A penalized-likelihood image reconstruction method for emission tomography, compared to post-smoothed maximum-likelihood with matched spatial resolution

Johan Nuyts , Jeffrey A. Fessler

***Abstract*— Regularization is desirable for image reconstruction in emission tomography. A powerful regularization method is the penalized-likelihood reconstruction algorithm (or equivalently, maximum-a-posteriori reconstruction), where the sum of the likelihood and a noise suppressing penalty term (or Bayesian prior) is optimized. Usually, this approach yields position dependent resolution and bias. However, for some applications in emission tomography, a shift invariant point spread function would be advantageous. Recently, a new method has been proposed, in which the penalty term is tuned in every pixel to impose a uniform local impulse response. In this paper, an alternative way to tune the penalty term is presented. We performed PET and SPECT simulations to compare the performance of the new method to that of the post-smoothed maximum-likelihood approach, using the impulse response of the former method as the post-smoothing filter for the latter. For this experiment, the noise properties of the penalized-likelihood algorithm were not superior to those of post-smoothed maximum-likelihood reconstruction.**

***Index Terms*— Tomography, Bayesian reconstruction, Regularization, PET, SPECT.**

I. Introduction

DUE to the low tracer dosage and the limited acquisition time, clinical emission data (positron emission tomography (PET) or single photon emission tomography (SPECT)) are usually strongly affected by Poisson noise. Even with optimal (according to the maximum-likelihood criterion) use of the data in statistical reconstruction, the noise propagation results in unacceptable noise levels in the reconstructed images. Several regularization methods have been proposed. A powerful method is to replace the maximum-likelihood criterion with a maximum-a-posteriori (MAP) criterion, by combining the likelihood with a Bayesian prior that encourages local smoothness [1–5]. Often, the prior is modeled as a Gibbs distribution of the form

$$\text{prior}(\lambda) = \frac{1}{Z} \exp\left(-\sum_j \sum_{k \in N_j} \beta \Phi(\lambda_j, \lambda_k) \right), \qquad (1)$$

J Nuyts is with the department of Nuclear Medicine, K.U.Leuven, Herestraat 49, B3000 Leuven, Belgium, e-mail: Johan.Nuyts@uz.kuleuven.ac.be. J. A. Fessler is with the department of Electrical Engineering and Computer Science, 4240 EECS, University of Michigan, Ann Arbor, MI 48109-2122, e-mail: fessler@umich.edu. This work is supported by the grants OT/00/32 of K.U.Leuven and G.0174.03 of the Flemish Fund for Scientific Research (FWO). J. A. Fessler's effort was supported in part by NSF grant BES-9982349.

where Z is a normalization constant, N_j is the set of neighbors of pixel j, Φ is a function operating on pairs of neighboring pixels [6] and β is a constant that specifies the relative strength of the prior. Usually, Φ is chosen as a shift invariant function that penalizes differences between neighboring pixels. The approach is attractive because it allows one to include the regularization in the reconstruction (so the final reconstructed image is directly verified against the raw data), and because the Gibbs-framework accepts a wide range of functions that can be optimized for particular purposes. However, because the prior is shift-invariant and the likelihood is not, the maximum-a-posteriori image has position dependent (and image dependent) bias and resolution. For some applications, this is an undesirable feature. For example, in tracer kinetic modeling, the time activity curves should only reflect changes in tracer concentration, and changes due to varying spatial resolution will cause errors. Similarly, when applying semi-quantitative analysis based on standard uptake values [7], it is important that the bias does not change with position and image contents.

Fessler and Rogers [8] have proposed to use a position dependent prior: they replace β in (1) with $\sqrt{\beta_j \beta_k}$ and tune these parameters to impose position independent resolution. This makes the "prior" data dependent, so it can no longer be regarded as a Bayesian prior; the authors call it a penalty term and their method penalized-likelihood reconstruction. With the position dependent penalty, the resolution was more uniform, but there was still position dependent asymmetry of the local impulse response function. Stayman and Fessler have proposed a further sophistication of the method, by replacing β with β_{jk} in (1) and optimizing the parameters to eliminate the asymmetry and even obtain an optimal fit of the local impulse response to a predefined target point spread function [9], [10]. Interestingly, similar work is being done for the "expectation-maximization-smooth" (EMS) algorithm [11], [12], which yields position dependent resolution if the smoothing between iterations is position independent. Mustafovic et al [13] have shown that with position dependent filters, it is possible to obtain uniform resolution with EMS as well.

An alternative method to obtain uniform resolution is to post-smooth the reconstruction obtained after many iterations of a maximum-likelihood reconstruction algorithm [14], [15]. Applying a sufficiently high number of iterations ensures a nearly bias-free reconstruction, so after post-smoothing, the spatial resolution is uniform and the point spread function is (nearly) identical to the smoothing filter. Note that the number of iterations needed depends on the application; for some, several hundreds of iterations may be required. In this paper, a new penalized-likelihood method is proposed to obtain a symmetric and shift invariant point spread function. The performance of this new algorithm is compared to that of post-smoothed maximum-likelihood reconstruction.

This paper is organized as follows. In the following section, we first derive an approximate expression for the "natural" shape of the local impulse response function associated with a quadratic penalty term. The rest of the section discusses how the certainty of the likelihood can be estimated and be used to tune the penalty term. In section III the setup of the simulation experiments is discussed. The main experiment is a comparison of signal-to-noise ratio at matched resolution, between post-smoothed maximum-likelihood and the new penalized-likelihood algorithm. Section IV presents the results, which are discussed in section V.

II. Theory

A. The local impulse response with the quadratic prior and uniform likelihood

Consider a one dimensional image, and assume that for every pixel exactly one measured value is available. Assume that the measurements are independent, and subject to Gaussian noise with constant and known variance, equal to 1. Then the logarithm of the likelihood $L(y, \lambda)$ equals

$$L(y, \lambda) = \sum_j L_j = -\frac{1}{2} \sum_j (\lambda_j - y_j)^2, \tag{2}$$

where y_j and λ_j are the measurement and the image values for pixel j. We also introduce an a-priori probability distribution. The logarithm of this Bayesian prior equals:

$$P_1(\lambda) = -\frac{1}{4} \sum_j \left(w(\lambda_j - \lambda_{j-1})^2 + w(\lambda_j - \lambda_{j+1})^2 \right). \tag{3}$$

Here w is the weight assigned to the difference between a pixel and its neighbor. This prior favors smooth images and reaches its maximum when the image is perfectly uniform. As in image reconstruction from projections, the maximum-a-posteriori (MAP) image is obtained by maximizing $L + P_1$. To study the local impulse response of the MAP-image, we assume that the measured values for all pixels are zero, except for a single pixel $j = 0$, for which it equals $A > 0$. For a pixel with $j \neq 0$, the MAP-image satisfies the following relation:

$$\begin{aligned} 0 &= \frac{\partial (L + P_1)}{\partial \lambda_j} \\ &= -(1 + 2w)\lambda_j + y_j + w\lambda_{j-1} + w\lambda_{j+1}. \end{aligned} \tag{4}$$

Because $y_j = 0$ for $j \neq 0$ we obtain:

$$\lambda_j = \frac{w}{1 + 2w}(\lambda_{j-1} + \lambda_{j+1}). \tag{5}$$

Substituting $\lambda_j = ab^j$ produces a quadratic equation in b:

$$wb^2 - (1 + 2w)b + w = 0, \tag{6}$$

with the following solution:

$$b = \frac{1 + 2w \pm \sqrt{1 + 4w}}{2w}. \tag{7}$$

Note that the product of the two solutions for b equals 1. It follows that $\lambda_j = a \exp(-|\ln(b)j|)$ is a solution. The local impulse response has an exponential shape for this 1D problem. The value of a can be determined by requiring that the sum (over all pixels) of the impulse response equals the sum of the impulse. The same result has been derived earlier by Unser et al [16] using the z-transform representation.

