Accurate EM-TV Algorithm in PET with Low SNR

Alex Sawatzky, Christoph Brune, Frank Wübbeling, Thomas Kösters, Klaus Schäfers and Martin Burger

***Abstract*—PET measurements of tracers with a lower dose rate or short radioactive half life suffer from extremely low SNRs. In these cases standard reconstruction methods (OSEM, EM, filtered backprojection) deliver unsatisfactory and noisy results. Here, we propose to introduce nonlinear variational methods into the reconstruction process to make an efficient use of a-priori information and to attain improved imaging results. We illustrate our technique by evaluating cardiac $H_2{}^{15}O$ measurements. The general approach can also be used for other specific goals allowing to incorporate a-priori information about the solution with Poisson distributed data.**

***Index Terms*—PET, $H_2{}^{15}O$ measurements, Total variation, Expectation maximization algorithm, Regularization techniques.**

I. Introduction

Image reconstruction in positron emission tomography (PET) is an inverse problem, which consists of computing an estimate u of the unknown object corresponding to given measurements f. Mathematically, the problem can be formulated as the solution of a linear equation

$$Ku = f \tag{1}$$

where K is an operator that transforms the spatial distribution of the radioactive agent into sampled signals on the detectors. Determining u by direct inversion of K is not suitable, since (1) is ill-conditioned. In such cases regularization techniques are needed to produce reasonable reconstructions.

A frequently used way to realize regularization techniques is the Bayesian model, whose aim is the computation of an estimate u of the unknown object maximizing the a-posteriori probability density $p(u|f)$. The latter is given according to Bayes formula

$$p(u|f) \sim p(f|u)\, p(u) \ . \tag{2}$$

This approach is called maximum a-posteriori probability (MAP) estimation. If the measurements f are given, we describe the density $p(u|f)$ as the a-posteriori likelihood function which depends on u only. The Bayesian approach (2) has the advantage that it allows to incorporate additional information about u via the prior probability density $p(u)$. The most frequently used a-prior densities are Gibbs functions

$$p(u) \sim e^{-\alpha R(u)} \tag{3}$$

where α is a positive parameter and R a convex energy functional. The usual models for the probability density $p(f|u)$ in (2) are exponentially distributed raw data f. In the canonical case of additive Gaussian noise

$$p(f|u) \sim e^{-||Ku-f||^2/(2\sigma^2)}$$

the minimization of the negative log likelihood function leads to classical Tikhonov regularization technique [1] based on minimizing a functional of the form

$$\frac{1}{2}||Ku - f||_2^2 + \alpha R(u) \ , \qquad u \geq 0 \ . \tag{4}$$

The first so-called data-fidelity term penalizes the deviation from the equality $Ku = f$ and R is the regularization as in (3). The regularization parameter α is a relative weight for both terms. The additional positivity constraint is usually necessary because in typical applications the unknown functions represent a density or intensity.

In PET data are not Gaussian but Poisson distributed [2] and the MAP estimation via the negative log likelihood function leads to the following variational problem [1]

$$\min_{u \geq 0} \int_\Sigma (Ku - f \log Ku)\ d\mu + \alpha R(u) \ . \tag{5}$$

A particular complication of (5) compared to (4) is the strong nonlinearity in the data fidelity term and resulting issues in the computation of minimizers. In the absence of regularization ($\alpha = 0$) the EM-algorithm (cf. [2]) has become a standard scheme, which is however difficult to be generalized to regularized cases. The robust solution of this problem for appropriate models of R is the main novelty of this paper.

