Total Variation Regulated EM Algorithm

V. Y. Panin, *Student Member IEEE*, G. L. Zeng, *Member IEEE* and G. T. Gullberg, *Senior Member IEEE*
Department of Radiology, University of Utah
729 Arapeen Drive, Salt Lake City, Utah 84108-1218, USA

Abstract

An iterative Bayesian reconstruction algorithm based on the total variation (TV) norm constraint is proposed. The motivation for using TV regularization is that it is extremely effective for recovering edges of images. This paper extends the TV norm minimization constraint to the field of SPECT image reconstruction with a Poisson noise model. The regularization norm is included in the OSL-EM (one step late - expectation maximization) algorithm. Unlike many other edge-preserving regularization techniques, the TV based method depends only on one parameter. Reconstructions of computer simulations and patient data show that the proposed algorithm has the capacity to smooth noise and maintain sharp edges without introducing over/under shoots and ripples around the edges.

I. INTRODUCTION

An inverse problem to recover an unknown function $f(\vec{x})$ for given noisy data $P(\vec{s})$ can be solved by utilizing a Bayesian paradigm. The goal is to maximize the posteriori probability distribution according to Bayes equation:

$$post(f|P) \propto prob(P|f)prior(f). \qquad (1)$$

Usually, the likelihood distribution $prob(P|f)$ is assumed to be in a Gaussian form with a linear relationship $Af=P$, where A is a linear operator, a normalization factor Z_σ is independent from f: $prob(P|f) = \exp(-\|P-Af\|^2/2\sigma^2)/Z_\sigma$, and the prior has the Gibbs form $prior(f) = \frac{1}{Z}e^{-U(f)}$, where $U(f)$ is the prior energy functional. Maximization of this posteriori probability distribution (the maximum *a posteriori* estimate) is equivalent to the Tiknonov regularization problem [1]

$$min_f\left\{\frac{1}{2}\|Af-P\|^2 + \beta U(f)\right\}, \qquad (2)$$

where $U(f)$ serves as a regularization functional and β is a regularization parameter that controls the weight given to the minimization of the regularization term relative to the minimization of the residual norm.

The regularization functional incorporates *a priori* assumptions about the size and smoothness of the desired solution. The use of Total Variation in a constrained minimization problem was first introduced by Rudin *et al* [2].

The research work presented in this manuscript was partially supported by NIH Grant R29 HL51462 and Picker International.

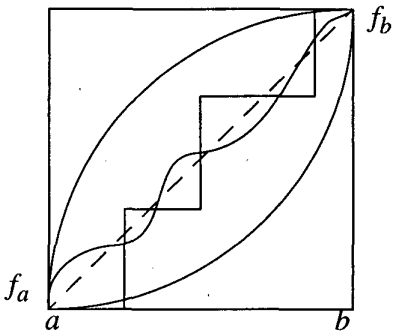


Figure 1. Functions with the same total variation in one dimension.

In that study, the TV norm was used as an *a priori* regularization functional to process noisy, blurred images. In the case of 2*D* functions the TV functional is defined as:

$$TV(f) = \int|\nabla f|d\vec{x} = \iint\sqrt{f_x^2 + f_y^2}\,dxdy. \qquad (3)$$

Note that f is not required to be continuous. It has been shown that the TV norm measures the total variation or "jumps" of f, even if f is discontinuous [3]. We reproduce the results of a one dimensional variational problem based on [3] here.

Let $S = \{f|f(a) = f_a, f(b) = f_b\}$

For the total variation norm, we have that $min_{f\in S}TV(f) = |f_a - f_b|$ is achieved by any monotonic, not necessarily continuous functions, $f \in S$. See Fig. 1.

However, if we use the H^1 semi-norm, we have that $min_{f\in S}\int_a^b(f'(x))^2dx$ has a unique solution $f(x) = (f_a(b-x) + f_b(x-a))/(b-a)$.

The TV functional allows many functions, including discontinuous ones, to approximate a given noisy function. When one solves problem (2) with the TV norm as a regularization functional, no particular monotonic function is preferred over another. Therefore edges can be preserved because the TV norm does not force the result to be continuous. However, contrast may be reduced because a bias can be introduced. For comparison, the regularization functional that is defined similarly to the TV norm but with the square of function gradient under the integral, prefers the linear function over others and leads to smoothed edges. Even though the TV norm preserves edges it is not biased in favor of edges. It will not introduce artificial jumps in the function

0018-9499/99$10.00 © 1999 IEEE

although a staircasing effect may occur in the presence of noise. A distinct *advantage* of utilizing the total variation constraint is that it can smooth the noise and keep sharp edges without causing over/under shooting and ripples. Also, according to (3), the TV functional is defined globally and independent of any additional parameters.

The minimization problem (2) with $U(f)$ given by (3) suffers from computational difficulties and its corresponding optimum Euler-Lagrange equations are highly non-linear [2][4]. Nonlinear partial differential equations are most often used to solve minimization problem (2). Many researchers in different fields of engineering are currently working on implementation of the TV regularization [3] [5]-[7].