A simple approximate expression for the 2D case can be obtained, under the assumption that the local impulse response is circularly symmetric, and that effects of the pixel grid can be ignored. For many applications, circular symmetry is desirable, and experience shows that it can be achieved with good approximation using a 4 or 8 pixel neighborhood. Assume that the local impulse is centered at pixel $j = 0$, and that λ_j represents the pixel value at a distance of j pixels from the center. For simplicity, we also assume that the neighbors of a pixel at distance j are all located on the circles with radii $j-1$, j and $j+1$. The neighbors at distance j all have the same value and contribute a zero term to the quadratic prior for λ_j. The circle with radius $j+1$ contains more pixels than the circle with radius $j - 1$, so pixel j has more neighbors at distance $j+1$ than at distance $j-1$ from the center. We will assume that the number of neighbors at distance j is proportional to $j + \epsilon$, where ϵ is a small positive constant, reflecting the finite size of the pixels (there is a finite pixel at distance zero ($j = 0$) from the center). With these approximations, the two-dimensional problem can be described by modifying the weights in (3):

$$\begin{aligned} P_2(\lambda) &= -\frac{1}{4} \sum_j \{ w\frac{j - 1 + \epsilon}{j + \epsilon}(\lambda_j - \lambda_{j-1})^2 \\ &\quad + w\frac{j + 1 + \epsilon}{j + \epsilon}(\lambda_j - \lambda_{j+1})^2 \}. \end{aligned} \tag{8}$$

As before, the prior is combined with the likelihood (2), where we assume that $y_j = 0$, and the maximum of $L + P_2$ is computed by setting the first derivative to zero:

$$\begin{aligned} 0 &= \frac{\partial (L + P_2)}{\partial \lambda_j} \\ &= -w\frac{j - 1 + \epsilon}{j + \epsilon}(\lambda_j - \lambda_{j-1}) \\ &\quad -w\frac{j + 1 + \epsilon}{j + \epsilon}(\lambda_j - \lambda_{j+1}) - \lambda_j \end{aligned} \tag{9}$$

Rearranging yields:

$$\begin{aligned} (2w + 1)(j + \epsilon)\lambda_j - w(j - 1 + \epsilon)\lambda_{j-1} & \\ -w(j + 1 + \epsilon)\lambda_{j+1} &= 0 \end{aligned} \tag{10}$$

Substitution of $\lambda_j = ab^j/(j+\epsilon)$ produces a quadratic equation in b, which is identical to (6). Consequently, we find that a maximum of $L + P_2$ is obtained for

$$\lambda_j \simeq \frac{a}{(j + \epsilon)} e^{-|\ln(b)|j}. \tag{11}$$

The main conclusion is that the local impulse response of the quadratic prior has an exponential shape which is rather different from that of typical low pass filters used in nuclear medicine. This is important when comparing the performance of penalized-likelihood methods to that of filter-based methods. Unless the filter is matched to the local impulse response of the penalized-likelihood method, it will be unclear if performance differences are due to intrinsic properties of the algorithms, or only to the different characteristics of the impulse responses.

B. Emission tomography

In emission tomography, the log-likelihood function can be written as [14]:

$$L(y,\lambda) = \sum_i \{y_i \ln(r_i) - r_i\} \tag{12}$$

$$r_i = \sum_j c_{ij}\lambda_j + q_i \tag{13}$$

where y_i is the measured photon count in detector i, λ_j is the estimated radioactivity in pixel j, c_{ij} is the probability that a photon emitted in j is detected in i, q_i is the expected number of counts contributed by such processes as scatter and randoms, and terms independent of λ have been dropped.

In the analysis above, the certainty provided by the likelihood was the same for every pixel. In contrast, the certainty provided by emission tomography is different for every pixel. When the non-uniform likelihood is combined with a uniform penalty term, position dependent smoothing results. In [8], an algorithm is presented to impose approximately uniform spatial resolution by tuning the weights w_{jk} of a quadratic penalty of the form:

$$P(\lambda) = \frac{1}{4}\sum_j \sum_k w_{jk}(\lambda_j - \lambda_k)^2, \tag{14}$$

where the weights w_{jk} are zero except when pixels j and k are neighbors, and $w_{jk} = w_{kj}$. Based on the analysis of an explicit expression for the local impulse response function, the authors propose to choose the weights as follows:

$$w_{jk} \sim \sqrt{\left(\sum_i \frac{c_{ij}^2}{\bar{y}_i}\right)\left(\sum_i \frac{c_{ik}^2}{\bar{y}_i}\right)}, \tag{15}$$

where $\bar{y}_i$ is the measurement mean for detector i. The factors between parentheses are the j-th and k-th diagonal elements of the Fisher information matrix [17], which can be regarded as a measure for the certainty provided by the likelihood. So (15) prescribes that the weight used to penalize the difference between two pixels should be proportional to the geometric mean of the certainties of the two pixels. The measurement mean $\bar{y}_i$ is not available, but the measurements y_i or the calculated projections $\hat{y}_i$ are useful approximations. We will denote this algorithm as CPL, which stands for "Certainty based Penalized-Likelihood reconstruction". Although this algorithm makes the resolution more uniform, the resulting local impulse response is asymmetric, and the asymmetry is still position dependent. Stayman and Fessler [9], [10] have extended the algorithm to reduce the asymmetry as well. Their approach is based on an explicit expression for the local impulse response function, and they optimize the weights w_{jk} to obtain a best fit between this computed local impulse response and a predefined target impulse response.

Here we follow a slightly different approach. The analysis presented above suggests that the shape of the local impulse response may be an intrinsic property of the quadratic penalty. For that reason, and also in an attempt to obtain a simpler algorithm, we do not use a target impulse response: we will accept any shape, as long as the impulse response is symmetric and position independent.

The objective function that must be maximized is $Q = L + P$, where L is given by (12) and P by (14). Assuming unconstrained maximization (and therefore ignoring the usual non-negativity constraint), the reconstruction λ maximizing Q must satisfy $\partial L/\partial\lambda_j = -\partial P/\partial\lambda_j$ or

$$\sum_i \left(c_{ij}\frac{y_i}{r_i} - c_{ij}\right) = \sum_k w_{jk}(\lambda_j - \lambda_k). \tag{16}$$

To compute the local impulse response, the value in a single pixel p is changed by adding a small impulse u_p. As a result, the new measurement y' and reconstruction λ' become:

$$\begin{aligned} y_i' &= y_i + c_{ip}u_p \\ \lambda_k' &= \lambda_k + \Delta\lambda_k \\ r_i' &= r_i + \Delta r_i \\ \Delta r_i &= \sum_k c_{ik}\Delta\lambda_k, \end{aligned} \tag{17}$$

where u_p is the impulse and $\Delta\lambda_k$ is the impulse response. The posterior is now maximized when

$$\sum_i \left(c_{ij}\frac{y_i + c_{ip}u_p}{r_i + \Delta r_i} - c_{ij}\right) = \sum_k w_{jk}(\lambda_j + \Delta\lambda_j - \lambda_k - \Delta\lambda_k). \tag{18}$$

Subtracting (16) from (18) yields:

$$\sum_i c_{ij}\frac{c_{ip}u_p r_i - y_i\Delta r_i}{(r_i + \Delta r_i)r_i} = \sum_k w_{jk}(\Delta\lambda_j - \Delta\lambda_k). \tag{19}$$

Since u_p is very small, Δr_i is also very small and we have that $r_i + \Delta r_i \simeq r_i$ (we are only interested in the impulse response within an active object, so it is reasonable to assume that $r_i > 0$). In addition, we assume that the penalty is not too strong, such that the calculated and measured projections are very similar, and as a result

$$\sum_i c_{ij}\frac{y_i\Delta r_i}{r_i^2} \simeq \sum_i c_{ij}\frac{\Delta r_i}{r_i}. \tag{20}$$

With these assumptions, (19) can be simplified to

$$\sum_i c_{ij}\frac{c_{ip}u_p - \Delta r_i}{r_i} = \Delta\lambda_j\sum_k w_{jk} - \sum_k(w_{jk}\Delta\lambda_k). \tag{21}$$

This result is equivalent to the expression for the local impulse response (equations (18) and (19)) obtained by Fessler and Rogers [8] [2].