The specific choice of regularization functional R in (5) is important for the way a-priori information about the expected solution is incorporated into reconstruction process. Smooth, in particular quadratic regularizations have attracted most attention in the past, mainly due to the simplicity in analysis and computation. However, such regularization approaches always lead to blurring of the reconstructions, in particular they cannot yield reconstructions with sharp edges. Recently singular regularization energies, in particular those of ℓ^1 or L^1-type have attracted strong attention. In this work, we introduce an approach which uses total variation (TV) as regularization functional. TV regularization was derived as a denoising technique in [3] and generalized to various other imaging tasks subsequently. The exact definition of TV [4] is

$$|u|_{BV} := \sup_{\substack{g \in C_0^\infty(\Omega;R^d) \\ ||g||_\infty \leq 1}} \int_\Omega u \ \mathrm{div} g \ , \tag{6}$$

which is formally (true if u is sufficiently regular)

$$|u|_{BV} = \int_\Omega |\nabla u| \ .$$

K. Schäfers is with the Department of Nuclear Medicine, University Hospital of Münster, D-48149 Münster, Germany, e-mail: schafkl@wwu.de.

All other authors are with the Department of Mathematics and Computer Science, University of Münster, D-48149 Münster, Germany, e-mail: {alex.sawatzky, christoph.brune, thomas.koesters, frank.wuebbeling, martin.burger} @ wwu.de.

http://wwwmath.uni-muenster.de/num/Arbeitsgruppen/ag_burger/

The function space of functions of bounded (total) variation is described by $BV(\Omega)$. For further properties and details on functions with bounded variation we refer to [4], [5]. The motivation for using TV is that it suppresses noise effectively and realizes almost homogeneous regions with sharp edges. These features are particularly attractive for several PET modes, when the goal is to identify the shape of objects that are separated by sharp edges.

In the past such methods have been suggested [6], [7] for use in PET, but still with limited success due to strong computational difficulties in the minimization of (5) with total variation as a regularization functional, and due to remaining blurring caused by using approximations of TV by differentiable functionals

$$\int_\Omega \sqrt{|\nabla u|^2 + \varepsilon} \; , \qquad \varepsilon > 0 \; .$$

The schemes suggested in [6], [7] are realized as elementary modifications of the Expectation-Maximization (EM) algorithm, with fully explicit or semi-implicit treatment of TV in the iteration. A major disadvantage of these approaches is that the regularization parameter α needs to be chosen very small, since otherwise the positivity of solutions is not guaranteed and the EM-type algorithm cannot be continued. Due to an additional parameter dependence on ε these algorithms are even less robust. Here, we propose a robust algorithm without approximation of TV, i.e. we use (6) respectively a dual version. This allows to realize cartoon reconstructions with sharp edges. Another advantage of our approach is that it can be performed equally well for large regularization parameter. Thus it is favourably applicable for problems with low SNRs.

The challenges in this work are that the Poisson based negative log likelihood functional in (5) is highly nonlinear and, due to total variation, the variational problem is non differentiable. We propose a forward-backward splitting approach, which can be realized by alternating classical EM-steps with almost standard TV minimization steps as encountered in image denoising. Furthermore, the proposed EM-TV algorithm is compared with different application strategies of post smoothing. To illustrate the behaviour of this methods we use the evaluation of cardiac $H_2\,^{15}O$ measurements, which suffer from very low SNRs.

II. Methods

In general, two types of reconstruction methods are used: analytic (direct) and algebraic (iterative) methods. A classical example for direct methods is the Fourier-based filtered backprojection (FBP). Although FBP is well understood and computationally efficient, iterative type methods obtain more and more attention in PET. The major reason is the high noise level (low SNR) and the type of statistics found in PET measurements, which cannot be taken into account by direct methods. We shall give here a short review of the Expectation-Maximization (EM) algorithm [2], [8], which is a popular iterative algorithm to maximize the likelihood function $p(u|f)$ in problems with incomplete data. Then we proceed to the presentation of the proposed EM-TV algorithm.

