Adaptive Bilateral Filtering for PET

László Papp, Gábor Jakab[1], Balázs Tóth, and László Szirmay-Kalos[2]

***Abstract*—Filtering the activity estimate during the iterative Positron Emission Tomography (PET) reconstruction process has several benefits. It works as a regularization eliminating high frequency components mainly due to overfitting, and it also suppresses noise due to the numerical computation of projections. A proper filtering scheme should maintain the true activity, not blur sharp edges, but eliminate noisy artifacts. A filter meeting this requirement is the bilateral filter, but it requires the knowledge of the variance of the local activity, which is not available. To attack this problem, we propose a statistical filtering method that automatically sets this variance locally based on the current activity distribution. We demonstrate that the new filter outperforms Gaussian filtering both for a simple 2D PET model and also for fully-3D human PET reconstruction. The presented model is built into the TeraTomo$^{\mathrm{TM}}$ system.**

I. Introduction

In iterative PET reconstruction forward and back projections alternate. Forward projection computes the expected number of hits in detector pairs called LORs, $\tilde{\mathbf{y}} = (\tilde{y}_1, \ldots, \tilde{y}_{N_{\mathrm{LOR}}})$, from the current tracer density estimation $x(\vec{v})$ defining the number of decays in unit volume around point $\vec{v}$, while back projection corrects this estimation. The tracer density is expressed by voxel coefficients $\mathbf{x} = (x_1, \ldots, x_{N_{\mathrm{voxel}}})$. The correspondence between voxels and LORs is established by the system matrix $\mathbf{A}_{LV}$ where an element gives the probability that LOR L detects a decay happening in voxel V.

The ML-EM scheme searches the voxel coefficients that maximize the probability of measurement results $\mathbf{y} = (y_1, \ldots, y_{N_{\mathrm{LOR}}})$, iteratively updating voxel estimates $x_V^{(n)}$ of iteration step n by scaling factors s_V obtained from the measured and expected LOR values [8]:

$$x_V^{(n+1)} = x_V^{(n)} \cdot s_V, \quad \text{where} \quad s_V = \mathcal{B}(\tilde{\mathbf{y}}) = \frac{\sum_L \mathbf{A}_{LV} \frac{y_L}{\tilde{y}_L}}{\sum_L \mathbf{A}_{LV}}. \quad (1)$$

Operator $\mathcal{B}$ is called back projection. The reconstruction process can also be interpreted as a control loop (Fig. 1), including the back projection and the following forward projection

$$\tilde{y}_L = \mathcal{F}(\mathbf{x}) = \sum_{V=1}^{N_{\mathrm{voxel}}} \mathbf{A}_{LV} x_V.$$

This loop is stabilized when $\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)}$, i.e. when scaling factors s_V are 1, which means that the iteration loop solves the following equation for $\mathbf{x}$:

$$\mathcal{B}(\mathcal{F}(\mathbf{x})) = 1. \quad (2)$$

Several authors proposed the inclusion of a voxel space filtering step in the reconstruction loop [9], [3] and it turned out that it is equivalent to the method of sieves that seeks to constrain the EM solution to a bandwidth limited subspace of all possible solutions [10], [11], [16]. To see how this pre-filtering affects the reconstruction loop, let us include filtering operator $\mathcal{G}$ before forward projection $\mathcal{F}$ (Fig. 2). Filtering operator $\mathcal{G}$ maps the iteration result x_V to filtered voxel value $\hat{x}_V$. The modified system also stabilizes when the scaling factors s_V are 1, thus we get

$$s_V = \mathcal{B}(\mathcal{F}(\hat{\mathbf{x}})) = \mathcal{B}(\mathcal{F}(\mathcal{G}(\mathbf{x}))) = 1.$$

Note that this is the same equation for $\hat{\mathbf{x}}$ as the original one (Eq. 2) for $\mathbf{x}$, thus considering $\hat{\mathbf{x}}$ to be the output of the control system, the modified system would behave similarly to the original one if the result could be represented by the filtering of $\mathbf{x}$. However, in our case both x and $\hat{x}$ are represented by discrete samples, thus these functions must be band-limited, which prohibits the generation of arbitrary output function $\hat{x}$. Our modified loop can converge to solutions $\hat{x}$ that can be the filtered version of x that can be represented by the voxel samples. It means that the search space is limited and the inclusion of a filter acts as a regularization.