[2]To clarify the equivalence, equation (21) can be rewritten as $(A'D(\frac{1}{r})A + R)V = A'D(\frac{1}{r})AU$, where A is the system matrix, $D(\frac{1}{r})$ is a diagonal matrix with elements $D_{ii} = 1/r_i$, U is the impulse, V is the impulse response, and R is a matrix defining the penalty as $P(\Lambda) = \frac{1}{4}\Lambda' R\Lambda$.

Since the penalty penalizes only differences, it is expected that the mean count is preserved so $u_p = \sum_k \Delta\lambda_k$. Inserting this in (21) and using (17) yields:

$$\sum_i c_{ij} \frac{\sum_k (c_{ip} - c_{ik})\Delta\lambda_k}{r_i} = \Delta\lambda_j \sum_k w_{jk} - \sum_k (w_{jk}\Delta\lambda_k). \tag{22}$$

Switching the order of summations and rearranging a bit we obtain:

$$\sum_k \left(\sum_i \frac{c_{ij}(c_{ip} - c_{ik})}{r_i} + w_{jk} \right) \Delta\lambda_k = \Delta\lambda_j \sum_k w_{jk}, \tag{23}$$

which can be rewritten as:

$$\Delta\lambda_j = \frac{\sum_k (G_{jkp} + w_{jk})\Delta\lambda_k}{\sum_k w_{jk}} \tag{24}$$

$$G_{jkp} = \sum_i \frac{c_{ij}(c_{ip} - c_{ik})}{r_i} \tag{25}$$

If the parameters w_{jk} are large compared to the contribution of the likelihood G_{jkp}, then (24) states that the response in pixel j is a weighted average of the responses in the neighboring pixels, as can be expected from a smoothing penalty. The contribution of the likelihood G_{jkp} changes the weights in a position dependent way. Moreover, it also changes the total sum of the weights, as there is no contribution from the likelihood to the denominator. As a result, it is clear that with position independent parameters w_{jk}, the local impulse response strongly depends on position.

To reduce the position dependence, we will try to tune the parameters w_{jk} such that at least the sum of the weights in (24) becomes independent of the position. A somewhat simplistic way to obtain this would be to set

$$w_{jk} = \alpha G_{jkp}, \tag{26}$$

which would ensure that the sum of the weights in (24) would be equal to $1 + 1/\alpha$. This approach has two problems. First, G_{jkp} is a function of the position of the impulse u_p, while w_{jk} is not. It seems not trivial to optimize the response in j for all possible positions p of the impulse simultaneously. To avoid this problem, we concentrate on the response in j for a perturbation in j, i.e., we set $p = j$ in (24) and (26). The second problem is that for practical reasons, w_{jk} should be zero except for the pixels k that are close neighbors of pixel j, while the support of G_{jkp} is much larger. We hope that this problem can be ignored, because G_{jkp} is a (modified) backprojection, which decays rapidly with increasing distance to p. All these approximations yield the following recipe:

$$w_{jk} = \alpha \sum_i \frac{c_{ij}(c_{ij} - c_{ik})}{r_i}, \tag{27}$$

where parameter α defines the global strength of the penalty. According to [9], only the symmetric component of the design matrix determines the smoothing characteristics. This component is

$$\frac{w_{jk} + w_{kj}}{2} = \frac{\alpha}{2} \sum_i \frac{(c_{ij} - c_{ik})^2}{r_i}. \tag{28}$$

Equation (28) can be derived in a different way as well. The conclusion in [8] was that approximate uniform spatial resolution could be imposed by requiring that the weights w_{jk} were proportional to the Fisher information for estimating the pixel values in j and k. The Fisher information estimates the "resistance" of the likelihood against smoothing, and more smoothing is required if the resistance is higher. However, the Fisher information measures the certainty about the absolute pixel values, whereas the smoothing only penalizes differences between pixel values. So it seems meaningful to estimate the resistance against smoothing by computing the certainty about pixel *differences* provided by the likelihood. To do this for a particular pixel pair (j, k), we rewrite the likelihood (12) as a function of the difference and sum of these pixels:

$$r_i = \sum_{\xi \neq j, \xi \neq k} c_{i\xi}\lambda_\xi + c_{ij}\frac{s_{jk} + d_{jk}}{2} + c_{ik}\frac{s_{jk} - d_{jk}}{2} + q_i$$

$$d_{jk} = \lambda_j - \lambda_k$$

$$s_{jk} = \lambda_j + \lambda_k.$$

Now, the diagonal element of the Fisher information matrix corresponding to d_{jk} can be computed as

$$-E\left(\frac{\partial^2 L(y,\lambda)}{\partial d_{jk}^2}\right) = E\left(\sum_i \left(\frac{c_{ij} - c_{ik}}{2}\right)^2 \frac{y_i}{r_i^2}\right) = \frac{1}{4}\sum_i \frac{(c_{ij} - c_{ik})^2}{\bar{y}_i}, \tag{29}$$

where E is the expectation, and $\bar{y}_i$ is the expectation of y_i. Equation (29) reproduces (28) if we can assume that $\bar{y}_i \simeq r_i$.

Equations (28) or (29) have an interesting intuitive interpretation. For a projection line i intersecting both pixels j and k, we have $c_{ij} \simeq c_{ik}$, so this projection i does not contribute any certainty. In contrast, a projection line perpendicular to the line connecting j and k cannot intersect both pixels. Consequently, projection lines with this orientation and intersecting one of the pixels contribute a maximum amount of certainty. For example, the projection line i intersecting pixel j but not k has $c_{ik} = 0$, and its contribution is proportional to

$$\frac{(c_{ij} - c_{ik})^2}{\bar{y}_i} = \frac{c_{ij}^2}{\bar{y}_i}. \tag{30}$$

The diagonal element (j, j) of the Fisher information matrix for estimating λ from the likelihood equals

$$F_{jj} = \sum_i \frac{c_{ij}^2}{\bar{y}_i}. \tag{31}$$

Expression (30) is the i-th term of (31). So (28) or (29) suggest to compute the Fisher information not from all projections, but only from a subset containing projection lines which are approximately perpendicular to the line through j and k.

C. Imposing uniform resolution

A strong reduction of the complexity and the computation burden is obtained by introducing the approximation suggested in the previous section:

$$\sum_i \frac{(c_{ij} - c_{ik})^2}{r_i} \simeq \sum_{i \in S_{j-k}} \frac{c_{ij}^2}{r_i}, \tag{32}$$

where S_{j-k} is the subset of projections with projection line approximately perpendicular to the line connecting the centers of pixels j and k. We investigated a modified penalty for 2D reconstruction by inserting approximation (32) directly in (28). This approach only somewhat improved the resolution uniformity if the weights w_{jk} were computed using 8 neighbors in a 3x3 neighborhood. However, if only horizontal and vertical neighbors were used, good resolution performance in vertical and horizontal direction was observed. It seems that there is some interference between diagonal and vertical directions in the 8-neighborhood system, which is not captured by (28). Therefore, we redistributed the penalty weight values using the following heuristic modifications.

