# Comparison between MAP and post-processed ML for image reconstruction in emission tomography when anatomical knowledge is available.

Johan Nuyts, *Member, IEEE,* Kristof Baete, *Student Member, IEEE,* Dirk Bequé, *Student Member, IEEE,* and Patrick Dupont

***Abstract*— Previously, the noise characteristics obtained with penalized-likelihood reconstruction (or maximum-a-posteriori, MAP) have been compared to those obtained with post-smoothed maximum-likelihood (ML) reconstruction, for emission tomography applications requiring uniform resolution. It was found that penalized-likelihood reconstruction was not superior to post-smoothed ML. In this study a similar comparison is made, but now for applications where the noise suppression is tuned with anatomical information. It is assumed that limited but exact anatomical information is available. Two methods were compared. In the first method, the anatomical information is incorporated in the prior of a MAP-algorithm, and is therefore imposed during MAP-reconstruction. The second method starts from an unconstrained ML-reconstruction, and imposes the anatomical information in a post-processing step. The theoretical analysis was verified with simulations: small lesions were inserted in two different objects, and noisy PET data were produced and reconstructed with both methods. The resulting images were analyzed with bias-noise curves, and by computing the detection performance of the non-prewhitening observer and a channelized Hotelling observer. Our analysis and simulations indicate that the post-processing method is inferior, unless the noise correlations between neighboring pixels are taken into account. This can be done by applying a so-called prewhitening filter. However, because the prewhitening filter is shift variant and object dependent, it seems that MAP reconstruction is the more efficient method.**

## I. Introduction

In previous studies, the noise performance of post-smoothed maximum-likelihood (ML) reconstruction has been compared to that of penalized-likelihood reconstruction, for emission tomography applications requiring "uniform" spatial resolution [1], [2]. Here, resolution is considered uniform if the local impulse response has spherical symmetry and is independent of the object and of the position in the image. The results indicated that noise suppression with penalized-likelihood is not superior to that with post-smoothed ML for these applications; a theoretical argument is given in [2].

There is an increasing interest in the combination of anatomical and functional image information. Promising results have been obtained with maximum-a-posteriori (MAP) reconstruction of emission data, by using the anatomical information to modify the prior as a function of position [3], [4]. In such images, the resolution is strongly position and orientation dependent, and therefore, the conclusions of the studies mentioned above may not apply. In this work, we compare the noise characteristics obtained with MAP reconstruction using anatomical priors, with those obtained with post-processing the unconstrained ML reconstruction. The aim of this work is to identify which of the two approaches is most efficient for anatomy-based regularization. It is assumed that limited but exact anatomical information is already available, the problems of extracting the anatomical information and aligning it to the emission images are not studied here.

The following section describes the MAP and post-processing methods, designed to improve the reconstruction by using anatomical information. The theory requires a prewhitening operation in the post-processing method, and (approximate) prewhitening filters are derived. Because accurate prewhitening is rather complex, also a simplified, non-prewhitening method is studied. In the third section, two simulation experiments are described. The theory predicts inferior noise characteristics for the post-processing method, unless accurate prewhitening is applied. To verify this, the image quality obtained with the different methods must be quantified, which is not a trivial problem. We have studied image quality by plotting bias-noise curves, and by computing the performance of two different numerical observers: the non-prewhitening observer and a channelized Hotelling observer.

In the first experiment, the theory is verified with a very artificial phantom: this phantom experiment does not resemble any clinical application, but it allows us to study the performance of the different methods in carefully selected situations. The second experiment verifies the results in a more realistic experiment: a 2D positron emission tomography (PET) brain study is simulated, taking into account attenuation and finite detector resolution. The findings are discussed in section four.

## II. Theory

### A. ML-reconstruction

In emission tomography, the measured number of photons $y_i$ for detector $i$, can be regarded as a sample of a Poisson distribution, leading to the following expressions:

$$p(y_i|\bar{y}_i) = \frac{e^{-\bar{y}_i}\bar{y}_i^{y_i}}{y_i!} \quad (1)$$

$$\bar{y}_i = \sum_j a_{ij}\lambda_j + t_i \quad (2)$$

Work supported by K.U.Leuven grants OT-00/32 and IDO-99/005, and F.W.O. grant G.0174.03

Nuclear Medicine, K.U.Leuven, B-3000 Leuven, Belgium. (e-mail: Johan.Nuyts@uz.kuleuven.ac.be)

where $p(y_i|\bar{y}_i)$ is the probability of measuring $y_i$ photons when $\bar{y}_i$ are expected, $\lambda_j$ is the amount of radioactivity in pixel $j$, $a_{ij}$ is the probability that the radioactivity in $j$ contributes to the count level in detector $i$, and $t_i$ is the expected contribution due to e.g. randoms or scatter and is assumed to be known. In clinical practice, the available data have often been precorrected by the acquisition system. However, the data $y_i$, and the values of $a_{ij}$ and $t_i$ can often be manipulated to approximately restore the Poisson characteristics, such that (2) can still be applied [5], [6]. ML-reconstruction seeks the image with pixel values $\lambda_j$, that maximizes the likelihood (1). It is equivalent to maximize

$$\begin{aligned} L(Y,\Lambda) &= \sum_i y_i ln(\bar{y}_i) - \bar{y}_i \qquad (3) \\ &= \sum_i y_i \ln(\sum_j a_{ij}\lambda_j + t_i) - (\sum_j a_{ij}\lambda_j + t_i), \end{aligned}$$

which is obtained by taking the logarithm of (1) and deleting terms independent of $\lambda_j$ [7]. As argued in [1], [2], a post-smoothed version of the maximum-likelihood image has excellent resolution and noise characteristics, which can be obtained in practice by post-smoothing after many iterations of the EM [7] or OS-EM algorithm [8].

### *B. MAP-reconstruction*

As mentioned above, the problem of obtaining and aligning anatomical boundaries from another modality is ignored here. It is assumed that this knowledge is available, and we only study how it can be incorporated most effectively in the emission images. The particular choices of objects and anatomical priors that have been used in the experiments are already presented here. Note, however, that the theoretical arguments given below are more general, as they depend on the likelihood, but not on the choice of the anatomical prior.

To study MAP reconstruction with an anatomy-based prior, we consider a phantom consisting of 3 different tissue classes and one background class. The classes are known, the activity has to be reconstructed from noisy sinogram data. This phantom is inspired by ongoing work on brain imaging with PET in epilepsy [9], [10], where the three classes, gray matter, white matter and cerebrospinal fluid (CSF), are obtained from a segmented MR-image. The three classes will be denoted accordingly as $G$ for gray matter, $W$ for white matter and $C$ for CSF. Lesions of reduced tracer uptake are only expected in $G$. The expected tracer uptake for $C$ is known a-priori (for many tracers, there is no uptake in the CSF), and the tracer uptake in $W$ is expected to be uniform, but with unknown mean. Extension to applications with a different number of classes and different class features is (at least in principle) straightforward.

