Generalization of Median Root Prior Reconstruction

Sakari Alenius*, *Member, IEEE*, and Ulla Ruotsalainen, *Member, IEEE*

***Abstract*—Penalized iterative algorithms for image reconstruction in emission tomography contain conditions on which kind of images are accepted as solutions. The penalty term has commonly been a function of pairwise pixel differences in the activity in a local neighborhood, such that smooth images are favored. Attempts to ensure better edge and detail preservation involve difficult tailoring of parameter values or the penalty function itself. The previously introduced median root prior (MRP) favors locally monotonic images. MRP preserves sharp edges while reducing locally nonmonotonic noise at the same time. Quantitative properties of MRP are good, because differences in the neighboring pixel values are not penalized as such. The median is used as an estimate for a penalty reference, against which the pixel value is compared when setting the penalty. In order to generalize the class of MRP-type of priors, the standard median was replaced by other order statistic operations, the L and finite-impluse-response median hybrid (FMH) filters. They allow for smoother appearance as they apply linear weighting together with robust nonlinear operations. The images reconstructed using the new MRP-L and MRP-FMH priors are visually more conventional. Good quantitative properties of MRP are not significantly altered by the new priors.**

***Index Terms*—*A priori* penalty terms, iterative reconstruction, median root prior, order statistics.**

I. Introduction

In ITERATIVE reconstruction, the reconstructed image is a solution of a maximization of an objective function. In maximum-likelihood expectation maximization (MLEM) [1], [2], the maximization is a data-fitting process, such that the solution is the image that makes the measured data most likely to occur. The problem in emission tomography is that with noisy data the reconstructed image estimate is also noisy. An ill-posed reconstruction problem, such as MLEM, can be made a well-posed one by controlling which solutions are most favorable. The image is required to fit with measured data to some extent, and also be consistent with additional regularizing criteria that are set independently of the data. Depending on the point of view, these restrictions can be considered as penalty functions, Bayesian priors, or as Tikhonov regularization [3]–[5]. They push the solution toward an assumption about the nature of the true image. If the image is thought to be smooth, the prior can be called a smoothing prior. Many proposed priors or penalty functions are further tailored such that large edges are not much blurred [6], [23]. It is important in emission tomography that the reconstructed image is quantitatively accurate and noise is reduced sufficiently at the same time.

The median root prior (MRP) algorithm [7] was developed based on a general assumption of the unknown image: *the desired image is locally monotonic*. Pixel values are spatially nonincreasing or nondecreasing in a local neighborhood. This is accomplished using median filtering, as described in Section II-B. As the smoothing prior does not penalize a flat image, MRP passes any locally monotonic image without penalty, which is a more realistic description for the true image. This assumption of local monotonicity of the true object is reasonable for both emission and transmission. The monotonicity includes sharp step edges, shallow ramp edges, and arbitrarily smooth edges of any height. MRP contains implicitly a general description of the unknown tracer concentration. No special knowledge of the appearance of the true image is required.

Unlike most priors, MRP does not distinguish noise from signal based on amplitude differences of neighboring pixels, but on the spatial size of a detail. As the median is used as an estimator of a penalty reference for each pixel, MRP penalizes only details smaller than a certain spatial size: individual pixels or too few pixels of different amplitude from their neighborhood do not pass median filtering. If they are many enough, it is more probable that they form a real detail and that passes median filtering. The edge preservation is a built-in feature in median filtering, independent of the height of the edge. It is important that no quantitative assumptions are made. Any differences in activity between details of sufficient spatial size are allowed independently of the type or the height of the spatial transition from one activity level to other. This makes it possible, at least in principle, to reveal small changes in activity with simultaneously suppressing the surrounding noise.

MRP is general, robust in use, and quantitatively accurate [8], [9]. However, it tends to generate sharp edges of small height onto flat and noisy areas. Although the noise is effectively reduced, the remaining blocks may be disturbing to the human eye. This effect, called streaking [10], arises when the output of median filter happens to be the same for adjacent window locations.

The aim here is to generalize the MRP prior to a class of priors where the penalty is set according to a local reference that is produced by a robust order-statistic estimator [10]. A specific goal is to replace the median with such an operation, that produces more conventional visual appearance but preserves the quantitative accuracy. A similar goal was aimed in a study applying a three-dimensional (3-D) median in volume space, and those results are reported elsewhere [11]. Visually, the reconstructed images using new priors from the MRP family and smoothing

Manuscript received July 16, 2001; revised June 20, 2002. This work was supported by the Academy of Finland. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was J. Fessler. *Asterisk indicates corresponding author.*

*S. Alenius is with the Institute of Signal Processing, Tampere University of Technology, P.O. Box 553, FIN-33101 Tampere, Finland (e-mail: sakari.alenius@tut.fi).

U. Ruotsalainen is with the Tampere University of Technology, Institute of Signal Processing, FIN-33101 Tampere, Finland.

Digital Object Identifier 10.1109/TMI.2002.806415

priors are quite similar, and that is intentional. The main difference in the new penalties is that they are based on order statistics and they are more robust with respect to the parameter values, thus, preserving the quantitative quality.

The form of the selected algorithms is such that they differ only in the way the penalty reference is computed. That makes the comparison straightforward, but may result in nonstandard forms of priors.

