

PET Image Reconstruction using Anatomical Information through Mutual Information Based Priors

Sangeetha Somayajula, Evren Asma, and Richard M. Leahy

Signal and Image Processing Institute, Univ. of Southern California, Los Angeles, CA 90089

Abstract—We propose a non-parametric method for incorporating information from co-registered anatomical images into PET image reconstruction through priors based on mutual information. Mutual information between feature vectors extracted from the anatomical and functional images is used as *a priori* information in a Bayesian framework for the reconstruction of the PET image. The computation of mutual information requires an estimate of the joint density of the two images, which is obtained by using the Parzen window method. Preconditioned conjugate gradient with a bent Armijo line-search is used to maximize the resulting posterior density. The performance of this method is compared with that using a Gaussian quadratic penalty, which does not use anatomical information. Simulation results are presented for PET and MR images generated from a slice of the Hoffman brain phantom. These indicate that mutual information based penalties can potentially provide superior quantitation compared to Gaussian quadratic penalties.

I. INTRODUCTION

The uptake of radioactive PET tracers typically results in a spatial density that reflects underlying anatomical morphology. This results in a strong correlation between the structure of the anatomical and functional images. Hence, the incorporation of anatomical information from high resolution MR/CT images into PET reconstruction algorithms can potentially improve the quality of low resolution PET images. This anatomical information is readily available from multimodality imaging equipment that is often used for acquiring data and can be incorporated into the PET reconstruction algorithm in a Bayesian framework through the use of priors. Previous work on the use of anatomical priors can be broadly classified into: (i) Methods based on anatomical boundary information, which encourage boundaries in functional images that correspond to anatomical boundaries [1],[2] and (ii) methods that use anatomical segmentation information, which encourage the distribution of tracers in regions corresponding to anatomical regions [3],[4]. In both of these approaches, the goal is to obtain a functional image that has a structure that is similar to the anatomical image. Mutual information (MI) is a measure of the amount of information contained by one random vector about the other and therefore 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. A parametric method is used where each class conditional prior is assigned a gamma distribution.

We propose 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. In this paper our feature vectors consist of the intensity, local mean in a neighborhood, and the horizontal and vertical gradients at each pixel, because we expect the boundaries in the two images to be similar and the intensities to follow similar homogeneous distributions within the boundaries. This method does not require anatomical segmentation information and hence the algorithm is not constrained by the accuracy of segmentation.

II. METHODS AND RESULTS

A. MAP Reconstruction with Mutual Information Prior

Let \mathbf{f} represent the functional image, let \mathbf{a} denote the anatomical image and \mathbf{g} denote the sinogram data. The maximum a-posteriori (MAP or equivalently penalized likelihood) estimate of \mathbf{f} is given by,

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

where $p(\mathbf{g}|\mathbf{f})$ is the Poisson likelihood function and $p(\mathbf{f})$ is a prior on the functional image. Let the N feature vectors extracted from the functional and anatomical images be represented as \mathbf{x}_i and \mathbf{y}_i , respectively for $i = 1, 2, \dots, N$. These can be considered as realizations of the random feature vectors \mathbf{X} and \mathbf{Y} . Mutual information $I(\mathbf{X}, \mathbf{Y})$ is defined as [6]:

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

where $H(\mathbf{X})$ denotes the entropy of \mathbf{X} given by:

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

and $H(\mathbf{X}, \mathbf{Y})$ denotes the joint entropy between \mathbf{X} and \mathbf{Y} :

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

where $d\mathbf{X}$ is shorthand for $dX_1 \dots dX_N$. Our mutual information based prior is then defined as

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

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

where Z is a normalization constant and μ is a positive constant. Let there be m features in each feature vector such that $\mathbf{X} = [X_1, X_2, \dots, X_m]$. By assuming independence between the features, $I(\mathbf{X}, \mathbf{Y})$ can be computed from the mutual information between individual features as follows:

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

$I(\mathbf{X}, \mathbf{Y})$ is maximized when the joint density $p(\mathbf{X}, \mathbf{Y})$ is sparse with localized peaks, implying that the two vectors are highly correlated and knowledge of one reduces the uncertainty (entropy) in the other.