Low pass filters, like the Gaussian, are good for regularization because high frequency components caused by overfitting are automatically eliminated, but they also blur sharp edges. This problem can be attacked by *Bilateral filters* [15], [7], which is a non-linear, edge-preserving and noise-reducing smoothing filter for images. Unlike in linear filters, the weights

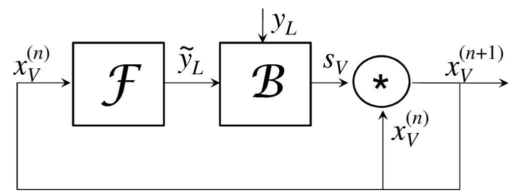

Fig. 1. The reconstruction as a control loop. Forward projection $\mathcal{F}$ takes the actual voxel values $x_V^{(n)}$ and computes the expectation of LOR events $\tilde{y}_L$. Back projection $\mathcal{B}$ calculates a correction ratio s_V for every voxel from the expected LOR events $\tilde{y}_L$ and the measured LOR hits y_L.

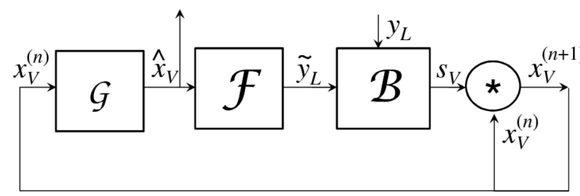

Fig. 2. The modified reconstruction loop. Forward projection $\mathcal{F}$ computes the expected LOR hits $\tilde{y}_L$ from the filtered voxel values $\hat{x}_V$ that are computed as by applying filter $\mathcal{G}$ to the result of previous iteration $x_V^{(n)}$. Back projector $\mathcal{B}$ calculates scaling factor s_V for each voxel from ratios $y_L/\tilde{y}_L$.

[1]: Mediso Ltd. (laszlo.papp@mediso.hu).
[2]: Budapest University of Technology and Economics (e-mail: szirmay@iit.bme.hu).

depend not only on Euclidean distance of voxels, but also on the range differences, which is the tracer density differences in PET. This preserves sharp edges by systematically looping through each voxel and adjusting weights to the adjacent voxels accordingly.

Bilateral filtering of input $x(\vec{v})$ is defined as

$$\hat{x}(\vec{v}') = \frac{\int G_\sigma(||\vec{v} - \vec{v}'||)G_\xi(x(\vec{v}) - x(\vec{v}'))x(\vec{v})\mathrm{d}\vec{v}}{\int G_\sigma(||\vec{v} - \vec{v}'||)G_\xi(x(\vec{v}) - x(\vec{v}'))\mathrm{d}\vec{v}}$$

with G_λ denoting the one-dimensional Gaussian function of standard deviation λ. The amount of blur is controlled by spatial variance parameter σ, while the amount of detail kept is determined by intensity variance parameter ξ. The intensity-dependent Gaussian weight ensures that neighboring voxels placed on the same side of a step-like signal as the centering voxel $\vec{v}'$ get higher weights while voxels from the other side of the edge give less contribution to the filter output, better preserving an edge. However, the appropriate value of ξ is not straightforward to find, since it is given in intensity space which is object dependent.

This paper proposes and adaptive statistical filter to solve this problem and automatically finds the intensity variance parameter for the bilateral filter.

II. Methods

The filtering step of the new method is a sequence of operations to determine the local intensity variance and finally executing the bilateral filter with locally varying parameter (the steps are illustrated by Fig. 3 for a 1D step-like function corrupted by noise):

1) Spatial activity average $a(\vec{v})$ is determined by separable Gaussian convolutions (denoted by $*$):
$$a(\vec{v}) = x(\vec{v}) * G_\sigma.$$
Resulting spatial average $a(\vec{v})$ is free from high frequency noise but the edges and the true signal are also blurred.
2) Standard deviation of the difference between the average and the signal is computed again with Gaussian convolutions:
$$d(\vec{v}) = \sqrt{(x(\vec{v}) - a(\vec{v}))^2 * G_\sigma - ((x(\vec{v}) - a(\vec{v})) * G_\sigma)^2}.$$
This standard deviation is roughly constant everywhere but around edges and significant changes of the true signal, thus it can be used to find these features.
3) Maximum
$$d_{\max} = \max d(\vec{v})$$
is obtained.
4) The spatial standard deviation and its maximum are used to obtain the local intensity smoothness factor:
$$i(\vec{v}) = \left(1 - \frac{d(\vec{v})}{d_{\max}}\right)^\alpha * G_\sigma$$
where α controls edge preservation. The smoothness factor is in the $[0, 1]$ interval and has value close to zero where the most significant changes are in the signal and the value is close to one where the true signal is smooth and is close to constant.
5) Finally, the bilateral filtering step is executed setting the intensity variance $\xi(\vec{v}) = \beta d(\vec{v}) i(\vec{v})$ to the product of spatial standard deviation $d(\vec{v})$, local intensity smoothness $i(\vec{v})$ and a user defined factor of filter strength β. Note that spatial standard deviation $d(\vec{v})$ would be the standard deviation if the process were stationary. However, it is an overestimate where the true signal changes, so we scale this with the smoothness parameter.

