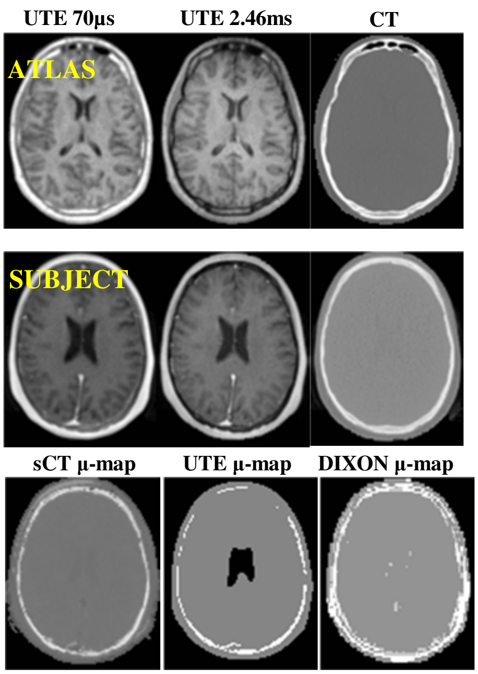
Accurate PET Reconstruction for PET/MR Scanners using Synthetic CT

Snehashis Roy[1], Wen-Tung Wang[1], John Anthony Butman[2], and Dzung Pham[1]

[1]*Center for Neuroscience and Regenerative Medicine, Henry Jackson Foundation, Bethesda, MD, United States,* [2]*Diagnostic Radiology, National Institute of Health, Bethesda, MD, United States*

INTRODUCTION: Integrated PET(Positron Emission Tomography)/MR systems are becoming increasingly used in clinical and research applications. However, proper PET reconstruction requires accurate computation of electron density attenuation coefficient maps (μ-maps). The challenge of PET/MR in contrast to PET/CT systems lies in the accurate computation of μ-maps from MR images. Since MRI does not provide any information related to electron density, the attenuation coefficients are computed using either segmentations[1] of the MR images or using atlases[2,4]. In computing μ-maps from head MRI, Ultra-short Echo Time (UTE) MR imaging sequences are preferred to regular T_1-weighted sequences, because UTE can produce improved contrast for bone, making it viable for synthesizing CT from a pair of MR images, one with ultra-short echo time (~100μs) and one with a longer echo time. Previous methods based on UTE images[3] involve simple threshold based segmentation of bone, soft tissue and air using difference between the dual-echo images. In this work, we show that UTE dual-echo images can effectively be used to synthesize realistic looking CT via patch matching from subject to an atlas. We show that PET reconstruction using synthetic CT (**sCT**) based μ-maps are very close to that obtained with original CT based μ-maps. We also demonstrate that **sCT** provides more accurate PET reconstruction than DIXON and Siemens product UTE based μ-maps.

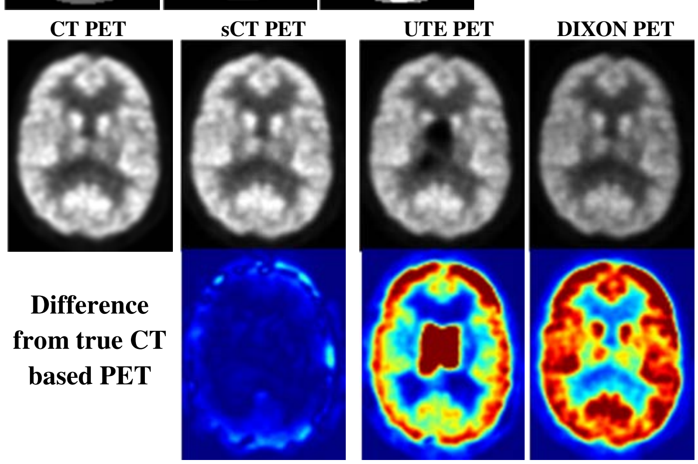
METHODS: We use the idea of patch matching[4,5] from subject to atlas to synthesize CT images using a pair of UTE images. An atlas is a registered triplet of images $\{\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3\}$, a UTE with T_E=70μs, a GRE T_1-w image with T_E=2.46ms and corresponding CT, respectively. A subject contains registered UTE images,$\{\boldsymbol{s}_1,\boldsymbol{s}_2\}$. The subject and the atlas are first decomposed into patches. The subject patch collection (or *cloud*) is matched to the atlas patch cloud using a number of Gaussian mixture models. In the ideal scenario, for every patch-pair (or equivalently, a *patch*) from $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$, a best matching patch can be found from $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. Then a corresponding patch from $\boldsymbol{a}_3$ can be used as the synthetic CT patch ($\hat{\boldsymbol{s}}$), since all the atlas images are co-registered. However, it is quite likely that a linear combination of a few atlas patches matches a subject patch better than a single atlas patch. Thus, we consider all convex combinations of *n*-tuples of atlas patches. We then postulate that a subject patch is a random vector whose probability density is multivariate Gaussian with mean given by an unknown convex combination of *n* nearby atlas patches. Expectation-Maximization is used to estimate the unknown combinations. Finally the synthetic CT patch is estimated by the convex combination of the atlas patches.



