Spline Based Inhomogeneity Correction for ^{11}C-PIB PET Segmentation Using Expectation Maximization

Parnesh Raniga[1,3], Pierrick Bourgeat[1], Victor Villemagne[2], Graeme O'Keefe[2], Christopher Rowe[2], and Sébastien Ourselin[1]

[1] BioMedIA Lab, e-Health Research Centre, CSIRO ICT Centre, Brisbane, Australia
[2] Department of Nuclear Medicine and Centre for PET, Austin Hospital, Melbourne, Australia
[3] School of Electrical and Information Engineering, The University of Sydney, Sydney, Australia

Abstract. With the advent of biomarkers such as ^{11}C-PIB and the increase in use of PET, automated methods are required for processing and analyzing datasets from research studies and in clinical settings. A common preprocessing step is the calculation of standardized uptake value ratio (SUVR) for inter-subject normalization. This requires segmented grey matter (GM) for VOI refinement. However ^{11}C-PIB uptake is proportional to amyloid build up leading to inhomogeneities in intensities, especially within GM. Inhomogeneities present a challenge for clustering and pattern classification based approaches to PET segmentation as proposed in current literature.

In this paper we modify a MR image segmentation technique based on expectation maximization for ^{11}C-PIB PET segmentation. A priori probability maps of the tissue types are used to initialize and enforce anatomical constraints. We developed a Bézier spline based inhomogeneity correction techniques that is embedded in the segmentation algorithm and minimizes inhomogeneity resulting in better segmentations of ^{11}C-PIB PET images. We compare our inhomogeneity with a global polynomial correction technique and validate our approach using co-registered MRI segmentations.

1 Introduction

With the recent development of Pittsburg Compound B (^{11}C-PIB) [1], a PET biomarker that binds to beta amyloid plaques (Aβ), it is now possible to observe *in-vivo* one of the major histopathological landmarks of Alzheimer's disease (AD). ^{11}C-PIB PET is now being tested worldwide in large studies aimed at tracking and better understanding the pathogenesis of AD with the hope that it could lead to early diagnosis. Processing and analyzing the large datasets generated by the studies is too resource intensive to perform manually and is inherently irreproducible prompting the need for automated techniques.

N. Ayache, S. Ourselin, A. Maeder (Eds.): MICCAI 2007, Part I, LNCS 4791, pp. 228–235, 2007.

However, the acquisition and reconstruction of PET images place many constraints and pose many challenges. PET image resolution is comparatively lower than traditional imaging modalities such as Magnetic Resonance Imaging (MRI) and Computed Tomography. The resulting partial volume effects (PVE) cause blurring of anatomical regions, loss of detail and of strong edge information. Alongside these factors the distribution of Aβ plaques varies considerably between individuals producing large inhomogeneities within the grey matter (GM) in ^{11}C-PIB PET images.

Because of the physiological and chemical differences between the different tissue types of the brain, quantitative analysis in neuroimaging relies on an accurate segmentation of brain tissues into its constituents, GM, white matter (WM) and cerebrospinal fluid (CSF). Notably for the analysis of ^{11}C-PIB PET images where AD pathologies and hence tracer uptake is confined to GM, an accurate segmentation is crucial for further analysis. Segmentation results are used to perform inter tissue comparison such as the ratio of uptake in GM to WM and for refinement of volume or regions of interest (VOI or ROI). We use the segmentations to refine cerebellar GM masks for standardized uptake value ratio (SUVR) calculations. Although MR image segmentation can be used instead, MR images are not always acquired as they represent a substantial cost burden [2].

There have been several approaches proposed in literature for segmentation of neuro-PET images. Pattern classification approaches such as those proposed in [3], [4] and [5] use tissue time activity curves (TTAC) derived from dynamic Flurodeoxyglucose (^{18}F-FDG) PET images. However, inhomogeneity in ^{11}C-PIB PET tracer uptake as compared to the relatively homogeneous ^{18}F-FDG uptake cause misclassifications and an inaccurate segmentation. Mykkänen *et al.* used simplex meshes and dual surface minimization (DSM) for the segmentation of ^{18}F-FDG and ^{11}C-Raclopride [6]. Tohka [7] conducted a study of DSM and other methods including generalized gradient vector field based active contours. A reliance on strong GM-WM boundaries which are not present in ^{11}C-PIB due to the inhomogeneities will result in inaccurate segmentation using these algorithms.