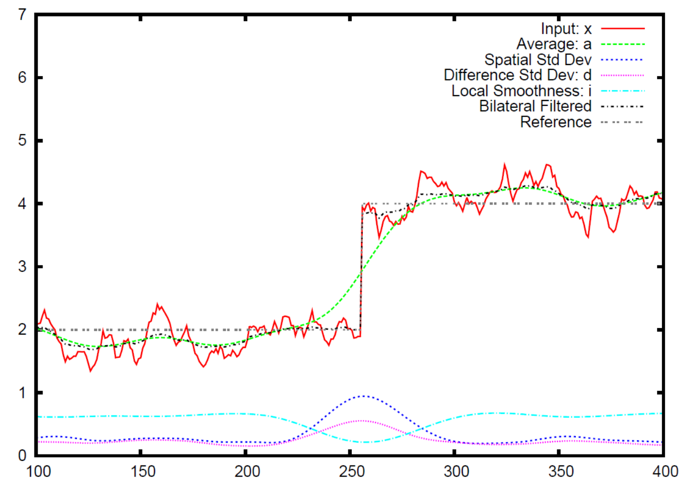

Fig. 3. Steps of the statistical filtering method applied for a 1D step function corrupted with Perlin noise [6]. We set $\sigma = 15$, $\alpha = 3$, and $\beta = 5$. The Average corresponds to standard Gaussian filtering, which reduces noise but also blurs the edge. The computed smoothness factor gets small around the step, prohibiting the blurring of the edge.

Note that steps 1, 2 and 4 apply the standard 3D Gaussian filter, which is separable and thus can be efficiently executed on the GPU [14]. The maximum in step 3 is also obtained with a parallel algorithm. The more expensive step here is the bilateral filter, which is not separable.

III. Results: 2D PET model

First we examine a simple 2D PET model [13] where $N_{\mathrm{LOR}} = 2115$ and $N_{\mathrm{voxel}} = 1024$ (Fig. 4).

We considered four phantoms, the *Three Squares* where each square has 64 Bq activity, the *Three Pyramids* where constant squares are replaced by linearly changing activity, the *Point Source* of 20 Bq activity, the *Homogeneity* of $2 \cdot 10^4$ Bq activity and using a Monte Carlo particle transport method, we simulated a 5 sec long measurement for all these phantoms (Fig. 5).

It means that the Three Squares and Three Pyramids phantoms are projected with about 1000 photon pairs, the Point source with 100 photon pairs, and the Homogeneity with 10^5 photon pairs, resulting in measured data having 1.21 Signal-to-Noise ratio (SNR) for the Three Squares, 1.07 SNR for the Point source and 1.69 SNR for the Homogeneity. Only geometric effects were simulated, we ignored scattering and absorbtion in the phantom. The Three Squares phantom is formed by three active squares of increasing size, and

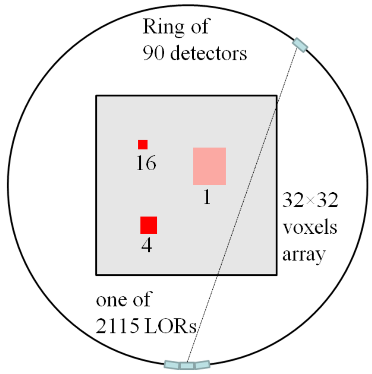

Fig. 4. 2D tomograph model: The detector ring contains 90 detector crystals and each of them is of size 2.2 in voxel units and participates in 47 LORs connecting this crystal to crystals being in the opposite half circle, thus the total number of LORs is $90 \times 47/2 = 2115$. The voxel array to be reconstructed is in the middle of the ring and has 32×32 resolution, i.e. 1024 voxels. The ground truth voxel array of the Three Squares phantom has three hot squares of activity densities 1, 4, and 16 and of sizes 8^2, 4^2, and 2^2.

