PET IMAGE RECONSTRUCTION USING ANATOMICAL INFORMATION THROUGH MUTUAL INFORMATION BASED PRIORS: A SCALE SPACE APPROACH

Sangeetha Somayajula[1], *Anand Rangarajan*[2] *and Richard M. Leahy*[1]

[1]Signal and Image Processing Institute, Univ. of Southern California, Los Angeles, CA 90089
[2]Dept. Computer and Info. Science and Engineering, University of Florida, Gainesville, FL 32611.

ABSTRACT

We propose a mutual information based prior for incorporating information from co-registered anatomical images into PET image reconstruction. The prior uses mutual information between feature vectors that are extracted from the anatomical and functional images using a scale space approach. We perform simulations on a realistic 3D phantom generated by replicating a 2-D autoradiographic cross section of a mouse labelled with F18-FDG. A digital photograph of the cryosection of the same slice is used to generate the anatomical image. The images are registered using mutual information based rigid registration. PET data are then simulated from the autoradiography based phantom. We use a preconditioned conjugate gradient algorithm to compute the PET image that maximizes the posterior density. The performance of this method is compared with that using a Gaussian quadratic penalty, which does not use anatomical information. Simulation results indicate that the mutual information based prior can achieve reduced standard deviation at comparable bias compared to the quadratic penalty.

Index Terms— positron emission tomography, anatomical priors, scale space theory, mutual information

1. INTRODUCTION

The uptake of radioactive tracers in vivo results in a spatial density that reflects underlying anatomical morphology. Hence, anatomical information from high resolution MR/CT images can potentially be used to improve the quality of low resolution PET images. This anatomical information can be incorporated into PET image reconstruction in a Bayesian framework through the use of priors, which encourage the functional image to be similar in structure to the anatomical image [1] - [4]. Mutual information (MI) between two random vectors is a measure of the amount of information contained by one random vector about the other, and can be used as a similarity metric between the two images [5]. In [4], a Bayesian joint mixture model is formulated such that the solution maximizes MI between class labels of PET and anatomical images.

In [6] we described a non-parametric method that uses MI between feature vectors extracted from the anatomical and functional images to define a prior on the functional image. The features were chosen as intensity, directional gradients, and local mean, based on the assumption that boundaries in the functional image are correlated with anatomical boundaries and that the activity within boundaries tends to be more homogeneous. We explored a best-case scenario through 2D simulations wherein the PET and anatomical images had identical structure. The results showed less noisy images with improved boundary definition, but the estimates had high variance.

In this paper we propose a more automated method of choosing relevant feature vectors, with the goal of reducing the high variance in the estimate obtained in [6]. Scale space theory provides a framework for the analysis of images at different levels of detail. It is based on generating a one parameter family of images by blurring the image with Gaussian kernels of increasing width (the scale parameter) and analyzing these blurred images to extract structural features [7]. Scale space images at one value of scale parameter and their derivatives were used as features in [8] for MI based registration. In this work, we define the features for MI based priors as the original image, images blurred at different scales, and the Laplacians of the blurred images. These features automatically emphasize the stronger boundaries that delineate important anatomical structures and attach less importance to internal detail that is less likely to be correlated in the two images. We perform simulations on a realistic 3D phantom generated by replicating a 2-D autoradiographic cross section of a mouse labelled with ^{18}FDG. The anatomical phantom was independently obtained from a digital photograph of the cryosection used to collect the autoradiograph. The phantom images were registered using MI-based rigid registration.

2. METHODS AND RESULTS

Let $\mathbf{f}$ represent the functional image, $\mathbf{a}$ the anatomical image, and $\mathbf{g}$ the sinogram data. The maximum *a posteriori* (MAP)

This work was supported in part by Grant R01 EB000363 from the National Institute of Biomedical Imaging and Bioengineering

estimate of $\mathbf{f}$ is defined as

$$\hat{\mathbf{f}} = \arg\max_{\mathbf{f}\geq 0} \frac{p(\mathbf{g}|\mathbf{f})p(\mathbf{f})}{p(\mathbf{g})}, \quad (1)$$

where $p(\mathbf{g}|\mathbf{f})$ is the Poisson likelihood function and $p(\mathbf{f})$ is a prior on the functional image.