The computation of mutual information requires knowledge of the joint density $p(\mathbf{X}, \mathbf{Y})$. Using a non-parametric approach, this joint density is estimated from \mathbf{x}_i and \mathbf{y}_i , which are the realizations of \mathbf{X} and \mathbf{Y} . Differentiable density estimates for both the functional and anatomical images can be obtained using Parzen windows [8] in the form of:

$$\hat{p}(\mathbf{X}) = \sum_{i=1}^N \phi\left(\frac{\mathbf{X} - \mathbf{x}_i}{\sigma}\right), \quad (7)$$

where ϕ is a Gaussian window and σ determines the width of the window. The window width σ is taken as a design parameter. Taking the log of the posterior density and dropping constants, our objective function becomes:

$$L(\mathbf{f}) = \log(p(\mathbf{g}|\mathbf{f})) + \mu I(\mathbf{X}, \mathbf{Y}) \quad (8)$$

This objective function can be maximized with a preconditioned conjugate gradient procedure with an Armijo line-search [7]. When the non-negativity constraint was violated, we used a second line-search between the current estimate and the projection of the estimate resulting from the first line-search onto the constraint set to impose non-negativity (i.e. a bent Armijo line-search). The objective function is a non-convex function of \mathbf{f} , so optimization of this function using gradient based techniques requires a good initial estimate to converge to the correct solution. We used 10 iterations of the expectation maximization (EM) maximum likelihood algorithm to initialize the conjugate gradient method. Note that we used the Armijo rule rather than the Newton-Raphson line search because of the cost involved in computing the second derivative and also because of the nonconvexity of the objective function.

B. Extraction of Feature Vectors

The feature vectors extracted from the images should be chosen such that they are correlated in the anatomical and functional images and accurately describe the common structure of the two images. In this paper, the following features were considered:

- 1) Intensity: The intensities in the anatomical and functional images follow similar distributions, though the actual values are not similar.
- 2) Horizontal and Vertical gradients: The edges in the anatomical and functional images are strongly correlated,

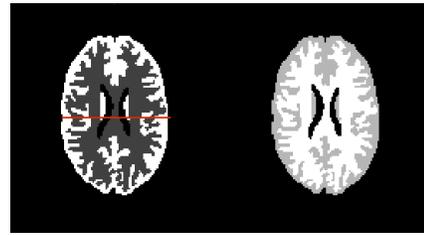


Fig. 1. The phantoms used in the simulations to represent the PET image (left) and the MR image (right). The line in the PET image represents the position of the transaxial profile in Figure 4.

since an edge in one image often corresponds to an edge in the other. The horizontal gradient at each pixel is computed by taking a left difference between the pixel intensities and the vertical gradient at each pixel is computed by taking the difference between the intensity at that pixel and the one below it.

- 3) Local mean: We expect generally small variations in image intensity within anatomical regions compared to variations from one region to another. Consequently we would expect to see similarity in the local means in the anatomical and functional images. The local mean at each pixel is computed by taking the average of intensities in a neighborhood of that pixel. This also provides spatial information, since the value of local mean at each pixel depends on the intensities of the pixels around it.

In the results presented below, we compare performance of MI-based reconstruction for the case where we use only image intensity with that where we also use the image gradient and local mean in the feature vector. Since the former uses no contextual information, while the latter is specifically dependent on the local image structure, we expect to find superior performance using the full feature vector.

C. Simulation Results

We used a 128×128 slice of the Hoffman brain phantom as our functional image and scaled the three different regions (gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF)) to generate our anatomical image. Both images are shown in Figure 1. The feature vectors extracted from the true PET image are shown in Figure 2. The simulations are based on a single ring of the microPET Focus 220 scanner, for which the sinogram dimensions are 288×252 . The sinogram data had approximately 800K counts. We reconstructed all images using 50 iterations of preconditioned conjugate gradient, initialized by the image reconstructed after 10 iterations of maximum likelihood expectation maximization (MLEM) algorithm.

