A Phantom Study of the Quantitative Behavior of Bayesian PET
Reconstruction Methods *

Erkan Ü. Mumcuğlu †, Richard M. Leahy‡, Zhenyu Zhou † and Simon R. Cherry †
† Crump Institute for Biological Imaging, UCLA School of Medicine, Los Angeles
‡ Signal & Image Processing Institute, University of Southern California, Los Angeles

Abstract

We examine the behavior of a Bayesian protocol for reconstruction of PET images in terms of the quantitative accuracy of the resulting Region-of-Interest (ROI) statistics. Our protocol uses a rapidly converging algorithm to compute a MAP estimate of the 3D PET image and includes novel methods for accurate attenuation and scatter modeling. The image is modeled using a Gibbs prior whose hyperparameters are estimated simultaneously with the PET image so that no critical user specified parameters are required. We computed ROI statistics over reconstructions of multiple frames of a computer generated brain phantom and an experimental chest phantom scanned using a Siemens/CTI ECAT931 scanner. Results indicate significant improvements in both bias and variance when compared to a standard filtered backprojection (FBP) based protocol.

1 Measurement Model and Calibration

We have previously described a fast and straightforward method for computing MAP estimates of 3-D PET images from 2-D data using Gibbs priors [4] [5]. In this paper, we will address the issue of whether this approach can lead to quantifiable improvements in image quality when compared to a standard imaging protocol.

The Bayesian reconstruction method under investigation uses a statistical model for the data that specifically includes randoms and scatter components as well as detector efficiency, dead-time and attenuation correction. The mean of randoms and scatter components are estimated before reconstruction of the PET image. The random component is estimated using a separate randoms sinogram (acquired simultaneously with the standard emission sinogram) through maximum likelihood estimation of the singles rate at each detector [6]. The scatter sinogram is computed by tracing scatter paths through a preliminary reconstruction of the emission image using a Bayesian reconstructed transmission image and the Klein-Nishina formula [6]. Efficient ray tracing techniques, the smoothness of the scatter profile, and a single scatter assumption are used to achieve fast computation. The attenuation correction factors (ACFs) are computed by reprojecting a transmission image which has been reconstructed using the Bayesian method described in [5], with simultaneous hyperparameter estimation [7].

The emission data are modeled as a set of independent Poisson random variables \( x_i \) with mean \( \bar{x}_i \) equal to the sum of three components: true coincidences \( \bar{c}_i \), randoms \( \bar{r}_i \), and scatter \( \bar{s}_i \). The only unknown parameters (\( A_j \) = emission intensities in each pixel) are related to the mean as follows

\[
\bar{x}_i = \bar{c}_i + \bar{r}_i + \bar{s}_i, \quad \bar{c}_i = \sum_j p_{ij} A_j.
\]

where \( \bar{c}_i \) and \( \bar{r}_i \) are first estimated using new methods described in [6]. The elements \( p_{ij} \) of the matrix \( P \) denote the probability of detecting an emission from pixel site \( j \) at detector pair \( i \). Sparsity of the \( P \) matrix is exploited by partitioning the detector efficiency, deadtime, geometric factors and attenuation into different submatrices [5]. In this way forward and backprojection can be computed efficiently. The model can be extended to also include positron range and angular separation effects, and inter-crystal penetration and scatter. However, these factors are second order in comparison to those included for the Siemens/CTI ECAT 931 scanner used in this work.

Using the model (1), we can write the log-likelihood function as

\[
\ln p(x|\lambda) = \sum_i \{-\bar{x}_i + x_i \ln(\bar{x}_i)} \}, \quad \bar{x}_i = \bar{c}_i + \bar{s}_i + \bar{r}_i
\]

where constants have been dropped for notational simplicity.

2 Bayesian Image Estimation

Our Bayesian method combines the likelihood model above with a Gibbs prior and computes a maximum a posteriori (MAP) estimate from the resulting posterior density. Our Gibbs priors have the form:

\[
p(\lambda) = \frac{1}{Z} e^{-\beta \sum_{i,j} k_{ij} V_{ij}(\lambda_i, \lambda_j)}
\]

where \( \beta \) is the hyperparameter that influences the degree of smoothness of the estimated images, and \( Z \) is the normalizing constant or partition function. The \( V_{ij}(\cdot) \) are pairwise

*This work was supported by a grant from the Whitaker Foundation and by the National Cancer Institute under Grant No. ROI CA57974.
potential functions defined on an 8-nearest neighbor system for 2-D images and a 26-nearest neighbor system for 3-D images. The $N_i$ denote the set of neighbors of pixel $i$. The constants $k_{ij}$ are the inverse of the Euclidean distance between the pixel centers for each pixel pair. In our previous work [5] [4], we have used primarily the Geman and McClure potential function and the quadratic potential function. Here, we consider the Huber potential functions as well:

$$V_{ij}(\lambda_i, \lambda_j) = \begin{cases} 
\frac{|\lambda_i - \lambda_j|}{\delta}, & |\lambda_i - \lambda_j| > \frac{\delta}{2} \\
\frac{\delta}{2}, & |\lambda_i - \lambda_j| \leq \frac{\delta}{2}.
\end{cases} \quad (4)$$

This prior provides less smoothing than the quadratic potential while avoiding the instabilities that are sometimes encountered with the Geman and McClure potential.