To compensate for the lack of detail and strong edge information, a priori data such as prior probability maps can be used similarly to the approach of Van Leemput *et al.*, which has been validated on MR images [8]. We adopted and extended this approach by including an intensity inhomogeneity correction step that takes into account local changes within the GM. Anatomical priors are used as initialization and constraints in the algorithm improving robustness and accuracy and are advantageous over other initialization techniques such as thresholding and clustering for ^{11}C-PIB as they are data independent.

The inhomogeneity correction is performed using bi-cubic Bézier spline patches. Using spline patches allows for constraints such as smoothness and spatial consistency to be easily incorporated into the spline model. Splines provide a general framework within which tracer and physiology models can be embedded for greater accuracy and robustness. A similar approach, using thin plate splines,

was taken by [9] for inhomogeneity correction in MR images with the assumption of low frequency global variations. Sample voxels, through which the spline was fit were determined either by an observer or through a pre-classification step. As discussed later in Section 2.2 these are not valid and are illustrated by our results (see Section 4). Validation of our approach was performed using Dice similarity coefficient (DSC) [10] for the obtained PET segmentations as compared to segmented co-registered MR images of the same patient.

2 Method

2.1 Data and Acquisition

For testing and validation of our algorithm we used a subset of scans from a study the application of ^{11}C-PIB PET for early diagnosis. The subset consists of 35 patients of which 10 were clinically diagnosed with AD, 12 with Mild Cognitive Impairment (MCI) and 13 were Normal Controls (NC). The AD group consisted of 5 males and 5 females, mean age 74±9 yrs, the MCI group consisted of 6 males and 6 females, mean age 72±10 yrs and the NC group consisted of 5 males and 8 females, mean age 72±6 yrs. Each patient underwent a 20 min PET scan, 40 minutes post injection and an SPGR T1-weighted MR scan. All PET scans were 128 by 128 voxels and 90 slices.

2.2 Tissue Classification

In van Leemput's tissue classification algorithm for MR images [8], voxel's intensities for the different tissue classes are modeled as Gaussian distributions in a Gaussian finite mixture model. Parameters of the model are estimated from the data and a priori probability maps in an iterative manner within a maximum likelihood - expectation maximization (ML-EM) framework. A priori probability maps for GM, WM and CSF, which are registered to the image, are used as initialization and as anatomical constraints. A Markov random field (MRF) is included within the algorithm to ensure consistent results as voxels in the neighborhood of a particular voxel contribute to its classification.

A bias field correction scheme is used in the algorithm to correct for the intensity inhomogeneities that are present in MR images due to fluctuations in magnetic fields of scanners. These inhomogeneities are modeled as smoothly varying polynomials the coefficients of which are calculated by fitting intra-class residuals using least squares. When fourth order polynomials are used, only slow, global intensity changes can be modeled and corrected. The use of higher order polynomials would result in extreme computation costs. Correction of local, high intensity changes such as those found in GM of ^{11}C-PIB requires higher degrees of freedom which are accommodated by piecewise spline functions. As MR bias fields affect intensities in all the tissue types equally, coefficients for the polynomials are calculated using information from all the tissue types, however ^{11}C-PIB inhomogeneities are confined only to GM, therefore at each iteration our spline based approach only considers voxels with a high probability of being GM.

2.3 Inhomogeneity Correction

Considering both local and global changes, we modified van Leemput's algorithm by replacing the bias field correction step with an inhomogeneity correction step based on bi-cubic Bézier patches. A set of patches are tiled on each slice of the image and serve to subdivide the image and to interpolate correction coefficients, it should be noted that the patches may be set arbitrarily and this setting simplifies implementation. This allows us to approximate variability to a much higher degree than the polynomial based correction (PBC) used for MR image bias field correction.

A bi-cubic Bézier patch consists of 16 control points in three-dimensional space forming a surface that intersects each of the points. Intermediate points lying on the surface are calculated by interpolation. We use the patches to divide each slice of the image into regions and the third dimension of the control points is used to interpolate correction coefficients. It should be noted that although we use spline surfaces to decrease computation, this is still a fully three-dimensional approach as we use information from neighbors in three dimensions in calculating correction coefficients. Control points of each of the patches is uniformly spaced within the patch, with neighboring patches sharing control points along their common edges thus preserving continuity and smoothness.