The goal of this work is to extend TV regularization to the field of SPECT image reconstruction. TV regularization, when applied to SPECT, can provide reconstruction images that have attractive features, including the identification of distinguishable sharp edges. Instead of using approaches mentioned previously, we used a Poisson noise model and used the TV norm as the energy functional U in the Gibbs prior distribution. The additional challenge, in comparison to (2), is that the Poisson based likelihood functional is also nonlinear. We have shown that the fast Green's One Step Late (OSL) algorithm was effectively able to incorporate the TV regularization. Computer simulations were performed to verify properties of the proposed algorithm incorporated with the TV regularization. The TV regularization method was also compared with other commonly used ML-EM regularization methods. Experimentally acquired data were used to verify the suggested reconstruction method.

II. METHODS

Hereafter we deal with a 2D function $f(\vec{x})$ as a set of discrete values f_i or $f_{k,l}$. Index i represents the pair of indices k and l that correspond to the x and y coordinates of a pixel position, respectively. Due to this, we also use the term energy function instead of functional.

A. Likelihood Distribution of ML-EM algorithm and I-Divergence

The Maximum Likelihood Expectation Maximization ML-EM) algorithm explores Poisson based likelihood distribution. The goal behind the algorithm is to maximize the following functional:

$$L = \sum_j [p_j \ln \lambda_j - \ln(p_j!) - \lambda_j], \text{ with } \lambda_j = \sum_i a_{ji} f_i \qquad (4)$$

where f is the image to be recovered, p is the projection data, and a_{ji} is the contribution from pixel i in f to projection bin j.

It was shown [8] that if function f is non-negative, then minimization of Csiszar's I-divergence is the only "consistent" choice of solution for $\boldsymbol{Af=P}$. Csiszar's I-divergence between two sets of data V and U is defined as:

$$I(V||U) = \sum_j \left(v_j \ln \frac{v_j}{u_j} - v_j + u_j \right). \qquad (5)$$

Using Stirling's formula [9]

$$\ln(p!) = \left(p + \frac{1}{2}\right)\ln(p) - p + \ln(\sqrt{2\pi}) + \frac{\theta}{12p}, \qquad (6)$$

where $p > 0$, $0 < \theta < 1$, (4) takes the form of

$$\begin{aligned} L &= -\sum_j \left(p_j \ln \frac{p_j}{\lambda_j} - p_j + \lambda_j \right) \\ &\quad + \sum_j \left\{ \frac{1}{2}\ln(p_j) - \ln(\sqrt{2\pi}) + \frac{\theta_{p_j}}{12p_j} \right\} \\ &= -I(P||Af) + \sum_j C(p_j). \end{aligned} \qquad (7)$$

Comparing (5) and (7), it is clear that the ML-EM algorithm minimizes Csiszar's I-divergence with respect to f. This fact was mentioned before [10]. Even though the ML-EM algorithm is based on a Poisson model it can be applied to solve any linear inverse problem when f is non-negative. For example, the ML-EM algorithm was widely applied to the reconstruction of transmission images, even when the data noise was not Poisson in nature.

B. Maximum A Posteriori (MAP) Estimate

The ML-EM algorithm is designed to be used in the reconstruction of images with SPECT data that satisfy a Poisson noise model. However regularization is still needed because of low-count projections. The MAP reconstruction in SPECT has been shown to have significant advantages over the ML-EM algorithm [11]. The MAP estimate is able to remove the divergence in quantitative accuracy at higher iteration numbers that is often seen in the ML-EM algorithm due to noise. We used an iterative OSL MAP-EM algorithm for the Gibbs distribution prior, which is a modification of the ML-EM algorithm and has been derived by Green [12]:

$$f_i^{new} = \frac{f_i^{old}}{\sum_j a_{ij} + \beta \frac{\partial}{\partial f_i} U(f^{old})} \sum_j \frac{a_{ij} p_j}{\sum_k a_{kj} f_k^{old}}, \qquad (8)$$

where β is a regularization parameter and should be relatively small to ensure the non-negativity constraint property of the algorithm. The restriction on β is a disadvantage of this algorithm, however, we found that the allowed range of values of β was sufficient for our aims. When $\beta = 0$ (8) is reduced to the usual ML-EM algorithm. The OSL MAP-EM algorithm shares the desirable performance properties of the ML-EM algorithm.

Many Bayesian energy functions U have been suggested for MAP reconstructions based on SPECT [12]-[15]. For example, the energy function U can be calculated locally in a clique. In order to preserve edges, different types of cliques are

used to calculate U, which leads to additional complications. We restricted our attention to analytically defined energy functions, which depend on few or no additional parameters, as in the case of the TV norm.