The MAP estimate of the image $\lambda$ is found as the maximizer of the log posterior density:

$$\ln p(\lambda|x) = \ln p(x|\lambda) - \beta \sum_i \sum_{j \in N_i} k_{ij} V_{ij}(\lambda_i, \lambda_j). \quad (5)$$

Once the calibration procedures have been performed to estimate the randoms and scatter sinograms, the only remaining un-specified variables in the log-posterior (5) are the image $\lambda$, and any parameters of the prior. The MAP estimate is computed using the gradient projection preconditioned conjugate gradient (GPPCG) method described in [4][6]. This method has similar convergence behavior to the penalized preconditioned conjugate gradient method described in [5] but has the advantage that no user specified parameters are required. The hyperparameter $\beta$ can be chosen heuristically (as $\beta$ is increased the MAP image estimate becomes increasingly smooth), or can be simultaneously estimated with the MAP images using the approximate maximum likelihood method described in [7]. In the case of the Huber potential, the variable $\delta$ must also be specified. The method in [7] can be modified to also estimate this parameter. However $\delta$ is not a critical variable as it affects only the rate of increase rather than the magnitude of the potential function and can be chosen heuristically. We use $\delta$ equal to $1/8$ of the image maximum, where we compute the image maximum using a preliminary FBP reconstruction.

3 Quantitative performance

3.1 Computer generated phantom

Realistic Poisson data was generated using a single plane of the 3-D Hoffman brain phantom, Figure (1). ACFs were computed assuming uniform attenuation within the head boundary. A 10% uniform randoms distribution was added based on the assumption that the singles distribution was uniform across the detectors. A scatter sinogram was computed by using the transmission and emission images in a Klein-Nishina based Compton scatter model. Scattered counts were set to approximately 10% of trues. To model variable detector sensitivity, the mean sinogram data were scaled using values for intrinsic sensitivity typical in a clinical scanner. The ACFs, randoms and scatter mean values were assumed to be known. The sinogram, detector, and image sizes were chosen to simulate the Siemens ECAT 831 PET scanner: sinogram size: $128 \times 160$, 320 detectors on a 64cm diameter ring, detector size: 0.628cm, image pixel size: 0.174cm. Perfect parallel projections were produced by accurately computing the areas of intersection of pixels and the rectangular bins formed by joining detector pairs. The original phantom image has 256 x 256 pixels with each pixel size: 0.087cm. Images were reconstructed on a $128 \times 128$ array with pixel size: 0.174cm. The phantom was scaled to produce an average of 1.6M (high count) or 400,000 (low count) coincidences per sinogram. Fifty realizations of the Poisson data were generated for both high and low counts.

PET images were reconstructed for each of the three priors. MAP images were estimated using the GPPCG algorithm for each of the fifty sets of data for both high and low count, for each of the priors and for a range of values of $\beta$.

Two ROIs, as marked in Figure (1), were selected for the quantitative study. One ROI is an entire structure in grey matter that includes boundary pixels, the other is an interior region of this structure and does not include boundaries. For these two regions, the bias and variance of the average intensity in the two ROIs were computed using an ensemble average over the 50 reconstructions. These statistics were computed using the three different priors for a range of hyperparameter values for the low and high count data.

For each prior, bias-squared versus variance were plotted, Figure (2), for all hyperparameter values. Similar plots were also made for images reconstructed using the FBP algorithm with Shepp filters of various roll-off frequency. We also indicate on each curve the values of ROI bias and variance corresponding to the reconstructions obtained using simultaneous MAP reconstruction and hyperparameter estimation. Note that the bias squared versus variance curves for the Bayesian method with Huber and quadratic poten-
tials lie below those for FBP. This indicates that for any bias level, the Bayesian methods give lower variance than FBP, or equivalently the total mean squared error for the Bayesian methods is always lower than for FBP. The Geman and McClure prior shows substantially higher variance than the other priors and generally performs worse in quantification than FBP. Finally, note that automatic selection of the hyperparameter for the Huber and quadratic potentials results in close to the minimum mean squared ROI quantitation error in Figure (2).

