Comparison of Atlas-based Bone Segmentation Methods in Whole-body PET/MRI

Hossein Arabi and Habib Zaidi†, Senior Member, IEEE

Abstract– The aim of this work is to compare bone extraction accuracy of commonly used atlas-based segmentation algorithms for generation of pseudo-CT images in whole-body PET/MRI. Eight different bone segmentation methods from whole-body MR images were implemented. The study comprised 23 patients who underwent sequential PET/CT and PET/MRI scans, thus enabling to produce pairs of aligned 3D Dixon MR and CT images. The voxel-wise weighting method, which employs local normalized cross-correlation similarity measure to give appropriate weights locally, outperformed other pseudo-CT generation approaches, yielding a Dice similarity index of 0.77±0.04 (P<0.05) compared to 0.60±0.02 (P<0.05) achieved by bone segmentation using simple averaging of all CT atlases. This approach is promising for MRI-guided attenuation correction in whole-body PET/MRI.

I. Introduction

A number of academic and corporate research groups focused their efforts on the development of combined PET/MR systems and addressing its many technical challenges and crucial issues, including the development of appropriate methodologies for MRI-guided attenuation correction of PET data [1]. Commercially available PET/MR scanners employ tissue classification methods which rely on segmenting an MR image into tissue classes and assigning uniform linear attenuation coefficients [2]. The major drawback of such methods, particularly in whole-body PET/MR imaging, is that bony structures are ignored and not considered as a separate tissue class. In order to take bone into account during attenuation correction (AC), few techniques have been proposed utilizing the concepts of atlas-guided attenuation map generation [3]. These methods primarily rely on prior information provided by registration of an atlas image, thus allowing the classification of bone tissue. The principles of atlas-based segmentation have been successfully applied to a wide variety of imaging modalities and segmentation tasks. In principle, each individual atlas image is converted to the coordinates of the target image. It has been proven that using the information provided by multiple atlas images leads to more robust and accurate results [4]. Information obtained from several atlas images can be pooled into an average atlas or into a so called probabilistic atlas. However, there is a trend to gain full advantage of multiple atlas images at hand via exploiting pattern recognition techniques in order to identify morphologically similar cases in atlas data set during the process of multi-atlas fusion which dramatically reduces non-systematic registration errors and improves the accuracy of resulting segmentation [5]. We based our measure for comparing various pseudo-CT generation approaches and atlas-based segmentation methods on the accuracy of bone extraction from whole-body MR images. To this end, we evaluate the performance of bone extraction from MRI of commonly used atlas-based segmentation techniques and pseudo-CT generation methods for the purpose of attenuation correction in whole-body PET/MRI. Eight different methods were evaluated using 23 pairs of aligned whole-body CT and MR Dixon images.

We aim to select the most promising approach among the evaluated methods for attenuation correction in whole-body PET/MRI.

II. Material and Methods

A. Data Preparation

The study population comprised 23 consecutive patients who underwent whole-body ^{18}F-FDG PET/MR and PET/CT examinations for staging of head and neck malignancies. ^{18}F-FDG PET/CT scans were performed on a Biograph 64 True Point scanner (Siemens Healthcare, Erlangen, Germany) whereas PET/MRI examinations were performed on the Ingenuity TF PET/MR (Philips Healthcare, Cleveland, OH). The MR Dixon pulse sequence is exploited for whole-body bone segmentation in this study. The whole body Dixon 3D volumetric interpolated T1-weighted sequence was acquired using the following parameters: flip angle 10°, TE_1 1.1 ms, TE_2 2.0 ms, TR 3.2 ms, 450×354 mm^2 transverse FOV, 0.85×0.85×3 mm^3 voxel size. In phase images are used for the assessment of whole-body bone segmentation. Usually MR images acquired using clinical protocols contain a relatively high level of noise and are corrupted due to the low frequency bias field and inter-patient intensity inhomogeneity. As it will be described later, the MR bone segmentation procedure entails direct handling of MR image intensity. As such, the presence of any of the aforementioned sources of intensity uncertainty in MR images might skew bone segmentation accuracy. In order to overcome these prospective sources of error, in-phase images of all patients underwent the following pre-processing.