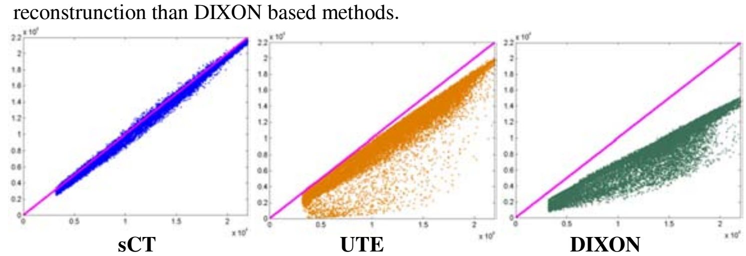
Unlike previous methods[2] where atlas and subject are needed to be registered properly, we do not require any registration, which can be problematic. Let $\mathbf{x}_i$ denote a patch-pair from the subject UTE images, and $\mathbf{y}_j$ denote a patch pair from the atlas UTE images. Also let $\mathbf{u}_i$ and $\mathbf{v}_j$ be the corresponding CT patches, where $\mathbf{u}_i$ is to be estimated from $\mathbf{x}_i$, $\mathbf{y}_j$, and $\mathbf{v}_j$. We postulate that the vector $\mathbf{p}_i=[\mathbf{x}_i\ \mathbf{u}_i]$ follows a Gaussian distribution with mean expressed as a convex combination on *n*-tuples, $\mathbf{q}_{j1}=[\mathbf{y}_{j1}\ \ \mathbf{v}_{j1}],\ldots,\ \mathbf{q}_{jn}=[\mathbf{y}_{jn}\ \ \mathbf{v}_{jn}]$. More formally, $\mathbf{p}_i \sim N(\alpha_1\mathbf{q}_{j1}+\ldots+\alpha_n\mathbf{q}_{jn},\Sigma)$, where α_1 to α_n denote the mixing coefficients of each atlas patch contributing to the subject patch p_i. The coefficients are estimated using expectation maximization. Once they are estimated, the corresponding CT patch is estimated via a linear combination of the atlas CT patches, $\mathbf{u}_i = \alpha_1\mathbf{v}_{j1}+\ldots+\alpha_n\mathbf{v}_{jn}$. More details of the method can be found here[4].

RESULTS: We synthesized **sCT** image for two subjects scanned on a Siemens Biograph mMR and compared the corresponding FDG-PET with the reconstruction using a CT, and scanner generated DIXON and UTE based μ-maps. For one of the subjects, the scanner failed to generate the μ-map using UTE. Dual-echo UTE images and the true CT of one subject, constituting the atlas, are shown in Fig.1. The μ-maps from DIXON, scanner UTE, and **sCT** are also shown in Fig.1. Visually, the **sCT** μ-map more closely resembles the original CT μ-map. It also has better bone to soft tissue discrimination than the other two, indicating the possibility of more accurate PET reconstruction.

Reconstructed PET images from the three μ-maps are shown in Fig.2. Assuming the true CT reconstructed PET as the ground truth, the **sCT** provides the closest reconstruction to the truth compared to DIXON and UTE based PET reconstructions and its difference from the ground truth CT. Visually, **sCT** produces a very similar reconstruction inside the brain. Quantitatively, the PSNR of DIXON and UTE reconstructed PET is 23dB and 15dB within brain, while it is 33dB for **sCT** based PET, indicating significant improvement in reconstruction. Scatter plots showing CT reconstructed PET intensities vs. MR reconstructed PET intensities at each voxel of the PET images indicate that the **sCT** method produces more accurate reconstruction and is less biased. Magenta lines indicate unit slope. Evidently, for UTE and DIXON, all the points lie below the line, indicating PET intensities are significantly lower than the truth. This is also indicated by R^2=0.99,0.89,0.65, for **sCT**, UTE and DIXON. All numbers are computed on brain region, while the brainmask is computed from UTE images.



CONCLUSION: We have presented a framework to reconstruct PET from MRI, where CT is usually not available in PET/MR systems. We synthesize CT images based on an atlas containing UTE dual-echo images. Since UTE contains signal from bone, unlike traditional T1, it is successfully used to produce better quality PET reconstrunction than DIXON based methods.



REFERENCES: : [1] Martinez-Moller A, *et. al*, J. Nuclear Medicine, 2009;50(4):520-526. [2] Ourselin S, *et.al*., MICCAI, 2013;147-154. [3] Keereman V, *et. al*, J. Nuclear Medicine, 2010;51(5):812-818. [4] Roy S, *et.al.*, MBIA, 2003;51-62. [5] Hoffman et. al. J. Nuclear Medicine, 2008;49(11):1875-1883.