II. Methods

A. Priors Penalizing Neighboring Pixel Differences

In MLEM, the objective function is based on the conditional probability $f(\boldsymbol{y}|\boldsymbol{\lambda})$, where $\boldsymbol{y}$ is the measured data, and $\boldsymbol{\lambda}$ is the emission image. The penalized solution $\hat{\boldsymbol{\lambda}}$ is of the form

$$\hat{\boldsymbol{\lambda}} = \arg\max_{\boldsymbol{\lambda}\geq 0}[L(\boldsymbol{\lambda}) + P(\boldsymbol{\lambda})] \tag{1}$$

where $L()$ is the log likelihood and $P()$ is the penalty term. In Bayesian terms, (1) represents the *a posteriori* probability $f(\boldsymbol{\lambda}|\boldsymbol{y}) \propto f(\boldsymbol{y}|\boldsymbol{\lambda})f_p(\boldsymbol{\lambda})$ and $P()$ is the prior term resulting from the *a priori* probability $f_p()$. A common Bayesian prior is the Gibbs distribution of the form [3]

$$\begin{aligned} f_p(\boldsymbol{\lambda}) &= Ce^{-\beta U(\boldsymbol{\lambda})} \\ &= Ce^{-\beta \sum_b U(\boldsymbol{\lambda}, b)} \end{aligned} \tag{2}$$

where β is the Bayes weight of the prior and C is a normalizing constant. When the image $\boldsymbol{\lambda}$ meets the prior assumptions, the energy function $U(\boldsymbol{\lambda})$ has its minimum and the prior has its maximum. $U(\boldsymbol{\lambda}, b)$ is the notation for the value of the energy function $U()$ evaluated on $\boldsymbol{\lambda}$ at pixel index b. A common choice for $U()$ in (2) is an energy function computed using a potential function $v()$ of the differences between pixels in the neighborhood N_b

$$\beta U(\boldsymbol{\lambda}, b) = \beta \sum_{i\in N_b} w_{bi} v(\lambda_b - \lambda_i) \tag{3}$$

where w_{bi} is the weight of pixel i in the neighborhood of pixel b [6], [23].

The one-step-late (OSL) algorithm uses the current image $\boldsymbol{\lambda}^{\langle k\rangle}$ when calculating the value of the derivative of the energy function $U()$ [12]. The OSL solution is

$$\begin{aligned} \lambda_b^{\langle k+1\rangle} &= \frac{\lambda_b^{\langle k\rangle}}{\sum_d p_{db} + \beta \frac{\partial}{\partial \lambda_b} U(\boldsymbol{\lambda}, b)\Big|_{\boldsymbol{\lambda}=\boldsymbol{\lambda}^{\langle k\rangle}}} \underbrace{\sum_d \frac{y_d p_{db}}{\sum_{b'} \lambda_{b'}^{\langle k\rangle} p_{db'}}}_{c_b^{L\langle k\rangle}} \\ &= \lambda_b^{\langle k\rangle} c_b^{P\langle k\rangle} c_b^{L\langle k\rangle} \end{aligned} \tag{4}$$

where $\lambda^{\langle k\rangle}$ is the kth iteration emission image, b is the pixel index, d is the sinogram bin index, p_{db} is a system matrix element, and $\boldsymbol{y}$ is the measured sinogram. The current image estimate $\lambda^{\langle k\rangle}$ is updated using two factors: $c_b^{P\langle k\rangle}$, which changes the pixel value such that prior assumptions are better met, and $c_b^{L\langle k\rangle}$ for better data fitting.

The potential function $v()$ in (3) defines the properties of the prior. With the quadratic choice, $v(r) = r^2$, the derivative in (4) is $2r$ and the penalty term is linear with respect to the pixel difference $r = \lambda_b^{\langle k\rangle} - \lambda_i^{\langle k\rangle}$, and the penalty factor to be used in (4) is

$$\begin{aligned} c_b^{P\langle k\rangle} &= \frac{1}{\sum_d p_{db} + 2\beta \sum_{i\in N_b} w_{bi}\left(\lambda_b^{\langle k\rangle} - \lambda_i^{\langle k\rangle}\right)} \\ &= \frac{1}{\sum_d p_{db} + 2\beta\left(\lambda_b^{\langle k\rangle} - \sum_{i\in N_b} w_{bi}\lambda_i^{\langle k\rangle}\right)} \\ &= \frac{1}{\sum_d p_{db} + 2\beta\left(\lambda_b^{\langle k\rangle} - A_b\right)} \end{aligned} \tag{5}$$

assuming $\sum_{i\in N_b} w_{bi} = 1$. The penalty is set with respect to the sum $A_b = \sum_{i\in N_b} w_{bi}\lambda_i^{\langle k\rangle}$, against which the current pixel $\lambda_b^{\langle k\rangle}$ is compared. Effectively, the penalty reference A_b is the output of a linear finite-impulse-response (FIR) filter [13] with filter coefficients w. With the proposed set of w [6], [23], the FIR is a low-pass filter [13], and nonsmooth and noisy images are penalized. Images that are close to the result of the reference FIR filter are smooth and they are not much penalized. This results in blurring of sharp edges between different activity concentrations in the tissues. The prior (5) can, therefore, be called the smoothing prior. In order the prevent excessive edge blurring, other choices of $v()$ or extra tuning parameters can be used so that pixel differences that are large enough for representing a real edge are smoothed less [6], [23]. However, such a threshold is difficult to set in general [14].

If β in (5) is replaced by $\beta/2A_b$, a relative form of the smoothing prior is formed

$$c_b^{P\langle k\rangle} = \frac{1}{\sum_d p_{db} + \beta\frac{\lambda_b^{\langle k\rangle} - A_b}{A_b}}. \tag{6}$$

Then β is less dependent on quantitative values and it is easier to set.

B. MRP

Median filtering does not change locally monotonic images, which are called roots [10]. The penalty can be restricted to only nonmonotonic structures in a neighborhood by comparing the pixel against the local median. Using this constraint in the term $U()$ of (4), the penalty factor $c_b^{P\langle k\rangle}$ of (4) for MRP is

$$c_b^{P\langle k\rangle} = \frac{1}{\sum_d p_{db} + \beta\frac{\lambda_b^{\langle k\rangle} - M_b}{M_b}} \tag{7}$$

where the penalty reference $M_b = Med(\lambda^{\langle k\rangle};\ b)$ is estimated as the median of the pixels in a neighborhood N_b. $0 < \beta \leq 1$ is the weight of the prior [7], [9], [15]. Deviations of the center pixel from the median of the pixels in the neighborhood are penalized. The difference between adjacent pixel values is not penalized in MRP. The penalty term $(\lambda_b^{\langle k\rangle} - M_b)/M_b$ in (7) is piecewise linear with respect to the difference between the center pixel and local median. That difference can be nonzero only if the image is not locally monotonic. Thus, noise is penalized because it is nonmonotonic inside a small neighborhood (usually, 3×3 or 5×5 pixels). The locally monotonic part

represents the signal. Large as well as small quantitative differences between two areas of different activity pass the prior without penalty, if the areas are spatially large enough so that center pixels are close to the running median. This is important concerning the sensitivity. There is no need to suppress the penalty function for large differences as they can only occur if the center pixel value is an outlier. The median operation preservers edges without data-dependent tuning of the derivative of the energy function.