Gradient anisotropic diffusion filtering using the following parameters: conductance= 4, iteration=10 and time step=0.01. This algorithm anisotropically diffuses an image. That is, it blurs over regions of an image where the gradient magnitude is relatively small (homogenous regions) but diffuses little over areas of the image where the gradient magnitude is large (i.e., edges). Therefore, objects are blurred, but their edges are blurred by a lesser amount. N4 bias field correction: Bspline

H. Arabi is with the Division of Nuclear Medicine & Molecular Imaging, Geneva University Hospital, CH-1211, Geneva, Switzerland (e-mail: hossein.arabi@etu.unige.ch).

†H. Zaidi is with the Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, CH-1211 Geneva, Switzerland, Geneva, Neuroscience Center, Geneva University, CH-1205 Geneva, Switzerland, and Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, Groningen, Netherlands (e-mail: habib.zaidi@hcuge.ch).

978-1-4799-6097-2/14/$31.00 ©2014 IEEE

grid resolution=400, Number of iteration=200 (at each grid resolution), Convergence threshold = 0.001, Bspline order = 3, Spline distance=400, Number of Histogram= 256 and Shrink factor= 3. Histogram matching: Histogram level = 1024 and match points = 128. In order to get the best result from histogram matching, it is recommended to exclude the background air voxels of the both reference and target images before processing. Normalization to the average water intensity: In order to overcome the inter-patient MR intensity non-uniformity, in addition to the histogram matching, the average intensity of the water images is used to normalize the corresponding in-phase images. To do so, fuzzy C-mean clustering algorithm was employed to segment each of Dixon water-only images into 3 distinct classes. The obtained mask of third class on water-only image (which is regarded to be almost purely water) was used to calculate the mean intensity value on the corresponding in-phase image; thereafter the calculated mean intensity value was used to normalize the in-phase image.

B. Atlas-based bone segmentation approaches

Gray level arithmetic averaging (GLA): This approach consists of simple averaging of all aligned CT images belonging to the atlas data set. In the first step, non-rigid registration is performed on each atlas MR image to match the target MR image individually. Thereafter, the obtained transformation field is applied to corresponding CT images. In the last step, gray level averaging is performed on all aligned CT images (Eq. 1) and a binary bone mask created using a threshold of 180 HU.

$$GLA_t(x) = \frac{1}{n}\sum_{i=1}^{n} T_{it}(CT_i(x)) \quad (1)$$

where T is the transformation matrix, n is the number of patients (n=14), t is an index for the target image and x is an index of image voxels.

Single average atlas (SAA): In this approach, an average atlas from all patients (leave-one-out) is produced offline and non-rigid registration is applied on the average atlas to match the target image. An average atlas is built for both MR and CT images. The MR atlas is used for final registration to the target MR image and its transformation field applied to the average CT atlas (Eq. 2).

$$SAA_t(x) = T_t\left(\frac{1}{n}\sum_{i=1}^{n} T_{ia}\left(CT_i(x)\right)\right) \quad (2)$$

where a is the index for the average atlas. This method is faster than the previous one due to the offline nature of atlas construction, thus requiring only one online registration.

Majority voting (MV): This technique is similar to gray level averaging where all patients are non-rigidly registered to the target image and the resulting transformation fields applied to the corresponding bone mask images. Majority voting assigns to each voxel the binary value that most registered atlas images agree with [5].

$$MV_t(x) = \max\left[f_l(x)\right] \quad (3)$$

where $f_l(x) = \sum_{i=1}^{n} w_{il}(x)$

$$W_{il}(x) = \begin{cases} 1 & if\ \ i = BM(x) \\ 0 & otherwise \end{cases}$$

where BM(x) is the binary image of bone mask.