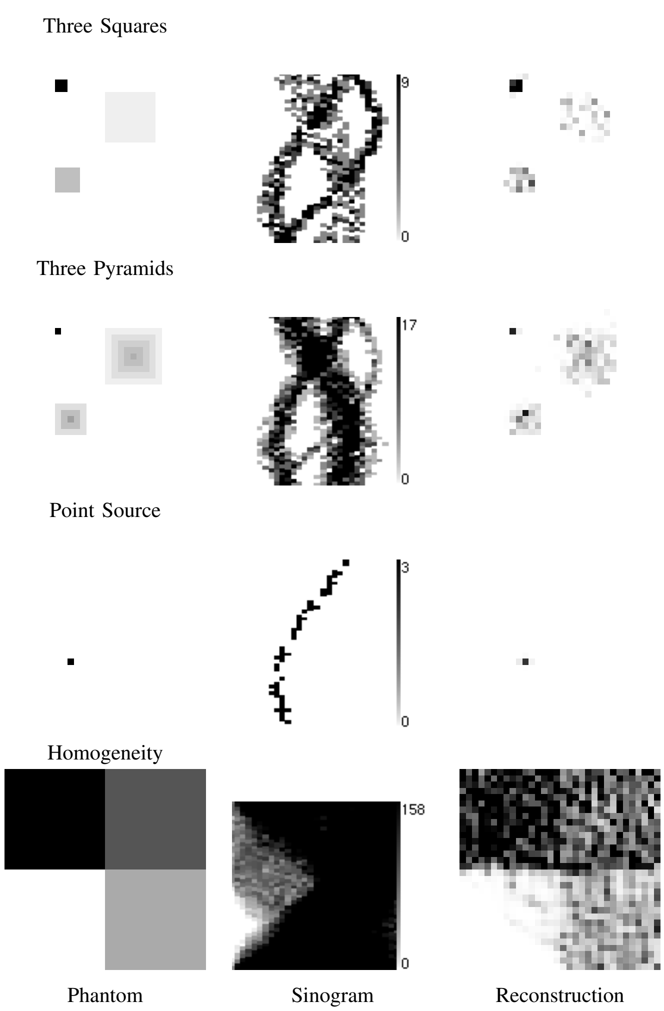

Fig. 5. The four phantoms used in the experiments, their random projections in sinogram parametrization, and the reconstructions without regularization.

represents a realistic example. The Point Source and the Homogeneity represent two extreme cases. Point Source has a high variation since it has just a single voxel where the activity is non-zero and is well determined by the measurement. Thus, the reconstruction of Point Source would not need regularization, and regularization would just slow down the convergence. Homogeneity is formed by four constant activity squares, so the activity distribution is rather flat and the measurement is quite noisy. Such cases badly need regularization.

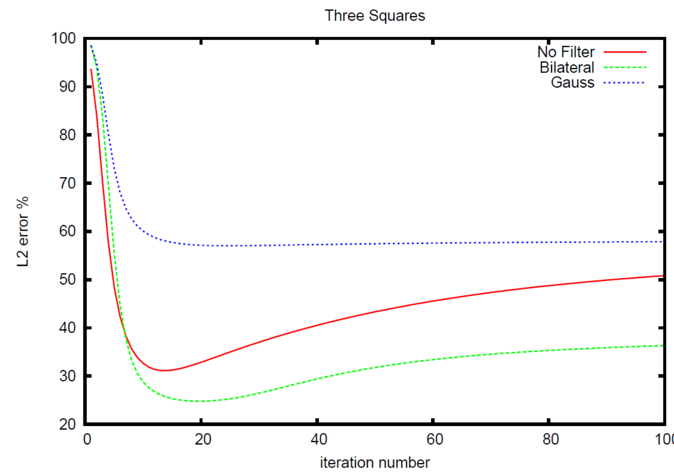

Fig. 6. L_2 error curves of the Three Squares reconstruction, $\sigma = 1$, $\alpha = 2$, $\beta = 5$.

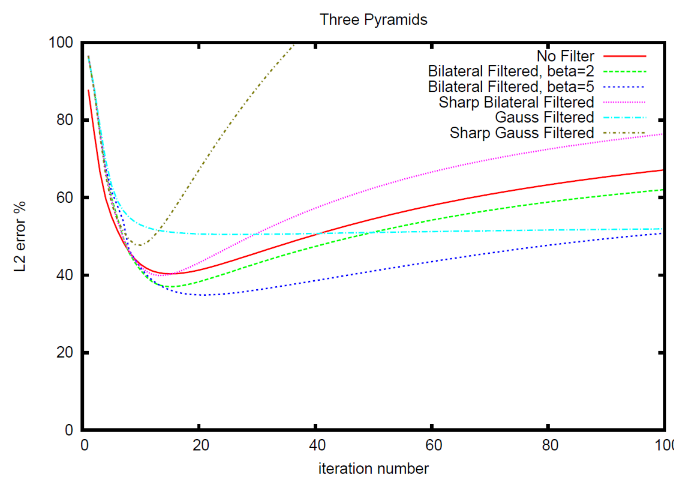

Fig. 7. L_2 error curves of the Three Pyramids reconstruction, $\sigma = 1$, $\alpha = 1$.