3.2 Experimental Chest Phantom

A phantom study was performed to evaluate the quantitative behavior of the reconstruction method described above. In this experiment, 40 frames of 1min emission data were acquired using an axially and transaxially asymmetric chest phantom, with cylindrical inserts of different activity levels. The data were collected using the 15 plane, Siemens/CTI ECAT931 scanner. The system was configured to acquire a simultaneous randoms sinogram; the emission data were not randoms corrected. I and 30 minute transmission scans were collected using an external ring source. Again, a randoms sinogram was acquired separately. In this experiment all 15 planes were reconstructed for each data set.

The relative activities in the compartments in the phantom were accurately measured (by counting samples) immediately before the experiment. Activities and placement of the compartments were as follows: comp. #1 between planes 1-15 (relative activity = 1.0); comp. #2:1-15 (1.293); comp. #3:1-15 (1.128); comp. #4:1-13 (1.743), comp. #5:11-13 (3.820); comp. #6:4-12 (0.0 - water); comp. #7:1-9 (0.0 - air); In all reconstructions, the image size is 128*128, the pixel size is 0.36 cm by 0.36 cm, and the axial sample interval is 0.675 cm. Several ROIs were specified, at least one for each compartment. Each ROI was chosen as a cylindrical volume, in the interior of one compartment, of size 75 pixels (25 pixels in each plane and covering 3 planes). The volume of each ROI is 6.561cm³ (3.24cm² by 2.025cm). The locations of the ROIs as they appear in the 8th and 11th reconstruction planes are shown in Figure (3).

Our purpose here was to compare the quantitative behavior of the Bayesian method with the standard clinical procedure for FBP-based quantitation. In the FBP reconstructions we used a ramp filter as this was found to give a lower mean squared error in ROI quantitation than when a smoothing filter was used. The scatter component of the sinogram was estimated in the same way as for the MAP estimator and subtracted from the data (randoms subtracted directly without smoothing). Attenuation factors were computed using the ratio of the smoothed blank and transmission scans, with different amounts of smoothing applied depending on transmission data duration [2]. In the case where the 30min transmission data was used for attenuation correction, a 2-D Gaussian smoothing kernel of size 9 by 9 and FWHM of 8.85 mm was convolved twice with the detector-efficiency normalized transmission

![Figure 2: ROI variance versus squared bias plot. At each point the summation of the horizontal and vertical coordinates equals the total mean squared error. First two rows: high count test; Last two rows: low count test. The ROIs are marked in Figure (1). In the figure, solid line = MAP with quadratic prior; dot-dashed line = MAP with Huber prior; dashed line = MAP with Geman and McClure prior; dotted line = filtered backprojection. The star, circle and cross signs indicate locations corresponding to the hyperparameter estimate.](image-url)
and blank sinograms. For the 1min transmission data, a FWHM of 15.33 mm was used. This type of smoothing results in resolution mismatch between the emission data and ACFs [3]. It is possible that resolution matching would produce some improvement in FBP quantitation, however this is not standard clinical procedure and we have not done so here.

Sample reconstructions are shown in Figure (4). All figures correspond to the 11th plane of the phantom. The first column shows reconstructions from the 2nd 1min frame of data for FBP with 1min and 30min attenuation correction and MAP with quadratic and Huber priors with 1min attenuation correction (from top to bottom: FBP1, FBP2, MAP1, MAP2 respectively). The second column shows the mean for this plane. Visual inspection of the left and right column clearly indicates that the MAP reconstructions have much lower variance than the FBP images. (Note the similarity of right and left column in MAP reconstructions.) The mean values for FBP using the 1min transmission data for attenuation correction also show considerable bias in several areas in the form of shadowing and streaking artifacts.

To examine the quantitative behavior of the new method, we computed the mean and standard deviation of the ROI values by averaging over the values computed from the 40 independent images reconstructed from the 40 1min frames. Both the FBP and MAP images were compensated for the exponential decay of the source over the duration of the experiment before any statistics were computed.

Since only the relative activities in the ROIs was measured, we computed the mean and standard deviation of these relative values. In this case, the choice of the background ROI against which the other ROI values are compared can have a large influence on the resulting accuracy. In Table (1) and Table (2), we show the relative activities computed using ROIs 1a and 1b respectively. Note that the mean activities in regions 1a and 1b vary considerably when using FBP due to bias effects resulting from the inaccurate attenuation correction. Consequently we see

Figure 3: MAP emission reconstructions of the phantom showing the location of the ROIs in the 8th and 11th reconstruction planes.