We define a prior that is based on the mutual information between feature vectors extracted from the anatomical and functional images. Let N feature vectors $\mathbf{x}_i$ and $\mathbf{y}_i$, for $i = 1, 2, \cdots, N$ be extracted from the anatomical and functional images respectively. These can be considered as realizations of the random vectors $\mathbf{X}$ and $\mathbf{Y}$ respectively. Let m be the number of features in each feature vector such that $\mathbf{X} = [X_1, X_2, \cdots, X_m]^T$. Mutual information between $\mathbf{X}$ and $\mathbf{Y}$ is defined as [9],

$$I(\mathbf{X}, \mathbf{Y}) = H(\mathbf{X}) + H(\mathbf{Y}) - H(\mathbf{X}, \mathbf{Y}) \quad (2)$$

$$H(\mathbf{X}) = -\int p(\mathbf{X}) \log p(\mathbf{X}) d\mathbf{X} \quad (3)$$

$$H(\mathbf{X}, \mathbf{Y}) = -\int p(\mathbf{X}, \mathbf{Y}) \log p(\mathbf{X}, \mathbf{Y}) d\mathbf{X} d\mathbf{Y} \quad (4)$$

where $H(\mathbf{X})$ denotes the entropy of $\mathbf{X}$, $H(\mathbf{X}, \mathbf{Y})$ denotes the joint entropy between $\mathbf{X}$ and $\mathbf{Y}$, and $d\mathbf{X}$ is shorthand for $dX_1 dX_2 \cdots dX_m$. Our mutual information based prior is then defined as,

$$p(\mathbf{f}) = \frac{1}{Z} \exp(\mu I(\mathbf{X}, \mathbf{Y})), \quad (5)$$

where Z is a normalization constant and μ is a positive constant that controls the weight given to the prior. For simplicity, the m features are assumed to be independent so that $I(\mathbf{X}, \mathbf{Y})$ can be computed from the mutual information of the individual features as follows:

$$I(\mathbf{X}, \mathbf{Y}) = \sum_{i=1}^{m} I(X_i, Y_i). \quad (6)$$

The computation of mutual information requires knowledge of the joint density $p(\mathbf{X}, \mathbf{Y})$. Taking a non-parametric approach, the joint density is estimated from the realizations $\mathbf{x}_i$ and $\mathbf{y}_i$ using the Parzen window method [10] with a Gaussian window. The window width is treated as a design parameter.

Taking the log of the posterior density and dropping constants, our objective function now becomes:

$$L(\mathbf{f}) = \log(p(\mathbf{g}|\mathbf{f})) + \mu I(\mathbf{X}, \mathbf{Y}). \quad (7)$$

This objective function can be maximized using preconditioned conjugate gradient (PCG) with a bent Armijo line search technique. The objective function is non-convex so that the local PCG search will be sensitive to the initial estimate. In the results presented below we initialize using the early iterations of an MLEM algorithm.

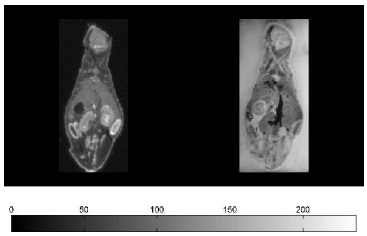

Fig. 1. (Left) 2D autoradiograph data used to generate the PET and (right) coregistered cryosection image used as the anatomical phantom.