Most similar patient: In this approach, after performing non-rigid registration to all patient images, the most similar patient to the target is selected through normalized mutual information (NMI) similarity measure applied to MR images (Eq. 4).

$$i = \arg\max_i\left[NMI(T_{it}(MR_i), MR_t)\right] \quad (4)$$
$$MSP_t(x) = T_{it}\left(CT_i(x)\right)$$

Gaussian process regression (GPR): This is the technique proposed by Hofmann et al. [3] for generation of pseudo-CT images to be used for attenuation correction in PET/MRI. This algorithm employs a combination of pattern recognition and atlas registration in order to estimate patient-specific attenuation map.

Weighted CT averaging (WCA): In this approach, the attenuation map is generated through voxel-wise weighting of the aligned CT images using the local normalized cross-correlation (LNCC) in order to alleviate miss-registration errors [6].

Sorted majority voting (SMV): In order to segment bone from MR images, the aligned CT images are sorted using normalized mutual information (Eq. 5) and a threshold is applied on r_i to discard dissimilar and miss-registered images. Thereafter, majority voting algorithm is used to yield a binary image of bone [7].

$$r_i = \frac{NMI(T_{it}(MR_i),MR_t) - min_k[NMI(T_{kt}(MR_k),MR_t)]}{max_k[NMI(T_{kt}(MR_k),MR_t)] - min_k[NMI(T_{kt}(MR_k),MR_t)]} \quad (5)$$

Shape Based Averaging (SBA): In this algorithm, a signed distance map is computed on each individual image before averaging the aligned binary images of bony regions. Then the combined output for each voxel is determined by minimizing the mean distance over all labels.

C. Quantitative evaluation

The validation of extracted bone from the various atlas-based segmentation strategies described in 2.4 section, was carried out on the basis of comparison with the segmented bone on the corresponding reference CT images using five volume/distance-based measures: Dice similarity (DSC), relative volume difference (RVD), Jaccard similarity (JC), sensitivity (S) and mean absolute surface distance (MASD).

$$DSC(A,M) = \frac{2|A\cap M|}{|A|+|M|}\ ,\quad RVD(A,M) = 100\times\frac{|A|-|M|}{|M|}\ ,$$
$$JC(A,M) = \frac{|A\cap M|}{|A\cup M|}\ ,\quad S(A,M) = \frac{|A\cap M|}{|M|}$$
$$MASD(A,M) = \frac{d_{ave}\ (S_A,S_M) + d_{ave}\ (S_M,S_A)}{2}$$

where, A is the segmented bone from the reference CT image and M denotes the extracted bone applying the atlas-based segmentation. $D_{ave}(S_A,S_M)$ is the average direct surface distance from all points on the reference bone surface S_A and to the segmented bone surface S_M.

The Shapiro-Wilk test was used to examine null hypothesis that measured SUVs follow a normally distributed population and the calculated p-values were reported for each individual segmentation scheme. The results were considered statistically significant if the p-value was less than 0.05.

Table 1. Comparison of validation measures (mean±SD), including Dice similarity (DSC), relative volume distance (RVD), Jaccard similarity (JC), Sensitivity (S) and mean absolute surface distance (MASD) for all combinations of atlas-based segmentation methods.

Methods	DSC	RVD(%)	JC	S	MASD(mm)
GLA	0.60±0.02	-46.0±02.4	0.43±0.02	0.46±0.02	10.7±03.9
SAA	0.60±0.02	-46.4±02.5	0.44±0.02	0.47±0.02	11.1±03.9
MV	0.59±0.02	-49.8±02.4	0.42±0.02	0.44±0.02	09.8±03.6
MSP	0.58±0.09	-39.2±08.1	0.41±0.10	0.52±0.11	06.2±02.0
GPR	0.60±0.02	-45.5±02.4	0.42±0.02	0.45±0.02	10.1±03.3
SBA	0.56±0.05	-15.3±04.7	0.39±0.05	0.52±0.06	11.1±04.1
WCA	0.77±0.04	-17.6±04.8	0.63±0.04	0.70±0.04	05.2±01.5
SMV	0.64±0.06	-36.5±05.6	0.47±0.05	0.53±0.06	06.4±01.5

III. Results

Whole-body bone segmentation through non-weighting averaging was carried for varying number of atlases chosen randomly across the whole data set. Figure 1 illustrates the accuracy of extracted bone in terms of DSC and RVD validation measures using the both intensity averaging (GLA) and majority voting (MV). The bars show the standard deviation at each measured point. The rising part of DSC plot in figure 1(1 to 14) based on intensity averaging frame work, can be justified as unsystematic misalignment reduction due to uncorrelated error between added atlases [5]. Adding more atlas images to the process of non-weighting atlas fusion led to degrading segmentation accuracy because of high level of smoothness and lack of the patient specific details.