The *Three Squares* phantom is reconstructed with different methods setting $\sigma = 1$, $\alpha = 2$, $\beta = 5$ (Figs. 6 and 10). Note that the unfiltered reconstruction is very noisy while Gaussian filtering causes strong blurring. Bilateral filter is a good compromise between these cases. The error curves also indicate that the unfiltered solution reduces the error quickly at the beginning but later the error increases due to overfitting. Gauss filter slows down and even stops convergence but can avoid diverging reconstruction. The speed of initial convergence of the Bilateral filter is similar to the unfiltered case, but the error

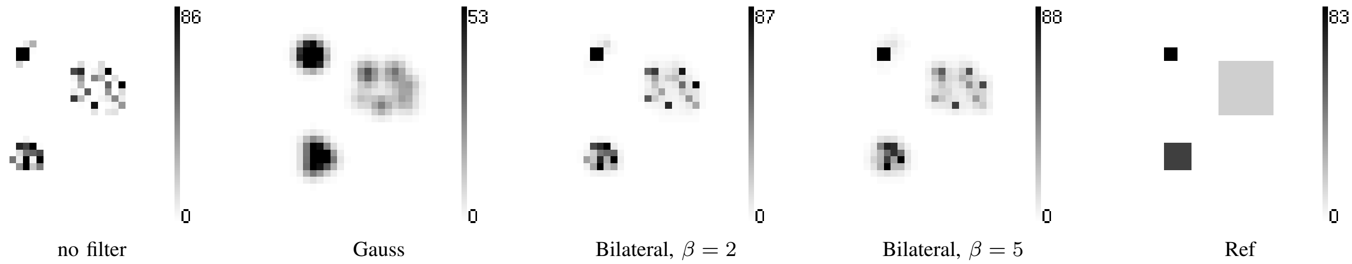

Fig. 10. Reconstructions of the Three Squares phantom.

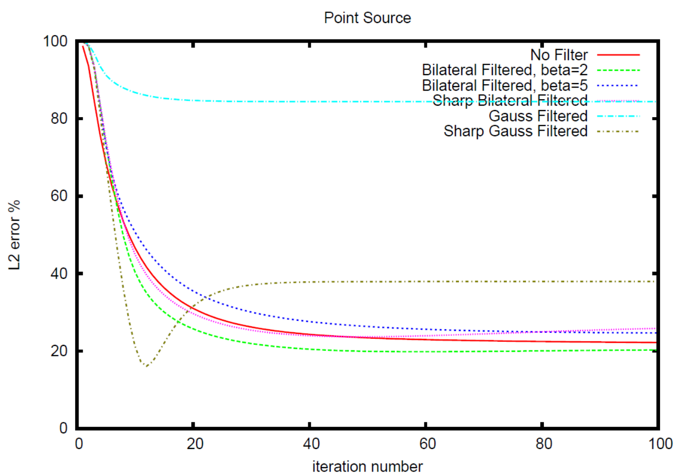

Fig. 8. L_2 error curves of the Point Source reconstruction, $\sigma = 1$, $\alpha = 1$.

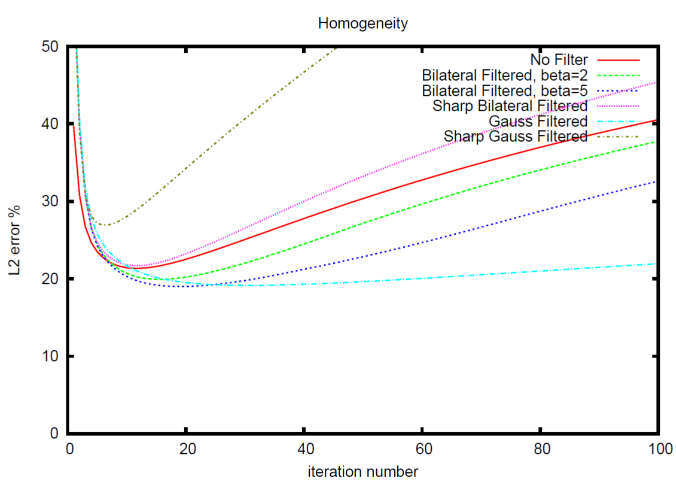

Fig. 9. L_2 error curves of the Homogeneity reconstruction, $\sigma = 1$, $\alpha = 1$.

gets reduced more and the diverging part due to overfitting is also less significant.

The *Three Pyramids*, *Point*, and the *Homogeneity* phantoms are reconstructed with different methods setting $\sigma = 1$, $\alpha = 1$, and with two different β parameters (Figs. 7–9 and 11–13). Note that Gaussian removes noise but also blurs the true signal, while the Bilateral filter preserves edges while reducing noisy artifacts. The error curves also show the error of the reconstructed signal before filtering. Note that this is not a valid output but can be considered as a sharpened version of the reconstruction. This sharpened version converges faster at the beginning but later behaves poorer than the filtered signal. The Point is a high statistics measurement where no regularization is needed, so Gaussian results in large blurring and in stopping the convergence after a few steps. Bilateral filter, however, provides similar results as the unfiltered case. Both the Three Pyramids and the Homogeneity need regularization, so even Gaussian helps on the longer run, but its initial convergence is slower than that of the unfiltered solution or the Bilateral filtering. As these are worse in eliminating overfitting, if the iteration is not stopped in time, Gaussian can become better later, which can be observed when the Homogeneity phantom is reconstructed.

