MRI segmentation algorithms using simulated and clinical brain MR data was performed, and subsequently their impact on PVC in ^{18}F-FDG and ^{18}F-DOPA brain PET imaging was assessed. Two algorithms, the first is bundled in the Statistical Parametric Mapping (SPM2) package while the other is the Expectation Maximization Segmentation (EMS) algorithm, incorporate *a priori* probability images derived from MR images of a large number of subjects. The third, here referred to as the HBSA algorithm, is a histogram-based segmentation algorithm incorporating an Expectation Maximization approach to model a four-Gaussian mixture for both global and local histograms. Simulated under different combinations of noise and intensity non-uniformity, MR brain phantoms with known true volumes for the different brain classes were generated. The algorithms' performance was checked by calculating the kappa index assessing similarities with the "ground truth" as well as multiclass type I and type II errors including misclassification rates. The impact of image segmentation algorithms on PVC was then quantified using clinical data. The segmented tissues of patients' brain MRI were given as input to the region of interest (RoI)-based geometric transfer matrix (GTM) PVC algorithm, and quantitative comparisons were made. The results of digital MRI phantom studies suggest that the use of HBSA produces the best performance for WM classification. For GM classification, it is suggested to use the EMS. Segmentation performed on clinical MRI data show quite substantial differences, especially when lesions are present. For the particular case of PVC, SPM2 and EMS algorithms show very similar results and may be used interchangeably. The use of HBSA is not recommended for PVC. The partial volume corrected activities in some regions of the brain show quite large relative differences when performing paired analysis on 2 algorithms, implying a careful choice of the segmentation algorithm for GTM-based PVC.

Keywords: MRI; PET; Brain imaging; Segmentation; Partial volume effect

Introduction

During the last decade, neuroimaging has advanced elegantly in the medical and research arenas. Molecular brain imaging using positron emission tomography (PET) plays a valuable role in the assessment of cellular targets, thus providing clinicians and neuroscientists with relevant information in various pathologies and neurological disorders (Zaidi and Montandon, 2006). Nevertheless, PET is obviously not the only major non-invasive tool for the assessment of brain disease. Major new technologies, such as spiral computed tomography (CT), high-field magnetic resonance imaging (MRI), bioluminescent and fluorescent imaging, and many other technologies, have now blurred the artificial distinction that once set PET apart as a "functional" rather than "anatomic" imaging modality. Nonetheless, PET maintains an exclusive standing in the delivery of targeted therapies, but its superior picomolar sensitivity is being challenged by competing technologies.

The high contrast of MRI makes the method of choice to detect abnormalities in the brain in addition to offering the possibility of partitioning the brain into its main classes. Automated medical image segmentation is becoming an increasingly important image processing step for a number of clinical and research applications including but not limited to brain volumetry, treatment planning in radiation therapy, surgical planning, and image-guided intervention procedures. Moreover, modern molecular brain imaging using PET relies on high-resolution segmented anatomical data for anatomically guided statistical image reconstruction (Baete et al., 2004), attenuation compensation (Zaidi et al., 2003), and partial volume effect correction (PVC) (Rousset and Zaidi, 2005; Rousset et al., 1998). The performance of those algorithms depends largely on the quality of the segmentation output, and thus special attention

---

* Corresponding author. Fax: +41 22 372 7169.
E-mail address: habib.zaidi@hcuge.ch (H. Zaidi).

1053-8119/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2006.05.031

should be given to the segmentation procedure and its algorithmic implementation.

Image segmentation has been identified as the key problem of medical image analysis and remains a popular and challenging area of research. A wide variety of brain MR image segmentation techniques including a number of promising approaches were devised and are described in the literature (Clarke et al., 1995; Suri et al., 2002; Thacker et al., 2004). This includes thresholding, region growing, classifiers, clustering, edge detection, Markov random field models, artificial neural networks, deformable models, atlas-guided, and many other approaches (Boudraa and Zaidi, 2005). The algorithms devised specifically for segmenting the cortex from 3D MR images fall within two broad categories: voxel classification and deformable models (Duncan et al., 2004). Automated segmentation approaches have proven sufficiently accurate for volumetric display and may also be adequate for brain volumetry. The inclusion of bias field correction and partial volume effects in the segmentation paradigm has especially proven valuable (Viergever et al., 2001). Nevertheless, more comprehensive studies of brain anatomy and physiology will necessitate more sophisticated segmentation approaches. More recent atlas-based approaches include *a priori* models of lesion growth to allow accurate brain tissue segmentation even in pathological brains and when space-occupying lesions are present (Cuadra et al., 2004; Pollo et al., 2005). Some investigators performed comparative assessment studies of MRI segmentation approaches focusing mainly on cortex segmentation and tissue volume computation using simulated and real data (Grau et al., 2004; Cuadra et al., 2005).

