Appears in *ISVC 2011, Part I, LNCS 6938, pp. 33-42, 2011*

EM-Type Algorithms for Image Reconstruction with Background Emission and Poisson Noise*

Ming Yan

Department of Mathematics, University of California, Los Angeles
520 Portola Plaza, Los Angeles, CA 90095-1555, United States
yanm@math.ucla.edu

Abstract. Obtaining high quality images is very important in many areas of applied sciences. In this paper, we proposed general robust expectation maximization (EM)-Type algorithms for image reconstruction when the measured data is corrupted by Poisson noise. This method is separated into two steps: EM and regularization. In order to overcome the contrast reduction introduced by some regularizations, we suggested EM-Type algorithms with Bregman iteration by applying a sequence of modified EM-Type algorithms. The numerical experiments show the effectiveness of these methods in different applications.
Keywords: background emission, expectation maximization, image reconstruction, Poisson noise.

1 Introduction

Obtaining high quality images is very important in many areas of applied sciences, such as medical imaging, optical microscopy and astronomy. Though for some applications such as positron-emission-tomography (PET) and computed tomography (CT), analytical methods for image reconstruction are available. For instance, filtered back projection (FBP) is the most commonly used method for image reconstruction from CT by manufacturers of commercial imaging equipment [1]. However, it is sensitive to noise and suffers from streak artifacts (star artifacts). An alternative to analytical reconstruction is the use of iterative reconstruction technique, which is quite different from FBP. The main advantages of iterative reconstruction technique over FBP are insensitivity to noise and flexibility [2]. The data can be collected over any set of lines, the projections do not have to be distributed uniformly in angle, and the projections can be even incomplete (limited angle). We will focus on the iterative reconstruction technique.

Image reconstruction in the applications mentioned above can be formulated as linear inverse and ill-posed problems,

$$y = Ax + b + n. \tag{1}$$

* This work was supported by the Center for Domain-Specific Computing (CDSC) under the NSF Expeditions in Computing Award CCF-0926127.

Here, y is the measured data (vector in $\mathbf{R}^M$ for discrete case), A is a compact operator (matrix in $\mathbf{R}^{M\times N}$ for the discrete case), which is different for different applications, x is the desired exact image (vector in $\mathbf{R}^N$ for discrete case), b is the background and n is the noise. We will consider only the case with background emission ($b \neq 0$) in this paper. In astronomy, this is due to sky emission [3, 4]. In fluorescence microscopy, it is due to auto-fluorescence and reflections of the excitation light. The computation of u directly by finding the inverse of A is not reasonable because (1) is ill-posed and n is unknown. Therefore regularization techniques are needed for solving these problems efficiently.

One powerful technique for applying regularization is the Bayesian model, and a general Bayesian model for image reconstruction was proposed by Geman and Geman [5], and Grenander [6]. The idea is to use *a priori* information about the image x to be reconstructed. In the Bayesian approach, we assume that measured data y is a realization of a multi-value random variable, denoted by Y and the image x is also considered as a realization of another multi-value random variable, denoted by X. Therefore the Bayesian formula gives us

$$p_X(x|y) = \frac{p_Y(y|x)p_X(x)}{p_Y(y)}. \tag{2}$$

This is a conditional probability of having $X = x$ given that y is the measured data. After inserting the detected value of y, we obtain a posterior probability distribution of X. Then we can find x^* such that $p_X(x|y)$ is maximized, as maximum a posterior (MAP) likelihood estimation.

In general, X is assigned as a Gibbs random field, which is a random variable with the following probability distribution

$$p_X(x) \sim e^{-\beta J(x)}, \tag{3}$$

where $J(x)$ is a given energy functional, and β is a positive parameter. There are many different choices for $J(x)$ depends on the applications. Some examples are instance quadratic penalization $J(x) = \|x\|_2^2/2$ [7, 8], quadratic Laplacian $J(x) = \|\nabla x\|_2^2/2$ [9], total variation (TV) $J(x) = \||\nabla x|\|_1$ [10–12], and Good's roughness penalization $J(x) = \||\nabla x|^2/x\|_1$ [13].