A simple, well-known regularization method using a quadratic energy function regularization, assumes that images are globally smooth and enforces a roughness penalty on the solution. As a result it yields oversmooth solutions. We now consider two energy functions in this class to compare with the TV regularization. The first energy function is defined similarly to the TV norm (3), except that it uses the square of the function gradient under the integral. We denote it as the Square Gradient (SG) energy function. Taking into account the two-point difference approximation of partial derivatives with respect to x and y, the SG energy function can be written as

$$U_{SG} = \frac{1}{2}\sum_{k,l}[(f_{k+1,l}-f_{k,l})^2 + (f_{k,l+1}-f_{k,l})^2]. \quad (9)$$

The second energy function corresponds to the Gaussian prior distribution with a weighted average, where the central pixel is excluded from the average. We denote it as the Gaussian Average (GA) energy function:

$$U_{GA} = \frac{1}{2}\sum_{i}\left(f_i - \frac{1}{8}\sum_{j\in N_i} f_j\right)^2 \quad (10)$$

where N_i denotes the set of indices associated with those at most eight pixels which share a vertex with the ith pixel. Because the central pixel is excluded from the average, the operation of the GA regularization is similar to bandstop filtering. This particular prior function biases image toward image after applying 3x3 filter kernel, where the central element is 0 and the others are equal to each other. According to the Fourier spectrum of this filter kernel, it is a band-stop filter. Note that if the central pixel was included in the average, the operation of the regularization would be similar to low-pass filtering.

The advantage of quadratic energy functions is that their derivatives are linear in f. This makes it simple to solve the maximization equations and achieve high performance from modified ML-EM algorithms such as [16], in which use of any regularization parameter β is possible. Herman *et al.* [17] gave an example of the application of the GA potential to this modified ML-EM algorithm. In addition to this, the quadratic energy functions are global and independent of additional parameters. The disadvantage of these methods lies in their linear nature. They may blur the edges, due to global smoothing, as in the case of the SG energy function. Also, if they are similar to linear filtering, as in the case of the GA energy function, ripples can be presented in the recovered image. However, quadratic energy functions considered above are only special cases in the family of all possible quadratic energy functions.

Another regularization model assumes that images are made of smooth regions separated by sharp edges. An edge preserving regularization requires a non-quadratic energy function. The main goal behind using these functions is to suppress the noise while maintaining the edges. One function design approach is to modify a quadratic energy function so that it has a non-quadratic tail if values of adjacent pixels have large differences (gradient values). Such modified energy functions depend on an additional parameter δ. This parameter δ determines the threshold. Gradient values greater than the threshold are supposed to be representative of the edges and therefore retained, otherwise they are set to zero because the influence of the noise must be lowered. Therefore, this type of energy function presents local smoothing, and selection of δ is critical for the appearance of the final reconstruction. In addition to this, the best choice of δ is sensitive to the choice of regularization parameter β. Careful parameter tuning is necessary, however it is generally a difficult task. In the comparison study in Section III, we will use the GM (Geman and McClue [13]) energy function which is considered to be the most useful [11], [18]:

$$U_{GM} = \frac{1}{2}\sum_{k,l}\frac{((f_{k+1,l}-f_{k,l})/\delta)^2}{1+((f_{k+1,l}-f_{k,l})/\delta)^2} + \frac{((f_{k,l+1}-f_{k,l})/\delta)^2}{1+((f_{k,l+1}-f_{k,l})/\delta)^2}. \quad (11)$$

Here, the limiting behavior of the tail of the energy function approaches a constant value.

Note that high performance algorithms designed for this type of edge-preserving energy function exist for minimization problem (2) based on the Gaussian likelihood functional [18]. However, an additional non-negativity constraint is required in SPECT reconstruction [19].

C. Proposed TV-EM Algorithm

The TV norm is calculated globally and is independent of any additional parameters. It therefore possesses the advantage of global smoothing quadratic energy functions. On the other hand, the TV norm is also an edge preserving regularization energy function. The TV norm differs from edge enhancement filters in that it is not designed to favor edges. The TV norm preserves the edges by "ignoring" them in such a way that when the image is "filtered" the edges are not smoothed. The "filtering" performed with TV minimization is done to smooth the image without introducing ripples, known as Gibbs phenomenon. Unfortunately, the TV regularization may cause some degree of bias and reduction of contrast.

The TV norm is a nonlinear functional and Green's OSL algorithm is used to implement the TV regularization because of its strong performance. The disadvantage of using the OSL MAP-EM algorithm is the restricted range of regularization parameter β. However, we found that this range was sufficient for our purposes.

In implementation of algorithm (8) with TV energy potential, the partial derivatives of the current estimate of the TV norm must be evaluated. Using a two point difference to

evaluate the derivatives with respect to x or y, functional (3) was evaluated as follows:

$$U_{TV} \approx \sum_{k,l} \sqrt{(f_{k+1,l} - f_{k,l})^2 + (f_{k,l+1} - f_{k,l})^2 + \varepsilon^2}$$

$$= \sum_{k,l} u(k,l) \quad . \tag{12}$$

Adding an artificial parameter ε is necessary to ensure the differentiability of U with respect to $f_{k,l}$ at a point where $u(k,l)=0$. This parameter is only necessary to ensure stable evaluations of derivatives and its value should be less than the typical value that is defined by the first two terms under the square root in (12). The parameter ε was equal to approximately 1% (or less) of the expected maximum value of f. We will show in Section III that reconstructions utilizing different ε in a large range are nearly indistinguishable, until ε is sufficiently small. This is consistent with the analytical analysis [4] and was previously observed experimentally [6]. Too large an ε, however, will smooth out the edges in the image because (12) approximates a quadratic energy function in this case. The partial derivatives in (8) were evaluated as follows:

$$\frac{\partial U}{\partial f_{k,l}} = \frac{f_{k,l} - f_{k-1,l}}{\sqrt{(f_{k,l} - f_{k-1,l})^2 + (f_{k-1,l+1} - f_{k-1,l})^2 + \varepsilon^2}}$$

$$+ \frac{f_{k,l} - f_{k,l-1}}{\sqrt{(f_{k+1,l-1} - f_{k,l-1})^2 + (f_{k,l} - f_{k,l-1})^2 + \varepsilon^2}}$$

$$- \frac{f_{k+1,l} + f_{k,l+1} - 2f_{k,l}}{\sqrt{(f_{k+1,l} - f_{k,l})^2 + (f_{k,l+1} - f_{k,l})^2 + \varepsilon^2}}$$

$$= \frac{f_{k,l} - f_{k-1,l}}{u(k-1,l)} + \frac{f_{k,l} - f_{k,l-1}}{u(k,l-1)}$$

$$- \frac{f_{k+1,l} + f_{k,l+1} - 2f_{k,l}}{u(k,l)} \quad . \tag{13}$$

We chose β by trial to achieve the best reconstruction using the TV-EM algorithm.

D. Figures of Merit

In the comparison studies in Section III we will use Bias and Variance over a Region of Interest (ROI) as figures of merit. The image values of the phantom under study is assumed to be constant over each ROI and equal to f_{true}. Let $\bar{f}_{ROI}$ be the mean image value over an ROI, then Bias and Variance over an ROI are defined as

$$Bias = \frac{\bar{f}_{ROI} - f_{true}}{f_{true}}, \tag{14}$$

$$Variance = \frac{1}{N-1} \sum_{i \in ROI} (f_i - \bar{f}_{ROI})^2, \tag{15}$$

respectively, where N denotes the number of pixels in the ROI. Note that these figures of merit are not used in the sense which is standard in estimation theory.

III. RESULTS

A. Computer Simulations

A Shepp-Logan phantom was used in the computer simulations. Projection data (i.e. line integrals) were generated in a 128x128 image array for 120 views over [0,π). Poisson noise was added to the projections. The total projection count was approximately 1,700,000.

Figure 2 verifies the practical convergence of the TV-EM algorithm by monitoring the change in the log-posterior $L - \beta U_{TV}$ distribution. The change in the log-posterior for the TV-EM, and the Log-likelihood L for the ML-EM are similar and monotonic.

Figure 3 demonstrates the stability of the TV-EM algorithm in comparison to the ML-EM algorithm. It shows the dependence of the standard deviation (square root of *Variance* of the whole image) of the reconstructed image from the true image upon the number of iterations. Eventually, each

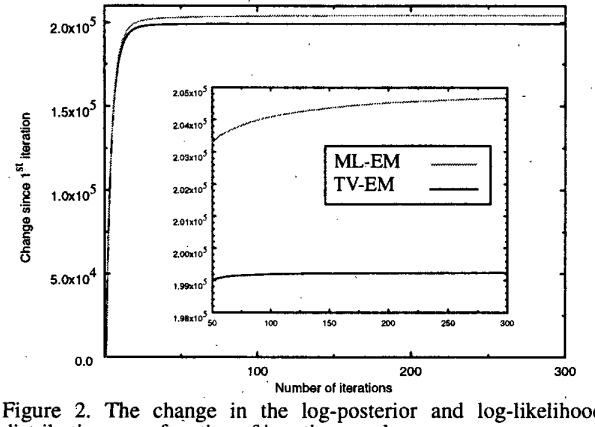


Figure 2. The change in the log-posterior and log-likelihood distributions as a function of iteration number.

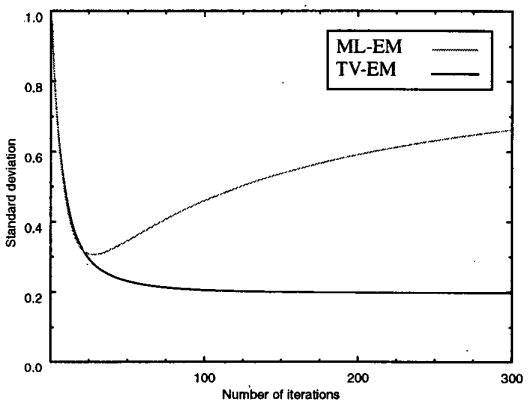


Figure 3. Dependence of difference between true and reconstructed image on the number of iterations.