An intuitive interpretation for the prior probability $f_p()$ is a Gaussian located at M_b [7], [9]. This prior probability is approximative because of the difficulties in the analytical derivation of expressions involving the median. Strictly speaking, the priors based on (6) or (7) are not actual priors as they are defined in such a way that the hyperparameters A_b and M_b depend on the data and on the algorithm-dependent iterates $\lambda^{\langle k \rangle}$, thereby making it difficult to define an explicit prior function $P(\boldsymbol{\lambda})$. Other edge-preserving regularization methods in the literature [16] avoid this difficulty by defining an explicit $P(\boldsymbol{\lambda})$ and then deriving a convergent iterative algorithm to perform the maximization (1). However, MRP is interesting as the most favorable image is a root signal which can be something other than a flat image.

C. Generalization of MRP Family of Priors

In MRP (7), the local median M_b is used as a penalty reference. The edge preserving property of the standard median is very good for avoiding edge blurring, but due to streaking it tends to transform noise to clear-cut steps of small height in flat areas, especially if the neighborhood (filter) size n is small. These blocks of slightly different values of activity may disturb the visual quality, although the image is quantitatively accurate.

MRP can be generalized by using other robust order-statistics estimators of the penalty reference in place of the median in (7). In order to make the visual appearance more conventional, the L-filter and FIR median hybrid (FMH) were used [17]. The new algorithms are named MRP-L and MRP-FMH, respectively.

1) L-Filter: As the median of set of pixels is the middle one of the ordered values, the L-filter [10] is a linear combination of the *ordered* values. This reduces streaking. Compared with median, it is less likely that L-filter outputs of adjacent window locations are the same. Using the L-filter, the penalty reference M_b in (7) is computed as

$$M_b = L(\lambda_i | i \in N_b;\ \boldsymbol{a}) = \sum_{i \in N_b} a_i \lambda_{(i)}. \tag{8}$$

The subscript in parentheses in $\lambda_{(i)}$ indicates the index of the ordered pixels. First, the neighborhood pixels are ordered according to their quantitative values and the weight a_i in the sum is taken according to their rank. The spatial location of the pixels is ignored. If all a_is are equal and $\sum a_i = 1$, the resulting filter is sample mean; if the middle weight is one and all the rest are zeros, then the resulting filter is median. Thus, the L-filter provides a compromise between median and mean. The idea is to put more weight on the center of the ordered values (as in median) and also include the rest to the sum (as in mean). This combines robustness with smoothness.

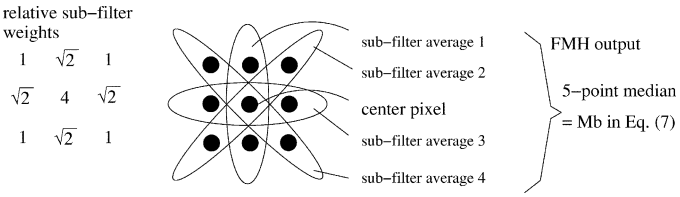


Fig. 1. FMH filter structure. The three-point subfilters are in the vertical and horizontal directions. The weights of each subfilter are normalized for $\sum_i w_i = 1$.

2) FMH: FMH is the median of multiple linear subfilters [10]. The filters used as subfilters in FMH are simple weighted averages, or FIR filters: $F = \sum_i w_i \lambda_i$. There are many possible choices for the FMH operation. The FMH filter used in this paper computes four weighted three-point averages and takes a five-point median of those subfilters and the center pixel. This combines linear operations with the median. The penalty reference in (7) is computed as

$$\begin{aligned} M_b &= \mathrm{FMH}(\lambda_i | i \in N_b;\ \boldsymbol{w}) \\ &= \mathrm{median}\{F1,\ F2,\ \lambda_b,\ F3,\ F4\}. \end{aligned} \tag{9}$$

The subfilters $F1$–$F4$ and the FMH filter structure are depicted in Fig. 1. The choice of the given structure of FMH filter and its weights $\boldsymbol{w}$ was made heuristically, based on experimental data.

D. Pixel and Filter Sizes

The size of the median filter has an effect on the visual quality of images produced by MRP. If the filter size is increased from 3×3 to 5×5, or to 7×7, there are more pixel values in the set where the median is selected. This makes the appearance of the filtered image smoother. When the filter size is increased, the physical size of the neighborhood covered by the filter window is increased. In order to keep the physical size of the neighborhood the same, the pixel size of the image can be decreased at the same time when the filter size is increased. (This is shown in the first row of Fig. 10.) Unfortunately, image sizes of 200–300 pixels in width are impractical.

The new priors are different from MRP in the sense that they produce smoother images also for small filter and image sizes (see Fig. 10). This is because the output of the filter is not limited to the given set of pixel values.