If the random variable Y of the detected values y follows a Poisson distribution [14] with an expectation value provided by $Ax + b$, we have

$$y_i \sim \text{Poisson}\{(Ax+b)_i\}, \quad i.e. \quad p_Y(y|x) \sim \prod_i \frac{(Ax+b)_i^{y_i}}{y_i!} e^{-(Ax+b)_i}. \tag{4}$$

By minimizing the negative log-likelihood function ($-\log p_X(x|y)$), we obtain the following optimization problem

$$\underset{x\geq 0}{\text{minimize}} \ \sum_i \big((Ax+b)_i - y_i \log(Ax+b)_i\big) + \beta J(x). \tag{5}$$

In this paper, we will focus on solving (5). It is easy to see that the objective function in (5) is convex if $J(x)$ is convex. Additionally, with suitably chosen regularization $J(x)$, the objective function is strictly convex, and the solution to this problem is unique.

The work is organized as follows. In section 2, we will give a short introduction of expectation maximization (EM), or Richardson-Lucy, used in image reconstruction with background emission. In section 3, we will propose a general EM-Type algorithm for image reconstruction when the measured data is corrupted by Poison noise. This is based on the maximum a posteriori likelihood estimation and EM. However, for some regularizations such as TV, the reconstructed image will lose contrast, EM-Type algorithm with Bregman iteration is introduced in section 4. Some numerical experiments are given in section 5 to show the efficiency of the EM-Type algorithm. We will end this work by a short conclusion section.

2 Expectation Maximization (EM)

A maximum likelihood (ML) method for image reconstruction based on Poisson data was introduced by Shepp and Vardi [14] in 1982 for applications in emission tomography. In fact, this algorithm was originally proposed by Richardson [15] in 1972 and Lucy [16] in 1974 for astronomy. Here, we consider the special case without regularization term, i.e. $J(x)$ is a constant, we do not have any *a priori* information about the image. From equation (4), for given measured data y, we have a function of x, the likelihood of x, defined by $p_Y(y|x)$. Then a ML estimate of the unknown image is defined as any maximizer x^* of $p_Y(y|x)$.

By taking the negative log-likelihood, one obtains, up to an additive constant

$$f_0(x) = \sum_i \big((Ax+b)_i - y_i \log(Ax+b)_i\big), \tag{6}$$

and the problem is to minimize this function $f_0(x)$ on the nonnegative orthant, because we have the constraint that the image x is nonnegative. In fact, we have

$$f(x) = D_{KL}(Ax+b, y) \equiv \sum_i \Big(y_i \log \frac{y_i}{(Ax+b)_i} + (Ax+b)_i - y_i\Big) = f_0(x) + C,$$

where $D_{KL}(Ax+b, y)$ is the Kullback-Leibler (KL) divergence of $Ax+b$ from y, and C is a constant independent of x. The KL divergence is considered as a data-fidelity function for Poisson data. It is convex, nonnegative and coercive on the nonnegative orthant, so the minimizers exist and are global.

The well-known EM algorithm, or Richardson-Lucy algorithm is

$$x_j^{k+1} = \frac{\sum_i \Big(a_{ij}\big(\frac{y_i}{(Ax^k+b)_i}\big)\Big)}{\sum_i a_{ij}} x_j^k. \tag{7}$$

Shepp and Vardi showed in [14] that when $b = 0$, this is an example of the EM algorithm proposed by Dempster, Laird and Rubin [17].