It is assumed that the available anatomical knowledge can be represented with a single label in every pixel $j$. A pixel belongs entirely to a single class, partial volume effects will not be considered here. The anatomical information is exploited by combining the likelihood with a prior. The logarithm of the prior equals $M_G + M_W + M_C$:

$$\begin{aligned} M_G(\Lambda) &= -\beta_G \sum_{j\in G}\sum_{k\in G} \alpha_{j-k} \frac{(\lambda_j - \lambda_k)^2}{\lambda_j + \lambda_k + \gamma|\lambda_j - \lambda_k|} \qquad (4) \\ M_W(\Lambda) &= -\beta_W \sum_{j\in W} (\lambda_j - \text{mean}_W(\Lambda))^2 \qquad (5) \\ M_C(\Lambda) &= -\beta_C \sum_{j\in C} (\lambda_j - \lambda_C)^2, \qquad (6) \end{aligned}$$

where $G$, $W$ and $C$ are the classes, $\beta_G$, $\beta_W$ and $\beta_C$ are parameters determining the prior strength for each class, and $\lambda_C$ is the expected mean of class $C$, which is assumed to be known. The weight $\alpha_{j-k}$ defines the neighborhood; we used $\alpha_{j-k} = 1$ for horizontal and vertical neighbors $j$ and $k$, $\alpha_{j-k} = 1/\sqrt{2}$ for diagonal neighbors and 0 otherwise. The function $\text{mean}_W$ returns the mean value over the class $W$. $M_G$ is a prior favoring smoothness by penalizing relative differences between pixels. The parameter $\gamma$ determines the point where the prior becomes somewhat edge tolerant [11]. We used $\gamma = 10$, resulting in edge tolerance for relative differences larger than 20%. $M_W$ favors smoothness by penalizing deviations from the mean of class $W$, and $M_C$ penalizes differences from the expected uptake value $\lambda_C$. The reconstruction $\Lambda$ is obtained by maximizing

$$L_{\text{MAP}} = L(Y,\Lambda) + M_G(\Lambda) + M_W(\Lambda) + M_C(\Lambda). \qquad (7)$$

In the experiments described below, this is done with the gradient ascent algorithm described in [11]. Finally, we arbitrarily set $\beta_W = \beta_C = \beta_G/3 = \beta$, and varied the value of $\beta$ over about two orders of magnitude to evaluate the effect of the prior.

### *C. Post-processing the ML-image*

The post-processing problem will be treated as an optimization problem; this allows for comparison to the MAP-approach (which is also an optimization problem). In addition, as will be explained below, this approach allows us to take into account the noise covariance between neighboring pixels in the ML-image. Consequently, the objective function should have the following form

$$L_{\text{PML}} = \Phi(X,\Lambda) + M_G(\Lambda) + M_W(\Lambda) + M_C(\Lambda), \qquad (8)$$

where the first term describes the similarity between the input image $X$, the ML-image, and the output image $\Lambda$, the post-processed image. The last three terms are the same priors defined above. This approach allows us to suppress the noise with the same priors, either during the reconstruction (MAP), or after the ML reconstruction (post-processed ML).

Ideally, the function $\Phi(X,\Lambda)$ should represent the logarithm of the probability to obtain the ML-image $X$, given that the true tracer distribution is $\Lambda$. An approximate expression for that probability can be obtained by investigating the characteristics of the ML-algorithm at the maximum of the likelihood, as proposed by Fessler [12]. To study mean and variance, a second order Taylor series expansion of $L(Y,\Lambda)$

is computed at $\Lambda = X$. The first derivative of $L(Y, \Lambda)$ with respect to $\Lambda$ equals (see (3)):

$$\frac{\partial L(Y,\Lambda)}{\partial \lambda_j} = \sum_i a_{ij} \frac{y_i - \bar{y}_i}{\bar{y}_i}. \quad (9)$$

The second derivatives of $L(Y, \Lambda)$ with respect to $\Lambda$ equal

$$\frac{\partial^2 L(Y,\Lambda)}{\partial \lambda_j \partial \lambda_k} = -\sum_i \frac{y_i a_{ij} a_{ik}}{{\bar{y}_i}^2}. \quad (10)$$

By definition, the Fisher information matrix $F$ equals minus the expectation (over $Y$) of the second derivatives, so we have:

$$F(j,k) = \sum_i \frac{a_{ij} a_{ik}}{\tilde{y}_i} \simeq -\frac{\partial^2 L(Y,\Lambda)}{\partial \lambda_j \partial \lambda_k}\bigg|_{\Lambda = X}, \quad (11)$$

where $\tilde{y}_i$ denotes the expectation of $y_i$. It is convenient to write the expansion in matrix notation. The sinograms $Y$ and $\bar{Y}$, and the images $X$ and $\Lambda$ are represented as column matrices. $A$ is the system matrix with elements $a_{ij}$ and is assumed to have maximum rank. $C_y$ is the covariance matrix of the data; because the measurements $y_i$ are independent Poisson variables, $C_y$ is diagonal with $C_y(i,i) = \tilde{y}_i$. With this notation, and using (11) the second order Taylor expansion becomes

$$\begin{aligned} L(Y,\Lambda) &\simeq L(Y,X) + (\Lambda - X)' A' C_y^{-1} (Y - \bar{Y}) \\ &\quad -\frac{1}{2}(\Lambda - X)' F (\Lambda - X). \end{aligned} \quad (12)$$

Setting the first derivative of (12) to zero, one can solve for $\Lambda - X$:

$$\Lambda - X = F^{-1} A' C_y^{-1} (Y - \bar{Y}), \quad (13)$$

which yields for the covariance matrix of $\Lambda$ (due to the noise on $Y$):

$$C_\lambda = F^{-1}\ A'\ C_y^{-1}\ C_y\ C_y^{-1}\ A F^{-1} = F^{-1}, \quad (14)$$

because $F = A' C_y^{-1} A$.

Now consider the expansion at the solution $X$, evaluating the derivatives at $\Lambda = X$, and setting $Y$ equal to the measured sinogram. Because $L$ has a maximum at $X$, it follows that

$$0 = \frac{\partial L(Y,\Lambda)}{\partial \Lambda}\bigg|_{\Lambda = X} = A' C_y^{-1} (Y - \bar{Y}) \quad (15)$$

and as a result, the expansion (12) reduces to

$$L(Y,\Lambda) \simeq L(Y,X) - \frac{1}{2}(\Lambda - X)' F (\Lambda - X). \quad (16)$$

Thus, given $Y$ and $X$, (16) approximates the likelihood $L(Y, \Lambda)$ as the logarithm of a multivariate Gaussian with mean $X$ and covariance matrix $F^{-1}$. Consequently, setting $\Phi(X, \Lambda)$ of (8) equal to the right hand side of (16) ensures that the post-processing method takes into account the noise propagation from the raw data into the ML-image $X$.