After each slice has been subdivided into patches, the following steps are taken at each iteration :

- Mean intensity ($\mu \mathrm{P}_{i,j,k}$) of voxels that have high probability of belonging to GM are calculated for each patch.
- The correction coefficient (CC) of each patch is set to the difference between the global GM mean (μG) and the GM mean of the patch($\mu \mathrm{P}_{i,j,k}$) that is to (μG - $\mu \mathrm{P}_{i,j,k}$).
- Intensity correction control points of each patch are set to weighted average of the CCs in 3×3×3 neighborhood, with patches geometrically closer having a linearly higher weighting.
- For each voxel that has high probability of belonging to GM, a voxel correction coefficient (VCC) is calculated by interpolating the control points of the corresponding patch and added to the voxels intensity.

These steps have the effect of normalizing means of local regions and since we restrict the correction to GM only, it normalizes GM means within local regions reducing the inhomogeneities present in ^{11}C-PIB images leading to better segmentations.

2.4 Experiments and Validation

Validation was performed using the DSC [10] between the PET Segmentation and the corresponding MR Segmentation as used in [11]. The PET segmentations considered were the spline based correction (SBC), the polynomial based correction (PBC) and the PBC constrained to only GM (PBC-GM). The effects of parameters such as the number of spline patches per slice were also tested

with segmentations conducted using grids of 8×8, 16×16 and 32×32 patches per slice. For registration of the a priori probability maps, we used a robust, multi-modal block-matching algorithm [12].

3 Results

Results of the experiments and validation are presented below. Table 1 presents mean DSC for GM PET segmentations as compared to co-registered MR approaches using SBC with grids of 8 by 8, 16 by 16 and 32 by 32 patches per slice. Mean DSCs for the PBC and GM-PBC for GM segmentations are also presented. Table 2 presents the DSCs for WM segmentation. The results show that using grids with fewer patches results in lower mean GM DSCs and slightly higher mean WM DSCs. Compared to PBC and GM-PBC, mean GM DSCs for SBC were higher, significantly so for a grid of 32 by 32 patches per slice. WM mean DSC were comparable between the approaches but those using a 32 by 32 grid were lower. SBC segmentations are also more consistent with similar DSCs for AD, MCI and NC patient types.

Table 1. Mean DSC ± standard deviation of the ^{11}C-PIB PET GM segmentations with regards to the MR segmentation for SBC using 8×8, 16×16 and 32×32 grids and PBC and GM-PBC

Classification	**SBC 8×8**	**SBC 16×16**	**SBC 32×32**	**PBC**	**GM-PBC**
AD	0.55±0.05	0.58±0.05	0.6±0.05	0.53±0.05	0.52±0.04
MCI	0.57±0.05	0.60±0.05	0.62±0.05	0.56±0.05	0.56±0.05
NC	0.57±0.04	0.59±0.05	0.62±0.04	0.56±0.04	0.55±0.04

Table 2. Mean DSC ± standard deviation of the ^{11}C-PIB PET WM segmentations with regards to the MR segmentation for SBC using 8×8, 16×16 and 32×32 grids and PBC and GM-PBC

Classification	**SBC 8×8**	**SBC 16×16**	**SBC 32×32**	**PBC**	**GM-PBC**
AD	0.66±0.02	0.65±0.02	0.63±0.03	0.65±0.02	0.66±0.01
MCI	0.67±0.03	0.66±0.03	0.63±0.05	0.67±0.04	0.67±0.03
NC	0.66±0.03	0.65±0.03	0.63±0.03	0.66±0.03	0.67±0.01

Figures 1 and 2 presents axial slices from MR and ^{11}C-PIB PET scans along with segmented GM and WM masks of AD and NC patients using SBC with a 32 by 32 grid. Figure 3 presents MR and ^{11}C-PIB PET segmentations using SBC, PBC and GM-PBC. Notice the over segmentation of WM and under segmentation of GM using PBC and GM-PBC. Also, notice the similarity between PBC and GM-PBC. Segmentations were also performed without correction but as both the quantitative (DSC around 0.3 for GM and 0.5 for WM) and qualitative results significantly lower, they have been excluded for brevity.

