Tissue Classification for PET/MRI Attenuation Correction Using Conditional Random Field and Image Fusion

Zhengyi Yang, Jeiran Choupan, Farshid Sepehrband, David Reutens, and Stuart Crozier

Abstract—Photon attenuation correction is a challenging task in the emerging hybrid PET/MRI medical imaging techniques because of the missing link between tissue attenuation coefficient and MRI signal. MRI-based tissue classification methods for attenuation correction have difficulties caused by the significantly different abilities of photon absorption in tissues with similar MRI signal, such as bone and air. We proposed a novel method of integrating the information from MRI and PET emission data to increase the tissue classification accuracy. A classifier based on conditional random field was trained using features extracted from fused MRI and uncorrected PET images. The efficacy of the proposed method was validated quantitatively on synthetic datasets. It was found that the inclusion of PET data improved the classifier’s performance in terms of classification accuracy and PET image reconstruction quality.

Index Terms—Attenuation correction, conditional random field, tissue classification, PET/MRI.

I. INTRODUCTION

Hybrid PET/MRI systems, which combine the capabilities of positron emission tomography (PET) to obtain metabolic information with high sensitivity, and magnetic resonance imaging (MRI) for structural and functional imaging without the additional radiation exposure associated with X-ray computerized tomography (CT), are currently under extensive research and development [1]. Among other technical hurdles, MRI-based correction for photon attenuation is the major issue unsolved in the development of quantitative PET/MRI [2]. Inadequate attenuation correction has serious implications, such as inaccurate cancer staging or failure to detect tumours [3].

In PET, tissue concentration of the injected radiotracer is estimated from the detection of photons emitted by the annihilation of positrons from the radiotracer and electrons in adjacent tissue. Some photons are absorbed by tissue prior to detection and this attenuation effect has to be corrected to generate PET images during image reconstruction. The degree of photon attenuation can range from 50% to 95% and is greater in larger patients [4]. Without attenuation correction, the regional tissue radiotracer concentration may be underestimated, leading to lesions becoming invisible on PET images [3].

Attenuation correction relies on the calculation of a voxelized attenuation coefficient map (μmap) in PET image domain representing the tissue-specific photon attenuation coefficient of each voxel, based on which attenuation correction factor map (ACFmap) in PET emission data (i.e. sinogram) domain is computed. ACFmap and μmap can be obtained using CT in PET/CT scanners and transmission PET scans in PET alone systems, respectively. Nevertheless, patients suffer from exposure to ionizing radiation and additional scan time. These drawbacks can be overcome using the simultaneously acquired MR image for attenuation correction on hybrid PET/MRI systems.

MRI-based correction, however, is much more challenging than with CT because the attenuation coefficient and MR signal intensity are not directly linked as is the case with CT. Photon attenuation is mainly caused by tissues with high concentrations of elements of high atomic number, such as calcium in cortical bone, which are not imaged well by commonly used MR sequences. For example, as illustrated in Fig. 1, both bone and air have low intensity on T1 weighted images but their attenuation coefficients differ greatly, which has been a main limitation in the ability to correct PET emission data using MR images.

Current MRI-based attenuation correction methods can be broadly classified into two categories: template-based [5] and tissue classification-based [6] techniques, as reviewed in [7]. Template-based methods rely on registration between patient images to a population-based atlas and may not account satisfactorily for individual differences in normal anatomy, such as bone defects and surgical resection, and variations due to lesions. In tissue classification-based methods, image voxels are allocated to classes with distinct attenuation coefficients (mainly bone, soft tissue, sinuses, lung and air), which are assigned the corresponding attenuation coefficient to create the μmap. The classification accuracy significantly determines the quality of attenuation correction and PET image reconstruction. The elimination of bony structures from the μmap will cause a local error over 3% and an underestimation of 11% in Standard Uptake Value (SUV) in clinical PET studies [8].

Fig. 1. Bone and air are indistinguishable in sinuses regions on T1-weighted

(a) T1-weighted MR image
(b) ‘Ground truth’ classification

T1 fused with PET
Existing methods can result in PET quantification errors of 10% for head imaging [5], [9] and 19% for torso imaging [10]. Recently, tissue classification based on ultrashort-echo-time (UTE) MRI sequences was proposed for attenuation correction [6]. An overall voxel classification accuracy (compared with CT) of 81.1% was achieved but bone segmentation was inaccurate in regions of bone/air interface such as the paranasal sinuses. Although UTE sequence provides better differentiation between bone and air, it has safety issues of high radio-frequency power to be delivered into patient body. Therefore, routinely acquired MR images in clinical settings, such as T1-weighted MRI, is preferable for attenuation correction.