For 2D reconstruction and with a penalty term defined in a 3x3 neighborhood, eight weights w_{jk} per pixel j must be defined. Requiring that $w_{jk} = w_{kj}$ reduces the number to four. Therefore, we assume that there are only four smoothing directions: horizontal, vertical and the two diagonal ones. In addition, we assume that the smoothing can be considered as consisting of two components, a uniform component and a component in one of the four directions. Finally, we assume that the uniform component can be implemented using only the weights in the horizontal and vertical neighbors, and that the directional component can be tuned independently by adjusting the two weights corresponding to that direction.

These heuristics yielded the following recipe: For each of the four axes, an image is generated that estimates the Fisher information along that axis. These four images are computed as:

$$F_j^\theta = \alpha_{\theta'} \sum_{i \in S_{\theta'}} \frac{c_{ij}^2}{\hat{y}_i}, \tag{33}$$

where θ equals 0, 45, 90 or 135 degrees, $\theta' = \theta + 90^o$, S_θ is the subset of projections with projection lines between $\theta - 22.5^o$ and $\theta + 22.5^o$, $\alpha_0 = \alpha_{90} = 1$ and $\alpha_{45} = \alpha_{135} = 1/\sqrt{2}$. Note that θ is used to define an axis, not a direction, so operations on θ are modulo 180^o. Then, for every pixel j, we define θ_{max} as the axis with the largest value F_j^θ = Fmax. The likelihood provides the strongest certainty along this axis θ_{max}, so a stronger penalty weight along this axis is needed to impose uniform resolution. The uniform smoothing component is estimated by taking the minimum of F_j^θ over the four angles, denoted as Fmin. The four images are then modified to implement the two components, by applying the following steps

$$\begin{aligned} F_j^{45} = F_j^{135} &= 0 \\ F_j^{0} = F_j^{90} &= \text{Fmin} \\ F_j^{\theta\text{max}} &= F_j^{\theta\text{max}} + \text{Fmax - Fmin} \end{aligned}$$

The resulting images F^θ are convolved with a 2D Gaussian, to avoid possible artifacts near abrupt changes of Fmax, and normalized to ensure that total strength of the penalty in each pixel (as estimated by summing (33) over the four images θ) is not changed by the heuristic manipulation and Gaussian convolution. Finally, inspired by equation (15), we compute the weights w_{jk} as follows:

$$w_{jk} = \sqrt{F_j^{\theta(j,k)} F_k^{\theta(j,k)}}, \tag{34}$$

where the axis $\theta(j,k)$ is parallel to the line connecting pixels j and k. The additional computational burden of this method is small compared to that of traditional penalized-likelihood reconstruction with a quadratic penalty. Computation of the four images F^θ involves backprojection for four subsets, so the work is equivalent to a single backprojection. The rest are simple pixel operations, and (34) is computed every time w_{jk} is needed. Of course, the method increases the memory load, because the images F^θ must be precomputed and kept in memory.

This new algorithm is actually a straightforward extension of the CPL-algorithm (15). The essential difference is that in the new algorithm, the Fisher information is split in different components, which represent the information about pixel differences along different orientations. It is convenient to give it a name, so we will denote the new algorithm as OCPL, "Orientation dependent Certainty Penalized Likelihood".

After designing the penalty function using (34), we are ready to maximize the penalized-likelihood objective function: the sum of (12) and (14). One could apply any of the many iterative algorithms in the literature to this optimization problem. For the results given in the following section, we have applied a gradient ascent algorithm. The algorithm is obtained as a simple modification of the classical maximum-likelihood expectation-maximization algorithm, and has been described elsewhere [5].

III. Experiments

A. The shape of the local impulse response

To assess the accuracy of the approximate equation (11), the two-dimensional uniform likelihood problem has been simulated, using an 8-pixel neighborhood, a weight of 1 for direct neighbors and of $1/\sqrt{2}$ for diagonal neighbors and a strong global weight for the penalty term. Two hundred iterations of a gradient ascent algorithm were applied. The horizontal row containing the center of the impulse response was extracted to obtain a one dimensional profile, and the three parameters of (11) were computed with least squares fitting.

B. Evaluation of the new method

Two simulation experiments were performed to assess the performance of the new method. The first experiment was designed to evaluate the resolution uniformity obtained with OCPL, comparing with a quadratic penalty, with the CPL-method and with post-smoothed maximum-likelihood expectation-maximization (MLEM). The main purpose was to verify that extending CPL to OCPL leads to more uniform resolution. In the second experiment, the noise characteristics of OCPL and post-smoothed MLEM were compared by computing the signal-to-noise ratio in a few points.

The simulations were carried out starting from a digital description of the object (an activity image and an attenuation map). We only considered a single slice. PET and SPECT projections were computed taking into account the dominating physical effects: attenuation for both, and for SPECT also collimator blurring (implemented with Gaussian diffusion [18]). We performed multiple Poisson noise realizations to estimate the variance for computing the SNR. The reconstructions were computed using the same system matrix that was used for computing the projections. It is clear that the simulation is a simplification compared to true life. However, the results are useful because the dominating effects have been taken into account and the algorithms were evaluated using exactly the same data.

For the second experiment, it was essential to ensure that the two methods had a (virtually) identical impulse response. Otherwise, differences in the signal-to-noise ratio could be attributed to the impulse response rather than to the reconstruction algorithm. The following procedure was applied to ensure a close match of the impulse responses. First, a second digital phantom was produced by increasing the activity value of a single pixel. This is the impulse. Then, two sets of projections were computed, one for the original phantom, and another one for the phantom with the impulse. Both were reconstructed with OCPL, subtraction yields the local impulse response. This local impulse response was then used as the post-smoothing filter in post-smoothed MLEM. This ensures a close resolution match at the position of the impulse, if MLEM was iterated close to convergence. Assuming that OCPL is successful in imposing uniform resolution, there should also be a good resolution match in the other pixels. We verified this by measuring the OCPL impulse response at a few other pixels as well. In the following paragraphs, the experiments are described in more detail.

1) Resolution uniformity with the new method: The OCPL method was implemented and evaluated with two-dimensional PET and SPECT simulations. Figure 1 shows the activity distribution of the 2D software phantom. The object consisted of a uniform low activity background disk containing circles of higher activity. The disk and circles had identical and uniform attenuation. The background activity was 2, the activity of the circles was 10. The diameter of the attenuating disk was 28 cm for SPECT and 36 cm for PET. In both cases, the attenuation was set to 0.095 per cm. For the SPECT simulation, the collimator had a full width at half maximum of 2 cm at 30 cm distance and the camera had an intrinsic resolution of 4 mm. A circle is useful to evaluate orientation dependent smoothing, since recovery of the circular activity is sensitive to smoothing in any direction. The asymmetric position of the circles ensures strong position and orientation dependence of the certainties provided by the likelihood.

For PET, attenuated projections with 100 detector bins were computed for 80 angles, assuming perfect resolution (except for the blurring due to interpolation in the projection software). For SPECT, 60 attenuated projections of 66 bins per projection were computed, simulating an orbit of 180 degrees with a parallel hole collimator. The gamma camera started at the top and rotated in clockwise direction. No noise was added.

In both cases, reconstructions were computed with a uniform quadratic penalty, with the CPL-algorithm (15), with the new OCPL-method (34) and with post-smoothed maximum-likelihood expectation-maximization (MLEM). The reconstructed image size was 100×100 for PET and 66×66 for SPECT. For the smoothing kernel in post-smoothed MLEM, we used the impulse response of the OCPL method as described above (the impulse response was measured in the center of the image). With the quadratic penalty and the CPL-method, the impulse response is not symmetrical and a close match with the other methods cannot be imposed. An approximate match was achieved by tuning the penalty aiming at similar mean signal recovery along the circle. A high number of iterations was applied: 200 for PET and 450 for SPECT. We used a higher number for SPECT, because the inclusion of collimator blurring slows down convergence.