Two different MI priors were considered and the strength of each prior was controlled by varying the smoothing parameter μ .

- 1) MI-Intensity: Mutual information between the intensities of the MR and PET images.
- 2) MI-Spatial: Mutual information between feature vectors consisting of intensity, local mean of 8 nearest neighbors

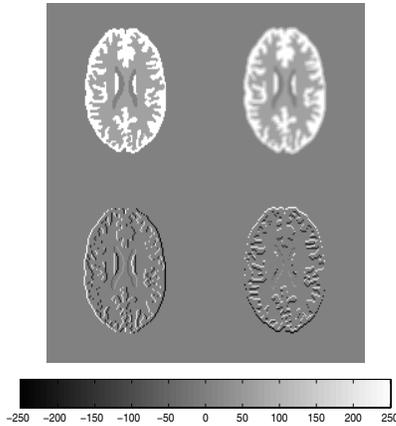


Fig. 2. Feature vectors extracted from the true PET image. Top Row: Intensity and Local Mean. Bottom Row: Horizontal and vertical gradients.

and horizontal and vertical gradients at each pixel.

The performance of these priors was compared to the quadratic prior (QP) with a neighborhood defined by 8 nearest neighbors, which does not use any anatomical information. The regularization parameter β for the QP was varied to control the resolution/noise trade-off. The reconstructed images using MI-Intensity, MI-Spatial and QP are shown in Figure 3. The transaxial profile of the true and reconstructed images at the position shown in the PET image of Figure 1 is shown in Figure 4. The maximum likelihood estimate is typically noisy. Using the quadratic penalty we reduce noise at the cost of reduced resolution as we increase β . It can be seen that the images obtained using MI priors have well-defined edges and noise is reduced. The transaxial profiles show the images using MI priors following region boundaries more closely than the QP images. The reconstructions using the MI-Intensity prior have some high intensity pixels at random locations in the image. With the addition of spatial information through local mean and gradients, this occurrence is reduced and the images have a more homogeneous distribution of intensities within the boundaries. Normalized root mean squared errors between the true and reconstructed images are plotted as a function of iteration number in Figure 5. It can be seen that the overall normalized mean squared error is considerably lower for the MI priors in comparison with QP and ML.

The joint density estimate of the anatomical and true PET image is shown in Figure 6(a) and the joint density estimates of anatomical and reconstructed PET images using the MI and QP priors are shown in Figure 6(b)-(d). The x-axis in each of these figures corresponds to the PET image intensity and the y-axis corresponds to the anatomical image intensity. In Figure 6(a), the GM, WM and CSF regions in the true images are seen as three distinct peaks in the joint density. The smooth variation in intensities in the images obtained using QP are seen as long clusters in Figure 6(b). The random high pixel intensities seen in the MI-Intensity reconstructions are reflected in Figure 6(c) as small, isolated peaks away from the clusters corresponding to

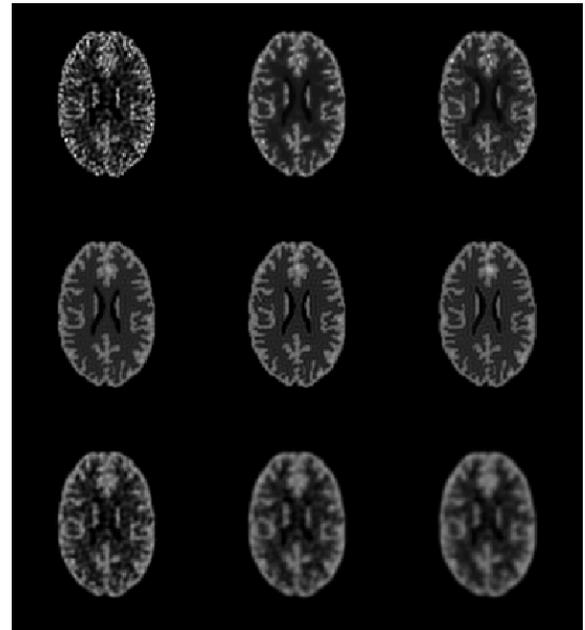


Fig. 3. Reconstructed images with the MI-Intensity prior (top) for $\mu=0, 5000$ and $1e4$, with the MI-Spatial prior (center) for $\mu=3000, 5000$, and 7000 , and QP(bottom) for $\beta=0.1, 0.5$, and 1.0