Fig. 2 and 3 depict representative slices of segmented bone for different approaches along with error distance map corresponding to the reference bone extracted from CT images. Visual inspection demonstrated higher quality of extracted bone using voxel-wise weighing approach (WCA) (Fig 3.D).

According to Table 1, the WCA approach results in the best bone extraction performance with 0.77±0.04 and -17.6±04.8 ($P<0.05$) Dice similarity measure. On the other hand, SAA is the fastest approach since it requires only one online registration while the GPR algorithm is the most time consuming approach since its needs to develop a Gaussian regression kernel almost for local patch to predict the corresponding surrogate CT value. Due to presence of local miss-registration errors, the WCA approach, which uses LNCC morphological similarity criterion to localy discard poorly aligned images, outperformed the other algorithms. SMV assigns weighting factor for each transformed atlas images globally. Global atlas weighting strategy exhibited moderate improvement compared to the general averaging (GLA) method as transformed atlases with large misalignment error are excluded or at least are given relatively small weights in the process of averaging. According to the table 1, SMV method which employed mutual information morphological similarity measure resulted in bone segmentation accuracy of 0.64±0.06. The similar approach was exploited by Ying et al. [7] for automated bone segmentation from MR images of the hip joint by using the NMI similarity criterion which resulted in DSC=0.95. The marked difference between our result and what is reported in this study stems from different MR image sequence and quality, registration algorithm and on top of all the smaller field of view which leads to better registration outcome. Still there is dramatic distance between voxel-wise and global weighting strategy (WCA) with 0.77±0.04. The transformed atlas image might be aligned almost perfectly in one region whereas other region has massive misalignment errors, for instance because of anatomical variability, which local weighting strategies can properly deal with it.

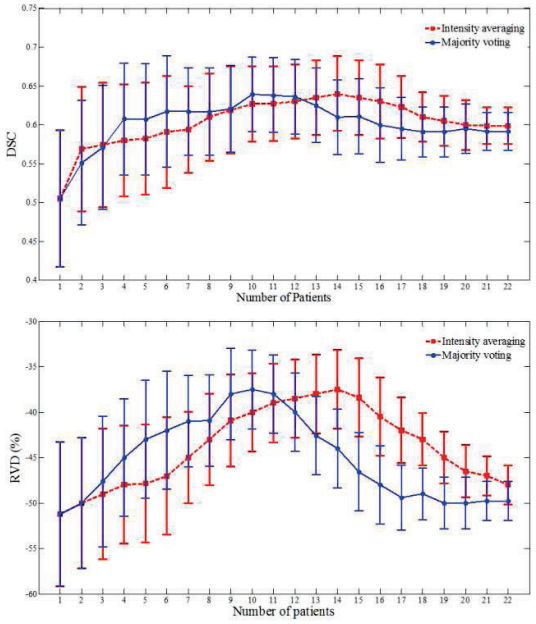

Fig 1. DSC (top) and RVD (bottom) similarity measures vs. number of patient using the Gray level arithmetic intensity averaging (GLA) and majority voting (MV) frameworks.

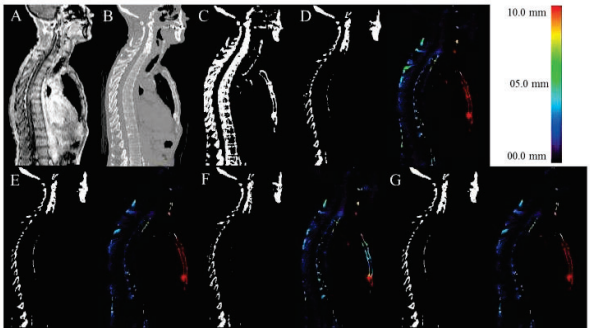

Fig 2. Representative slice of bone segmentation from MR image along with corresponding error distance map. (A) In-phase MR image, (B) corresponding CT image, (C) binary image of reference bone, (D) GLA, (E) SAA, (F) MV, (G) GPR.