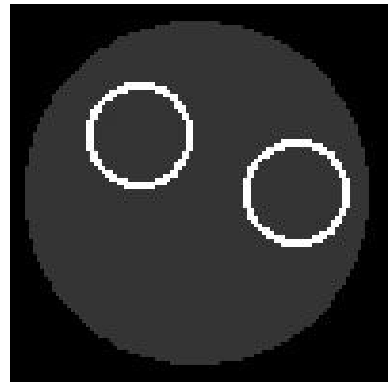

Fig. 1. Simulation object to evaluate the new uniform resolution penalized-likelihood with PET and SPECT.

2) Signal-to-noise comparison with post-smoothed MLEM: The aim of this experiment was to compare the signal-to-noise ratio obtained with the OCPL algorithm to that obtained with post-smoothed MLEM. The elliptical object, shown in figure 2 was used. It has uniform activity and uniform attenuation. First, a single hot pixel was inserted in the image (see figure 2) and noise-free attenuated PET-projections were computed (128 projections with 80 bins per projection). An OCPL-reconstruction was computed using 200 iterations. The very same procedure was applied again, but this time without the hot pixel. The difference between the two images is the local impulse response. This local impulse response was captured in a filter mask (15 x 15 pixels), for later use as the smoothing filter in post-smoothed MLEM.

Subsequently, two more hot pixels were inserted as shown in figure 2, and attenuated PET-projections were computed. These were used as the mean of a Poisson distribution, and 400 noise realizations were generated. In addition, 400 noise realizations in absence of the hot pixels were produced. From all these simulated projections, images were reconstructed with three different algorithms:

1) 200 iterations of the new OCPL-algorithm;
2) 200 iterations of the MLEM algorithm, followed by post-smoothing with the local impulse response function determined in the first step
3) 6 iterations of iterative filtered backprojection (IFBP), followed by post-smoothing with the same impulse response.

We used IFBP, because with regular (non-iterative) filtered backprojection, a small amount of smoothing due to interpolation is hard to avoid. This smoothing is eliminated after a few iterations of IFBP, resulting in a sharper impulse response.

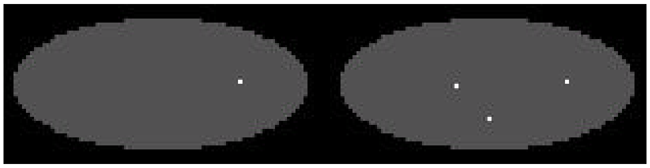

Fig. 2. Simulation objects to compare uniform resolution penalized-likelihood reconstruction to post-smoothed MLEM. Left: the object used to determine the local impulse response. Right: two more points were added for the signal-to-noise ratio measurement.

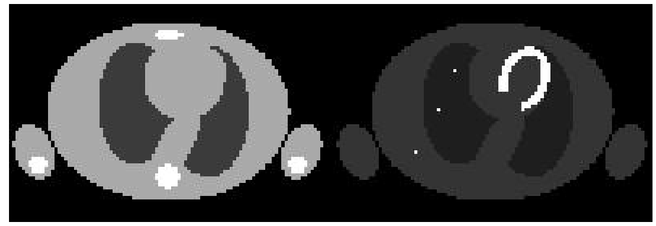

Fig. 3. Attenuation map (left) and activity distribution (right) for the simulated thorax phantom. The points are numbered from bottom to top, the first point (in tissue) is used the determine the local impulse response.

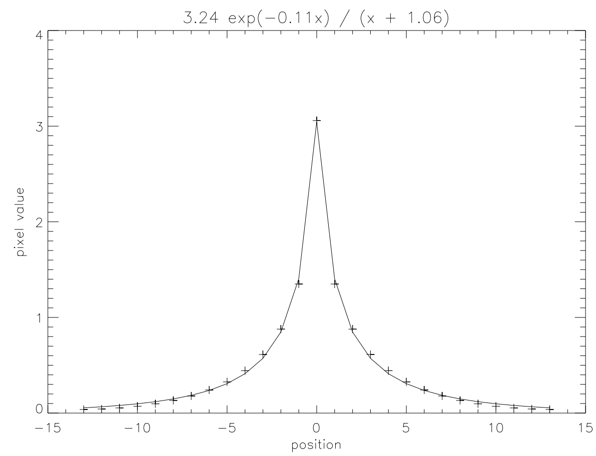

Fig. 4. Horizontal profile through the impulse response (+) for a two-dimensional image with uniform likelihood, with the fitted function (solid line) using expression (11).

The iterative FBP-algorithm applies the following scheme:

$$\lambda^{\text{new}} = \lambda^{\text{old}} + \text{FBP}(y' - \text{proj}(\lambda^{\text{old}})), \quad (35)$$

where y' is the measurement precorrected for attenuation, λ is the reconstruction, and "proj" denotes non-attenuated projection. We used 200 iterations of MLEM and 6 iterations of IFBP to ensure that the impulse response of the unsmoothed reconstructions was very close to an ideal impulse. Consequently, after post-filtering, both reconstructions should have nearly exactly the same impulse response as the penalized-likelihood algorithm.

From the 400 noise realizations with and 400 realizations without signal, the signal-to-noise ratio is computed as follows:

$$\text{SNR}_j = \frac{\text{mean}(\lambda_j^1 - \lambda_j^0)}{\sqrt{(\text{var}(\lambda_j^1) + \text{var}(\lambda_j^0))/2}}, \quad (36)$$

where j is the position of one of the three hot pixels, λ^1 represents the reconstruction with the hot pixels and λ^0 the reconstruction without the hot pixels.

For visual inspection, also the mean and variance images were computed for each of the reconstruction algorithms.

The results were verified using a second, very different simulation object, shown in figure 3. It is a simplified simulation of a PET-study of the thorax. Three hot pixels were inserted, two in the lungs and one in the tissue. The point in the tissue was used to define the post-smoothing filter. The image has 100×100 pixels, 128 projections were computed, assuming a contribution of randoms and scatter (q_i in (12)) of 28%. Due to the asymmetry of the attenuation, the local impulse response function is very asymmetric if a uniform penalty is used [9]. For this image, 200 MLEM iterations did not yet produce a sufficiently sharp impulse response function. Therefore, the equivalent of about 500 iterations were computed using ordered subsets acceleration (OSEM) [19]. We used a decreasing number of subsets (16, 8, 4, 2, 1) and applied 16 iterations for each of those. The same was done for the OCPL algorithm, and 10 iterations of IFBP were applied. For the rest, the processing was identical as for the elliptical phantom.

IV. Results

A. The shape of the local impulse response

Applying 200 iterations of a simple gradient ascent algorithm seemed sufficient to reach convergence (more iterations did not produce visible changes). Figure 4 shows the horizontal profile extracted from the image, together with the curve produced by fitting (11) to the profile. The impulse had a value of 100, the fitted parameters were a = 3.24, $\ln(b)$ = 0.11 per pixel and ϵ = 1.06 pixels.