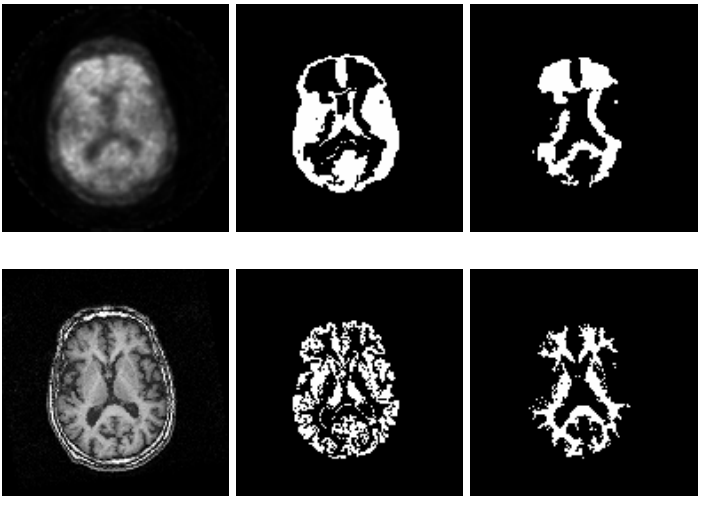

Fig. 1. Transaxial slices of ^{11}C-PIB PET (top row) and MR (bottom row) scans with original scan (column one) GM (column two) and WM (column three) segmentations of an AD patient

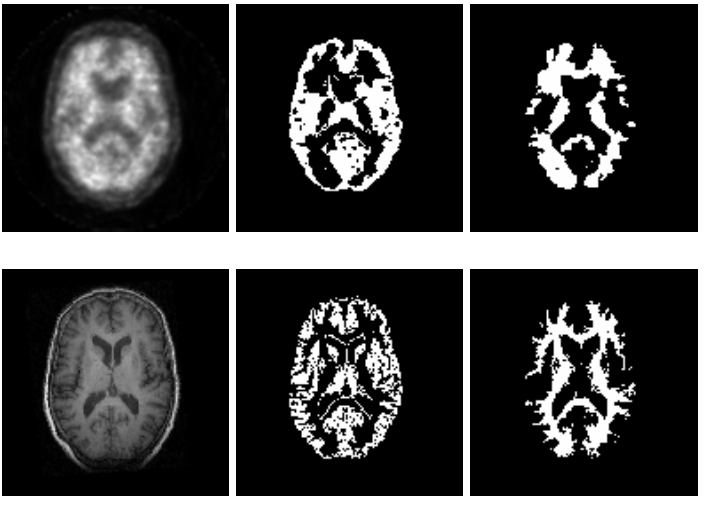

Fig. 2. Transaxial slices of ^{11}C-PIB PET (top row) and MR (bottom row) scans with original scan (column one), GM (column two) and WM (column three) segmentations of a NC patient

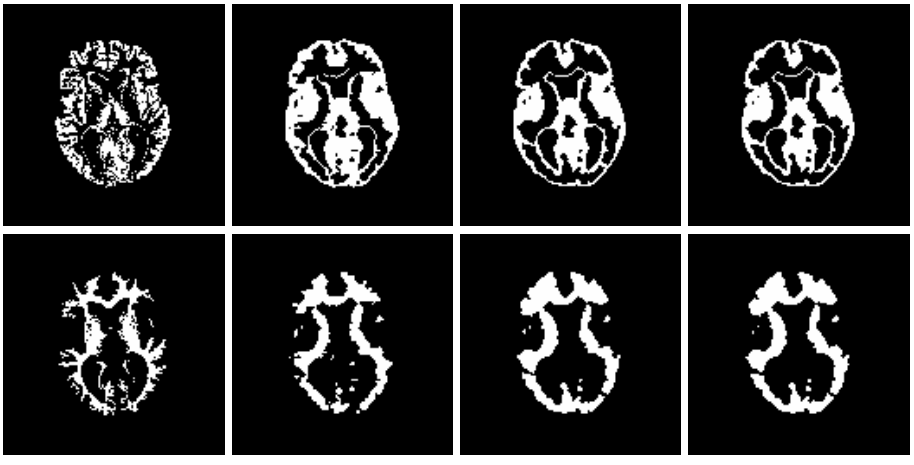

Fig. 3. Transaxial slices of ^{11}C-PIB PET and MR scans of an AD patient, with GM (top row) and WM (bottom row) segmentations. Column one are the MR segmentations, column two are PET segmentations using SBC and a 32 by 32 grid, column three are PET segmentations using PBC and the fourth column are PET segmentations using GM-PBC.