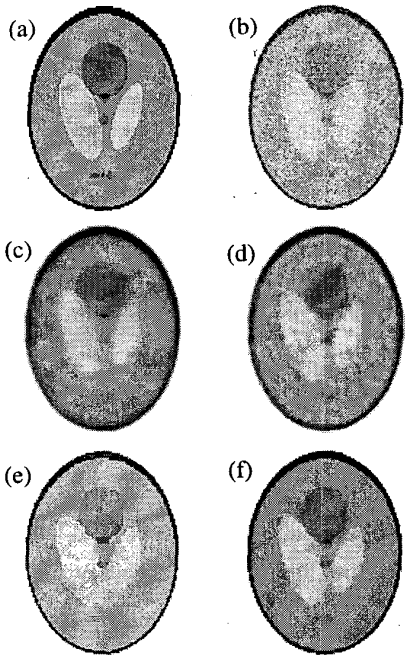


Figure 4. Computer simulations. Negative images: (a) Simulated Shepp-Logan Phantom; (b) ML-EM reconstruction, 50 iterations; (c) SG-EM reconstruction, β=8; (d) GA-EM reconstruction, β = 15; (e) GM-EM reconstruction, β = 1, δ = 0.3; and (f) TV-EM reconstruction, β =1.

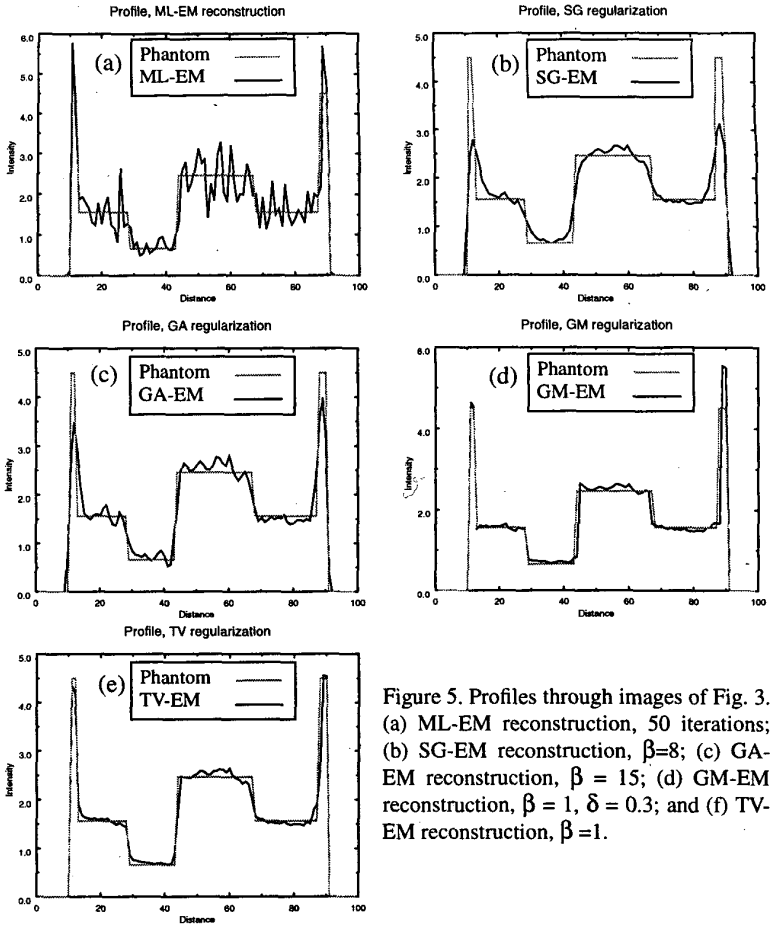


Figure 5. Profiles through images of Fig. 3. (a) ML-EM reconstruction, 50 iterations; (b) SG-EM reconstruction, β=8; (c) GA-EM reconstruction, β = 15; (d) GM-EM reconstruction, β = 1, δ = 0.3; and (f) TV-EM reconstruction, β =1.

	ML-EM	SG-EM	GA-EM	GM-EM	TV-EM
Bias, ($x10^{-2}$)	-1.40	-2.85	-1.49	-1.17	-1.72
Variance, ($x10^{-2}$)	30.9	3.76	4.94	5.91	2.72

Figure 6. Bias and Variance in an ROI of the higher activity. The ROI contains 655 pixels.

	ML-EM	SG-EM	GA-EM	GM-EM	TV-EM
Bias, ($x10^{-2}$)	11.9	16.55	7.71	7.01	9.71
Variance, ($x10^{-2}$)	5.44	3.11	3.65	2.49	1.99

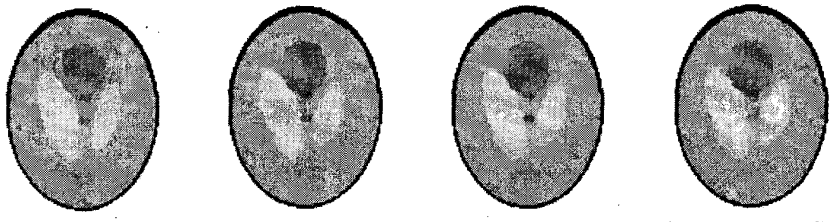
Figure 7. Bias and Variance in an ROI of the lower activity. The ROI contains 1275 pixels.

	ML-EM	SG-EM	GA-EM	GM-EM	TV-EM
Bias, ($x10^{-2}$)	0.65	8.28	5.65	-0.15	2.45
Variance, ($x10^{-2}$)	16.5	9.69	13.1	3.44	3.99

Figure 8. Bias and Variance in an ROI of the medium activity. The ROI contains 5419 pixels.