2.1. Extraction of feature vectors

The feature vectors should be chosen such that they are correlated in the anatomical and functional images. Comparison of autoradiographs (which can be viewed as equivalent to high resolution PET scans) and cyrosection images (a surrogate for structural MRI or CT images) reveals a high degree of similarity in the morphology of the images, even though the intensities within each region can be quite different. Consequently, the feature vectors should reflect the morphology of the images. Since the mutual information metric is naturally insensitive to differences in intensity between the two images, the combination of local measures of image morphology and mutual information should facilitate reconstruction of PET images whose structure is similar to that of the registered anatomical image.We use a scale space theory based approach to define the morphological feature vectors as:

1. The original image
2. The image blurred by a Gaussian kernel of width σ_1
3. The image blurred by a kernel of width σ_2, where $\sigma_2 > \sigma_1$
4. Laplacian of the image obtained in (2)
5. Laplacian of the image obtained in (3).

Since the Laplacian and blurring operations are linear, the gradients of these features can be easily computed from the gradient of the original image. By analyzing the image at two different scales, we are giving more emphasis to those boundaries that remain in the image at the higher scale. In this paper, we choose the kernel widths as $\sigma_1 = 0.2$ and $\sigma_2 = 0.5$. We used a 3×3 Laplacian kernel given by $\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$. More features could be added by using different sizes of kernels, but we restrict our analysis to two scales to limit complexity.

2.2. Simulation Results

We generated a 256x256x7 PET phantom from one slice of an autoradiograph of a mouse collected shortly after a PET scan using ^{18}F-FDG. We used a digital photograph of the same slice of the cryosectioned mouse to generate the anatomical image. For both images, the chosen slice was replicated in planes 2 to 7 and the rest of the image was zero padded. To

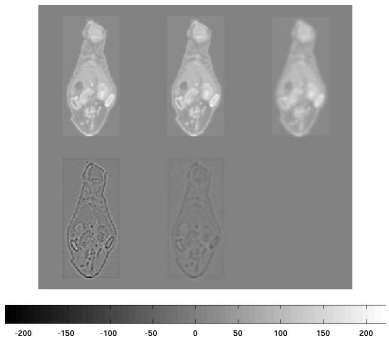

Fig. 2. Feature vectors extracted from true PET image. Top: Original image, image blurred by kernel with $\sigma_1 = 0.2$, image blurred by kernel with $\sigma_2 = 0.5$. Bottom: Laplacian of blurred images.

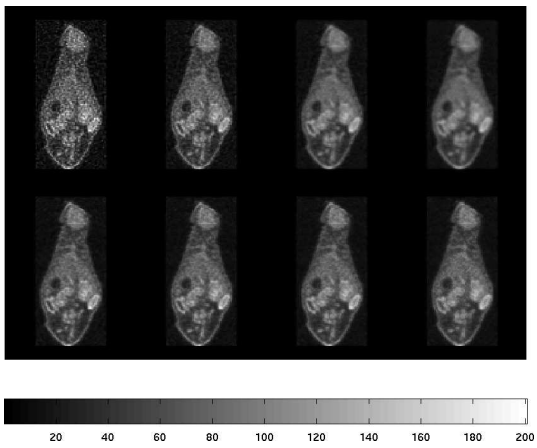

Fig. 3. Reconstructed images with the QP (top) for β=0, 0.01, 0.05, and 0.1, and MI prior (bottom) for μ=1000, 5000, 1e4, and 25000.

ensure alignment of the two phantom images we applied a rigid mutual information registration algorithm. One plane of the PET and anatomical phantoms is shown in Figure 1.

The simulations are based on a 4 ring microPET Focus 220 scanner, for which the sinogram is of size 288×252×13. The sinogram data had approximately 30 million counts. We reconstructed images in a 128x128x7 voxel volume. We used 50 iterations of preconditioned conjugate gradient (PCG), initialized by the image reconstructed after 1 iteration of the maximum likelihood expectation maximization (MLEM) algorithm. We compare the performance of the MI based prior to a quadratic prior (QP) using 26 nearest neighbors. The feature vectors extracted for the true PET image are shown in Figure 2.