Dropping the constant factor and term of (16), we have:

$$\begin{aligned} \Phi(X,\Lambda) &= -(\Lambda - X)' F (\Lambda - X) \\ &= -(F^{\frac{1}{2}}\Lambda - F^{\frac{1}{2}} X)' (F^{\frac{1}{2}}\Lambda - F^{\frac{1}{2}} X). \end{aligned} \quad (17)$$

$F$ is positive definite, so $F^{1/2}$ exists [13]. $F^{1/2}$ acts as a prewhitening filter: it decorrelates the noise. After decorrelation, the discrepancy between the two images is simply computed as the sum of squared pixel differences. In emission tomography, $F^{1/2}$ is a low-pass filter.

The relation $F^{-1} = C_\lambda$ only holds for efficient estimators [13], [14]. MLEM is not an efficient estimator, and the relation is only an approximation. However, this approximation has been used previously with good results [1], [12], [15], [16].

The derivation assumes that $X$ is a maximizer of the (unconstrained) likelihood function. This assumption is only reasonable if a sufficiently high number of MLEM iterations has been applied to obtain good convergence.

### *D. Post-processing with approximate prewhitening*

The prewhitening filter $F^{1/2}$ is difficult to implement, because it is a shift variant filter. We propose a few additional approximations to obtain a filter that is tractable in practice.

*1) Local optimization (filter $F_1$):* If the position of the lesion is approximately known a-priori, one can attempt to achieve good local performance by computing a shift invariant filter $F_1$, that is (nearly) identical to $F$ at that particular position in the image, as proposed in [17], [16]. This is obtained by replacing all rows (or columns) of the matrix $F$ with shifted versions of the row (or column) $j$, where $j$ is the pixel index corresponding to the center of the lesion. Then, $F_1$ or $F_1^{1/2}$ can be computed using the Fourier transform, since $F_1$ is shift invariant. Direct application of (11) gives the following recipe to compute $F_1$: 1) make an image with zeros everywhere and set pixel $j$ to 1 ($j$ is the lesion center); 2) compute the projection of that image (taking into account attenuation, detector blurring etc); 3) divide by the noise-free sinogram (or an estimate of that sinogram) on a pixel by pixel basis; 4) backproject this sinogram. The resulting blurred image is the point spread function of the filter $F_1$. Filtering images with $F_1$ or $F_1^{1/2}$ (as prescribed in (17)) can be done with FFT.

In our experiments, the solution $\Lambda$ maximizing $L_{\text{PML}}$ was found with a gradient ascent algorithm. The gradient of the function $\Phi(X, \Lambda)$ equals:

$$\nabla \Phi_1(X,\Lambda) = F_1^{1/2}(F_1^{1/2} X - F_1^{1/2}\Lambda) = F_1(X - \Lambda), \quad (18)$$

where the subscript of $\Phi_1$ was added because $F$ was approximated as $F_1$. In [18], we have illustrated with simulations that $F_1^{1/2}$ is indeed effective as a prewhitening filter, and much more so than a Gaussian low-pass filter.

*2) Lesion independent approximation (filter $F_2$):* In clinical practice, the position of the lesion is usually unknown. In addition, it is much easier to work with a predictable, object-independent filter. For that purpose, we calculate the shape of the filter in absence of attenuation and with a uniform sinogram $\tilde{y}_i$. In that case $\tilde{y}_i$ becomes an irrelevant constant in (11). The computation is the same as above, but without the division by $\tilde{y}_i$, and without position dependencies such as attenuation and position dependent blurring in the system matrix $A$. A profile through the point spread function of $F_2^{1/2}$ is shown in figure 1 for a projector with ideal resolution. It is sharp in the center, but decays only slowly to zero. The full width at half maximum (FWHM) is about 1.5 pixels. If finite (position independent) detector resolution is incorporated, the point spread function will become wider.

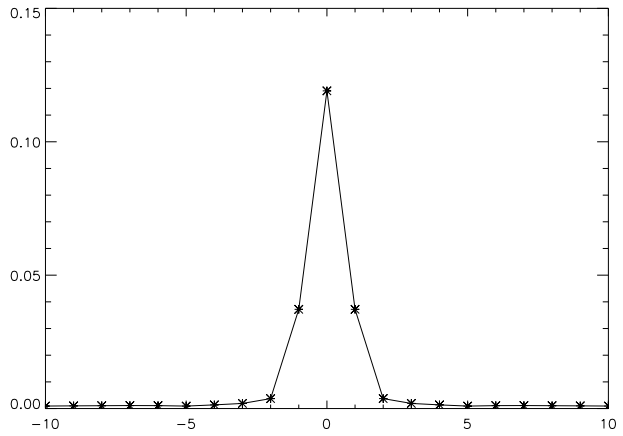


Fig. 1. Radial profile of the 2D convolution kernel of the shift invariant and data independent prewhitening kernel $F_2^{1/2}$. The FWHM is 1.5 pixels.

*3) Small convolution kernel (filter $F_3$):* A third approximation is obtained by extracting a small convolution kernel from the point spread function of $F_2^{1/2}$. With this kernel $F_3^{1/2}$, (18) can be computed with convolutions instead of with filtering in the Fourier domain. However, because of the truncation, its prewhitening performance is expected to be less than that of $F_2^{1/2}$. In our experiments, we used a kernel of 7 × 7 pixels.

### *E. Post-processing without prewhitening (NPW)*

It was shown in [1] that a quadratic smoothing prior of the form $\Lambda' R \Lambda$ combined with a likelihood of the form $(\Lambda - X)'(\Lambda - X)$ produces a predictable and position independent local impulse response. In other words, it is equivalent to smoothing with a predefined low-pass kernel. Similarly, we assume that "straightforward" post-processing, obtained by restricting a smoothing operator to within an anatomical boundary, corresponds to applying (8) with $\Phi(X, \Lambda) = (\Lambda - X)'(\Lambda - X)$, which is (17) without the prewhitening filter. Note that for non-quadratic priors, as in our experiments, the smoothing cannot in general be described as a simple restricted convolution, even in this "straightforward", non-prewhitening approach.