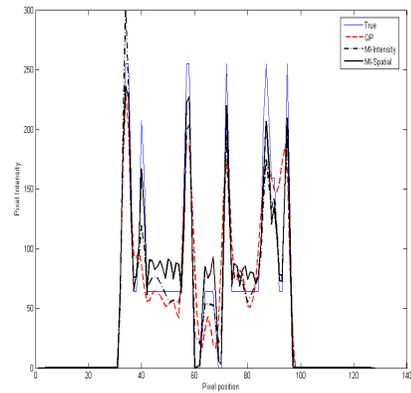


Fig. 4. Transaxial Profile of true and reconstructed PET images

the three regions. These isolated peaks are almost negligible in the MI-Spatial prior, where the joint density estimate is closest to that of the true images, with long clusters in the GM and WM regions, corresponding to a smooth variation in intensities in those regions.

To evaluate the performance of the MI priors in quantifying uptake, Monte Carlo simulations were performed. Forty datasets were generated with 800K counts each using the same scanner model as described above. The bias, standard deviation and root mean squared (RMS) error of the images reconstructed using MI-Intensity and MI-Spatial priors with $\mu = 5000$ and the QP with $\beta = 0.5$ are shown in Figure 7. The MAP estimate using QP has high bias along the boundaries and uniform standard deviation. The RMS error is also high along

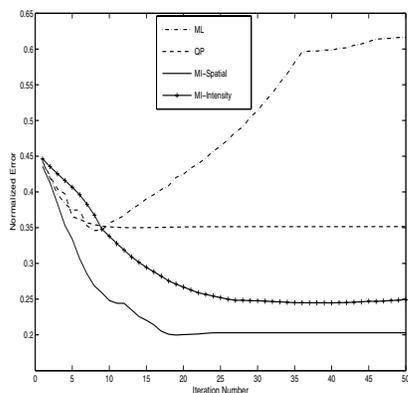


Fig. 5. Normalized mean squared errors as a function of iteration number for ML, QP with $\beta = 0.5$, MI-Intensity and MI-Spatial with $\mu = 5000$.

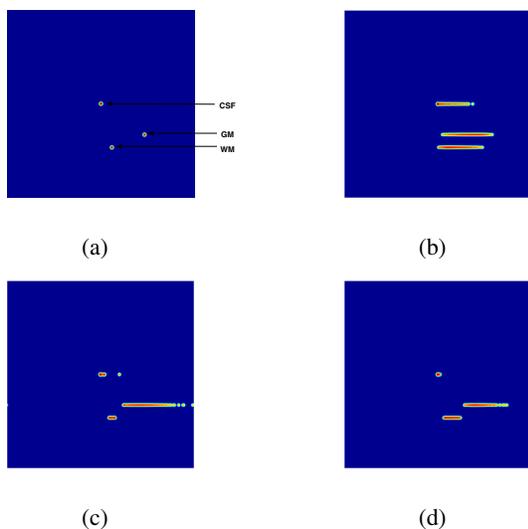


Fig. 6. Joint density estimates of anatomical and (a) true PET image (b) reconstructed PET image using QP with $\beta = 0.5$ (c) reconstructed PET image using MI-Intensity with $\mu = 5000$ (d) reconstructed image using MI-Spatial with $\mu = 5000$

the boundaries. The MI priors have lower bias than the QP, but it is distributed within the boundaries. The MI-Intensity prior has a structured standard deviation and RMS error that is higher in the GM, where the random high intensity pixels were seen. The standard deviation and RMS error for the MI-Spatial prior are low close to the boundaries and higher away from them, which is consistent with the fact that the MI-Spatial prior follows the true boundaries closely.