3 EM-Type Algorithms for Image Reconstruction

The method shown in last section is also called maximum-likelihood expectation-maximization (ML-EM) reconstruction, because it is a maximum likelihood approach without any Bayesian assumption on the images. If additional *a priori* information about the image is given, we have maximum a posteriori probability (MAP) approach [18, 19], which is the case with regularization term $J(x)$. Again we assume here that the detected data is corrupted by Poisson noise, and the regularization problem is

$$\begin{cases} \underset{x}{\text{minimize}} & E^p(x) \equiv \beta J(x) + \sum_i \left((Ax+b)_i - y_i \log(Ax+b)_i\right), \\ \text{subject to} & x_j \geq 0, \quad j = 1, \cdots, N. \end{cases} \tag{8}$$

This is a convex constraint optimization problem and we can find the optimal solution by solving the Karush-Kuhn-Tucker conditions:

$$\begin{aligned} \beta \partial J(x)_j + \sum_i \left(a_{ij}(1 - \frac{y_i}{(Ax+b)_i}) \right) - s_j = 0, &\qquad j = 1, \cdots, N, \\ s_j \geq 0, \quad x_j \geq 0, &\qquad j = 1, \cdots, N, \\ s^T x = 0. & \end{aligned}$$

Here s_j is the Lagrangian multiplier corresponding to the constraint $x_j > 0$. By the positivity of $\{x_j\}$, $\{s_j\}$ and the complementary slackness condition $s^T x = 0$, we have $s_j x_j = 0$ for every $j = 1, \cdots, N$. Thus we obtain

$$\beta x_j \partial J(x)_j + \sum_i \left(a_{ij}(1 - \frac{y_i}{(Ax+b)_i}) \right) x_j = 0, \qquad j = 1, \cdots, N,$$

or equivalently

$$\beta \frac{x_j}{\sum_i a_{ij}} \partial J(x)_j + x_j - \frac{\sum_i \left(a_{ij}(\frac{y_i}{(Ax+b)_i}) \right)}{\sum_i a_{ij}} x_j = 0, \qquad j = 1, \cdots, N.$$

Notice that the last term on the left hand side is an EM step (7). After plugging the EM step into the KKT condition[20], we obtain

$$\beta \frac{x_j}{\sum_i a_{ij}} \partial J(x)_j + x_j - x_j^{EM} = 0, \qquad j = 1, \cdots, N,$$

which is the optimality for the following optimization problem

$$\underset{x}{\text{minimize}}\ E_1^p(x, x^{EM}) \equiv \beta J(x) + \sum_j (\sum_i a_{ij}) \left(x_j - x_j^{EM} \log x_j \right). \tag{9}$$

Therefore we propose the general EM-Type algorithm in Algorithm 1. If $J(x)$ is constant, the second step is just $x^k = x^{k+\frac{1}{2}}$ and this is exactly the ML-EM in last section. For the case $J(x)$ is not constant, we have to solve an optimization problem for each iteration. In general, the problem can not be solved analytically, and we have to use iterative methods to solve it. In practical, we do not have to solve it exactly by stopping it after a few iterations and the algorithm will also converge without solving it exactly.

Input: Given x^0;
for *k=1:IterMax* **do**
  $x^{k-\frac{1}{2}} = EM(x^{k-1})$ using (7) ;
  $x^k = \text{argmin } E_1^p(x, x^{k-\frac{1}{2}})$ by solving (9);
end

Algorithm 1: Proposed EM-Type algorithm.

4 EM-Type Algorithms with Bregman Iteration

In previous section, the EM-Type algorithms are presented to solve the problem (8). However, the regularization may lead to reconstructed images suffering from contrast reduction [21]. Therefore, we suggest a contrast improvement in EM-Type algorithms by Bregman iteration, which is introduced in [22–24]. An iterative refinement is obtained from a sequence of modified EM-Type algorithms.

For the problem with Poisson noise, we start with the basic EM-Type algorithm, i.e. finding the minimum x^1 of (8). After that, variational problems with a modified regularization term

$$x^{k+1} = \underset{x}{\text{argmin}} \ \beta(J(x) - \langle p^k, x\rangle) + \sum_i ((Ax+b)_i - y_i \log(Ax+b)_i) \qquad (10)$$

where $p^k \in \partial J(x^k)$, are solved sequentially. From the optimality of (10), we have the following formula for updating p^{k+1} from p^k and x^{k+1}.

$$p^{k+1} = p^k - \frac{1}{\beta} A^T \left(1 - \frac{y}{Ax^{k+1}+b}\right). \qquad (11)$$

Therefore the algorithm with Bregman iteration is described in Algorithm 2.