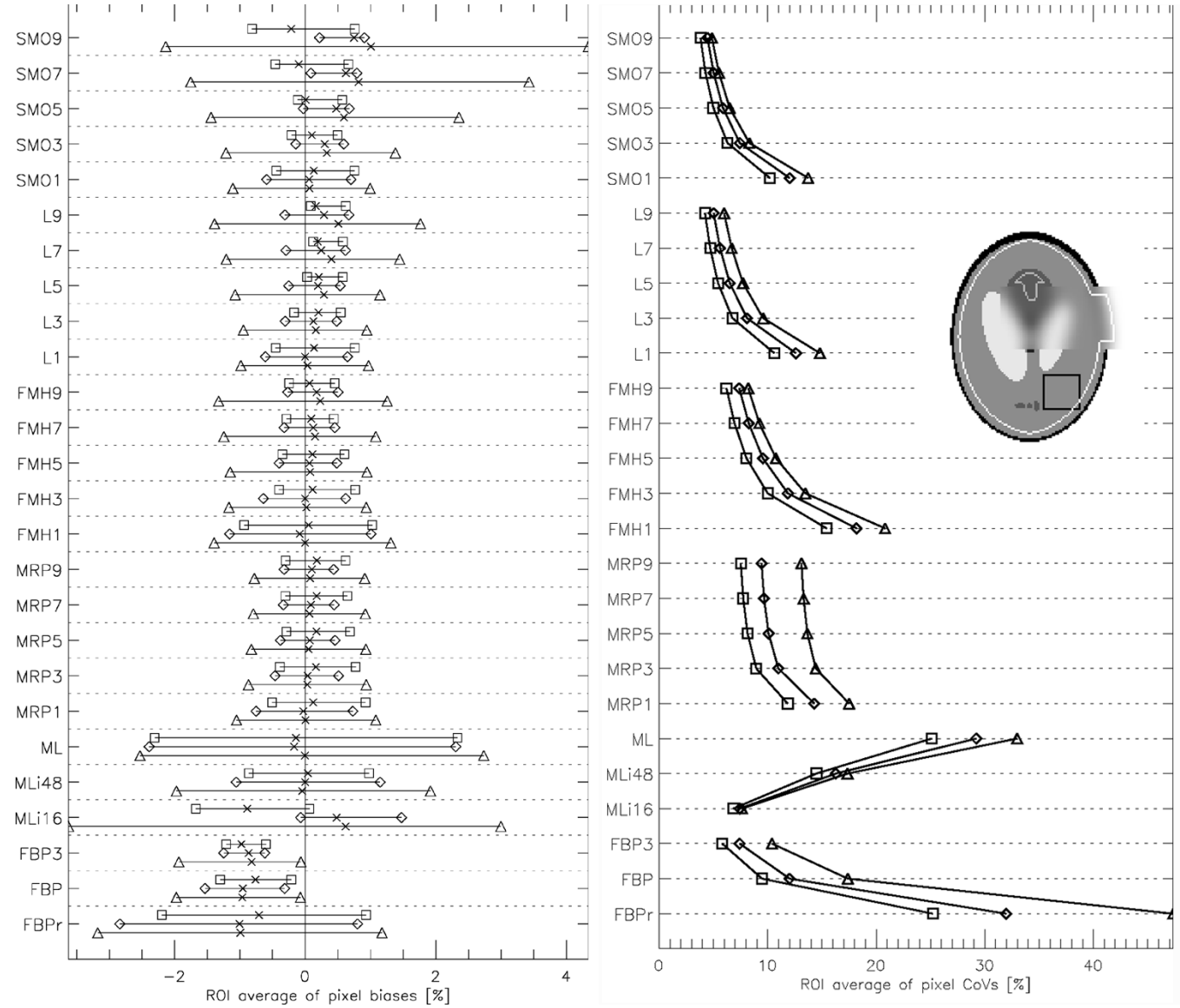
III. Results

A. Algorithms

The following algorithms were used in testing:

- filtered backprojection (FBP) with ramp window, Hann window, and Hann with cutoff of 0.3;
- MLEM [(4) with $c_b^{P\langle k \rangle} \equiv 1$];
- MRP [(4) with $c_b^{P\langle k \rangle}$ from (7)];
- MRP-L [(4) with $c_b^{P\langle k \rangle}$ from (8)];
- MRP-FMH [(4) with $c_b^{P\langle k \rangle}$ from (9)];
- smoothing prior using a weighted average filter [6], [23] [(4) with $c_b^{P\langle k \rangle}$ from (6)].

The L-filter weights a_i for a 3×3 neighborhood are listed in the Appendix. They were taken from [18]. They are optimized



(a) (b)

Fig. 2. The 1/4 and 3/4 quartiles of the (a) bias and (b) CoV of three ROIs of Shepp–Logan phantom with different reconstruction methods. The bias and CoV were computed for each pixel with respect to the 200 noise realizations. From these bias and CoV images, ROI averages were computed. Symbols △ (ROI1), □ (ROI2), and ⋄ (ROI3) indicate each ROI. The number of pixels are 6799 in ROI1 (white line), 140 in ROI2 (gray line), and 420 in ROI3 (black line). Increasing the value of the regularization parameter decreases CoV, but causes some extra bias. The bias and its spread is smaller for priors from the MRP family than for the smoothing prior.

for a Laplacian distribution, but this is not critical as more arbitrary choices did work well.

B. Accuracy and Noise

For testing the effect of the prior and its weight β on both to the quantitative accuracy and to the noise reduction capability, a set of modified noisy Shepp–Logan sinograms were generated by adding 200 Poisson noise realizations [19] with different deviations [20] to the noiseless sinogram. The Shepp–Logan phantom was modified by making one part of the phantom smooth in order to see how a nonpiecewise constant object is reconstructed. The bias and coefficient of variation (CoV, the ratio of the sample standard deviation to the sample mean) were calculated for each pixel with respect to the noise realizations. Two kinds of variability of the pixel values are shown in Fig. 2. Fig. 2(a) shows the spread of the pixelwise bias and Fig. 2(b) pixelwise CoV averaged over three regions of interest (ROIs) of the phantom for various reconstruction algorithms. The notations are as follows:

FBPr	FBP with ramp;
FBP	FBP with Hann;
FBP3	FBP with Hann and cutoff 0.3;
MLiN	MLEM with N iterations (if not indicated, the number of iterations is 150);
L	MRP-L;
FMH	MRP-FMH;
SMO	smoothing prior.

The number after the prior name is $\beta \times 10$. The spread of the bias values indicated by lower and upper quartiles in Fig. 2(a) shows the distribution of the bias inside a ROI. The average bias

is marked by the symbol $\times$. The CoV in Fig. 2(b) indicates the variability of reconstructed pixel values with different noise realizations. Preferably, both bias and CoV should be near zero at the same time, indicating robustness against different noise realizations. The pixel variances were used in order to emphasize the variability of each reconstructed pixel, which is important in computing pixel-by-pixel parametric images. We also examined the ROI bias of the various reconstruction methods and found that all of the statistical methods except SMO9 had ROI biases less than 0.5%.