A. Reconstruction method: EM algorithm

In the absence of prior knowledge any object u has the same relevance, i.e. the Gibbs a-priori density $p(u)$ in (3) is constant. We can then normalize $p(u)$ such that $R(u) \equiv 0$. Hence (5) reduces to the constrained minimization problem

$$\min_{u \geq 0} \int_\Sigma (Ku - f \log Ku) \; d\mu \; . \qquad (7)$$

A suitable iteration scheme for computing stationary points, which also preserves positivity, is the so called EM algorithm (cf. [9])

$$u_{k+1} = u_k \frac{K^*}{K^* 1} \left(\frac{f}{K u_k} \right) \; , \qquad k = 0, 1, \ldots \; . \qquad (8)$$

For the EM algorithm several convergence proofs can be found in the literature [10], [9], [11], [12]. But, as described in [11], the EM iterates show the following typical behaviour. The (metric) distance between the iterates and the solution first decays and then increases. The reason is again the ill-conditioning of $Ku = f$, which transfers to problem (7). As a consequence the noise is amplified during iteration process. This issue might be regulated by using appropriate stopping rules to obtain reasonable results. In [11] is shown that certain stopping rules indeed allow stable approximations. Another possibility to considerably improve reconstruction results are regularization techniques, which we shall discuss in the following.

B. Reconstruction method: Proposed EM-TV algorithm

The EM algorithm is currently the standard iterative reconstruction method in PET. However, in the case of tracer measurements with low SNR, like lower dose rate or tracer with short radioactive half life, it delivers unsatisfactory and noisy results even with early termination. Therefore we propose to introduce nonlinear variational methods into the reconstruction process to make an efficient use of a-priori information and to obtain improved results.

An interesting approach to the improvement of reconstructed images is the EM-TV algorithm. In the classical EM algorithm, the negative log likelihood functional (7) is minimized. In the EM-TV approach, we modify this functional by adding a weighted total variation (TV) term [3],

$$\min_{\substack{u \in BV(\Omega) \\ u \geq 0}} \int_\Sigma (Ku - f \log Ku) \; d\mu \; + \; \alpha \, |u|_{BV} \; . \qquad (9)$$

This is exactly (5) with TV as the regularization functional R. It means that images with smaller total variation are prefered in the minimization (have higher prior probability). The expected reconstructions are cartoon-like images, i.e. they will result in almost uniform (mean) tracer activities inside the different structures. Of course, such an approach cannot be used for studying activity inside the structures (which is anyway unrealistic given the low SNR), but it will be perfect for segmenting the different structures, e.g. inside the human heart.

For the solution of (9), we propose here a forward-backward splitting algorithm, which can be realized by alternating a classical EM steps with almost standard TV minimization steps as encountered in image denoising. The latter is solved by using duality [13] and one obtains a robust and efficient algorithm. For designing the proposed alternating algorithm, we consider the first order optimality condition of (9). Due to the total variation, this variational problem is not differentiable in the usual sense. However, since we can extend the data fidelity term to a convex functional (Kullback-Leibler functional, cf. [14]) without changing the stationary points, the minimization problem becomes convex. For such problems powerful methods from convex analysis are available, e.g. a generalized derivative called the subdifferential [15], denoted by ∂. This generalized notion of gradients and the Karush-Kuhn-Tucker (KKT) conditions [15, Theorem 2.1.4] yield the existence of a Lagrange multiplier $\lambda \geq 0$ such that

$$\left\{ \begin{array}{lcl} 0 & \in & K^*1 \; - \; K^* \left(\dfrac{f}{Ku} \right) \; + \; \alpha \, \partial |u|_{BV} \; - \; \lambda \\ 0 & = & \lambda \, u \end{array} \right\} \; . \quad (10)$$

With the formal definition of total variation, $|u|_{BV} = \int_\Omega |\nabla u|$, and $\nabla u \neq 0$ the optimality condition has the form

$$0 \; = \; K^*1 \; - \; K^* \left(\frac{f}{Ku} \right) \; - \; \alpha \, \mathrm{div} \left(\frac{\nabla u}{|\nabla u|} \right) \; - \; \lambda \; .$$