5 Numerical Experiments

In this section, we will illustrate the proposed EM-Type algorithms for image reconstruction (more specifically, image deblurring). The regularization we used is total variation (TV) regularization, and we present some deblurring results with the proposed EM-TV algorithm and the Bregman version of it. The first image used is a synthetic 200×200 phantom. It consists of circles with intensities 65, 110 and 170, enclosed by a square frame of intensity 10. For the experiment, we choose the background $b = 20$. Firstly, we consider the case without noise.

Input: Given x^0, δ, ϵ, $k = 1$ and $p^0 = 0$;
while $k \leq$ *Num_outer* & $D_{KL}(Ax^{k-1} + b, y) < \delta$ **do**
  $x^{temp,0} = x^{k-1}$; $l = 0$;
  while $l \leq$ *Num_inner* & $\|x^{temp,l} - x^{temp,l-1}\| \leq \epsilon$ **do**
    $l = l + 1$;
    $x^{temp,l-\frac{1}{2}} = EM(x^{temp,l-1})$ using (7) ;
    $x^{temp,l} = \underset{x}{\operatorname{argmin}}\ E_1^p(x, x^{temp,l-\frac{1}{2}})$ with $J(x) - \langle p^{k-1}, x\rangle$;
  end
  $x^k = x^{temp,l}$;
  $p^k = p^{k-1} - \frac{1}{\beta} A^T \left(1 - \frac{y}{Ax^k + b}\right)$; $k = k + 1$;
end

Algorithm 2: Proposed EM-Type algorithm with Bregman iteration.

The blurred image is obtained from original image using a Gaussian blur kernel K with standard deviation $\sigma = 100$.

To illustrate the advantage of Bregman iterations, we show the results in Figure 1. The RMSE for 1(a), 1(b) and 1(c) are 11.9039, 5.0944 and 2.5339, respectively, and the corresponding KL distances $D_{KL}(Ax + b, y)$ are 93.0227, 0.8607 and 0.0172, respectively.

How to choose a good β is important for algorithm 1, but not for algorithm 2 with Bregman iteration. For this example, though β is not chosen to be optimal for algorithm 1, the result of Bregman iteration shows that we can still obtain good result after some iterations. From the lineouts we can see that the result with Bregman iteration fits the original image very well.

Next experiment is to perform deconvolution on an image of a satellite (Figure 2(a)), and the point spread function (PSF) is shown in Figure 2(b). In order to make the algorithm fast, we choose the initial guess x^0 to be the result from solving $Ax = y - b$ using conjugate gradient (CG). The negative values are changed into zero before applying EM-TV algorithm. The corresponding RMSE for x^0 and the result are 13.6379 and 11.8127, respectively. By using the EM-TV with Bregman iteration, we get a better image with sharp edges and remove the artifacts.

Though convergence analysis of EM-Type algorithms can be shown if $J(x)$ is convex. When $J(x)$ is not convex, the functional may have several local minimums and the algorithm will converge to one of them. For the last experiment (Figure 3), we will try to separate the spare objects in lensfree fluorescent imaging [25] using EM-Type algorithm with a non-convex $J(x)$. Though EM (or Richardson-Lucy) method will separate the particles when the distance is large ($30\mu m$ and $20\mu m$), but it can not separate them when they are close to each other ($13\mu m$ and $9\mu m$ in this experiment). However, we can choose $J(x) = \sum_j |x_j|^p$ for $p \in (0, 1)$, and these two particles can be separated even when the distance