The reconstructed images using different values of hyperparameter for the QP and MI priors are shown in Figure 3. For QP, we trade off noise for resolution as we increase β. The MI prior images are less noisy, have sharper boundaries,

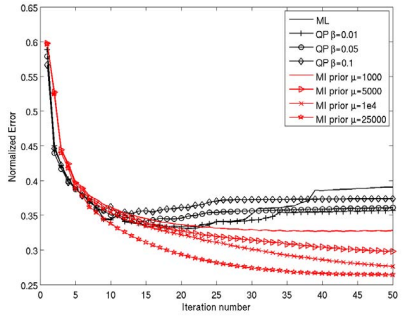

Fig. 4. Normalized mean squared errors as a function of iteration for QP and MI prior.

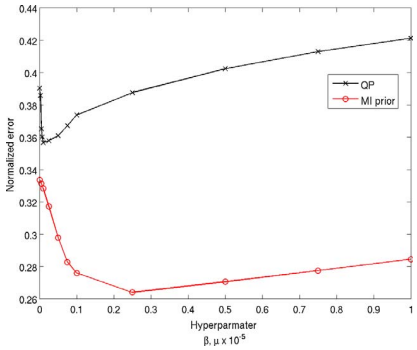

Fig. 5. Normalized mean squared error as a function of hyperparameter for QP and MI prior.

and have more contrast than the QP images. The normalized root mean squared (rms) error of the reconstructed image is plotted versus iteration number in Figure 4. The normalized rms error as a function of hyperparameter is shown in figure 5. It can be seen that the MI prior has lower error than the QP for all values of hyperparameter considered. To analyze the performance of the MI prior in comparison with the QP in quantifying uptake, we performed Monte Carlo simulations. We chose the operating point for these simulations as the hyperparameter corresponding to minimum error in Figure 5. For the QP this point is $\beta = 0.01$ and for the MI prior $\mu = 25000$. We generated 100 datasets of approximately 30 million counts each, and used 50 iterations of PCG for each reconstruction. The bias and standard deviation (SD) images for QP and MI priors are shown in Figures 6 and 7. It can be seen that the MI prior estimate has lower SD than for QP. However, the bias for the MI prior seems to be lower than QP in the outermost planes, but comparable overall in the central planes.

We computed the bias and variance in quantifying uptake in a small region of interest (ROI) within the kidneys, for the central planes 3, 4, and 5. The ROI considered is highlighted in Figure 8 for one plane. The bias and SD in estimating the mean activity in the ROI of the reconstructed image, nor-

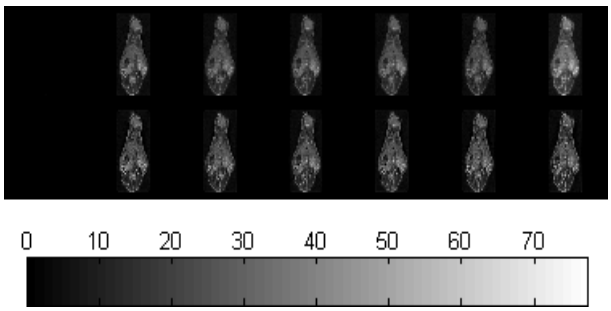

Fig. 6. Bias image (L to R: planes 1 to 7) for QP (top) and MI prior (bottom).

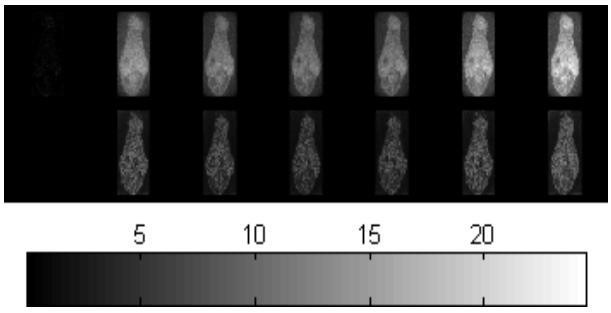

Fig. 7. SD image (L to R: planes 1 to 7) for QP (top) and MI prior (bottom).

malized by the mean activity in the ROI for the true image is shown in Table 1. It can be seen that the bias for the two priors is comparable, while the SD is lesser for MI prior although there is no significant difference between these two ROI estimates.