Multiplying (10) with u we can eliminate the Lagrange multiplier and derive the following semi-implicit iteration scheme

$$u_{k+1} \; - \; u_k \frac{K^*}{K^*1} \left(\frac{f}{K u_k} \right) \; + \; \tilde{\alpha} \; u_k \; p_{k+1} \; = \; 0 \qquad (11)$$

with $p_{k+1} \; \in \; \partial |u_{k+1}|_{BV}$ and $\tilde{\alpha} \; := \; \frac{\alpha}{K^*1}$. Interestingly, the second term in this iteration is the EM step from (8). Therefore, the method (11) for the solution of the variational problem (9) can be realized as a nested two step iteration,

$$\left\{ \begin{array}{lcll} u_{k+\frac{1}{2}} & = & u_k \, \dfrac{K^*}{K^*1} \left(\dfrac{f}{K u_k} \right) & \text{(EM step)} \\ u_{k+1} & = & u_{k+\frac{1}{2}} - \tilde{\alpha} \; u_k \; p_{k+1} & \text{(TV step)} \end{array} \right\} \; . \quad (12)$$

Thus, we alternate an EM step with a TV correction step. The complex second half step from $u_{k+\frac{1}{2}}$ to u_{k+1} can be realized by solving the following variational problem

$$u_{k+1} \; = \; \underset{u \in BV(\Omega)}{\arg\min} \left\{ \frac{1}{2} \int_\Omega \frac{(u - u_{k+\frac{1}{2}})^2}{u_k} + \tilde{\alpha} \, |u|_{BV} \right\} \; . \quad (13)$$

Inspecting the first order optimality condition confirms the equivalence of this minimization with the TV correction step in (12). Problem (13) is just a modified version of the Rudin-Osher-Fatemi (ROF) model, with weight $\frac{1}{u_k}$ in the fidelity term. This analogy creates the possibility to carry over efficient numerical schemes known for the ROF-model.

For the solution of (13) we use the exact definition of TV (6) with dual variable g and derive an iteration scheme for the dual problem similar to Chambolle [13]. The minimization in (13) can be written as a saddle point problem in u and g,

$$\inf_u \sup_{g, ||g||_\infty \leq 1} \left\{ \frac{1}{2} \int_\Omega \frac{(u - u_{k+\frac{1}{2}})^2}{u_k} + \tilde{\alpha} \int_\Omega u \, \mathrm{div} g \right\} \; . \quad (14)$$

After exchanging $\inf$ and $\sup$, the primal optimality condition for this problem is given by

$$u = u_{k+\frac{1}{2}} - \tilde{\alpha} \, u_k \, \mathrm{div} \tilde{g} \; . \qquad (15)$$

If the optimal dual variable $\tilde{g}$ is known then we can use (15) to obtain a solution u_{k+1} of (14). For the computation of $\tilde{g}$, we insert (15) in (14) and obtain a pure dual problem that depends on g only, i.e.

$$\begin{array}{c} \tilde{g} = \underset{g \in C_0^\infty(\Omega; R^d)}{\arg\min} \displaystyle\int_\Omega \frac{(\tilde{\alpha} \, u_k \, \mathrm{div} g - u_{k+\frac{1}{2}})^2}{u_k} \\ \text{s.t. } ||g||_\infty - 1 \leq 0 \end{array} \; . \qquad (16)$$

The dual problem here is a quadratic optimization problem with a nonlinear constraint. Similar to [13] we derive a semi-implicit (projected) gradient descent algorithm for solving (16). We initialize the dual variable g^0 with 0 (or the resulting g from the previous TV correction step) and for any $n \geq 0$ we compute the update

$$g^{n+1} \; = \; \frac{g^n + \tau \; \nabla(\tilde{\alpha} \, u_k \, \mathrm{div} g^n - u_{k+\frac{1}{2}})}{1 + \tau \; |\nabla(\tilde{\alpha} \, u_k \, \mathrm{div} g^n - u_{k+\frac{1}{2}})|}$$

with the damping parameter τ constrained by

$$0 \; < \; \tau \; < \; \frac{1}{4 \, \tilde{\alpha} \, u_k},$$

which ensures stability and convergence of the algorithm.