Sinogram before correction contains some information that can aid the \( \mu \)-map creation and Helgason-Ludwig Consistency Conditions (HLCC) was formulated to extract the information [11]. However, HLCC alone is not sufficient and sometimes not valid due to sparse data sampling.

In this study, we proposed a supervised machine learning framework to integrate the information about tissue classification from conventional structural MRI and PET sinogram for attenuation correction. A graphic model based on conditional random field (CRF) was trained and features extracted from the fused T1 MR images and partially corrected PET images used. The efficacy of the method was quantitatively validated on synthetic datasets using the accuracies of tissue classification, and reconstructed PET image as metrics, compared with ‘ground truth’ and results of using T1 images alone.

![Fig. 2. Flow chart of synthetic dataset creation. GT: ‘ground truth’.](image)

### II. METHOD

#### A. ExperimentalDatasets

As illustrated in Fig. 2, we created synthetic PET images and sinograms with and without attenuation correction based on the digital brain phantoms of 20 subjects downloaded from BrainWeb [12]. Each phantom consists of a T1-weighted MRI structural image and a priori tissue classification. T1 images were simulated using spoiled FLASH sequence with TR=22ms, TE=9.2ms, flip angle=30 degree and 1 mm isotropic voxel size. Each voxel was labeled as one of the 11 classes (background, cerebrospinal fluid, gray matter, white matter, fat, muscle, muscle/skin, skull, vessels, around fat, dura matter, and bone marrow). Two dimensional slices with 256x256 pixels were extracted from the T1 and label volumes and used as the raw datasets in this paper.

PET image contrast is generated by the difference in radiotracer uptake between different tissues. We merged the tissue labels into 4 groups: gray matter (GM), white matter (WM), other tissue, and air and set the uptake ratio to 40:10:2.0 to create the ‘ground truth’ radiotracer uptake images, i.e., the synthetic PET images. Denote the radiotracer uptake in a given pixel \((x,y)\) as \(\lambda(x,y)\). For a given line-of-response (LOR), which is defined by the line linking two crystal photon detectors of the PET scanner, the number of photon pairs detected \(i.e.\) emission data \(E\) by these two detectors can be calculated as line integral using forward projection,

\[
E_{GT}(s, \phi) = \int_{LOR(s, \phi)} \lambda(x,y)dt
\]

where, \(t\) is the coordinate along the LOR, \(s\) and \(\phi\) are the radial and angular variables used to parameterize the LOR, respectively. In our study, LOR was parameterized by \(\phi = \{0,1,2,...,179\}\) and \(s = \{-183,-182,...,183\}\) . The emission data for all the LORs were calculated to construct the 2-dimensional ‘ground truth’ sinogram as a \(180\times367\) matrix.

Three classes of tissues with distinct attenuation coefficients were used: air, soft tissue and bone to define the synthetic \(\mu\) maps. We generated the ‘ground truth’ classification by keeping the classes of background (air) and skull (bone) and pooling all the rest as soft tissue. The ‘ground truth’ \(\mu\) map was obtained by assigning uniform attenuation coefficient for each of the three classes: 0/cm for air, 0.009/cm for soft tissue and 0.015/cm for bone.

The ACF for each LOR was calculated as the exponential of forward projection of the \(\mu\) map,

\[
ACF(s, \phi) = \exp\left(\int_{LOR(s, \phi)} \mu(x,y)dt\right).
\]

The ACFs for all LORs formed the ACFmap, which was in sinogram domain. Then the uncorrected sinogram (sinogram without attenuation correction) was the entrywise division between the ‘ground truth’ sinogram and the ACFmap:

\[
E_{UC}(s, \phi) = E_{GT}(s, \phi) / ACF(s, \phi)
\]

PET image was then reconstructed from sinogram using filtered backprojection (FBP). In our machine learning algorithm, the uncorrected sinogram and the T1 MR image created from the 20 subjects were the samples, and the corresponding \(\mu\) maps were the labels.

#### B. Image Fusion

We observed that the PET image reconstructed from uncorrected sinogram, \(i.e.\), the uncorrected PET image, contains tissue classification information and hypothesized that a classifier working on fused T1 image and uncorrected PET image would give better classification result than that based on T1 image or uncorrected PET image alone. Dual-Tree Complex Wavelet Transform (DT-CWT) based
image fusion algorithm [13] was employed because of its good performance in both spatial and spectral domains achieved by resolving the shift variance and lack of directionality issues in real valued wavelet transforms.