B. Evaluation of the new method

Figure 5 shows the PET-images obtained with the four reconstruction programs. In figure 6 profiles along the circles are shown. They are computed by scanning the pixel positions on the circles in the true image (figure 1) and extracting the corresponding reconstructed pixel values. The profiles along the two circles are shown in the same plot. Ideally, the concatenated profiles should form a single flat curve, because the two circles have identical and constant intensity. Figure 7 and 8 show the corresponding results for the SPECT simulation. The uniform quadratic penalty produces a very non-uniform reconstruction, and the two profiles have a different mean value. With the CPL-algorithm, the non-uniformity is reduced and the mean values of the two profiles are now much closer, indicating that some sources of position dependent resolution have been removed. With the OCPL-algorithm, the profiles are more uniform, although still not as uniform as those produced by post-smoothed MLEM. Also some oriented artifacts near the object boundary are visible, in particular in figure 5 (one of them indicated with an arrow). They are most likely caused by imperfect transition from one of the four smoothing directions

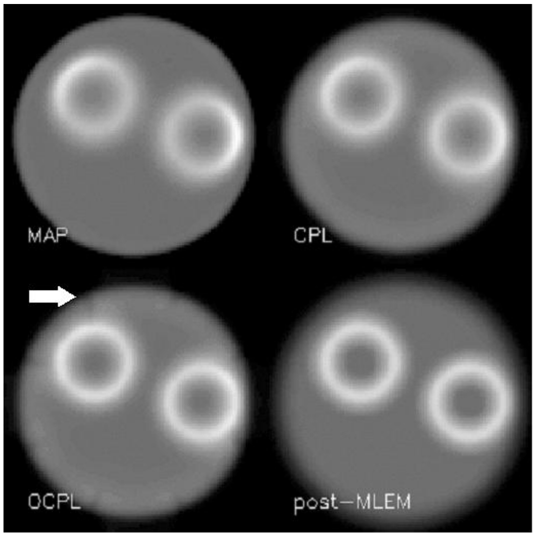

Fig. 5. The reconstructions of the PET simulations: the MAP-reconstruction with quadratic penalty, CPL-reconstruction, OCPL-reconstruction and post-smoothed MLEM-reconstruction. The arrow points at an artifact (see text)

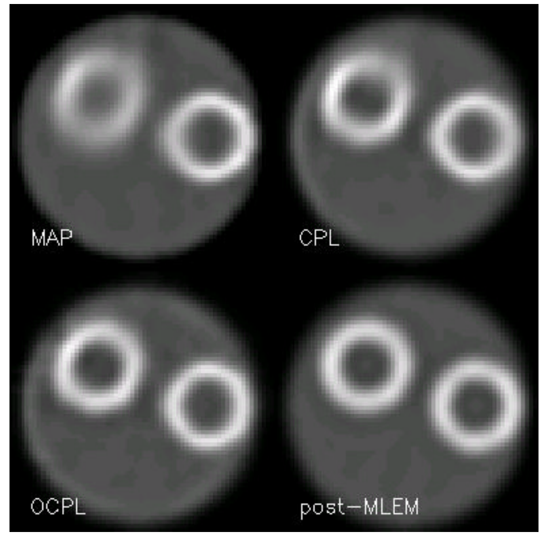

Fig. 7. The reconstructions of the SPECT simulations: the MAP-reconstruction with quadratic penalty, CPL-reconstruction, OCPL-reconstruction and post-smoothed MLEM-reconstruction.

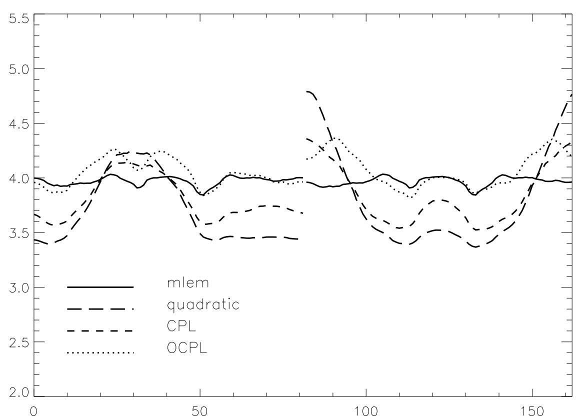

Fig. 6. Profiles along the circles in the PET-images of figure 5. The x-axis corresponds to the position on the perimeter of the circles, the y-axis is the reconstructed value at that position. Solid line: MLEM. Long dashes: uniform quadratic penalty. Short dashes: CPL method. Dotted line: OCPL method.

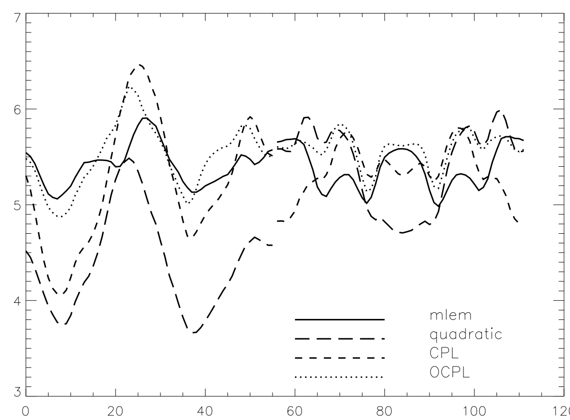

Fig. 8. Profiles along the circles in the SPECT-images of figure 7. The x-axis corresponds to the position on the perimeter of the circles, the y-axis is the reconstructed value at that position.

to the other.

C. Comparison to post-smoothed MLEM

Figure 9 and figure 11 show the variance and mean images computed from the 400 noise realizations, for each of the reconstruction algorithms. Because there was no non-negativity constraint in IFBP, this algorithm produces noticeable variance in the background. In the mean images, a small overshoot near the boundary of the object is seen for the OCPL-algorithm.

The mean image in absence of hot pixels was subtracted from the mean image with hot pixels, to generate the local impulse responses at the three hot pixel positions. For each local impulse response, four profiles (horizontal, vertical, and the two diagonal ones) were extracted by sampling along oriented straight line intervals through the center of the impulse response. The profiles are plotted in figures 10 and 12. The profiles for the three algorithms are nearly identical in all four directions, confirming that a close match of spatial resolution was achieved.

Table I shows the signal-to-noise ratios for each of the points. With 400 simulations, the relative error on the standard deviation should be about $\sqrt{1/(2*400)}$ = 3.5%. The error on the signal is smaller than that, so the signal-to-noise ratio has a relative error of about 3.5%. In each case, point 1 was the hot pixel that was used to define the local impulse response function. The signal-to-noise ratio was best for post-smoothed MLEM, but the performance differences are relatively small and position dependent.

Finally, figure 13 compares the coefficients of variation in every pixel, for the three algorithms and for the thorax phantom. Images are produced by setting a pixel to 1 if the ratio of standard deviation and mean in that pixel is lower with one algorithm than with the other. Of course, this figure

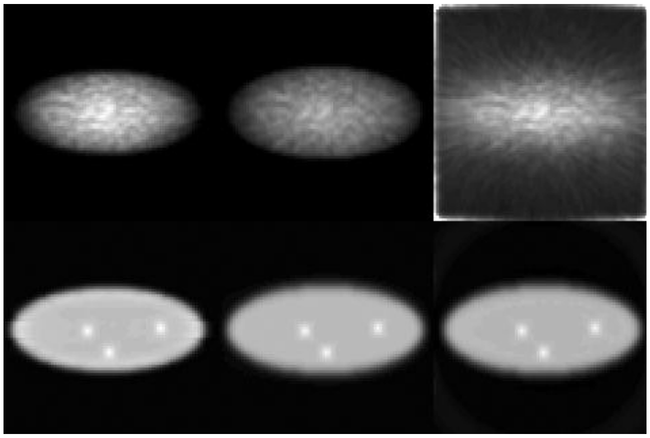

Fig. 9. The variance (top) and mean (bottom) images, computed from the 400 Poisson noise realizations of the elliptical object. Left: OCPL-reconstruction, center: post-smoothed MLEM, right: iterative filtered backprojection. The images on the same row are displayed with the same gray scale