III. Results

We demonstrate our technique by evaluating cardiac $H_2\,^{15}O$ measurements [16]. For the illustration of tracer intensity in the right and left ventricle we take a fixed 2D layer in two different time frames (25 seconds and 45 seconds after tracer injection in the blood circulation), see Fig. 1.

To illustrate the SNR problem we present reconstructions with the classical EM algorithm (A). As expected, the results suffer from unsatisfactory quality. We hence take EM reconstrutions with Gauss smoothing (B) as a reference. The next results (C - F) show different application strategies of TV smoothing with the (weighted) ROF model (13). The classical approach is a completed EM algorithm followed by a single denoising step. The result C demonstrates this approach with the standard ROF model, i.e. replacing u_k by one and $u_{k+\frac{1}{2}}$ by the final EM reconstruction in (13). The result D is generated similarly to C but with weighted ROF model, i.e. both u_k and $u_{k+\frac{1}{2}}$ are replaced by the final EM reconstruction in (13). Our approach with the nested EM-TV algorithm (12) is presented in F. The reconstruction E is the same result as in D, but the image is scaled to the maximum intensity of F, such that a comparison is possible also for quantitative values .

IV. Discussion

The study demonstrates that denoising techniques based on total variation minimization deliver better results than the standard Gauss smoothing approach in low SNRs situations. The advantage is that cartoon reconstruction with TV preserves desired structures, e.g. homogeneous regions with sharp edges, very well. Furthermore, the reconstructions in D and F with