A wide spread [a long error bar in Fig. 2(a)] of pixel-by-pixel bias values in a ROI indicates a potentially unreliable result when the ROI average is estimated from a single noise realization. MLEM is sensitive to the number of iterations, as the regularized algorithms were run 150 iterations without increase in noise. From Fig. 2(b) can be seen that with priors the reconstruction is less ill posed, that is, changes in data do not generate large changes in the values of a pixel. However, increasing β increases bias when the smoothing prior is used. This effect is less prominent when penalties from the MRP family are used.

Fig. 2(a) and (b) also shows no problems with the convergence of the regularized algorithms, as different noise realizations are reconstructed to a consistent set of images. Unfortunately, convergence analysis is difficult due to the lack of exact first- and second-order derivatives of functions containing order statistics.

C. Edge Width

A cylinder phantom with diameter of 20 cm filled with uniform solution of Ge-68 and water was used for testing edge sharpness. The conventional measure for resolution, full-width at half maximum (FWHM), was not applicable due to the fact that the point spread functions (PSFs) of the reconstruction methods are very different in nature and the FWHM measures do not characterize the resolution in a comparable way. FWHM is a good measure of spreading if the PSFs are of the same shape.

For a more pragmatic choice, the edge widths of a measured phantom were used as a resolution measure. This represents the problem encountered in practice. The edge width was computed as an average width of the edge transition solving the width w from

$$A_1 - A_2 = \pi(R + w/2)^2 - \pi(R - w/2)^2$$

as

$$w = \frac{A_1 - A_2}{2\pi R} \tag{10}$$

where A_1 and A_2 are the areas of thresholded images using thresholds $T_1 = 10\%$ and $T_2 = 90\%$ of mean cylinder activity, respectively. R is the average radius of the edge location in the image.

Fig. 3 plots the edge width versus CoV of the uniform area computed from the reconstructed image. Points near the lower left corner of the plot represent a small pixel variation in the uniform area with a small amount of blurring of the edge at the same time. MRP and MRP-L provide an efficient noise reduction with a good resolution. The edges of the cylinder images are shown in Fig. 4. The edges in MRP images are sharp but blocky. The smoothing prior and MRP-L produce smoothly curving edges but the blurring is smaller for MRP-L.

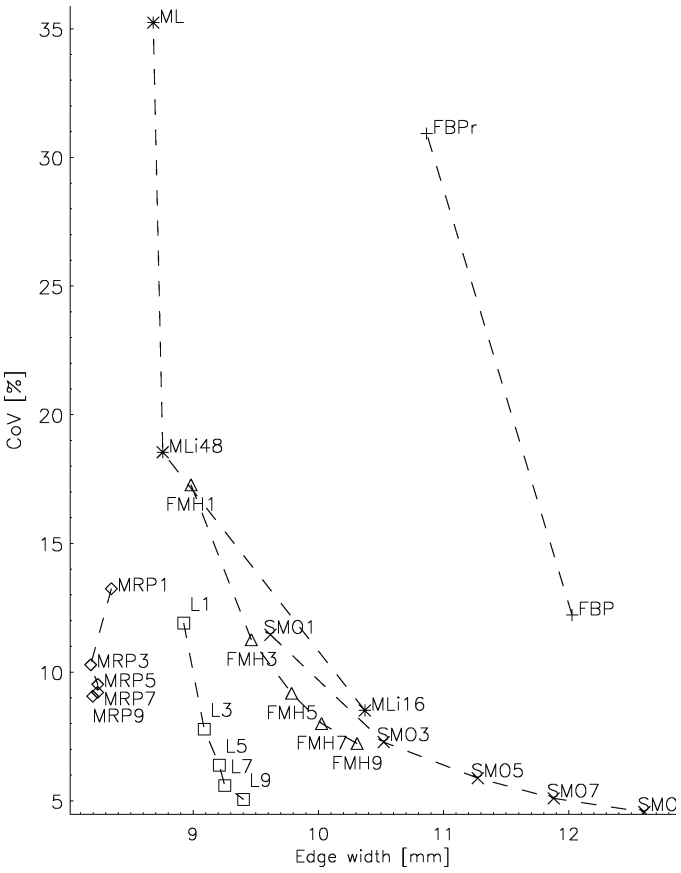


Fig. 3. CoV of the inner cylinder area versus edge width of reconstructed images of a cylinder phantom. The tradeoff between noise reduction and resolution is clearly seen for MLEM and for the smoothing prior but not for MRP.

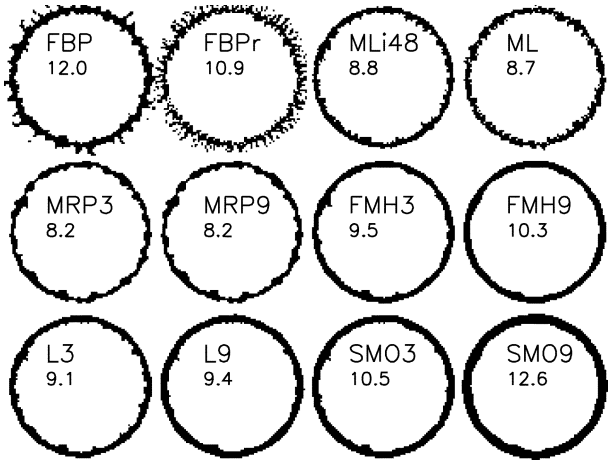


Fig. 4. Edges of reconstructed images of a cylinder phantom. Black pixels indicate the transition from 10% to 90% of the mean activity. The numerical values are the average width of the edge in millimeters (10).