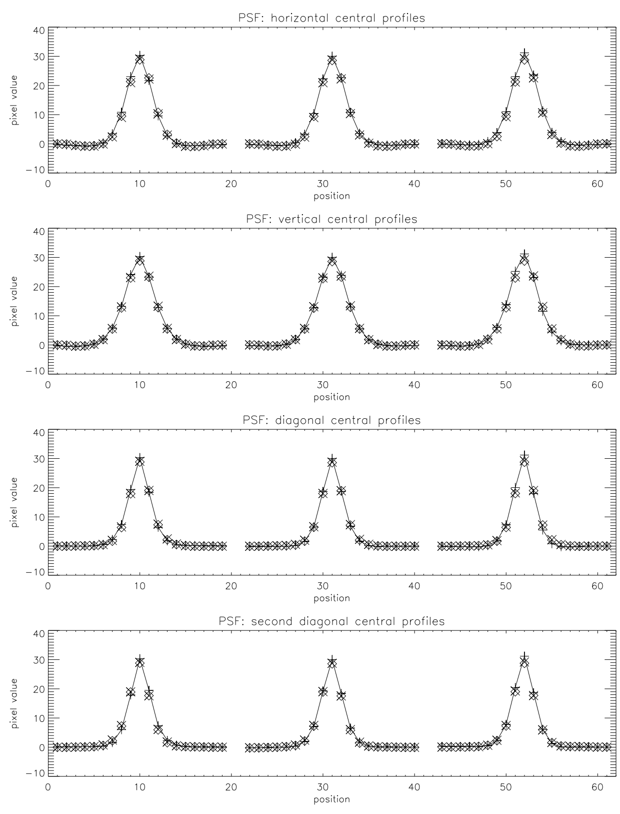

Fig. 10. Profiles along straight lines through the three impulse responses in the elliptical object. For each of the three points, the profile along the horizontal, vertical and the two diagonal axis was computed. Symbols: + for OCPL, x for post-smoothed MLEM and diamonds for IFBP.

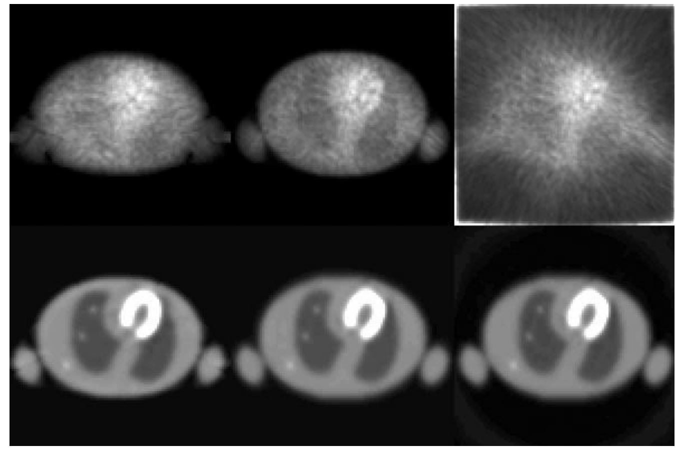

Fig. 11. The variance (top) and mean (bottom) images, computed from the 400 Poisson noise realizations for the thorax phantom. Left: OCPL-reconstruction, center: post-smoothed MLEM, right: iterative filtered back-projection. The images on the same row are displayed with the same gray scale

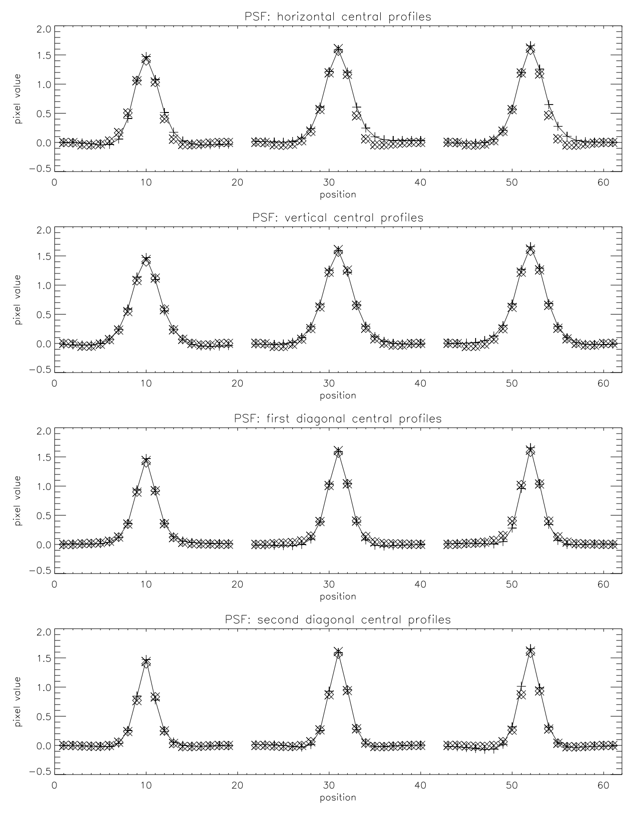

Fig. 12. Profiles along straight lines through the three impulse responses in the thorax phantom. For each of the three points, the profile along the horizontal, vertical and the two diagonal axis was computed. Symbols: + for OCPL, x for post-smoothed MLEM and diamonds for IFBP.

TABLE I

THE SIGNAL-TO-NOISE RATIO'S FOR THE THREE POINTS IN THE MONTE CARLO SIMULATION FOR THE THREE RECONSTRUCTION ALGORITHMS (OCPL, POST-SMOOTHED MLEM AND IFBP, AND FOR THE TWO SOFTWARE PHANTOMS.

Elliptic object			
point	OCPL	pMLEM	IFBP
1	17.2	18.4	15.8
2	14.1	15.4	13.7
3	16.9	18.0	17.1
Thorax phantom			
point	OCPL	pMLEM	IFBP
1	4.35	4.37	4.18
2	4.34	4.63	4.35
3	4.55	4.74	4.39

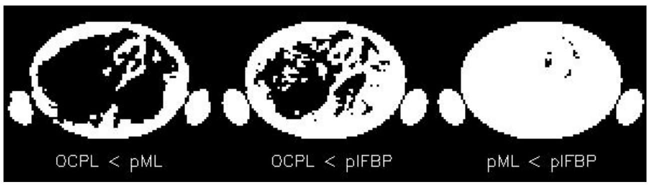

Fig. 13. Comparison of the coefficient-of-variation (cov) images. Left: pixels are set to white where OCPL-cov was lower than post-smoothed MLEM-cov. Center: OCPL-cov lower than post-smoothed IFBP-cov. Right: post-smoothed MLEM-cov lower than post-smoothed IFBP-cov.

provides no information about signal recovery or signal-to-noise ratios.

V. DISCUSSION

The first experiment confirms the derivation of the expression for the local impulse response function of a quadratic prior in combination with a shift-invariant likelihood function. In this simple denoising problem, the prior produces an exponential impulse response, with a narrow peak and relatively large extent. In [20], the impulse response function was studied for an idealized tomograph, where the sinogram has position independent noise properties. For a tomograph with ideal resolution, similar shapes were observed as reported here, but the shapes change if more realistic detector blurring is taken into account. These findings indicate that the penalized likelihood approach offers little control over the shape of the impulse response, which can be very different from that of the low-pass filters that are commonly used in nuclear medicine applications. For some applications, the freedom to choose any shape for the impulse response may be an advantage for post-smoothed MLEM over penalized-likelihood methods. Consequently, in studies comparing penalized-likelihood with traditional filtering, care must be taken to eliminate the influence of the different impulse responses of the methods.

The penalized-likelihood method (OCPL) can be applied to both PET and SPECT, and marked improvements were obtained on simulation studies for both modalities. The profiles show that the uniformity of signal recovery was similar, though still inferior to that obtained with post-smoothed MLEM after 200 iterations. This may be partly due to incomplete convergence: even at 200 iterations the MLEM algorithm was not fully converged, which is why the corresponding profiles are not flat. Convergence is slow, in particular for SPECT, where both attenuation and collimator blurring must be compensated. As noted in [8], the penalty improves the conditioning of the problem, which could be exploited to design faster optimization algorithms. Our algorithm [5] is a straightforward extension of MLEM and may not converge faster than MLEM. Moreover, because the OCPL method focuses on only four different smoothing axes, it is expected that some non-uniformity will persist at any iteration number. This is probably also the cause of the oriented artifacts near the object boundary in figure 5.