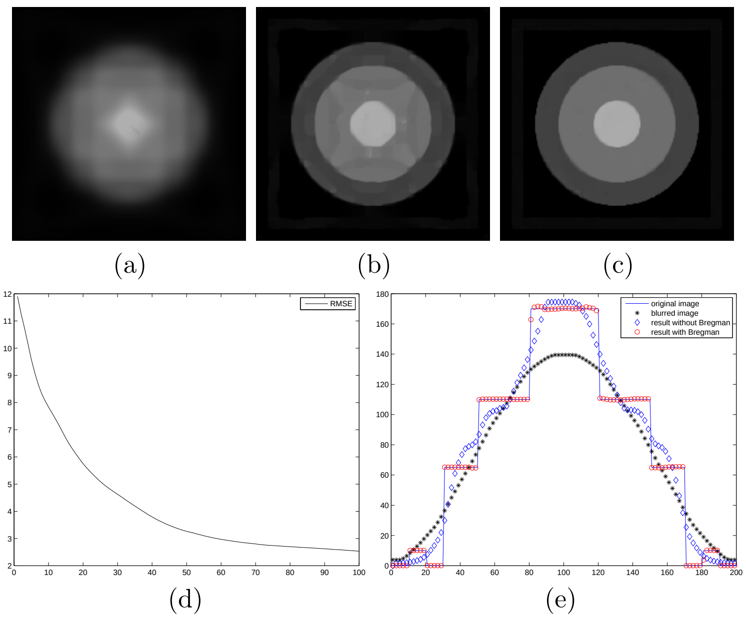

Fig. 1. (a) The result without Bregman iteration. (b) The result with 25 Bregman iterations. (c) The result with 100 Bregman iterations. (d) The plot of RMSE versus Bregman iterations. (e) The lineouts of original image, blurred image, the results with and without Bregman iterations. Some parameters chosen are $\beta = 0.001$, Num_inner $= 100$ and Num_outer $= 100$.

is very small. For the numerical experiment, top row show the lensfree raw images. As the distance between particles become smaller, their signatures become indistinguishable to the bare eye. The PSF is measured using small diameter fluorescent particles that are imaged at a low concentration. We choose the same numbers of iterations for EM-l_p and EM method, and the results show that with $p = 0.5$, we can obtain better results (them are separated).

6 Conclusion

In this paper, we proposed general robust EM-Type algorithms for image reconstruction when the measured data is corrupted by Poisson noise: iteratively performing EM and regularization in the image domain. If $J(x)$ is convex, the algorithm will converge to the global minimum of the object functional. However the problem with some regularizations such as total variation will lead to contrast reduction in the reconstructed images. Therefore, in order to improve the contrast, we suggested EM-Type algorithms with Bregman iteration by applying a sequence of modified EM-Type algorithms. To illustrate the effectiveness of these algorithms, we choose total variation (TV) regularization first, and EM-TV algorithm with Bregman iteration is applied to image deblurring. The results