## III. Experiments

### *A. Simulated objects*

PET studies of two different objects have been simulated. The first object is a Shepp-Logan phantom, shown in figure 2. This object was not intended to be similar to a clinical image; the main purpose was to analyze the reconstruction at different lesion locations. The relative intensities of the three classes $G$, $W$ and $C$ was 4, 2 and 1 respectively. Five lesions were used to assess algorithm performance. The lesions are small squares; they are indicated in figure 2. All lesion pixels are of class $G$ (class $G$ is depicted with the highest intensity in fig 2). In the "lesion absent" case, the intensity in the lesion equals the normal class $G$ value of 4. In the "lesion present" case, the intensity is reduced with 20%. Lesion 0 is an island

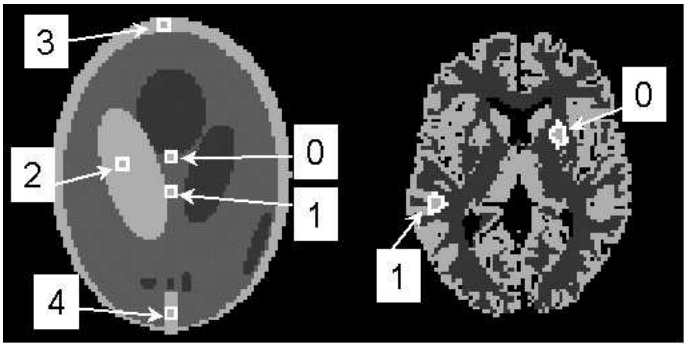


Fig. 2. The simulated objects, the lesion boundaries are indicated in white. Left: SHEPP phantom; right: BRAIN phantom.

of class $G$, therefore the boundary of the lesion coincides with the anatomical boundary. Three sides of lesion 1 match the boundary of class $G$ as well, the fourth side touches the large class $G$ ellipse. Lesion 2 is located near the center of the class $G$ ellipse, so its boundary is entirely invisible in the anatomy. Finally, lesions 3 and 4 are located near the object boundary, where the Fisher information is expected to be most asymmetrical, because both the total attenuation and total activity along the projection lines change strongly with projection angle. In these two lesions, the approximate prewhitening with the symmetrical filter $F_2$ is expected to be inferior to the prewhitening with the lesion dependent filter $F_1$.

The second object is a slice from the brain phantom, provided by the BrainWeb database at McGill University [19], [20]. The classes $G$, $W$ and $C$ were assigned relative activities of 4, 1 and 0, corresponding to the normal uptake of $^{18}$F-fluorodeoxyglucose. Two lesions were analyzed: the boundary of lesion 0 coincides with the anatomical boundary, while lesion 1 has only partial agreement with anatomy (fig 2). Again, the "lesion present" case corresponded to a 20% reduction of the activity in the lesion.

### *B. PET simulation and reconstruction*

Both phantoms were digitized at 100 × 100 pixels. PET sinograms of 100 lines of response per angle, and 144 projection angles were computed. Attenuation and position independent detector blurring were included in the system matrix $A$. The attenuation was that of water at 511 keV (0.096 $\text{cm}^{-1}$), and a pixel size of 1.5 mm was assumed for the Shepp phantom and 2.18 mm for the brain phantom. No randoms or scatter contribution was assumed: $t_i = 0$ in (2). The total number of counts was $1.3 \times 10^6$ for the brain phantom, and $0.68 \times 10^6$ for the Shepp phantom. For each object and each lesion position, 1 noise-free and 200 noisy sinograms (Poisson noise) with and without the lesion were computed, and reconstructed with 6 algorithms: MLEM with Gaussian post-smooth, MAP with anatomical prior, and MLEM post-processed for anatomy (pML) with filters $F_1$, $F_2$, $F_3$ and without prewhitening filter (NPW). The projection and backprojection operations in the ML and MAP reconstruction used the same system matrix $A$ that was used for the simulations. The ML and MAP algorithms were accelerated with ordered subsets: a decreasing

number of subsets was used to avoid convergence to a limit cycle: the scheme was (main iterations × number of subsets): $2\times36, 2\times24, 1\times16, 1\times12, 1\times8, 1\times4, 4\times1$, which is roughly equivalent to 165 regular iterations. For each algorithm, 6 different strengths of the regularization were applied: for ML, the width of the Gaussian post-smoothing kernel was varied, while for MAP and pML the value of $\beta$ was changed.

### C. Performance evaluation

As argued above, expression (17) takes the noise properties of the image $X$ into account. The strong covariances between neighboring pixels imply that neighboring pixels carry significant information about one another; the spread of that information is described by the Fisher information matrix (see e.g. fig. 1). Optimal prewhitening ensures that all information about a pixel is used when the final value of that pixel is computed. Applying poor or no prewhitening during post-processing is equivalent to reconstructing a pixel value using less information than available. Consequently, it is expected that image quality will become poorer when the post-processing is based on poorer approximations of the prewhitening filter $F$.

It is not trivial to quantify image quality. Because this paper is about reconstruction in PET, the quality should be measured either by a quantification task such as tracer kinetic modeling, or by a detection task, ideally the detection of realistic lesions in realistic images by human observers. To keep the evaluation feasible, several simplifications are introduced.

*1) Bias - noise curves:* One way to quantify PET images is to compute parametric images on a pixel by pixel basis. This task is sensitive to both bias and noise. For that reason, bias-noise curves are computed for the pixel in the center of the lesion, both in the lesion-present and lesion-absent images. The bias is computed as the difference with the true pixel value, the noise as the standard deviation of the noise realizations. The results are combined in a single curve per algorithm, that plots the root mean square of all biases versus the root mean square of all standard deviations. The summation is over all lesion positions and over lesion-present and lesion-absent cases, and the curve is obtained by changing the regularization parameter. Consequently, there are six points per curve, and one curve per algorithm.

*2) NPW and channelized Hotelling observers:* For the detection, we used the simplest paradigm, the "signal known exactly, background known exactly" or SKE/BKE task [21]. Two numerical observers have been applied: the non-prewhitening observer (NPWO) and a channelized Hotelling observer (CHO). Although detection tasks with random signals, background and location are more realistic, we preferred to exclude all other sources of randomness, focusing on propagation of noise from the data into the final image.