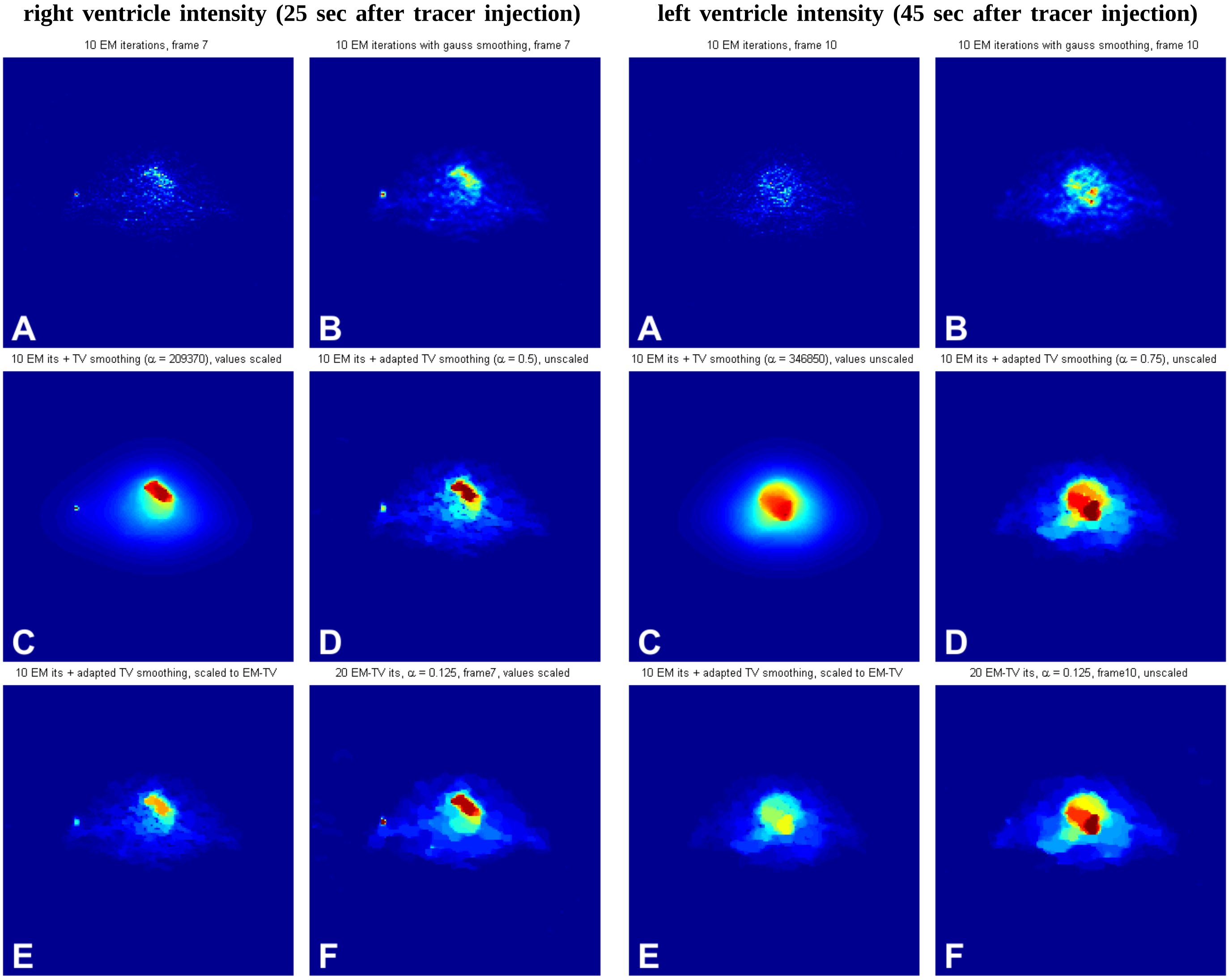

Figure 1. Cardiac $H_2\,^{15}O$ PET measurements: results of different reconstruction methods in two different time frames. **A:** classical EM algorithm (8). **B:** EM with Gauss smoothing. **C:** EM with following TV smoothing (13) where $u_{k+1/2}$ = EM reconstruction and $u_k = 1$. **D:** As C but with $u_{k+1/2} = u_k$ = EM reconstruction in (13). **E:** D scaled to maximum intensity of F. **F:** Nested EM-TV algorithm (12).

weighted ROF model have a qualitatively better resolution than the standard model, caused by the adaptation of the regularization parameter to the reconstruction u_k. For further use, e.g. segmentation for quantification of myocardial blood flow [17], the reconstructions with a single denoising step (D) are satisfactory. For quantitative conclusions the reconstructions with nested EM-TV algorithm delivers better interpretations, see Fig. 1 (E and F).

Beside the improved numerical results with nested iteration (12), the alternating algorithm has the advantage that we might control the interaction between reconstruction and denoising via a simple adaptation of the TV correction step. A possibility for improving convergence is a damped TV correction step

$$u_{k+1} = (1-\omega_k)u_k \; + \; \omega_k \, u_{k+\frac{1}{2}} \; - \; \omega_k \, \tilde{\alpha} \, u_k \, p_{k+1} \qquad (17)$$

with $\omega_k \in [0,1]$ relating the current EM step $u_{k+\frac{1}{2}}$ and the previous TV denoising result u_k . The damped half step (17) can also be realized in an analogous way to (13), namely by minimizing

$$u_{k+1} \;=\; \underset{u \in BV(\Omega)}{\arg\min} \left\{ \frac{1}{2} \int_\Omega \frac{(u - (\omega_k \, u_{k+\frac{1}{2}} + (1-\omega_k)\, u_k)\,)^2}{u_k} \right. \left. + \; \omega_k \, \tilde{\alpha} \, |u|_{BV} \right\} .$$

Damping might be necessary to attain a monotone descent of the regularized negative log likelihood functional (9) during the iteration. For $\omega_k = 1$, we recover the original TV correction step in (12).

It is well known that the reconstructed images resulting from denoising techniques based on total variation suffer from contrast reduction. In [18], we propose to extend the EM-TV algorithm by an iterative regularization to obtain a simultaneous contrast enhancement. We realize the contrast improvement via inverse scale space methods and Bregman iterations, derived in [19], [20], [21]. According to these methods, the iterative contrast correction will be realized by a sequence of modified EM-TV problems with similar form as (9).