	$\beta = 1, \delta = 0.1$	$\beta = 1, \delta = 0.3$	$\beta = 1, \delta = 0.9$
Bias, ($x10^{-2}$)	-0.085	-1.17	-1.11
Variance, ($x10^{-2}$)	10.52	5.91	15.86

Figure 9. GM-EM reconstruction strongly depends on parameter δ. The Bias and Variance in an ROI of the higher activity is presented.



	$\varepsilon=10^{-4}$	$\varepsilon=10^{-3}$	$\varepsilon=10^{-2}$	$\varepsilon=10^{-1}$
Bias, ($x10^{-2}$)	-1.72	-1.72	-1.73	-1.85
Variance, ($x10^{-2}$)	2.72	2.72	2.71	3.15

Figure 10. TV-EM reconstruction is not sensitive to parameter ε. The Bias and Variance in an ROI of the higher activity is presented.

additional iteration of the ML-EM algorithm shows an increase in the noise level when compared with the true image. The ML-EM algorithm converges to a maximum-likelihood image. However this is not the desired solution due to its high noise level. The ML-EM algorithm should be discontinued after a certain number of iterations to achieve a low-noise solution. In practice however, the stopping rule in the ML-EM algorithm is not easily defined. The TV-EM algorithm does not suffer from this problem and it converges to a low noise solution after several iterations. Note that stability is a general property of MAP estimates in contrast to ML-EM reconstructions.

Figure 4 shows MAP reconstructed images using different energy functions. The ML-EM reconstruction shows a significant noise level after 50 iterations, even though the ML-EM algorithm did not converge. Noise was suppressed by using the SG and GA regularizations, by applying edge-preserving GM and TV regularizations. Note that there is a free parameter β in the reconstruction and the quality of the images depends on its presence. We chose this parameter to achieve the best visual quality in each case. For example, in the GA regularization, the larger regularization parameter would lead to smoother ripples in the image, but would also lead to an increase in the standard deviation between the true and reconstructed image due to the smooth edges. Also, in the GM prior, parameters β and δ needed to be tuned. Our aim is not to provide a rigorous comparison between energy functions, but to demonstrate their main features and introduce TV regularization. The number of iterations was 150, when regularization energy functions were used to ensure convergence of the OSL algorithm. However, smaller numbers could be used.

Figure 5 presents corresponding profiles in the images. It is obvious that the TV regularization is able to preserve edges, as in the case of GM regularization. The SG regularization smoothed noise and edges. The GA regularization smoothed the edges and introduced ripples in the image. The GM and TV regularizations gave similar results. As can be seen in the middle part of the profile, GM regularization can result in an artificial jump in the image and "miss" a true boundary.

Figures 6-8 illustrate reconstruction bias and variance in the ROIs with different activities. After 50 iterations, the ML-EM reconstruction showed the highest level of noise in any ROI. At the same time, the level of bias was relatively small due to averaging. The converged ML-EM algorithm should display a high level of noise in the reconstructed image, but a small level of bias due to averaging. The SG-EM reconstruction led to a high bias due to the smoothing of al edges with high levels of noise. The GA-EM reconstruction displayed a much higher noise level in comparison to the TV-EM. The bias is less for two of the ROIs due to averaging over ripples. However, the results of the GA-EM reconstruction for an ROI with medium background activity were much worse than the results of the TV-EM reconstruction. The GM-EM algorithm reconstruction was much noisier and had a smaller bias than the TV-EM reconstruction for two of the ROIs. The

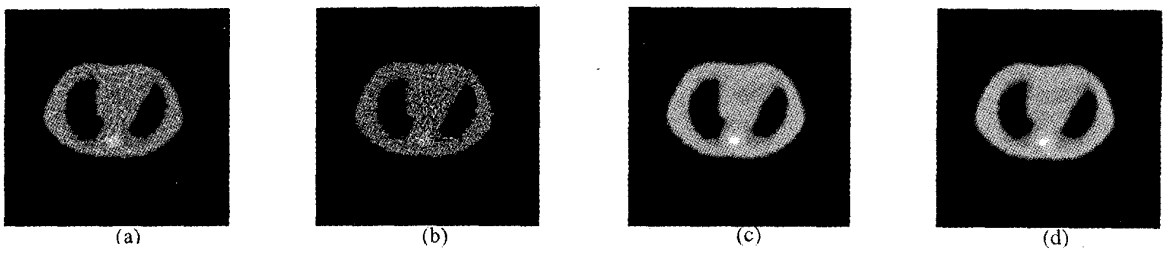


Figure 11. Phantom transmission SPECT study. (a) ML-EM reconstruction, 50 iterations, (b) ML-EM reconstruction, 300 iterations, (c) TV-EM reconstruction, β=1, 50 iterations, and (d) TV-EM reconstruction, β=1, 300 iterations.