To decide if a lesion is present or absent in the noisy image $\Lambda$, the NPWO computes the statistic $q_{\mathrm{npw}}$:

$$q_{\mathrm{npw}} = (\bar{\Lambda}_1 - \bar{\Lambda}_0)'\Lambda, \tag{19}$$

where $\bar{\Lambda}_1$ and $\bar{\Lambda}_0$ are the noise-free images with and without the lesion. If this value exceeds the decision threshold, NPWO decides that the lesion is present. The signal-to-noise ratio (SNR) of $q_{\mathrm{npw}}$ equals:

$$\mathrm{SNR}(q_{\mathrm{npw}}) = \frac{\left|\mathrm{mean}(q_{\mathrm{npw}}|1) - \mathrm{mean}(q_{\mathrm{npw}}|0)\right|}{\sqrt{(\mathrm{var}(q_{\mathrm{npw}}|1) + \mathrm{var}(q_{\mathrm{npw}}|0))/2}}, \tag{20}$$

where mean(.|$i$) and var(.|$i$) are mean and variance under the hypothesis $i$ (1 = lesion present, 0 = lesion absent). If $q_{\mathrm{npw}}$ obeys Gaussian statistics, then the area under the ROC curve is a monotonically increasing function of SNR($q_{\mathrm{npw}}$) [21]. Consequently, (20) is a measure of the performance of the NPWO; it can be readily computed from the 200 noisy lesion-present and lesion-absent images for each algorithm.

The design of the CHO was similar to that in [22]. It used three channels, located over the known signal position. The channels had circular symmetry, and each channel had a rectangular frequency band [22], [23]. The frequency bands were defined as: $[Bq^{-3}, Bq^{-2}]$, $[Bq^{-2}, Bq^{-1}]$ and $[Bq^{-1}, B]$, with B = 0.4 cycles per pixel and q = 2.3. Thus, the information in an image is reduced to a vector $f$, with only three elements, the channel outputs. The statistic computed by the CHO equals

$$q_{\mathrm{cho}} = (\bar{f}_1 - \bar{f}_0)' C_f^{-1} f, \tag{21}$$

where $\bar{f}_1$ and $\bar{f}_0$ are the channel bank outputs for the noise-free images with and without the lesion, $f$ is the channel bank output for the noisy image to be analyzed, and $C_f^{-1}$ is the inverse covariance matrix of $f$. The 3 × 3 matrix $C_f$ has only 6 different elements, which can be computed with sufficient accuracy from the 200 noise realizations. The SNR of $q_{\mathrm{cho}}$ can be computed as in (20). However, inserting (21) reveals that

$$\mathrm{SNR}^2(q_{\mathrm{cho}}) = (\bar{f}_1 - \bar{f}_0)' C_f^{-1} (\bar{f}_1 - \bar{f}_0). \tag{22}$$

In the expressions above, the noise-free image (the image obtained from noise-free data) was used to estimate the mean image. For MLEM, Wilson et al [24] verified that this is an excellent approximation. We have assumed that this holds also for all other algorithms applied here, and we verified that the difference between the noise-free image and the mean of the 200 noise realizations was indeed very small. Using the noise-free image is more convenient because it can be computed before the Monte Carlo is started. In addition, it is probably a better estimate of the ensemble mean, than the mean computed from 200 realizations.

*3) Error bars from bootstrapping:* To estimate the accuracy of values computed from the sets of noise realizations, the bootstrapping method of Efron was used [25]. For each estimate, 2000 bootstrap samples (with replacement) were drawn from the set of available noise realizations. Each bootstrap sample has the same number of realizations (200) as the original set. This yields 2000 estimates, from which the standard error on that estimate is computed. With a few hundred bootstrap samples virtually identical results were obtained, indicating that 2000 samples is largely sufficient. The standard error on values computed from noise-free images was assumed to be zero.

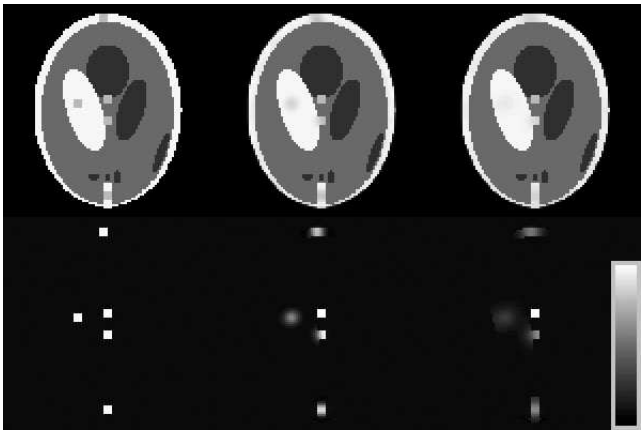

Fig. 3. The SHEPP phantom with the lesions, and its MAP reconstruction with a moderate and very strong anatomical prior. The second row shows the difference between signal-present and signal-absent reconstructions. The three images on the same row have the same gray value scaling. All lesions are present for illustration purposes, the simulations have been done for each lesion independently.

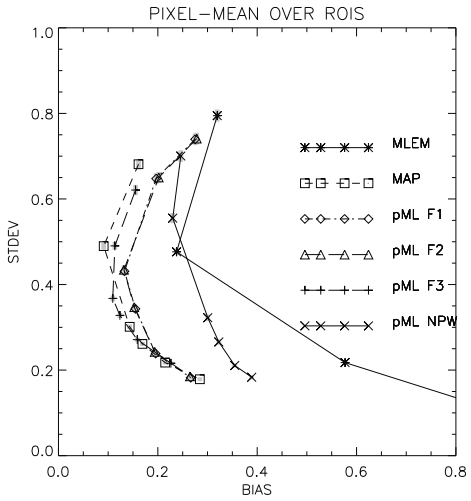


Fig. 4. Mean bias-noise curve (over 5 regions) for the pixel in the center of the lesions in the SHEPP phantom. **MLEM**: ML-reconstruction post-smoothed with Gaussian, **MAP**: maximum-a-posteriori with anatomy-based prior, **pML Fx**: ML post-processed with prewhitening using filter Fx, **pML NPW**: ML post-processed without prewhitening.

## IV. Results

Figure 3 shows two MAP reconstructions with different strengths of the anatomical priors. As expected, the smoothing is anisotropic and position dependent.

### A. Bias-noise curve

The average bias-noise curves for the central pixel in the lesion are plotted in figures 4 (Shepp phantom) and 5 (brain phantom). There are no horizontal error bars, because the bias is computed from noise-free images. Vertical error bars are plotted in gray, but they are very small, because the values have been obtained by averaging over many realizations (averaging over lesion-present and lesion-absent, and over all ROIs).

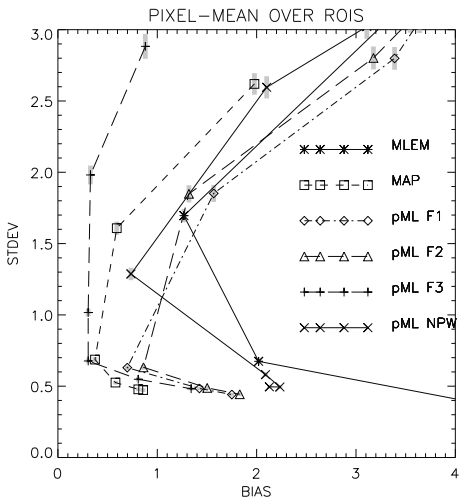


Fig. 5. Mean bias-noise curve (over 2 regions) for the pixel in the center of the lesion in the BRAIN phantom.