We also included the error curves of the "sharp" versions of the Gaussian and bilateral filters, which represent the data before the filter. Note that it is not the theoretically valid output, but a signal from which the output can be generated by a low pass filter, i.e. the sharpened version of the real output. Thus the error of the sharpened version is larger than the valid output after many iteration steps or may even diverge, but the initial behavior of the sharpened version is quite good, especially for the Point Phantom. It means that combining the sharpened version with the real output for the first few iteration steps can speed up the reconstruction.

IV. Fully-3D reconstruction

The proposed method has been implemented in CUDA and integrated into the TeraTomo$^{\mathrm{TM}}$ fully-3D system [1], [12]. Due to the high arithmetic performance and bandwidth of the GPU, the execution time of the filtering step is negligible comparing to the projection operators even for higher resolution volumes. Thus, our proposed method has practically no overhead.

The method is demonstrated with the reconstruction of Small animal IQ phantom projected by GATE [2] simulating the NanoPET/CT pre-clinical PET-CT system [5] (Fig. 14). Note that the noise has been eliminated by the Bilateral filter without the blurring effect of the Gaussian filter. The resolution of the reconstructed volume is $200 \times 200 \times 200$ voxels with edge size of 0.4 [mm].

The Derenzo phantom has also been projected with GATE assuming the NanoPET/CT pre-clinical PET-CT system and reconstructed without filtering, with Gaussian and finally with the adaptive bilateral filter (Fig. 15). The resolution of the reconstructed volume is $128 \times 144 \times 144$ voxels of 0.234 [mm].

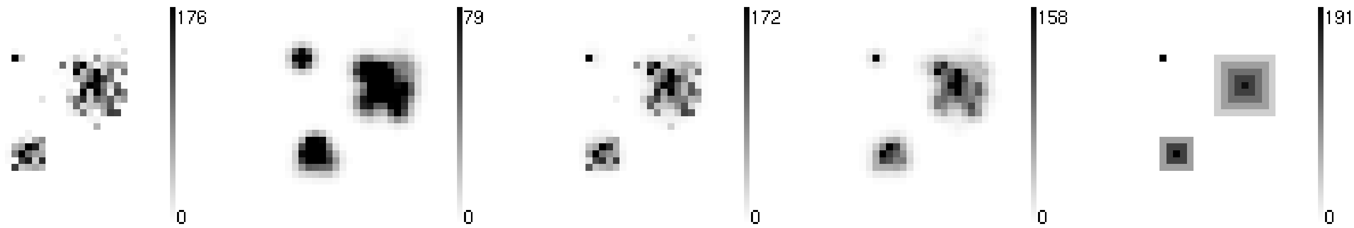

Fig. 11. Reconstructions of the Three Pyramids phantom.

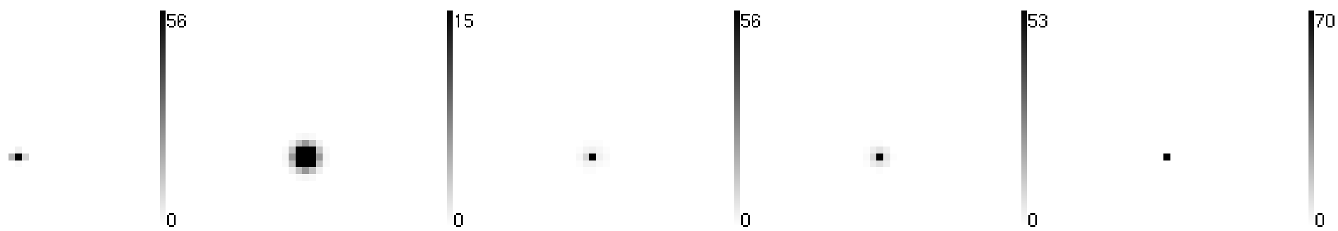

Fig. 12. Reconstructions of the Point Source phantom.