Figure 4: A single plane from the 3-D reconstructions of chest phantom. Left column: reconstruction from a single 1min frame; Right column: mean image averaged over the reconstructions from all 40 frames. First row: (FBP1) Ramp filter FBP reconstructions using ACFs from 1min transmission scan; Second row: (FBP2) Ramp filter FBP reconstructions using ACFs from 30min transmission scan. Third row: (MAP1) MAP reconstructions using quadratic prior and MAP ACFs from 1min transmission scan. Fourth row: (MAP2) MAP reconstructions using Huber prior and MAP ACFs from 1min transmission scan.
<table>
<thead>
<tr>
<th>ratio</th>
<th>true</th>
<th>$F B P_1$</th>
<th>$F B P_2$</th>
<th>$M A P_1$</th>
<th>$M A P_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/1a</td>
<td>1.293</td>
<td>1.04 ± 0.11</td>
<td>1.15 ± 0.14</td>
<td>1.33 ± 0.04</td>
<td>1.33 ± 0.03</td>
</tr>
<tr>
<td>3/1a</td>
<td>1.128</td>
<td>0.92 ± 0.10</td>
<td>1.02 ± 0.12</td>
<td>1.18 ± 0.04</td>
<td>1.20 ± 0.03</td>
</tr>
<tr>
<td>4/1a</td>
<td>1.743</td>
<td>1.71 ± 0.13</td>
<td>1.67 ± 0.13</td>
<td>1.75 ± 0.04</td>
<td>1.73 ± 0.04</td>
</tr>
<tr>
<td>5/1a</td>
<td>3.820</td>
<td>2.83 ± 0.21</td>
<td>3.07 ± 0.24</td>
<td>2.87 ± 0.07</td>
<td>3.31 ± 0.06</td>
</tr>
<tr>
<td>6/1a</td>
<td>0.000</td>
<td>-0.04 ± 0.07</td>
<td>0.01 ± 0.09</td>
<td>0.07 ± 0.02</td>
<td>0.08 ± 0.02</td>
</tr>
<tr>
<td>7/1a</td>
<td>0.000</td>
<td>0.18 ± 0.07</td>
<td>0.10 ± 0.07</td>
<td>0.15 ± 0.02</td>
<td>0.16 ± 0.02</td>
</tr>
</tbody>
</table>

Table 1: The mean and standard deviation of relative ROI values computed from reconstructions of the 40 1min frames. Here ROI 1a was used as the background activity against which the other activities were computed.

<table>
<thead>
<tr>
<th>ratio</th>
<th>true</th>
<th>$F B P_1$</th>
<th>$F B P_2$</th>
<th>$M A P_1$</th>
<th>$M A P_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/1b</td>
<td>1.293</td>
<td>1.64 ± 0.26</td>
<td>1.28 ± 0.21</td>
<td>1.32 ± 0.04</td>
<td>1.37 ± 0.04</td>
</tr>
<tr>
<td>3/1b</td>
<td>1.128</td>
<td>1.45 ± 0.25</td>
<td>1.14 ± 0.19</td>
<td>1.17 ± 0.04</td>
<td>1.23 ± 0.03</td>
</tr>
<tr>
<td>4/1b</td>
<td>1.743</td>
<td>2.68 ± 0.40</td>
<td>1.86 ± 0.26</td>
<td>1.73 ± 0.05</td>
<td>1.77 ± 0.04</td>
</tr>
<tr>
<td>5/1b</td>
<td>3.820</td>
<td>4.45 ± 0.64</td>
<td>3.42 ± 0.45</td>
<td>2.84 ± 0.08</td>
<td>3.39 ± 0.09</td>
</tr>
<tr>
<td>6/1b</td>
<td>0.000</td>
<td>-0.06 ± 0.11</td>
<td>0.01 ± 0.11</td>
<td>0.07 ± 0.02</td>
<td>0.08 ± 0.02</td>
</tr>
<tr>
<td>7/1b</td>
<td>0.000</td>
<td>0.29 ± 0.12</td>
<td>0.11 ± 0.08</td>
<td>0.15 ± 0.02</td>
<td>0.16 ± 0.02</td>
</tr>
</tbody>
</table>

Table 2: The mean and standard deviation of relative ROI values computed from reconstructions of the 40 1min frames. Here ROI 1b was used as the background activity against which the other activities were computed.

4 Conclusions

The results of applying the Bayesian protocol described above and in [6] appears to show significant quantitative improvements in comparison to the standard clinical FBP-based procedures. In comparing the Bayesian method with FBP-based methods, further investigation is necessary to take into account recent improvements in FBP-based methods [3] [1]. Our results support previous observations of qualitative improvements when using a Bayesian approach [4]. The differences between FBP and Bayesian methods are particularly significant when the transmission and emission data are collected with short frame durations. Consequently the method described here may be important for quantitative whole body studies or for dynamic studies requiring rapid frame rates.

References