C. Conditional Random Field

A probabilistic graphical model based on CRF theory was used as the classifier in this study [14]. Given the training set $X$, and classification label $y$, CRF models the conditional distributions $p(y | X)$ directly, instead of the joint distribution $p(y, X)$ as in hidden Markov models. Therefore, the priors $p(X)$ does not need to be modelled. Recent successes of CRF-based classification have been in the segmentation of natural scenes, text, brain MR images for tumour detection and optical coherence tomography. However, it has not been attempted in the context of generating attenuation maps for PET/MRI scanners.

The CRF is defined as a unidirectional graph $G \in \{V, E\}$, in which each node $v_i \in V$ is associated with a random variable $y_i$ over the tissue type space $L = \{l_1, ..., l_n\}$ describing the tissue type of the $i^{th}$ voxel. In this study, $L = \{\text{bone, soft tissue, air}\}$. A classification $y$ is the vector $y = (y_1, ..., y_m)$ for the $m$ voxels in a given image $X$. When conditioned on $X$, the random variable $y_i$ is assumed to be Markovian $G : p(y | X, y_j, i \neq j) = p(y | X, y_j, i \sim j)$, where $i \sim j$ indicates $i$ and $j$ are neighbors in $G$, which means each variable is independent on the entire graph when conditioned on its neighbors. Then $(X, y)$ is a CRF [14] and the conditional probability function for $y$ is

$$p(y | X; \theta) = \frac{\exp(\Psi(y, X; \theta))}{Z(X, \theta)}, \quad (4)$$

where $\theta$ is the model parameter vector and $Z(X, \theta)$ is a normalization factor. The potential function $\Psi$ is a summation of potential terms depending on clique variables that are defined according to the neighborhood. Each term has the form $\sum_C \phi_C (y_C, X, \theta)$, where $C$ is the set of cliques and $y_C = (y_i : i \in c)$. Cliques could be unary or pairwise, resulting in unary and pairwise potential terms. Instead of the popular Maximum a Posteriors (MAP) algorithm, we employed themarginalization-based parameter learning approach, which was found outperformed likelihood-based algorithms, to train the CRF models and the Tree-reweighted Belief Propagation method to make the inferences [15].

D. Feature Selection

The feature set used in classifiers plays a critical role and finding the relevant features to the learning task is often too expensive to explicitly enumerate and compare all the candidate feature subsets. Information criterion (e.g., Akaike’s Information Criterion and its variants) can be adapted to feature selection through a ‘wrapper’ algorithm. However, ‘wrappers’ algorithm suffers from high computational complexity and poor generalization to other classifiers. To reduce the computational time, we used a filter feature selection method based on mutual information using minimal-redundancy-maximal-relevance (mRMR) [16].

The candidate features are constant, pixel location, pixel intensity, n-neighborhood pixel intensities ($n=4,8,26$), histogram of oriented gradients (HoG) descriptors (number of bins=$4,8,16,32$), local binary patterns (LBP), differential threshold, Sobel edge, Canny edge, Prewitt edge, Roberts edge, Scharr feature, and Laplacian of Gaussian.

E. CRF Models Using Different Training Sets

To test the hypothesis that uncorrected PET image can be helpful in tissue classification, we trained a series of CRF models using T1 image alone, ‘ground truth’ PET image alone, uncorrected PET image alone, T1 image fused with ‘ground truth’ PET image, and T1 image fused with uncorrected PET image as training sets, respectively. Correspondingly, the trained models are called CRF-T, CRF-P, CRF-P uc, CRF-TP, and CRF-TP uc, respectively. To evaluate the classification performance of these models for unseen data, “Leave-One-Out Cross Validation” (LOOCV) was used. For each LOOCV run, the cross validation error was defined as the percentage of misclassified pixels. The mean and standard deviation of this percentage over the 20 runs were used to evaluate the model performance. Given the ‘ground truth’ classification was known, the Dice Similarity Coefficient (DSC) for each of the 3 tissue classes (bone, soft tissue and air) was computed for comparison as well,

$$DSC(A, B) = \frac{2 \text{Num}(A \cap B)}{\text{Num}(A) + \text{Num}(B)}, \quad (5)$$

where, $\text{Num}(\cdot)$ is the function to count the pixels in an image, and $A \cap B$ is the intersection of $A$ and $B$, which are ‘ground truth’ classification and the predicted classification, respectively.

To evaluate the effect of different tissue classification on the PET image reconstruction quality, Relative Error (RE) is defined as follow,

$$(\text{PET}_{\text{ac}} - \text{PET}) / \text{PET} \times 100\%,$$

where, $\text{PET}_{\text{ac}}$ is the corrected PET image and PET is the ‘ground truth’ PET image. The mean and standard deviation (over the 20 runs) of RE and the absolute value of RE are calculated for each model.