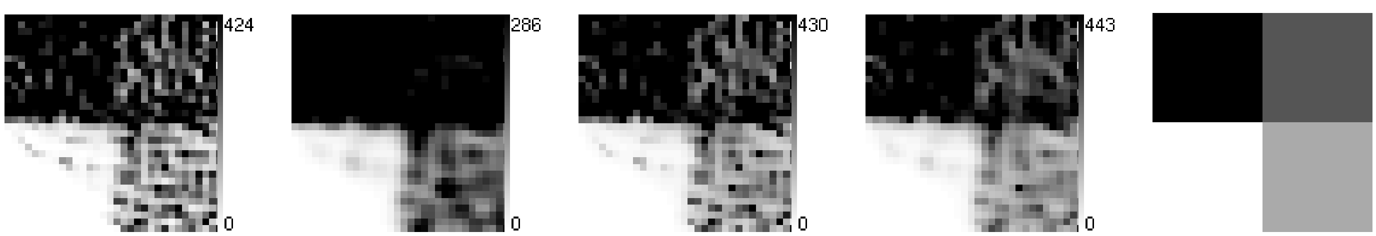

Fig. 13. Reconstructions of the Homogeneity phantom.

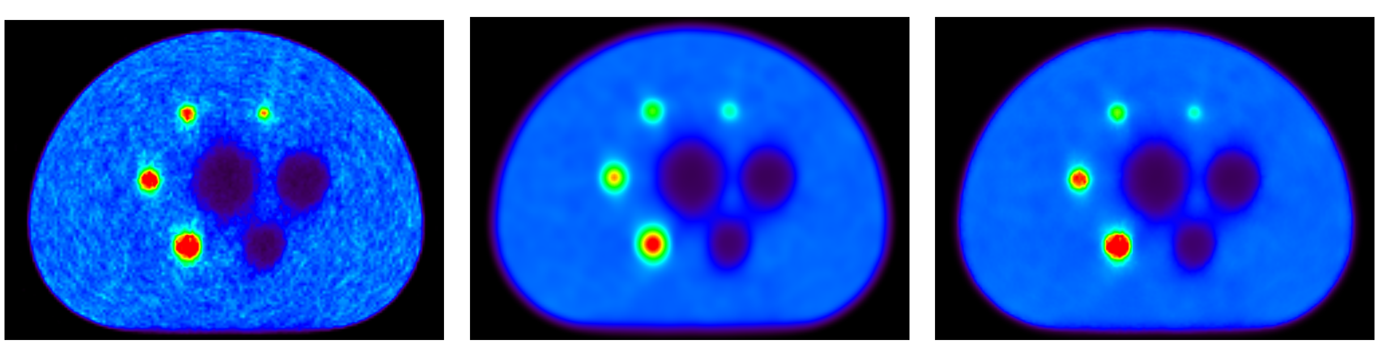

Fig. 17. Reconstructions of the Human IQ phantom.

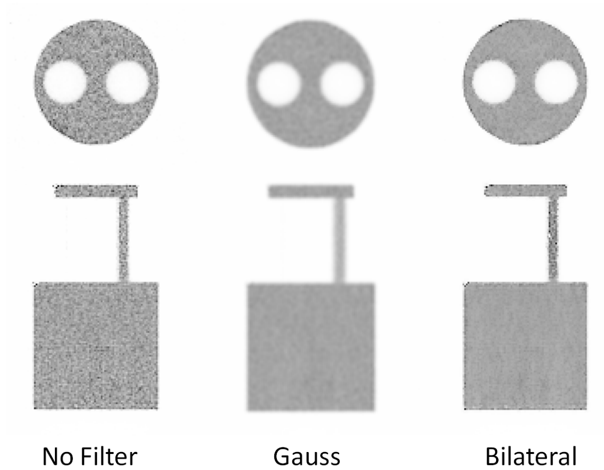

Fig. 14. Reconstruction of the Small Animal phantom

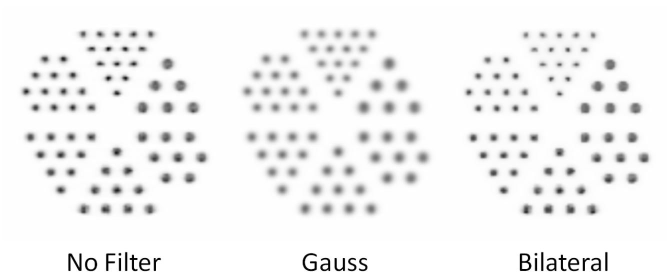

Fig. 15. Reconstruction of the Derenzo phantom

The method is also tested with the reconstruction of the NEMA NU2-2007 human IQ phantom projected with GATE [2] simulating the Mediso AnyScan PET/CT system [4]. The resolution of the reconstructed volume is $166 \times 166 \times 75$ voxels with edge size of 2 [mm]. The error curves and the reconstructions are in Figs. 16 and 17. This data requires regularization because the error of the unfiltered reconstruction diverges after 10 iteration steps. Both the Gaussian and Bilateral filtering can eliminate divergence, but the error curve of the Bilateral