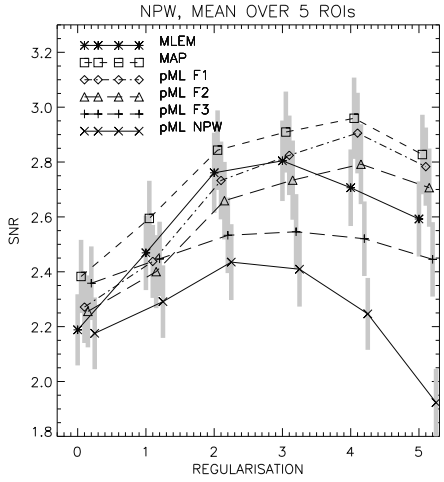


Fig. 6. Signal-to-noise ratio averaged overall all lesions, as a function of the degree of regularization (0 = minimum, 5 = maximum), for the NPWO and the SHEPP phantom. The error bars ($\pm$ 1 std err) are plotted in gray

### B. Numerical observers.

Figures 6, 7, 8 and 9 show the signal-to-noise ratio as a function of the regularization for each of the 6 algorithms and for both objects. Recall that 6 different degrees of regularization were obtained by varying the width of the Gaussian filter for MLEM, and the value of $\beta$ for the methods using the prior. These curves are obtained by averaging the SNR over all lesions. Figure 10 plots the SNR of the channelized Hotelling observer for each of the five lesions of the Shepp phantom, showing that the relative performance of the algorithms is very position-dependent.

## V. Discussion

An obvious advantage of the "background known exactly" assumption is its convenience. A disadvantage is that with this assumption, the observer can obtain valuable information from very low frequencies, which is impossible in real life due to anatomical variability, uncertainty about administered dose

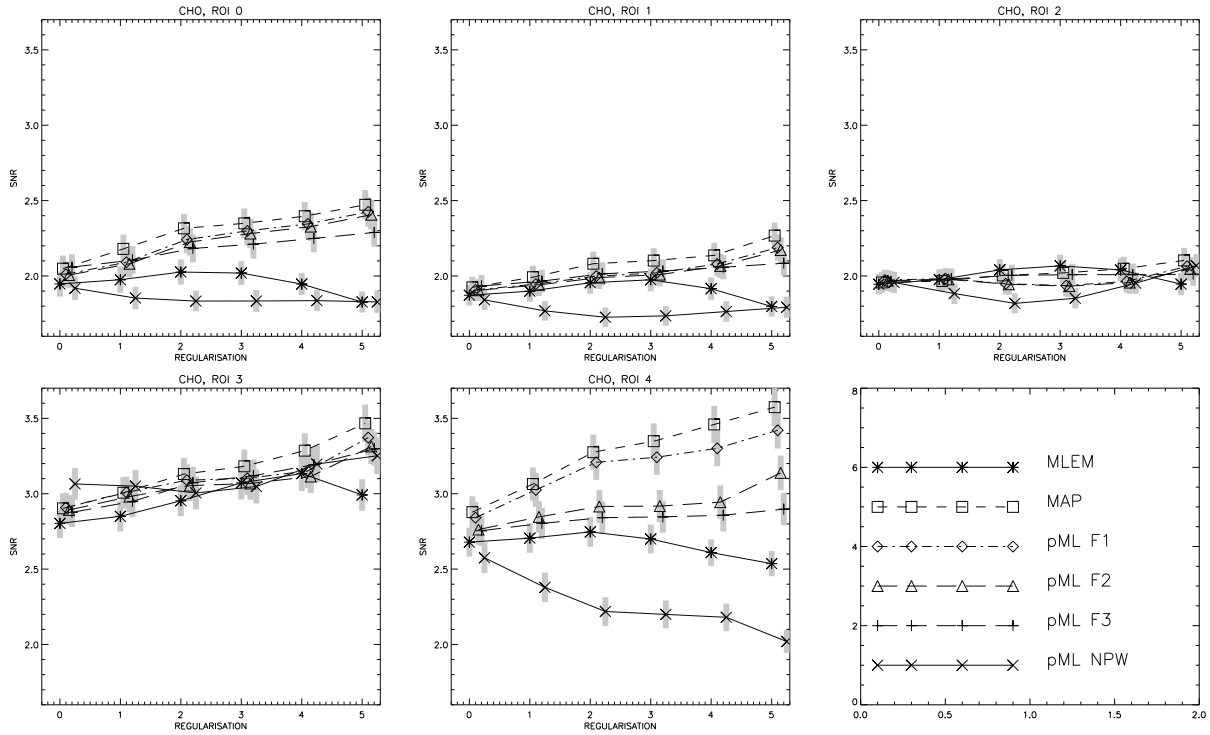


Fig. 10. CHO signal-to-noise ratio for each of the lesions in the SHEPP phantom (see fig 2 for the position of the lesions).

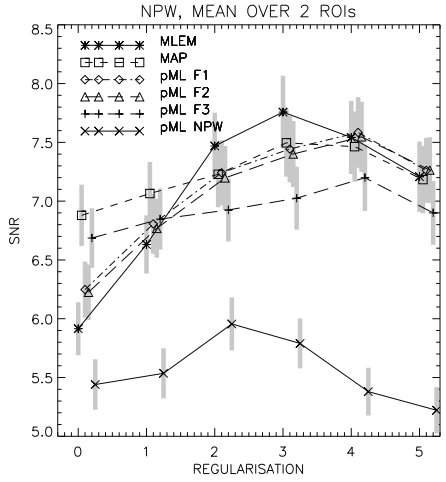


Fig. 7. Signal-to-noise ratio averaged overall all lesions, as a function of the degree of regularization (0 = minimum, 5 = maximum), for the NPWO and the BRAIN phantom

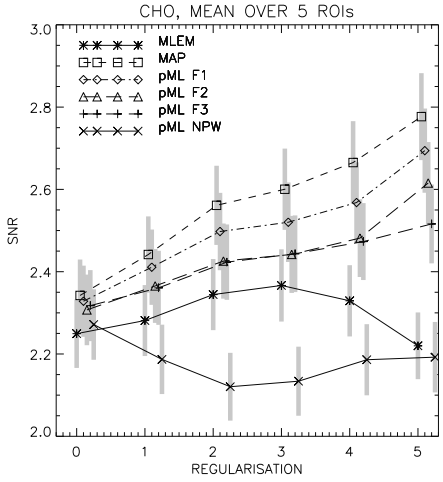


Fig. 8. Signal-to-noise ratio averaged overall all lesions, as a function of the degree of regularization (0 = minimum, 5 = maximum), for the CHO and the SHEPP phantom

etc.. The problem is alleviated when using the channelized observer, which discards the lowest frequencies. However, the addition of background variability would probably still reduce the information in the lowest frequency channel. This would result in a decrease of performance, particularly at high degrees of regularization. For our CHO-experiments, such a performance drop was not seen for the Shepp phantom, but a tendency toward lower performance at the highest regularization level was observed for the brain phantom.