All the calculations in this study were carried out in MatLab®. The CRF toolbox in [15] was adapted to train the models and make inferences. The mRMR toolbox in [16] was tailored for feature selection.

III. RESULTS AND DISCUSSION

We targeted on the lower skull region because that the complicated bone/sinuses/air interfaces cause problems in existing tissue classification methods. As illustrated in Fig. 1, we chose the 11th slice to createthe training and test datasets. Examples of reconstructed PET images and fused images are given in Fig. 3. It is obvious that the corrected PET image is uniform; while the uncorrected PET image has intensity inhomogeneity issue and the central part appears to be darker because of the photon attenuation effect. The bone/air interface is much clearer in uncorrected PET than in corrected PET. However, the bone and soft tissue boundaries
are weak in these PET images. The boundary between soft tissue and background air is poorly defined too. In the fused images, however, both the bone/air and bone/soft tissue contrasts are preserved to some extent.

Feature selection based on mRMR criterion highlighted these features: constant, pixel intensity, 26-neighborhood pixel intensities, HoG(16), HoG(32), LBP, differential threshold, and Scharr feature, which were used to train the models.

The LOOCV results for a representative subject (Subject 1) using the five CRF models are shown in Fig. 4. It can be observed that T1-based classification generated very poor result, while PET-based classification performed well in distinguishing bone and soft tissue but not in bone and air. The uncorrected PET based classifier achieved good accuracy in air/bone segmentation in nasal sinuses region but performed poorly in soft tissue/air interface around the nose. The results from the fused images showed a combinational effect. The classification results with the best and the worst accuracy obtained using each model are illustrated in Fig. 5.

The LOOCV error for each subject was plotted in Fig. 6. It can be seen that the classifier based on fused images outperformed that solely based on T1 or uncorrected PET images in all cases. The quantitative comparison between the trained models is shown in Table I. The classification based on ‘ground truth’ PET (CRF-P), which is not available when performing attenuation correction in practical settings, has the least mean and standard deviation of misclassification error. The classifier based on T1 fused with uncorrected PET (CRF-TPuc) achieved an accuracy very close to that based on ‘ground truth’ PET, and better than those based on T1 or uncorrected PET only.
TABLE I: QUANTITATIVE EVALUATION OF CRF MODELS FOR TISSUE CLASSIFICATION

<table>
<thead>
<tr>
<th>CRF-</th>
<th>LOOCV Error (%) Mean</th>
<th>LOOCV Error (%) Std.</th>
<th>Mean DSC (%) Bone</th>
<th>Mean DSC (%) Soft Tissue</th>
<th>Mean DSC (%) Air</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>7.05</td>
<td>1.11</td>
<td>12.57</td>
<td>90.08</td>
<td>97.65</td>
</tr>
<tr>
<td>Puc</td>
<td>4.39</td>
<td>1.09</td>
<td>72.81</td>
<td>93.48</td>
<td>98.19</td>
</tr>
<tr>
<td>TPuc</td>
<td>3.03</td>
<td>0.77</td>
<td>70.39</td>
<td>96.32</td>
<td>98.95</td>
</tr>
<tr>
<td>TP</td>
<td>3.13</td>
<td>0.83</td>
<td>70.16</td>
<td>96.19</td>
<td>98.91</td>
</tr>
<tr>
<td>P</td>
<td>2.83</td>
<td>0.59</td>
<td>72.90</td>
<td>96.65</td>
<td>98.94</td>
</tr>
</tbody>
</table>

There is no statistically significant correlation in LOOCV error between T1 based and PET based classifications, which can be argued to justify that the complementary information about attenuation contained in T1 and PET images should be combined to achieve better classification.

TABLE II: EVALUATION ON PET RECONSTRUCTION QUALITY. RE: RELATIVE ERROR; ARE: ABSOLUTE VALUE OF RELATIVE ERROR.