Early attempts to compensate for PVE date back to the time where they were first pointed out as a serious limitation in quantitative analysis (Hoffman et al., 1979). Some correction methods require only the PET emission data. This includes pioneering research attempting some sort of PVC through the application of recovery coefficients described in the reference above as well as those derived from Kessler's formulation (Kessler et al., 1984). A distinct class of correction methods requires the definition of the various objects being imaged in addition to the characterization of the scanner's point spread function (PSF). These include anatomy-based post-reconstruction correction methods that make use of concomitant high-resolution structural information from MRI or CT (Rousset and Zaidi, 2005). It is worth emphasizing that the geometric transfer matrix-based method (Rousset et al., 1998) belonging to this class of techniques and used in this work is among the most popular techniques in the field.

For the particular application of MRI-guided PVC in brain PET, the main segmentation task is the extraction of cerebral tissue classes, namely gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). The accuracy of the PVC algorithm is highly dependent on the degree of accuracy achieved by the MRI-PET realignment and the MRI segmentation procedures. Several studies reported the impact of segmentation errors on the accuracy of PVC in both phantom and simulation studies (Meltzer et al., 1999; Strul and Bendriem, 1999; Frouin et al., 2002; Quarantelli et al., 2004; Rousset and Zaidi, 2005), however, there is a lack of detailed investigations of the effect of the MRI segmentation algorithm on PVC using clinical data where the ground truth is unknown. The objectives of this work are two-fold: firstly to assess qualitatively and quantitatively the performance of three commonly used brain MR image segmentation algorithms using simulated phantoms under controlled noise and intensity non-uniformity conditions and clinical data acquired in realistic conditions; and secondly to assess the impact of segmentation algorithms on a region-based PVC technique using clinical ^{18}F-FDG and ^{18}F-DOPA PET studies by using the output of each segmentation technique as input to the PVC algorithm. The three segmentation methods assessed in this work include the segmentation algorithm bundled in the Statistical Parametric Mapping (SPM2) package (Ashburner and Friston, 1997), the Expectation Maximization Segmentation (EMS) algorithm (Van Leemput et al., 1999b), which incorporate *a priori* probability images derived from MR images of a large number of subjects, and the classical approach for brain extraction and automatic tissue segmentation of MR images using the Expectation Maximization (EM) algorithm to model a four-Gaussian mixture for both global and local histograms (Kovacevic et al., 2002) here referred to as histogram-based segmentation algorithm (HBSA).

Materials and methods

Image segmentation algorithms

Given the wide range of MRI brain segmentation methods, the two main criteria instigating the choice made by the authors are software reliability and availability to the neuroimaging community and its applicability to MRI-guided PVC in PET. SPM is now considered among the gold standard tools and is being used worldwide both in research and clinical neuroimaging investigations. EMS is an open source code which can be freely downloaded from the web site of its authors (following registration), whereas HBSA can be obtained from its authors. The three techniques consider the three tissue classes of interest for PVC in PET, namely GM, WM, and CSF. Although the algorithms consider several clusters including scalp, eyes, background, etc., the segmentation output is limited to three classes. SPM2 was used as toolbox, whereas the two other packages required slight modifications beforehand to match our needs and image formats. Among the three algorithms assessed, only HBSA takes into account partial volume mixture voxels whereas EMS has the particularity of considering local spatial information by means of a Markov Random Field.

SPM2

The approach used in the segmentation algorithm incorporated in SPM2 (Ashburner and Friston, 1997, 2000) has some similarities with the EMS technique in the sense that both of them are based on the statistical brain atlas incorporated in SPM2 and use a two-step EM algorithm (see below). The affine transformation implemented in SPM2 is used to map the images and templates of the same modality. Basically, a maximum-likelihood (ML) clustering algorithm is used, which takes advantage of *a priori* images containing the statistical probability for each voxel to belong to GM, WM, or CSF. This information is used to initialize the algorithm and is also involved in computing the posterior probability. The algorithm then repeats a set of steps to compute the probabilities and estimate the cluster parameters. The parameters for the Gaussian mixture model, i.e. the number of voxels belonging to each cluster, the mean intensity of voxels within their respective cluster, and the variance in intensity of each cluster are estimated. Based on normally distributed voxels in the GM, WM, and CSF classes, respectively, and their currently estimated parameters, it is possible to assign to each voxel in each class a probability to belong to this class. The