The channelized Hotelling observer has often been found to be a good predictor of human performance [21], [22], [26], [27]. Similar to others, we have used radially symmetrical channels. This may be justified when the signal and noise covariances are approximately radially symmetrical [26], as is the case in most studies. In our study, though, this was not the case (the use of anatomy results in asymmetrical smoothing) and oriented channels [28] may have been more appropriate. Moreover, Oldan et al [23] reported that, for good agreement with humans, the CHO should be handicapped with internal noise. Optimizing the CHO and background variability for this particular application was beyond the scope of this study. Consequently, we cannot claim that our channelized

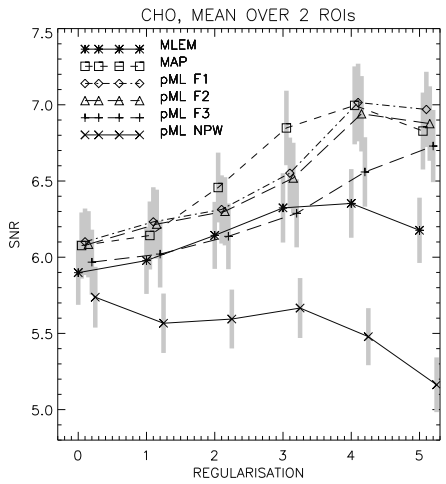


Fig. 9. Signal-to-noise ratio averaged overall all lesions, as a function of the degree of regularization (0 = minimum, 5 = maximum), for the CHO and the BRAIN phantom

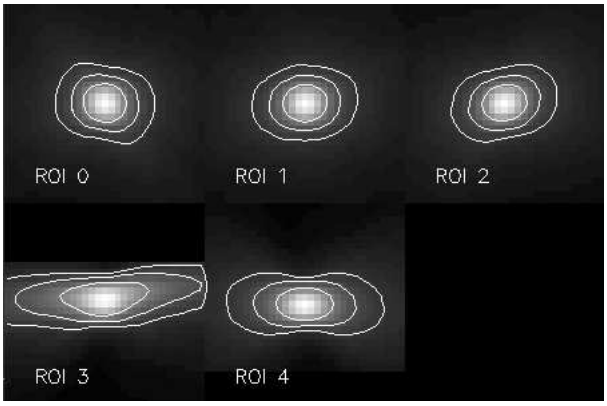


Fig. 11. The rows from the Fisher information matrix corresponding to the lesion (ROI) positions, are represented as images. From these images, a region of 31 $\times$ 31 pixels centered on the lesion position is displayed. Three isocontours are plotted in overlay.

Hotelling observer accurately predicts human performance here. However, the theoretical analysis suggests that in absence of prewhitening, information is lost, which makes us expect performance degradation for most tasks. For that reason, we have used three different measures of image quality: a channelized Hotelling observer, a non-prewhitening observer and a bias-noise analysis. As expected, MAP and post-processing with prewhitening are superior to post-processing without prewhitening, according to all three measures of image quality, and for both simulated objects. Thus, we expect similar results for human detection performance, but this remains to be verified.

Incorporating anatomical knowledge was helpful for the CHO, as it performed better with MAP than with MLEM post-smoothed with a Gaussian. This is in agreement with a human observer study comparing these algorithms based on more realistic simulations [10].

The ranking of the 6 algorithms is similar for all measures and both objects, except for that of prewhitening with filter $F_3$, the symmetrical 7 $\times$ 7 convolution kernel. The bias-noise curves from MAP and pML-F3 suggest similar performance, whereas MAP yielded a superior SNR for the numerical observers. Prewhitening with pML-F3 was not expected to be superior to more accurate prewhitening, we don't have a good explanation for this result. The observed bias is the result of several factors which causes complicated behavior: e.g. a tendency to underestimate the gray matter may partly compensate an overestimation of the (reduced) lesion activity.

The relative performance of the algorithms depends on the lesion. This is illustrated by figure 10 for the channelized Hotelling observer and the Shepp phantom. The boundaries of lesion 2 do not match any anatomical boundary. In addition, it is positioned centrally, so the Fisher information has nearly circularly symmetry. As a result, all algorithms produce a similar, relatively symmetrical smoothing effect, and the performance differences are small. In contrast, lesion 0 has a perfect match with anatomy. The pML-NPW algorithm ignores all information that is present (due to pixel covariances) at the other side of the boundary, resulting in more noise and reduced SNR. Post-smoothed MLEM suppresses the noise but reduces the signal by smoothing over the lesion boundary, again reducing the SNR. In contrast, MAP and the prewhitening post-processing methods do not smooth the signal over the lesion boundary, but they manage to reduce the noise by using also the information at the other side of the boundary. Because of the central position, the Fisher information is relatively symmetrical and the approximate filters $F_2$ and $F_3$ do not differ dramatically from $F_1$.

To explain the results for the eccentric lesion positions 3 and 4, it is useful to look at the Fisher information. The images in figure 11 are taken from the rows of the Fisher information matrix corresponding to the lesion positions. These figures show that for regions 3 and 4, the information is mostly spread horizontally. For region 3, the class $G$ is oriented horizontally as well (see fig 2). Consequently, smoothing withing class $G$ will bring together most of the information, resulting in similar performance for all algorithms. In contrast, for lesion 4, class $G$ is oriented vertically. Smoothing vertically within class $G$ will not result in gathering a significant amount of information. The prewhitening filter $F_1$ takes into account the asymmetric shape of the Fisher information, and as a result, it incorporates the information outside class $G$. The symmetrical filters $F_2$ and $F_3$ recover less information. Finally, post-processing without prewhitening results in very poor performance, because all the information outside class $G$ is ignored.

In agreement with the theory, the experiments indicate that the anatomical information can be incorporated either during reconstruction (MAP) or after reconstruction (post-processing unconstrained MLEM). However, to ensure optimal performance (that is, similar to that of MAP) at all positions in the image, the post-processing method should apply position dependent prewhitening. This is difficult to implement efficiently: with our approach it requires two FFTs per pixel. Although more efficient implementations can be devised [29], it seems unlikely that these will make the post-processing method competitive with the relatively simple MAP-approach. Moreover, the MAP-approach has the additional advantage of

converging faster than unconstrained MLEM, as indicated by simulations in [2].