III. DISCUSSION

We proposed a method for incorporating anatomical information in PET image reconstruction through mutual information based priors. The images reconstructed using MI priors had well defined boundaries and less noise in comparison with those obtained using ML and QP. The overall rms error was also found to be considerably lesser for the MI priors. The

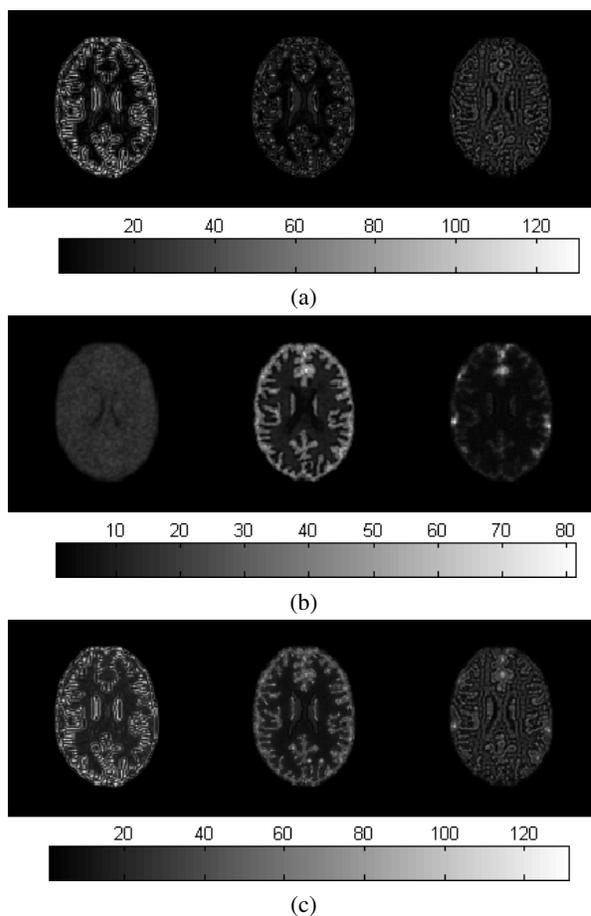


Fig. 7. (a) Bias, (b) standard deviation and (c) RMS error images of reconstructions using (L to R) QP with $\beta = 0.5$, MI-Intensity and MI-Spatial using $\mu = 5000$

MI-Spatial prior, which included spatial information in the feature vectors that were extracted from the anatomical and functional images performed the best, with accurate boundaries and homogeneous distribution of intensities within them.

These results represent a best-case scenario for the use of MI since the anatomical and functional images used in our simulation have identical boundary structure and are uniform within regions. We do not anticipate that this approach will prove useful for lesion detection in the form presented here, since the MI criteria seeks to maximize similarity between functional and anatomical images and can therefore not be expected to enhance lesions not visible in the anatomical image. The non-convexity of the MI prior makes it prone to local maxima, making it necessary to start with a good initial estimate. However, the results look promising and indicate that with a good choice of feature vectors that accurately describe the structure of the two images, this approach might be useful in applications where function follows anatomy, like those involving quantitation of uptake in specific organs.

REFERENCES

- [1] R. Leahy and X. Yan. Incorporation of MR data for improved functional imaging with PET. *Information Processing in Medical Imaging*, pp.105-

- 120, *Springer-Verlag*, 1991.
- [2] G.Gindi, M.Lee, A.Rangarajan and I.G.Zubal: Bayesian Reconstruction of functional images using anatomical information as priors. *IEEE Trans Med. Imag.*, vol.12(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(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. Imaging and Vision*, vol. 12, pp.199-217, 2000.
 - [5] W.Wells III, P.Viola, H.Atsumi, S.Nakjima and R.Kikinis. Multimodal volume registration by maximization of mutual information. *Medical Image Analysis*, vol.1(1), pp. 33-52, 1996.
 - [6] T.Cover and J. Thomas. *Elements of Information Theory*, Wiley Series in Telecommunications, 1991.
 - [7] D.Bertsekas. *Nonlinear Programming*, Athena Scientific, 1999.
 - [8] R.O. Duda, P.E. Hart and D.G. Stork, *Pattern Classification*, Second Edition, Wiley, 2001.