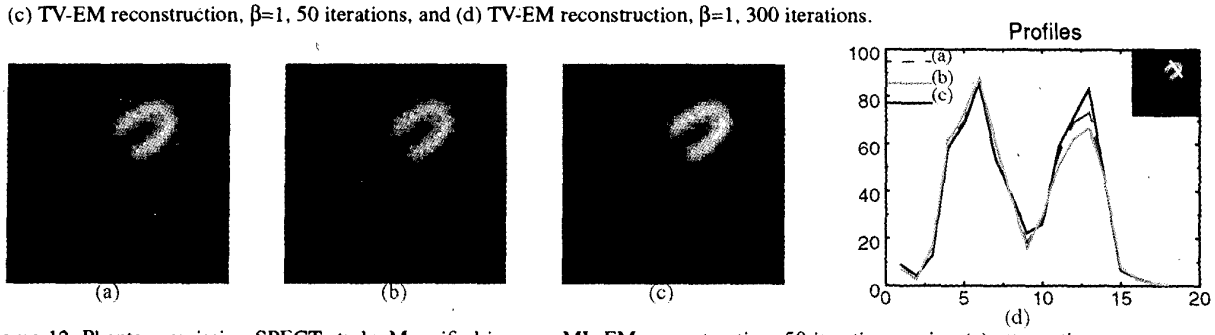


Figure 12. Phantom emission SPECT study. Magnified images. ML-EM reconstruction, 50 iterations, using (a) attenuation map from Fig 4(a), (b) attenuation map from Fig. 4(b), (c) attenuation map from Fig. 4(c), and (d) profiles.

medium background activity bias of the GM-EM reconstruction was much smaller at the same noise level in comparison to the TV-EM reconstruction. The TV regularization successfully suppresses noise, but can lead to some degree of bias and reduction of contrast.

Figures 9 and 10 demonstrate the important advantage of the TV regularization over the GM regularization. As shown in Fig. 9, the quality of the GM-EM reconstruction strongly depends on the parameter δ. Visually and quantitatively, these reconstructions were very sensitive to small changes of δ. For example, for a smaller δ, high amplitude noise peaks appear in the ROI of high activity. These high amplitude noise peaks were interpreted as an edge and were favored by the GM regularization. The reconstruction using a given δ can be improved by changing β. Therefore, to achieve the highest quality using the GM regularization, a two-parameter "optimization" is required. Parameter δ should be in the range of values of actual heights of edges in the image. This requires an investigation of the structure of the reconstructed image. In addition to this, it is not clear whether a single value of δ can accurately describe edges in the whole image. In practice, all edges in an image do not have a single size. This is a common difficulty in applying the GM and similar edge-preserving regularizations. On the other hand, the TV-EM method depends only on the parameter β. As shown in Fig. 10, the TV-EM reconstruction is not sensitive to the artificial parameter ε in a very large range, until ε is sufficiently small. This is consistent with the theoretical study about TV regularization [4]. Practically, it is not difficult to choose a sufficiently small ε before reconstruction because its actual value does not affect the results.

B. Physical Jaszczak Phantom Study

The Jaszczak torso phantom was scanned using Picker's

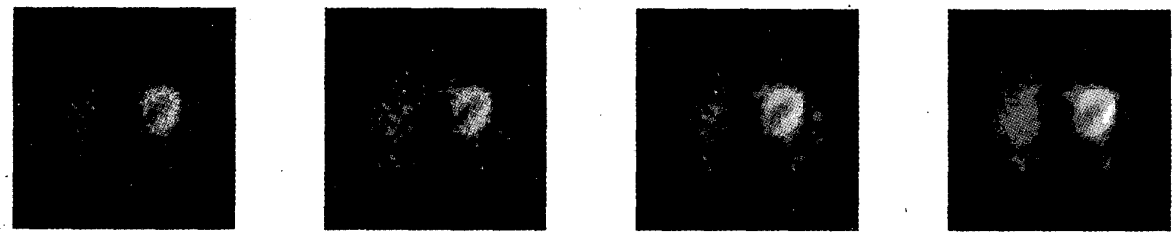


Figure 13. Patient emission SPECT. (a) ML-EM reconstruction, 50 iterations, (b) TV-EM reconstruction, 50 iterations, β=0.5, (c) TV-EM reconstruction, 50 iterations, β=1, (d) TV-EM reconstruction, 50 iterations, β=2.

PRISM 2000XP SPECT system with parallel-hole collimators and a scanning line-source. The number of views was 120 and the number of detector bins was 128. One slice of the Jaszczak phantom reconstruction from the transmission projections is shown in Fig. 11. The TV-EM reconstruction significantly improved the quality of the transmission image over images obtained using the ML-EM reconstruction. In addition to this, there was no limit on the number of iterations used in the TV-EM reconstruction which was not the case with the ML-EM reconstruction. The obtained reconstructed transmission maps were applied to attenuation correction for corresponding emission images (see Fig. 12). Noise from the ML-EM reconstructed attenuation maps was propagated to emission images. The TV-EM reconstructed attenuation map improved the emission images.

C. Emission patient study

The TV-EM algorithm can be applied to the reconstruction of the emission image itself and can be considered as a non-linear filter during reconstruction. A patient was scanned using the same imaging system used in our Jaszczak phantom study. Compared with the phantom study, the patient emission image had a higher noise level. Figure 13 represents the TV-EM reconstructed emission images at different values of regularization parameters without attenuation correction. The degree of "filtering" can be controlled by the regularization parameter β.