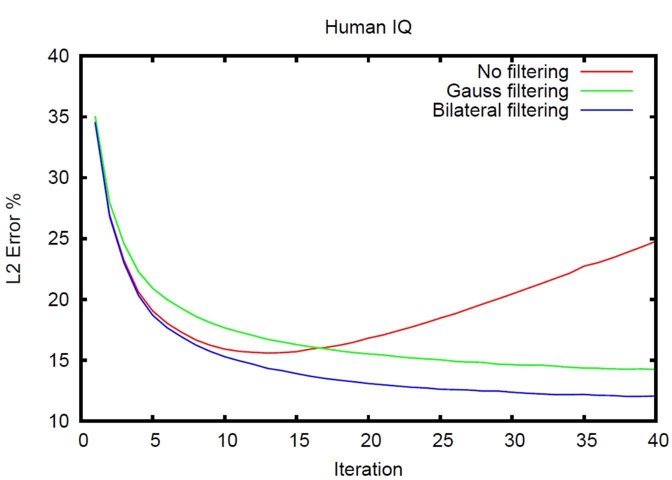

Fig. 16. L2 error curve of the reconstructions of the Human IQ phantom.

filter is below that of the Gaussian. The reconstruction results also explain this because the unfiltered reconstruction is noisy, the Gaussian is blurred, while the proposed adaptive bilateral scheme combines the low noise of Gaussian filtering and the high contrast of unfiltered reconstruction.

V. Conclusion

We studied the inclusion of non-linear bilateral filter into the PET reconstruction loop and concluded that this filter acts as a regularization that preserves edges but smoothes noise. The main problem of bilateral filters is the determination of the intensity range filter parameters, which are object dependent. This paper proposed an adaptive, statistical filtering method. Unlike in the classical bilateral filter, our method automatically estimates the intensity variance. All steps are implemented on the GPU where the added computational cost of filtering is negligible with respect to forward and back projection calculations.

Acknowledgement

This work has been supported by OTKA K–104476. The authors are grateful to Gergely Patay for the GATE simulations.

References

[1] M. Magdics et al. TeraTomo project: a fully 3D GPU based reconstruction code for exploiting the imaging capability of the NanoPET/CT system. In *World Molecular Imaging Congress*, 2010.

[2] S. Jan and et al. GATE: A simulation toolkit for PET and SPECT. *Physics in Medicine and Biology*, 49(19):4543–4561, 2004.

[3] M. Magdics, L. Szirmay-Kalos, B. Tóth, and T. Umenhoffer. Filtered sampling for PET. In *IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, 2012.

[4] http://www.mediso.com/products.php?fid=1,9&pid=73.

[5] http://www.mediso.com/products.php?fid=2,11&pid=86.

[6] K. Perlin. An image synthetisizer. In *Computer Graphics*, pages 287–296, 1985.

[7] I. Rodrigues, J. Sanches, and J. Bioucas-Dias. Denoising of medical images corrupted by poisson noise. In *15th IEEE International Conference on Image Processing, 2008. ICIP 2008.* , pages 1756–1759, 2008.

[8] L. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imaging*, 1:113–122, 1982.

[9] E.T.P. Slijpen and R. J. Beekman. Comparison of post-filtering and filtering between iterations for SPECT reconstruction. *IEEE Trans. Nuc. Sci.*, 46(6):2233–2238, 1999.

[10] D. L. Snyder and M.I. Miller. The use of sieves to stabilize images produced with the em algorithm for emission tomography. *IEEE Trans. on Nuc. Sci.*, 32(5):3864–3872, 1985.

[11] D. L. Snyder, M.I. Miller, L. J. Thomas, and D.G. Politte. Noise and edge artifacts in maximum-likelihood reconstructions for emission tomography. *IEEE Trans. on Med. Imaging*, 6(3):228–238, 1987.

[12] L. Szirmay-Kalos, M. Magdics, and B. Tóth. Multiple importance sampling for PET. *IEEE Trans. Med. Imaging*, 33(4):970–978, 2014.

[13] L. Szirmay-Kalos, M. Magdics, B. Tóth, and T. Bükki. Averaging and metropolis iterations for positron emission tomography. *IEEE Trans. Med. Imaging*, 32(3):589–600, 2013.

[14] L. Szirmay-Kalos, L. Szécsi, and M. Sbert. *GPU-Based Techniques for Global Illumination Effects*. Morgan and Claypool Publishers, San Rafael, USA, 2008.

[15] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proceedings of the Sixth International Conference on Computer Vision*, ICCV '98, pages 839–, 1998.

[16] E. Veklerov and J. Llacer. The feasibility of images reconstructed with the methods of sieves. *IEEE Trans. Nuc. Sci.*, 37(2):835–841, 1990.