IV. Conclusion

The primary aim of this work is to compare the performance of atlas-based bone segmentation approaches in whole-body PET/MRI. It appears that the voxel-wise atlas fusion strategy via weight assignment on the basis of morphological similarity between target image and atlas images led to promising results in the realm of whole-body bone segmentation. Future studies will focus on implementing other state-of-the-art bone segmentation algorithms to select the most promising approaches for attenuation correction in whole-body PET/MRI.

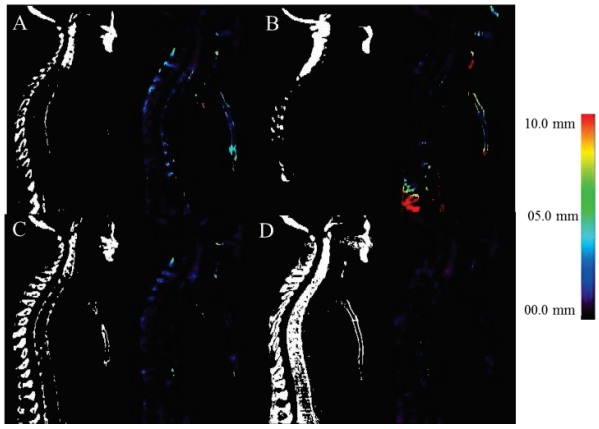

Fig 3. Representative slice of bone segmentation from MR image along with corresponding error distance map. (A) SMV, (B) SBA, (C) MSP, (D) SMV

V. Acknowledgments

This work was supported by the Swiss National Science Foundation under grant SNSF 31003A-149957.

References

[1] H. Zaidi and A. Del Guerra, "An outlook on future design of hybrid PET/MRI systems.," Med Phys, vol. 38, pp. 5667-5689, 2011.

[2] A. Martinez-Moller, M. Souvatzoglou, G. Delso, R. A. Bundschuh, C. Chefd'hotel, S. I. Ziegler, et al., "Tissue classification as a potential approach for attenuation correction in whole-body PET/MRI: Evaluation with PET/CT data.," J Nucl Med, vol. 50, pp. 520-526, Apr 2009.

[3] M. Hofmann, I. Bezrukov, F. Mantlik, P. Aschoff, F. Steinke, T. Beyer, et al., "MRI-based attenuation correction for whole-body PET/MRI: Quantitative evaluation of segmentation- and Atlas-based methods.," J Nucl Med, vol. 52, pp. 1392-1399, Sep 2011.

[4] J. M. P. Lötjönen, R. Wolz, J. R. Koikkalainen, L. Thurfjell, G. Waldemar, H. Soininen, et al., "Fast and robust multi-atlas segmentation of brain magnetic resonance images," NeuroImage, vol. 49, pp. 2352-2365, 2/1/ 2010.

[5] X. Artaechevarria, A. Munoz-Barrutia, and C. Ortiz-de-Solorzano, "Combination Strategies in Multi-Atlas Image Segmentation: Application to Brain MR Data," Medical Imaging, IEEE Transactions on, vol. 28, pp. 1266-1277, 2009.

[6] N. Burgos, M. Cardoso, M. Modat, S. Pedemonte, J. Dickson, A. Barnes, et al., "Attenuation Correction Synthesis for Hybrid PET-MR Scanners," in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013. vol. 8149, K. Mori, I. Sakuma, Y. Sato, C. Barillot, and N. Navab, Eds., ed: Springer Berlin Heidelberg, 2013, pp. 147-154.

[7] X. Ying, F. Jurgen, S. C. Shekhar, S. Raphael, E. Craig, and C. Stuart, "Automated bone segmentation from large field of view 3D MR images of the hip joint," Phys Med Biol, vol. 58, p. 7375, 2013.