IV. DISCUSSION

The total variation (TV) norm minimization is a relatively new, effective technique for image processing. It smooths the image while preserving the edges. A distinguished feature of the TV norm minimization is that it does not introduce over/under shoots and ripples at the edges. Most TV norm minimization methods are based on a least-squares model which is solved by using a set of partial differential equations.

This paper presented results obtained from applying the TV norm minimization to the field of image reconstruction in which a Poisson noise model was used. Green's EM algorithm was adopted with the TV norm being used as the energy function U. The implementation of the algorithm is advantageous partially because it is simple. Moreover, the TV-EM algorithm can be useful for solving linear inverse problems in other applications in which a non-negative result is desired. Therefore, the proposed algorithm may be applied to different fields of engineering in addition to nuclear medicine.

Reconstructions made with computer generated phantom, SPECT transmission, and emission data verify that the proposed algorithm is effective for suppressing noise and preserving edges. Finally, we suggest that the TV-EM algorithm can be used to reconstruct emission images with noisy projections, thereby avoiding the use of post-linear filtering. A post-linear filter (e.g., a Wiener filter or Metz filter) usually introduces ripples and artifacts. In addition, this non-linear TV regularization can be an alternative to existing linear ones that may have disadvantages of linear filtering and blurring of the edges in the reconstructed image. The TV regularization also does not require the setting of any edge height dependent parameters, as is the case for existing non-linear edge preserving priors.

ACKNOWLEDGMENTS

One of the authors, V.Y.P., would like to thank Roman Basko for bringing to his attention successful uses of TV based reconstructions in other fields of engineering. We would like to thank Sean Webb for editing the manuscript.

REFERENCES

[1] P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM Press, 1997.

[2] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D.*, vol. 60, pp. 259-268, 1992.

[3] P. Blomgren and T. F. Chan, "Color TV: Total Variation methods for restoration of vector-valued images," *IEEE Trans. Img. Proc.*, vol. 7, pp. 304-309, 1998.

[4] C. Acar and C.R. Vogel, "Analysis of total variation penalty methods," *Inv. Prob.*, vol.10, pp. 1217-1229, 1994.

[5] C. R. Vogel and M. E. Oman, "Iterative methods for total variation denoising," *SIAM J. Sci. Statist. Comput.*, vol. 17, pp. 227-238, 1996.

[6] C. R. Vogel and M. E. Oman, "Fast, robust total variation-based reconstruction of noisy, blurred images," *IEEE Trans. Img. Proc.*, vol. 7, pp. 813-824, 1998.

[7] T. F. Chan and C. K. Wong, "Total Variation Blind Deconvolution," *IEEE Trans. Img. Proc.*, vol. 7, pp. 370-375, 1998.

[8] I. Csiszar, "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems," *The Annals of Statistics*, vol. 19, no. 4, pp. 2032-2066, 1991.

[9] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards Applied Mathematics Series, 1965.

[10] D. L. Snyder, T. J. Schulz, and J. A. O'Sullivan, "Deblurring subject to nonnegativity constraints," *IEEE Trans Sign. Proc.*, vol. 40, pp. 1143-1150, 1992.

[11] D. S. Lalush and B. M. W. Tsui, "Simulation evaluation of Gibbs prior distributions for use in maximum *a posteriori* SPECT reconstructions," *IEEE Trans. Med. Imag.*, vol. 11, pp. 267-275, 1992.

[12] P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," *IEEE Trans. Med. Imag.*, vol. 9, pp. 84-93, 1990.

[13] S. Geman and D. McClure, "Bayesian image analysis: An application to single photon emission tomography," in *Proc.*

Amer. Statist. Assoc., Stat. Comp. Sect., pp. 12-18, 1985.

[14] T. Hebert and R. Leahy, "A generalized EM algorithm for 3D Bayesian reconstruction from Poisson data using Gibbs priors," *IEEE Trans. Med. Img.*, vol.8, pp. 194-202, 1989.

[15] S. Alenius, U. Ruotsalainen, and J. Astola, "Using local median as the location of the prior distribution in iterative emission tomography image reconstruction," I*EEE Trans. Nucl. Sci.*, vol. 3097, pp. 3097-3104, 1998.

[16] A. R. De Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," *IEEE Trans. Med. Img.*, vol. 14, pp. 132-137, 1995.

[17] G. T. Herman, A. R. De Pierro, and N. Gai, "On methods for maximum a posteriori image reconstruction with a normal prior," *J. Vis. Comm. Img. Repr.*, vol. 3, pp. 316-324, 1992.

[18] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Trans. Img. Proc.*, vol. 6, pp. 298-311, 1997.

[19] I. Laurette, J. Darcourt, L. Blanc-Feraud, P.M. Koulibaly and M. Barlaud, "Combined constraints for efficient algebraic regularized methods in fully 3D reconstructions," *Phys. Med. Biol.*, vol. 43, pp. 991-1000, 1998.