3. DISCUSSION

We proposed a scale space approach to extracting relevant features from anatomical and functional images for use in mutual information based priors for PET image reconstruction. We performed 3D simulations on a realistic PET phantom, with an anatomical image that was not identical in structure to the PET image. The reconstructed images using MI prior had less noise and sharper boundaries than QP. Monte Carlo simulations revealed that the MI prior has comparable bias and lesser standard deviation than QP. These results are preliminary and future in vivo studies validated using autoradiography, and phantom studies using a set of cryosection slices instead of replicating a single slice, are planned.

The scale space approach is well suited for mutual information based priors in the form presented here because it permits a multi-scale analysis of the images while maintaining the same spatial sampling as the original image. This is important because the accuracy of the non-parametric estimate of probability density is limited by the number of samples.

It is encouraging to see that the MI prior performs well in the reasonably realistic conditions described in this paper. We can expect similar performance for real data. We were suprised that there there was little difference in the ROI quantitation studies since we expected the increased resolution at the boundary with MI to reduce partial volume effects and improve quantitation. This issue will be explored further, both in simulations and in vivo studies.

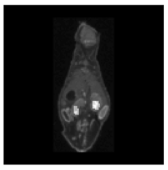

Fig. 8. The highlighted region is chosen as the ROI for planes 3, 4 and 5.

Prior	Bias	SD
QP	-0.2368	0.0042
MI-prior	-0.2329	0.0035

Table 1. Bias and SD of normalized mean uptake in region of interest.

4. REFERENCES

[1] R.Leahy and X.Yan, "Incorporation of MR data for improved functional imaging with PET.," *Inf. Proc. in Med. Imaging*, pp. 105–120, 1991.

[2] G.Gindi, M.Lee, A.Rangarajan, and I.G.Zubal, "Bayesian reconstruction of functional images using anatomical information as priors.," *IEEE Trans in Med. Imag.*, vol. 12, no. 4, Dec 1993.

[3] J.E.Bowsher, V.E.Johnson, T.G.Turkington, R.J.Jacszak, and C.E. Floyd Jr, "Bayesian reconstruction and use of anatomical *a priori* information for emission tomography.," *IEEE Trans. Med. Imag.*, vol. 15, no. 5, pp. 673–686, 1996.

[4] A.Rangarajan, I.-T.Hsiao, and G.Gindi, "A Bayesian joint mixture framework for the integration of anatomical information in functional image reconstruction.," *Journ. of Math. Imag. and Vision*, vol. 12, pp. 199–217, 2000.

[5] W.Wells III, P.Viola, H.Atsumi, S.Nakajima, S.Nakajima, and R.Kikinis, "Multimodal volume registration by maximization of mutual information," *Med. Image Analysis*, vol. 1, no. 1, pp. 35–51, 1996.

[6] S.Somayajula, E.Asma, and R.M.Leahy, "PET image reconstruction using anatomical information through mutual information based priors," *Conf. Rec: IEEE Nucl. Sci. Symp. and Med. Imag. Conf.*, pp. 2722–26, 2005.

[7] T.Lindeberg, *Scale-Space Theory in Computer Vision*, Kluwer Academic, Netherlands, 1994.

[8] M. Holden, L. D. Griffin, and D. L. G. Hill, "Multi-dimensional mutual information image similarity metrics based on derivatives of linear scale space.," in *Proc. of the APRS Workshop on Dig. Imag. Comp.*, 2005, pp. 55–60.

[9] T.Cover and J.Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications, 1991.

[10] R.O.Duda, P.E.Hart, and D.G.Stork, *Pattern Classification*, Wiley, second edition, 2001.