D. Images

Fig. 5 shows a simulation study and Figs. 6 and 7 show several positron emission tomography (PET) fluoro-2-deoxy-D-glucose (FDG) studies reconstructed using FBP, aborted MLEM,

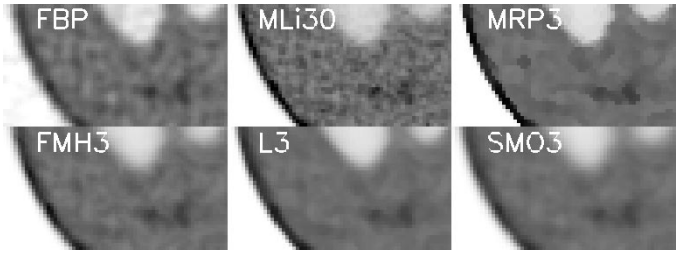


Fig. 5. Enlarged parts of Shepp–Logan phantom images.

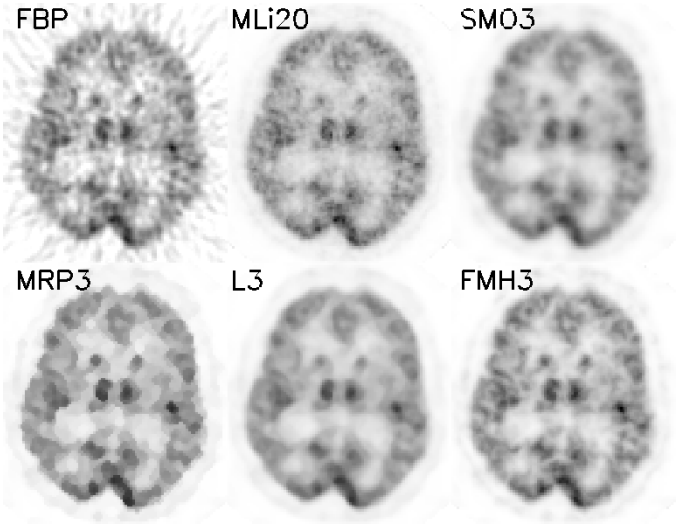


Fig. 6. FDG brain images.

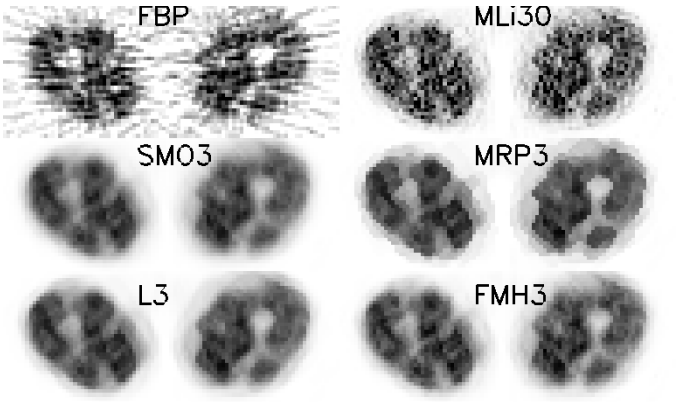


Fig. 7. PET thigh images, FDG. The muscle compartments can be seen more clearly when the noise reduction does not worsen the resolution.

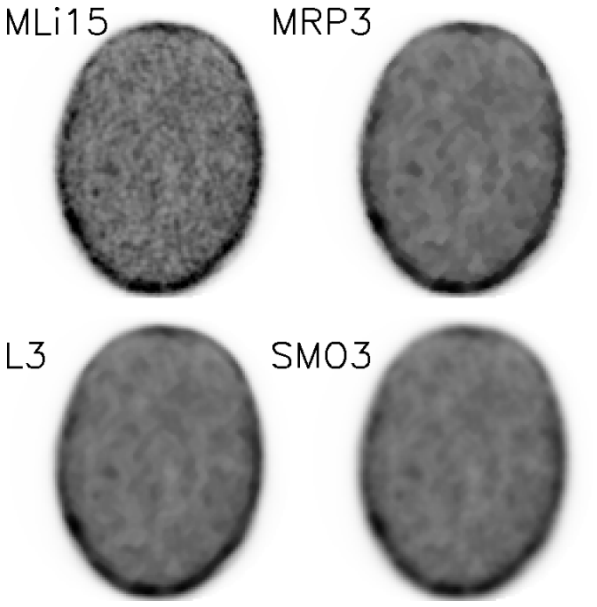


Fig. 8. SPECT transmission images. Regularized reconstructions were run with 45 iterations.

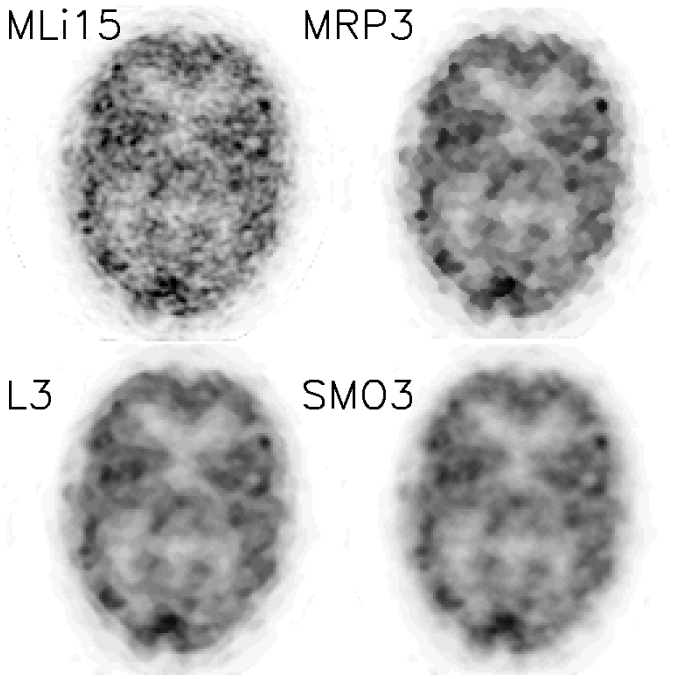


Fig. 9. SPECT attenuation corrected emission images. The corresponding transmission images from Fig. 8 were used as attenuation maps. Regularized reconstructions were run with 90 iterations.

smoothing prior, MRP, MRP-L, and MRP-FMH. The prior weight $\beta = 0.3$; the neighborhood size is 3×3. A total of 100 iterations were run with ordered subsets (OS) factor of two.