References

[1] M. Bertero, H. Lanteri, and L. Zanni, "Iterative image reconstruction: a point of view," in *Mathematical Methods in Biomedical Imaging and Intensity-Modulated Radiation Therapy (IMRT)*, ser. Publications of the Scuola Normale, CRM series, **8**, 2008.

[2] L. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Transactions on Medical Imaging*, vol. 1, no. 2, pp. 113–122, 1982.

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

[4] R. Acar and C. Vogel, "Analysis of bounded variation penalty methods for ill-posed problems," *Inverse Problems*, vol. 10, pp. 1217–1229, 1994.

[5] L. Evans and R. Gariepy, *Measure theory and fine properties of functions*. CRC Press: Boca Raton/New York/London/Tokyo, 1992.

[6] E. Jonsson, C.-S. Huang, and T. Chan, "Total variation regularization in positron emission tomography," *CAM Report 98-48, UCLA*, November 1998.

[7] V. Panin, G. Zeng, and G. Gullberg, "Total variation regulated EM algorithm [SPECT reconstruction]," *IEEE Trans. Nucl. Sci. Vol.*, vol. 46, pp. 2202–2210, 1999.

[8] A. Dempster, N. Laird, and D. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society, B*, vol. 39, pp. 1–38, 1977.

[9] F. Natterer and F. Wübbeling, *Mathematical methods in image reconstruction*. SIAM, 2001.

[10] A. Iusem, "Convergence Analysis for a Multiplicatively Relaxed EM Algorithm," *Mathematical Methods in the Applied Sciences*, vol. 14, pp. 573–593, 1991.

[11] E. Resmerita, H. Engl, and A. Iusem, "The expectation-maximization algorithm for ill-posed integral equations: a convergence analysis," *Inverse Problems*, vol. 23, pp. 2575–2588, 2007.

[12] Y. Vardi, L. Shepp, and L. Kaufman, "A Statistical Model for Positron Emission Tomography," *Journal of the American Statistical Association*, vol. 80, no. 389, pp. 8–20, 1985.

[13] A. Chambolle, "An Algorithm for Total Variation Minimization ans Applications," *Journal of Mathematical Imaging and Vision*, vol. 20, pp. 89–97, 2004.

[14] E. Resmerita and S. Anderssen, "Joint additive Kullback-Leibler residual minimization and regularization for linear inverse problems," *Math. Meth. Appl. Sci*, vol. 30, pp. 1527–1544, 2007.

[15] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms I, volume 305 of Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, 1993.

[16] K. P. Schäfers, T. J. Spinks, P. G. Camici, P. M. Bloomfield, C. G. Rhodes, M. P. Law, C. S. R. Baker, and O. Rimoldi, "Absolute Quantification of Myocardial Blood Flow with $H_2\,^{15}O$ and 3-Dimensional PET: An Experimental Validation," *Journal of Nuclear Medicine Vol.*, vol. 43, pp. 1031–1040.

[17] M. Benning, T. Kösters, F. Wübbeling, K. P. Schäfers, and M. Burger, "A Nonlinear Variational Method for Improved Quantification of Myocardial Blood Flow Using Dynamic $H_2\,^{15}O$ PET," 2008, IEEE Nuclear Science Symposium and Medical Imaging Conference, MIC Poster I, M06-507.

[18] A. Sawatzky, C. Brune, F. Wübbeling, T. Kösters, and M. Burger, "EM-TV Methods in Inverse Problems with Poisson noise," 2008, in progress.

[19] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation based image restoration," *Multiscale Modelling and Simulation*, vol. 4, pp. 460–489, 2005.

[20] M. Burger, G. Gilboa, S. Osher, and J. Xu, "Nonlinear inverse scale space methods," *Comm. Math. Sci.*, pp. 179–212, 2006.

[21] M. Burger, K. Frick, S. Osher, and O. Scherzer, "Inverse total variation flow," *SIAM Multiscale Modelling and Simulation*, pp. 366–395, 2007.