<table>
<thead>
<tr>
<th>Model CRF-</th>
<th>RE (%)</th>
<th>ARE (%)</th>
<th>RE (%)</th>
<th>ARE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>7.6</td>
<td>2.6</td>
<td>10.7</td>
<td>3.9</td>
</tr>
<tr>
<td>Puc</td>
<td>7.0</td>
<td>2.3</td>
<td>10.3</td>
<td>3.6</td>
</tr>
<tr>
<td>TPuc</td>
<td>6.5</td>
<td>1.9</td>
<td>9.9</td>
<td>2.3</td>
</tr>
<tr>
<td>TP</td>
<td>5.7</td>
<td>1.8</td>
<td>10.0</td>
<td>2.4</td>
</tr>
<tr>
<td>P</td>
<td>6.8</td>
<td>1.8</td>
<td>9.6</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Example reconstructed PET images using different tissue classification for a representative dataset (Subject 15) is shown in Fig. 7. The average (over 20 LOOCV runs) RE and absolute RE for each model are shown in Table II. It can be seen that the image quality of PET reconstructed with fused image based attenuation correction is better than that using a single modality based correction in terms of RE (6.5% v.s. 7.6% and 7.0%). As RE could be negative or positive reflecting an underestimation or overestimation in radiotracer uptake, the mean and standard deviation of the absolute value of RE (ARE) is a better evaluation criterion. Again, the inclusion of uncorrected PET image for classification yields better PET image reconstruction (9.9% for CRF-TPuc against 10.7% for CRF-T and 10.3% for CRF-Puc).

Fig. 7. Reconstructed PET images with ACFmaps created using different CRF models for attenuation correction.

In this study, the image reconstruction method FBP was chosen for its simplicity. More sophisticated reconstructed methods could be employed to improve the PET image quality, especially in the case of noisy data. Poisson noise was not introduced to the synthetic datasets because we wanted to elucidate the helpfulness of uncorrected PET in tissue classification. To test the efficacy of the proposed method on clinical dataset is our future work. The effect of different image fusion algorithms for noisy images on the classification accuracy is to be investigated.

IV. CONCLUSION

In this study, we demonstrated on synthetic data that PET image without attenuation correction could be useful for attenuation correction. To combine the information, we trained conditional random field based classifiers on fused images of T1 MRI and uncorrected PET. The classification results on the fused images were found to be better than that on T1 or uncorrected PET image alone.

REFERENCES


Zhengyi Yang is currently a research fellow working in School of Information Technology and Electrical Engineering at The University of Queensland, Australia. He holds degrees of PhD from The University of Hong Kong, Master of Science and Bachelor of Engineering from Sichuan University, China. His current research focuses on medical image processing and analysis.

Jeiran Choupan is currently a PhD candidate with Centre for Advanced Imaging at The University of Queensland, Australia. She did her Bachelor of Science and Engineering (2008) and Master of Science and Engineering in Information Technology at Sharif University of Technology, Iran (2010). Her current research focus is on brain decoding using functional MRI. Her area of expertise is on signal processing and machine learning methods.

Farshid Sepehrband is currently a PhD candidate with Centre for Advanced Imaging at The University of Queensland, Australia. After completing his undergraduate studies in hardware computer science and engineering at Tehran central branch, Azad University in 2008 and graduate studies in digital electronic at Sharif University of Technology, Iran in 2010, he started his PhD in the University of Queensland in December 2011. His current research focuses are on diffusion Magnetic Resonance Imaging, optimum acquisition design and micro-structural imaging.

Professor David Reutens was appointed as the inaugural director of the CAI in 2008. Prior to that he was the Professor of Neuroscience at Monash University and Director of Neurology at Southern Health. He is also a clinical neurologist specializing in epilepsy and is a senior staff specialist at the Royal Brisbane and Women’s Hospital. He directs the Australian Mouse Brain Mapping Consortium. Professor Reutens’ research involves the combination of imaging techniques such as MRI, EEG, transcranial magnetic stimulation and PET to study functions such as memory in the healthy brain, the mechanisms behind diseases such as epilepsy and stroke and how the brain responds to overcome injury. Some of the research studies involve patients with injuries or genetic mutations that affect brain function – these studies are able to provide privileged insights into how the brain works. With his collaborators, he uses imaging to study animal models of disease. Members of his group are working on ways of improving imaging technology such as the development of new diagnostic agents and new mapping methods that may improve the safety of brain surgery.

Stuart Crozier is currently Professor and Director of Biomedical Engineering at The University of Queensland. He is an ATSE Fellow and a Fellow of The Institute of Physics (UK). He holds a PhD and higher Doctorate (D.Eng.) in Biomedical Engineering. He has published over 190 journal papers and holds 30 patents in the field of medical imaging. He has supervised more than 20 PhD students to graduation in the field. In 2012 he was awarded the Clunies Ross medal for research with a societal benefit. His main contributions have been to the development of applications and engineering innovation in Magnetic Resonance Imaging (MRI). Several of his innovations have been adopted by industry. He is an Associate Editor of IEEE Transactions on Biomedical Engineering and also a Deputy Editor of Magnetic Resonance in Medicine.