As the quantitative properties of the new MRP-L and MRP-FMH were close to MRP, the visual appearance is different and, in fact, closer to the smoothing prior than MRP.

SPECT images are commonly pre- or postfiltered, because the visual quality is important. In Figs. 8 and 9, no pre- or post-filtering was applied; only the priors are used as noise reduction. The results are similar to those of PET. MRP-L is visually close to the smoothing prior, but MRP-L is more robust and preserves edges better.

Fig. 10 shows reconstructions of the Hoffman brain phantom using different image and filter sizes. For MRP, the visual quality becomes more conventional if large images and median filters are used. In the first row of Fig. 10, the image sizes are chosen with respect to the increased filter size ($128/214 \approx 3/5$; $214/298 \approx 5/7$). MRP-L makes a similar kind of appearance with standard image and filter sizes. Also, MRP-FMH does not need large image sizes.

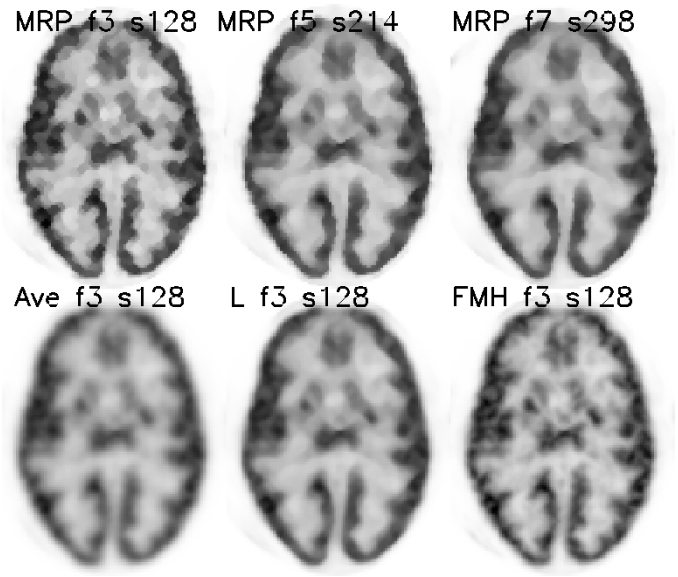


Fig. 10. Effects of different image and filter sizes. Images are 128 × 128 unless otherwise stated. MRP with 3 × 3 filter (top left). MRP with 5 × 5 filter and 214 × 214 image (top middle). MRP with 7 × 7 filter and 298 × 298 image (top right). Smoothing prior with 3 × 3 filter (bottom left). MRP-L with 3 × 3 filter (bottom middle). MRP-FMH with 3 × 3 filter (bottom right).

IV. Discussion

The results in Section III present some aspects of image reconstruction algorithms: quantitative accuracy of noise reduction methods in Section III-B; resolution or sharpness of a uniform object in Section III-C; and the resulting general image quality in Section III-D. The data presented by Fig. 2(a) and (b) demonstrate that the new variants of MRP (MRP-L and MRP-FMH) share the quantitative accuracy and robustness of MRP. Aborted MLEM or smoothing prior suffer from bias/noise tradeoff; an increase of noise reduction causes a local bias (early stopping of MLEM iterations or more effective smoothing cause a shift in the quantitative ROI value). MRP has been shown to behave in a more robust way [9], and Fig. 2(a) and (b) reproduce those results. The new MRP-based methods are also robust. An increase in β adds more weight on the prior, but the local bias is small.

A similar kind of robustness can be seen from Figs. 3 and 4 in terms of edge blurring. Usually, noise reduction increases edge smoothness, which is the case for FBP, MLEM, and smoothing prior. For MRP, increasing the noise reduction weight does not cause edge smearing, but keeps the transition sharp. MRP-FMH and MRP-L do not suffer from extensive edge blurring when the noise reduction becomes more intensive. This is because ranking of pixels in the algorithm assigns more weight to those pixels that are quantitatively close to the center pixel. Thus, the transition of the edge is both smooth and thin for MRP-L.

Figs. 5–9 show that the new MRP-L and MRP-FMH are *visually* quite different from MRP. Thus, the new variants, especially MRP-L, combine the quantitative accuracy and robustness of MRP with more conventional visual quality. The overall image quality is suitable for both automatic postprocessing and manual inspection. Notably, the studies are reconstructed using the same prior parameters.

In MRP, the noise reference M_b in (7) is selected by the median. When the noise varies, the value of median may or may not change from one value to another, thus, the abrupt small edges. But in MRP-L as the noise varies, the value of M_b changes more smoothly as it is a linear combination of pixels. However, the fact that filter weights are set according to pixel ranking ensures that outliers do not disturb the filter output. This makes MRP-L more robust than the smoothing prior, whose weights are fixed according to spatial locations.

MRP-FMH first computes the linear subfilters and then uses the median operation, which results in similar kind of robustness. Because the structure of FMH is designed by intuitive and experimental means, a general form of FMH may be hard to set up.

The choice of the optimal values of filter weights themselves is partly a matter of taste, but fortunately that is quantitatively not very critical. One approach to be carried out in the future is the optimization of the filter weights according to clinical tasks, e.g., by receiver operating characteristic (ROC) analysis.

There are no threshold-based parameters in the L- and FMH filters. Some priors use a parameter for tuning the edge height above which the pixel difference is encouraged to be preserved [6], [23]. The setting of such parameters is difficult [14], especially when small quantitative changes are of interest. Linear filters suffer from the fact that often the real signal and noise overlap in frequency space; threshold tuned filters or priors suffer from the fact that the amplitudes of the signal and noise overlap. In median, L-, and FMH filters the rank of the pixel defines how it is dealt with. The spatial frequency and the amplitude of pixel differences play no direct role.