4 Discussion

GM segmentations produced by SBC had higher DSCs than those produced by PBC and GM-PBC. However, the DSCs for WM segmentations were lower. On inspection of the data, we noticed that the PBC and GM-PBC approaches tended to under segment GM and over segment WM as illustrated in Figure 3. This was especially true of AD patients with plaque build up in the frontal lobes where GM structures with medium-high uptake were segmented as WM. The PBC approaches also have a tendency to converge to the a priori probability maps, as there is too much variability for correct classification.

The need for a more local approach is illustrated by the results of SBC using grids of 8 by 8 and 16 by 16 patches per slice. DSCs for GM segmentations increased as the grid sizes were increased and those of WM segmentations decreased. Inspection of the segmentations revealed that using fewer patches per slice resulted in over segmentation of WM and under segmentation of GM as was with PBC and GM-PBC. Since grids with more patches have more degrees of freedom, they are able to better model the local changes.

To further improve the results of the segmentation, models of tracer uptake based on patient type and/or neuropsychological tests can be integrated into the inhomogeneity correction step. Better a priori data and template will not only improve registration of the PET images to the templates but also the segmentation itself. The technique presented could be applied to other tracers where there is inhomogeneity in uptake in either the GM or WM.

5 Conclusion

We have presented a novel approach to inhomogeneity correction and segmentation of ^{11}C-PIB images and validated it against co-registered, segmented MR images of the same patient and compared it against a global polynomial based correction approach. Our approach produced good GM segmentations with similar DSCs irrespective of patient classifications. The polynomial based approach tended to under segment the WM and over segmenting GM and were sensitive to patient types thus sensitive to inhomogeneity. Using different number of patches per slice, we have illustrated the need for techniques that correct on a local scale, which our approach does.

Since we are using an iterative EM approach and splines for inhomogeneity correction, we can embed models for different patients types and will be exploring this in the future. We are also investigating the use of a priori data and templates that are suited to the different patient types.

References

1. Klunk, W.E., Engler, H., et al.: Imaging brain amyloid in Alzheimer's disease with Pittsburgh compound-B. Ann. Neurol. 55(3), 306–319 (2004)
2. Raniga, P., Bourgeat, P., et al.: PIB-PET segmentation for automatic SUVR normalization without MR information. In: ISBI, pp. 348–351. IEEE, Washington DC (2007)
3. Wong, K.P., Feng, D., et al.: Segmentation of dynamic PET images using cluster analysis. IEEE Trans. Nucl. Sci. 49, 200–207 (2002)
4. Brankov, J.G., Galatsanos, N.P., et al.: Segmentation of dynamic PET or fMRI images based on a similarity metric. IEEE Trans. Nucl Sci. 50, 1410–1414 (2003)
5. Chen, J.L., Gunn, S.R. et al.: Markov random field models for segmentation of PET images. In: IPMI, Davis, CA, USA, June 18-22, 2001 pp. 468–474 (2001)
6. Mykkänen, J., Tohka, J., et al.: Automatic extraction of brain surface and midsagittal plane from PET images applying deformable models. Comput Methods Programs Biomed 79(1), 1–17 (2005)
7. Tohka, J.: Surface extraction from volumetric images using deformable meshes: A comparative study. In: ECCV, pp. 350–364 (2002)
8. Van Leemput, K., Maes, F., et al.: Automated model-based tissue classification of MR images of the brain. IEEE Trans. Med. Imaging. 18(10), 897–908 (1999)
9. Dawant, B.M., Zijdenbos, A.P., Margolin, R.A.: Correction of intensity variations in MR images for computer-aided tissue classification. IEEE Trans. Med. Imaging. 12, 770–781 (1993)
10. Dice, L.: Measures of the amount of ecologic association between species. Ecology 26, 297–302 (1945)
11. Kim, J., Cai, W., Feng, D., Eberl, S.: Segmentation of VOI from multidimensional dynamic PET images by integrating spatial and temporal features. IEEE Trans Inf. Technol. Biomed. 10, 637–646 (2006)
12. Ourselin, S., Roche, A., et al.: Reconstructing a 3D structure from serial histological sections. IVC 19, 25–31 (2001)