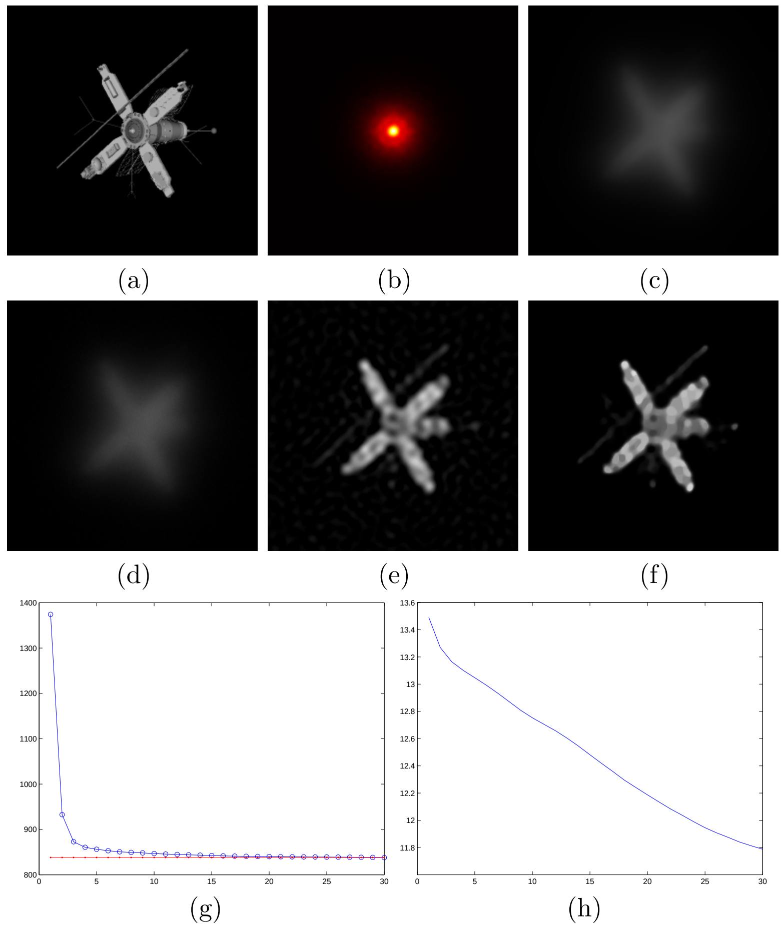

Fig. 2. (a) The original image. (b) The PSF image. (c) The blurred image. (d) The noisy blurred image. (e) Initial guess from CG. (f) The result of EM-Type algorithm with Bregman iteration. (g) The plot of KL versus Bregman iteration. (h) The RMSE versus Bregman iteration. Some parameters chosen are $\beta = 1$, Num_inner $= 200$ and Num_outer $= 30$.

show the performance of these algorithms. Also, when $J(x)$ is non-convex, this algorithm will also converge to a local minimum, and the numerical example in lensfree fluorescent imaging shows better results with l_p regularization ($p < 1$) than without regularization in separating two particles.

References

1. Shepp, L., Logan, B.: The Fourier reconstruction of a head section. IEEE Transaction on Nuclear Science **21** (1974) 21–34

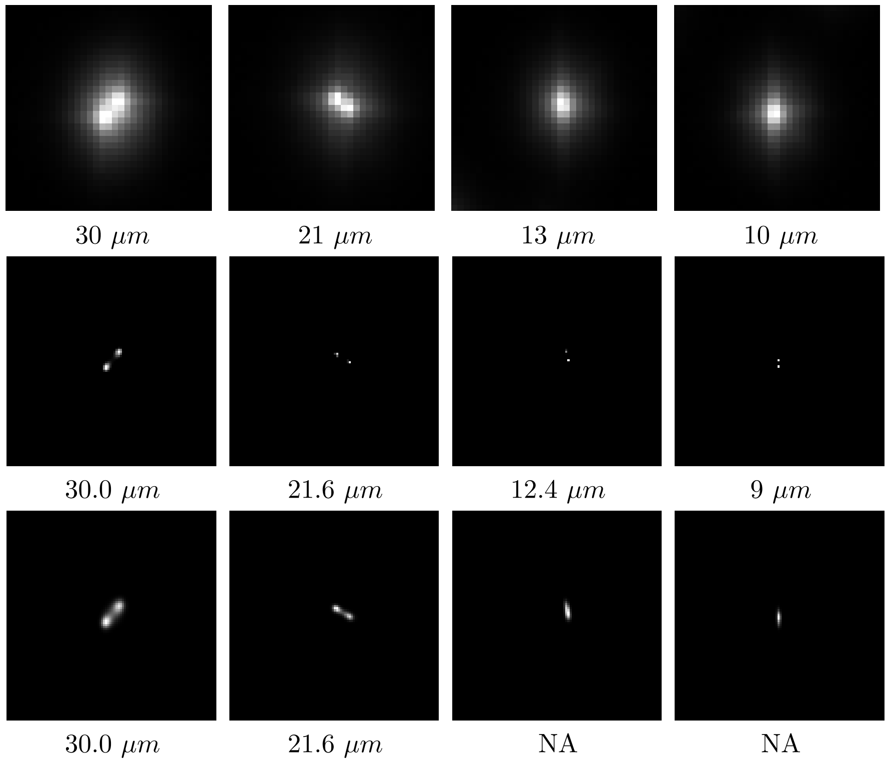

Fig. 3. Top row shows raw lensfree fluorescent images of different pairs of particles. The distances betweens these two particles are $30\mu m$, $21\mu m$, $13\mu m$ and $9\mu m$, from left to right. Middle row shows the results of EM-Type algorithm with $p = 0.5$, Bottom row shows the results for EM (or Richardson-Lucy) method.