As mentioned in Section II-D, it is possible to get visually more conventional MRP images by using large images and median filters. Then there are more pixel values (25 or 49 compared to nine) where the median is selected from, and quantitative changes when moving from one filter location to the next are less abrupt in noisy uniform regions. The drawback of that approach is that the image gets too large for practical computational tasks. In contrast, MRP-L and MRP-FMH can produce smooth images for an image size 128 × 128 and a filter size 3 × 3 (Fig. 10).

The computational load of the L-filter itself is somewhat larger than that of median of the same size. FMH has less sorting but due to subfilters the load is of the same magnitude. However, the computational load of these and most other priors is negligible compared to other tasks required during iterations, such as projections between image and sinogram spaces.

MRP-L and MRP-FMH are suitable for transmission image reconstruction as well, using the same approach as in [21] and [22].

V. Conclusion

In the family of MRP-type of priors, the penalty is set with respect to the difference between the pixel and an order statistic, such as median, of the local neighborhood. The noise regularization in the basic form of MRP is effective. The nonlinear median operation providing a penalty reference can be made smoother by adding linear operations together with order-statistic ones.

TABLE I
3×3 L-FILTER WEIGHTS [18]

a_0	a_1	a_2	a_3	a_4
a_8	a_7	a_6	a_5	
-0.01899	0.02904	0.06965	0.23795	0.36469

Both the L-filter and FMH give smoother results than the standard median, but still the edge preservation is better than in conventional linear filtering. The new priors are quantitatively accurate with a moderately wide range of the prior weight β. The combination of effective noise regularization and a good resolution suggests that the generalized MRP reconstruction can improve automatic segmentation of emission tomography images.

APPENDIX

Table I lists the L-filter weights used in (8). In the filter implementation, the weights were normalized such that $\sum a_i = 1$.

ACKNOWLEDGMENT

The authors would like to thank Turku PET Centre for providing the PET data and Kuopio University Hospital for the SPECT data.

REFERENCES

[1] K. Lange and R. Carson, "EM reconstruction algorithms for emission and transmission tomography," *J. Comput. Assist. Tomogr.*, vol. 8, no. 2, pp. 306–316, Apr. 1984.
[2] T. Moon, "The expectation maximization algorithm," *IEEE Signal Processing Mag.*, vol. 13, pp. 47–60, Nov. 1996.
[3] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 721–741, Nov. 1984.
[4] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[5] J. Biemond, R. L. Lagendijk, and R. M. Mersereau, "Iterative methods for image deblurring," *Proc. IEEE*, vol. 78, pp. 856–883, May 1990.
[6] K. Lange, "Convergence of EM image reconstruction algorithms with Gibbs smoothing," *IEEE Trans. Med. Imaging*, vol. 9, pp. 439–446, Dec. 1990.
[7] S. Alenius and U. Ruotsalainen, "Bayesian image reconstruction for emission tomography based on median root prior," *Eur. J. Nucl. Med.*, vol. 24, no. 3, pp. 258–265, Mar. 1997.
[8] G. Kontaxakis, L. G. Strauss, and G. van Kaick, "Optimized image reconstruction for emission tomography using ordered subsets, median root prior, successive substitutions and a web-based interface," in *Proc. IEEE Medical Imaging Conf.*, 1998, pp. 1347–1352.
[9] S. Alenius, "On noise reduction in iterative image reconstruction algorithms for emission tomography: Median root prior," Ph.D. dissertation, Tampere Univ. Technol., Tampere, Finland, Oct. 1999.
[10] J. Astola and P. Kuosmanen, *Fundamentals of Nonlinear Digital Filtering*. Boca Raton, FL: CRC, 1997.
[11] V. Bettinardi, E. Pagani, M. C. Gilardi, S. Alenius, K. Thielemans, M. Teräs, and F. Fazio, "Implementation and evaluation of a 3-D one-step late reconstruction algorithm for 3-D positron emission tomography brain studies using median root prior," *Eur. J. Nucl. Med.*, vol. 29, no. 1, pp. 7–18, 2002.
[12] P. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," *IEEE Trans. Med. Imaging*, vol. 9, pp. 84–93, Mar. 1990.
[13] A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[14] L. Kaufman, "Maximum likelihood, least squares, and penalized least squares for PET," *IEEE Trans. Med. Imaging*, vol. 12, pp. 200–214, June 1993.
[15] S. Alenius, U. Ruotsalainen, and J. Astola, "Using local median as the location of the prior distribution in iterative emission tomography image reconstruction," *IEEE Trans. Nucl. Sci.*, vol. 45, pp. 3097–3104, Dec. 1998.
[16] C. Bouman and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. Image Processing*, vol. 2, pp. 296–310, July 1993.
[17] S. Alenius and U. Ruotsalainen, "Improving the visual quality of median root prior images in PET and SPECT reconstruction," in *IEEE Nuclear Science Symp. Conf. Record 2000*, vol. 2, pp. 15/216–15/223.
[18] A. C. Bovik, T. S. Huang, and D. C. Munson, "A generalization of median filtering using linear combinations of order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1342–1350, Dec. 1983.
[19] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C, The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992.
[20] R. W. Rowe and S. Dai, "A pseudo-Poisson noise model for simulation of positron emission tomographic projection data," *Med. Phys.*, vol. 19, no. 4, pp. 1113–1119, July/Aug. 1992.
[21] S. Alenius, U. Ruotsalainen, and J. Astola, "Attenuation correction for PET using count-limited transmission images reconstructed with median root prior," *IEEE Trans. Nucl. Sci.*, vol. 46, pp. 646–651, June 1999.
[22] G. Kontaxakis, L. G. Strauss, and A. Dimitrakopoulou-Strauss, "Improved attenuation correction for the ECAT EXACT HR+ tomograph using iterative transmission image reconstruction," in *Proc. 46th Annu. Meeting Society of Nuclear Medicine*, vol. 40, May 1999, p. 301.
[23] K. Lange, "Correction to 'Convergence of EM image reconstruction algorithms with Gibbs smoothing'," *IEEE Trans. Med. Imaging*, vol. 10, p. 228, June 1991.