## VI. Conclusion

This work studied how detailed anatomical information can be most efficiently utilized to improve the reconstruction of emission data, assuming that such knowledge is available.

It was shown that this anatomical information can be incorporated either during reconstruction (MAP) or after the reconstruction (post-processed MLEM). However, post-processed MLEM only achieves optimal performance at all image locations if it applies a shift variant prewhitening. This makes the post-processing very complex. Consequently, we believe that for these applications, maximum-a-posteriori reconstruction should be the method of choice.

## References


[1] J Nuyts, JA Fessler, "A penalized-likelihood image reconstruction method for emission tomography, compared to post-smoothed maximum-likelihood with matched spatial resolution." *IEEE Trans Med Imaging*, vol 22, pp 1042-52, 2003.

[2] JW Stayman, JA Fessler, "Compensation for nonuniform resolution using penalized-likelihood reconstruction in space-variant imaging systems," *IEEE Trans Med Imaging*, vol 23, pp 269-84, 2004.

[3] C. Comtat, P.E. Kinahan, J.A. Fessler, T. Beyer, D.W. Townsend, M. Defrise, and C. Michel, "Clinically feasible reconstruction of 3D whole-body PET/CT data using blurred anatomical labels," *Phys Med Biol*, vol 47, pp 1-20, 2002.

[4] B Lipinski, H Herzog, E Rota Kops, W Oberschelp, HW Müller-Gärtner, "Expectation maximization reconstruction of positron emission tomography images using anatomical magnetic resonance information", *IEEE Trans Med Imaging*., vol. 16, pp. 129-136, 1997.

[5] M Yavuz, JA Fessler. "Statistical image reconstruction methods for randoms-precorrected PET scans. *Med Im Anal* vol 2, pp. 369-378, 1998.

[6] L Janeiro, C Comtat, C Lartizien, PE Kinahan, M Defrise, C Michel, R Trebossen, P Almeida. "NEC-scaling applied to FORE+OSEM", *IEEE Nuclear Science Symposium Conference Record*, vol 2, pp. 717 - 721, 2002,

[7] LS Shepp, Y Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans Med Imaging*, vol MI-1, pp. 113-122, 1982.

[8] MH Hudson, RS Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans Med Imaging* vol 13, pp. 601-609, 1994.

[9] K Baete, J Nuyts, W Van Paesschen, P Suetens, P Dupont. "Anatomical based FDG-PET reconstruction for the detection of hypo-metabolic regions in epilepsy." *IEEE Trans Med Imaging*, 2004: 23, (4): 510-519.

[10] K Baete, J Nuyts, K Van Laere, W Van Paesschen, S Ceyssens, L De Ceuninck, O Gheysens, A Kelles, J Van den Eynden, P Suetens, P Dupont. "Evaluation of anatomy based reconstruction for partial volume correction in brain FDG-PET." *NeuroImage*, vol 23, pp. 305-317, 2004.

[11] J Nuyts, D Bequé, P Dupont, L Mortelmans. "A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography." *IEEE Trans Nucl Sci*, vol 49, pp 56-60, 2002.

[12] JA Fessler. "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): applications to tomography". *IEEE Trans Image Processing*, vol 5, pp. 493-506, 1996

[13] KV Mardia, JT Kent, JM Bibby. "Multivariate analysis". Academic, San Diego, Calif., 1979.

[14] HH Barrett, JL Denny, RF Wagner, KJ Myers. "Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance". *J Opt Soc Am A*, vol 12, pp. 834-852, 1995.

[15] J Qi, RH Huesman. "Theoretical study of lesion detectability of MAP reconstruction using computer observers." *IEEE Trans Med Imaging,* vol 20, pp 815-822, 2001.

[16] J Qi, RM Leahy. "Resolution and noise properties of MAP reconstruction for fully 3-D PET." *IEEE Trans Med Imaging*, vol 19, pp. 493-506, 2000.

[17] JA Fessler, SD Booth. "Conjugate-gradient preconditioning methods for shift-variant PET image reconstruction". *IEEE Trans Image Processing*, vol 8, pp. 688-699, 1999.

[18] J Nuyts, K Baete, P Dupont. "Comparison between MAP and post-processed ML for incorporating anatomical knowledge in emission tomography." *Proceedings of the IEEE NSS-MIC Symposium* 2003, M5-2.

[19] BrainWeb database at McGill University (Montreal, Canada), http://www.bic.mni.mcgill.ca/brainweb/.

[20] D Collins, A Zijdenbos, V Kollokian, J Sled, N Kabani, C Holmes, A Evans. "Design and construction of a realistic digital brain phantom." *IEEE Trans Med Imag* vol 17, pp 463-468, 1998.

[21] HH Barrett, J Yao, JP Rolland, KJ Myers. "Model observers for assessment of image quality". *Proc Natl Acad Sci USA* vol 90, pp 9758-9765, 1993.

[22] HC Gifford, MA King, DJ de Vries, EJ Soares. "Channelized Hotelling and human observer correlation for lesion detection in hepatic SPECT imaging.". *J Nucl Med* vol 41, pp. 514-521, 2000.

[23] J Oldan, S Kulkarni, Y Xing, P Khurd, G Gindi. "Channelized Hotelling and Human Observer Study of Optimal Smoothing in SPECT MAP Reconstruction" *IEEE Trans Nucl Sci*, To appear, 2004.

[24] DW Wilson, BMW Tsui, HH Barrett. "Noise properties of the EM algorithm: II. Monte Carlo simulations". *Phys Med Biol*, vol 39, pp. 847-871, 1994.

[25] B. Efron. "Bootstrap methods: another look at the jackknife." *The Annals of Statistics* vol 7, pp 1-26, 1979.

[26] CK Abbey, HH Barrett. "Human and model observer performance in ramp-spectrum noise: effects of regularization and object variability". *J Opt Soc Am A*, vol 18, pp. 473-488, 2001.

[27] HC Gifford, MA King, MV Narayanan, PH Pretorius, MS Smyczynski, RG Wells. "Effect of block-iterative acceleration on Ga-67 tumor detection in thoracic SPECT". *IEEE Trans Nucl Sci,* vol 49, pp. 50-55, 2002.

[28] AB Watson. "Detection and recognition of simple spatial forms", in Physical and Biological Processing of Images (Braddic and Sleigh Eds), 1983, Berlin, Springer Verlag.

[29] S Mustafovic, K Thielemans. "Object dependency of resolution in reconstruction algorithms with interiteration filtering applied to PET data." *IEEE Trans Med Imag* vol 23, pp 433-446, 2004.