2. Kak, A., Slaney, M.: Principles of Computerized Tomographic Imaging. Society of Industrial and Applied Mathematics (2001)
3. Brune, C., Sawatzky, A., Wubbeling, F., Kosters, T., Burger, M.: An analytical view on EM-TV based methods for inverse problems with Poisson noise. Preprint, University of Münster (2009)
4. Politte, D.G., Snyder, D.L.: Corrections for accidental coincidences and attenuation in maximum-likelihood image reconstruction for positron-emission tomography. IEEE Transaction on Medical Imaging **10** (1991) 82–89
5. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence **6** (1984) 721–741
6. Grenander, U.: Tutorial in pattern theory. Lecture Notes Volume, Division of Applied Mathematics, Brown University (1984)
7. Conchello, J.A., McNally, J.G.: Fast regularization technique for expectation maximization algorithm for optical sectioning microscopy. In: Proceeding of SPIE Symposium on Electronic Imaging Science and Technology. Volume 2655. (1996) 199–208

8. Markham, J., Conchello, J.A.: Fast maximum-likelihood image-restoration algorithms for three-dimensional fluorescence microscopy. Journal of the Optical Society America A **18** (2001) 1052–1071
9. Zhu, D., Razaz, M., Lee, R.: Adaptive penalty likelihood for reconstruction of multi-dimensional confocal microscopy images. Computerized Medical Imaging and Graphics **29** (2005) 319–331
10. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Phys D, **60** (1992) 259–268
11. Dey, N., Blanc-Feraud, L., Zimmer, C., Roux, P., Kam, Z., Olivo-Marin, J.C., Zerubia, J.: Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution. Microscopy Research and Technique **69** (2006) 260–266
12. Yan, M., Vese, L.A.: Expectation maximization and total variation based model for computed tomography reconstruction from undersampled data. In: Proceeding of SPIE Medical Imaging: Physics of Medical Imaging. Volume 7961. (2011) 79612X
13. Joshi, S., Miller, M.I.: Maximum a posteriori estimation with Good's roughness for optical sectioning microscopy. Journal of the Optical Society of America A **10** (1993) 1078–1085
14. Shepp, L., Vardi, Y.: Maximum likelihood reconstruction for emission tomography. IEEE Transaction on Medical Imaging **1** (1982) 113–122
15. Richardson, W.H.: Bayesian-based iterative method of image restoration. Journal of the Optical Society America **62** (1972) 55–59
16. Lucy, L.B.: An iterative technique for the rectification of observed distributions. Astronomical Journal **79** (1974) 745–754
17. Dempster, A., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society Series B **39** (1977) 1–38
18. Hurwitz, H.: Entropy reduction in Bayesian analysis of measurements. Physics Review A **12** (1975) 698–706
19. Levitan, E., Herman, G.T.: A maximum a posteriori probability expectation maximization algorithm for image reconstruction in emission tomography. IEEE Transactions on Medial Imaging **6** (1987) 185–192
20. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press (2004)
21. Meyer, Y.: Oscillating Patterns in Image Processing and in some Nonlinear Evolution Equations. American Mathematical Society (2001)
22. Bregman, L.: The relaxation method for finding common points of convex sets and its application to the solution of problems in convex programming. USSR Computational Mathematics and Mathematical Physics **7** (1967) 200–217
23. Osher, S., Mao, Y., Dong, B., Yin, W.: Fast linearized Bregman iteration for compressed sensing and sparse denoising. Communications in Mathematical Sciences **8** (2010) 93–111
24. Yin, W., Osher, S., Goldfarb, D., Darbon, J.: Bregman iterative algorithms for l1- minimization with applications to compressed sensing. Journal on Imaging Sciences **1** (2008) 143–168
25. Coskun, A.F., Sencan, I., Su, T.W., Ozcan, A.: Lensless wide field fluorescent imaging on a chip using compressive decoding of sparse objects. Optics Express **18** (2010) 10510–10523