The performance of the OCPL-method degrades near the object boundaries. A small overshoot phenomenon is visible in the reconstruction, e.g. in the mean image of figure 9 and in figure 7. The corresponding variance image reveals a lower variance near the boundaries than for the other algorithms, suggesting that the boundaries are being oversmoothed. A similar decrease of performance was observed with the method of Stayman et al. [9], [21].

For the Monte Carlo simulation at matched resolution, post-smoothed MLEM achieved a better signal-to-noise ratio than post-smoothed IFBP and OCPL. The performance difference is different for each point, and seems to be higher when the asymmetry in detection probabilities is more pronounced. Figure 13 compares the coefficients of variation in every pixel. This is only meaningful if we can assume that the local impulse response function is uniform as intended, which can only be verified in the three hot pixels. This figure suggests that post-smoothed MLEM outperforms OCPL, which in turn outperforms IFBP, but as indicated by table I, the performance differences are relatively small. With their more sophisticated method, Stayman et al. [21] obtained identical noise performance for post-smoothed MLEM and their new method. Probably, the approximations made in the derivation of OCPL have resulted in somewhat degraded noise performance. However, comparison of the results is difficult because they were obtained for different configurations (SPECT in [21] and PET in our study). In any case, these studies suggest that post-smoothed MLEM has excellent noise characteristics, which are not improved by including the smoothing as a penalty in our penalized-likelihood methods. Moreover, the impulse response in MLEM tends to be more uniform than with penalized-likelihood methods, because the latter have a suboptimal performance near the object boundaries.

For application in clinical practice, several options exist. Post-smoothed MLEM has a very low implementation cost, since MLEM is now available in the system software of most emission tomography systems. Moreover, it allows free selection of the shape of the impulse response, in contrast to the penalized-likelihood method. So straightforward application of post-smoothed MLEM seems the obvious choice. However, as illustrated by our simulation experiments, a very high number of iterations is required to ensure that the MLEM impulse response is small compared to that of the target resolution. It is currently common practice to apply a few tens of MLEM-iterations (or OSEM-subiterations). This number should be raised to a few hundreds to ensure uniform resolution, in particular when the aim is to (partially) compensate for the loss of resolution due to the system response (e.g. collimator blurring in SPECT). In this work, we have used either pure MLEM-iterations, or OSEM-schemes in which the number of subsets gradually decreases to unity (pure MLEM). When a fixed and high number of subsets is used for stronger acceleration, OSEM converges to a limit cycle with inferior noise characteristics [22], so the conclusions of our paper cannot be extrapolated to such OSEM schemes.

As suggested by Stayman et al. [21], convergence speed may be a reason to use a penalized-likelihood approach as a kind of acceleration technique: the penalty improves the condition number, which can be exploited to obtain faster convergence than with unregularized MLEM. In order to avoid possible suboptimal response of the penalized-likelihood method, or to allow more freedom in selecting the shape of the impulse response, it could be combined with post-smoothing, or even with post-smoothed MLEM as a finishing touch.

Finally, it should be noted that we have only studied a quadratic penalty, applied to emission tomography. No conclusions can be drawn about the relation between non-quadratic penalties and linear or non-linear post-filtering. Similarly, the results cannot be extrapolated to transmission tomography, because there, in contrast to emission tomography, the measurements are a highly non-linear function of the parameters to be estimated.

VI. Conclusion

The impulse response typically produced by penalized-likelihood methods with a quadratic penalty tends to have a relatively sharp peak and wide extent.

Our simulation experiments confirm that the new penalized-likelihood method (OCPL) achieves nearly uniform resolution. However, its noise characteristics are not superior to that of post-smoothed MLEM. This finding calls for further study of the performance differences between post-smoothed MLEM and penalized-likelihood methods.

References

[1] DS Lalush, EC Frey, BM Tsui, “Fast maximum entropy approximation in SPECT using the RBI-MAP algorithm,” *IEEE Trans Med Imaging,* vol 19, 2000; pp. 286-294, 2000.

[2] SJ Lee, Y Choi, GR Gindi, “Validation of new Gibbs priors for Bayesian tomographic reconstruction using numerical studies and physically acquired data,” *IEEE Trans Nucl Sci* vol 46, pp. 2154-2161,1999.

[3] J Qi, RM Leahy, SR Cherry, A Chatziioannou, TH Farquhar, “High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner,” *Phys Med Biol* vol 43, pp. 1001-1013, 1998.

[4] S Alenius, U Ruotsalainen, “Bayesian image reconstruction for emission tomography based on median root prior,” *Eur J Nucl Med* vol 24, pp. 258-265, 1997.

[5] 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.

[6] S. Geman, D.E. McClure., “Statistical methods for tomographic image reconstruction,” *Bull. Int. Stat. Inst.* vol 52-4, pp. 5-21, 1987

[7] JF Vansteenkiste, SG Stroobants, PJ Dupont, PR De Leyn, EK Verbeken, GJ Deneffe, LA Mortelmans, MG Demedts. “Prognostic importance of the standardized uptake value on (18)F-fluoro-2-deoxy-glucose-positron emission tomography scan in non-small-cell lung cancer: An analysis of 125 cases.” *J Clin Oncol* vol 17, pp. 3201-3206, 1999.

[8] JA Fessler, WL Rogers, “Spatial resolution properties of penalized-likelihood image reconstruction: space-invariant tomographs,” *IEEE Trans Image Proc* vol 5, pp 1346-1358, 1996.

[9] JW Stayman, JA Fessler, “Regularization for uniform spatial resolution properties in penalized-likelihood image reconstruction,” *IEEE Trans Med Imaging* vol 19, pp 601-615, 2000.

[10] JW Stayman, JA Fessler, “Nonnegative definite quadratic penalty design for penalized-likelihood reconstruction,” *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, 2001.

[11] BW Silverman, MC Jones, JD Wilson, DW Nychka. “A smoothed EM approach to indirect estimation problems, with particular reference to stereology and emission tomography,” *J R Statist Soc, B,* vol 52, pp. 271-324, 1990.

[12] ETP Slijpen, FJ Beekman. “Comparison of post-filtering and filtering between iterations for SPECT reconstruction,” *IEEE Trans Nucl Sci* vol 46, pp. 2233-2238, 1999.

[13] S Mustafovic, K Thielemans, D Hogg, P Bloomfield, “Object dependency of resolution and convergence rate in OSEM with filtering,” *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, 2001.

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

[15] J Nuyts, “On estimating the variance of post-smoothed MLEM images,” *IEEE Trans Nucl Sci*, vol 49, pp. 714-721, 2002.

[16] M Unser, A Aldroubi, M Eden. “Recursive regularization filters: design, properties and applications,” *IEEE Trans Pattern Anal Machine Intell*, vol 13, pp 272-277, 1991.

[17] 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.

[18] AW McCarthy, MI Miller, “Maximum likelihood SPECT in clinical computation times using mesh-connected parallel computers”, *IEEE Trans Med Imaging,* vol 10, pp. 426-436, 1991.

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

[20] JA Fessler, “Spatial resolution properties of penalized weighted least-squares tomographic image reconstruction with model mismatch,” Technical Report 308, Comm. and Sign. Proc. Lab., Dept. of EECS, Univ. of Michicgan, Ann Arbor, MI, 48109-2122, Mar 1997.

[21] JW Stayman, JA Fessler, “Compensation for nonuniform resolution using penalized-likelihood reconstruction in space-variant imaging systems,” submitted to *IEEE Trans Med Imaging*.

[22] P Schmidlin, ME Bellemann, G Brix. “Subsets and overrelaxation in iterative image reconstruction.” *Phys Med Biol* vol 44, pp. 1385-1396, 1999.