

EM+TV Based Reconstruction for Cone-Beam CT with Reduced Radiation

Ming Yan¹, Jianwen Chen², Luminita A. Vese¹, John Villasenor², Alex Bui³,
and Jason Cong⁴

¹ Department of Mathematics, University of California, Los Angeles

² Department of Electrical Engineering, University of California, Los Angeles

³ Department of Radiological Sciences, University of California, Los Angeles

⁴ Department of Computer Science, University of California, Los Angeles
Los Angeles, CA 90095, United States

Abstract. Computerized tomography (CT) plays a critical role in modern medicine. However, the radiation associated with CT is significant. Methods that can enable CT imaging with less radiation exposure but without sacrificing image quality are therefore extremely important. This paper introduces a novel method for enabling image reconstruction at lower radiation exposure levels with convergence analysis. The method is based on the combination of expectation maximization (EM) and total variation (TV) regularization. While both EM and TV methods are known, their combination as described here is novel. We show that EM+TV can reconstruct a better image using much fewer views, thus reducing the overall dose of radiation. Numerical results show the efficiency of the EM+TV method in comparison to filtered backprojection and classic EM. In addition, the EM+TV algorithm is accelerated with GPU multicore technology, and the high performance speed-up makes the EM+TV algorithm feasible for future practical CT systems.

Keywords: expectation maximization, CT reconstruction, total variation, GPU medical image processing

1 Introduction

As a class of methods for reconstructing two-dimensional and three-dimensional images from the projections of an object, iterative reconstruction has many applications, including computerized tomography (CT), positron emission tomography (PET), and magnetic resonance imaging (MRI). Iterative reconstruction is quite different from the filtered back projection (FBP) method [1, 2], the algorithm most commonly used by manufacturers of commercial imaging equipment. The main advantages of iterative reconstruction over FBP are reduced sensitivity to noise and increased data collection flexibility [3]. For example, the data can be collected over any set of lines, the projections do not have to be distributed uniformly, and the projections can even be incomplete (limited angle).

There are many available algorithms for iterative reconstruction. Most of these algorithms are based on solving a system of linear equations $Ax = b$

where $x = (x_1, \dots, x_N)^T \in \mathbf{R}^N$ is the original unknown image represented as a vector; $b = (b_1, \dots, b_M)^T \in \mathbf{R}^M$ is the given measurement; A is a $M \times N$ matrix describing the direct transformation from the original image to the measurements. A depends on the imaging modality used; for example, in CT, A is the discrete Radon transform, with each row describing an integral along one straight line, and all the elements of A are nonnegative.

One example of iterative reconstruction uses an expectation maximization (EM) algorithm [4]. The noise can be presented in b as Poisson noise. Then if x is given and A is known, the conditional probability of b using Poisson distribution is $P(b|Ax) = \prod_i^M \frac{e^{-(Ax)_i} ((Ax)_i)^{b_i}}{b_i!}$. Given an initial guess x^0 , the EM iteration for $n = 0, \dots$, is as follows:

$$x_j^{n+1} = \frac{\sum_i (a_{ij} \left(\frac{b_i}{(Ax^n)_i}\right))}{\sum_i a_{ij}} x_j^n. \quad (1)$$

All the summation in i and j are from 1 to M and N , respectively.

The total-variation regularization method was originally proposed by Rudin, Osher and Fatemi [5] to remove noise in an image while preserving edges. This technique is widely used in image processing and can be expressed in terms of minimizing an energy functional of the form: $\min_x \int_{\Omega} |\nabla x| + \alpha \int_{\Omega} F(Ax, b)$, where x is viewed as a two- or three-dimensional image with spatial domain Ω , A is usually a blurring operator, b is the given noisy-blurry image, and $F(Ax, b)$ is a data-fidelity term. For example, for Gaussian noise, $F(Ax, b) = \|Ax - b\|_2^2$.

In the present paper we combine the EM algorithm with TV regularization. While each of these methods has been described individually in the literature, the combination of these two methods is new in CT reconstruction. The assumption is that the reconstructed image cannot have a large total-variation (thus noise and reconstruction artifacts are removed). For related relevant work, we refer to [6–10] and [11] for issues related to compressive sensing. A preliminary version of this work was presented in [10].

2 The Proposed Method (EM+TV)

In the classic EM algorithm, no prior information about the solution is provided. However, if we are given *a priori* knowledge that the solution has homogeneous regions and sharp edges, the objective is to apply this information to reconstruct an image with both minimal total-variation and maximal probability. Thus, we can consider finding a Pareto optimal point by solving a scalarization of these two objective functions, and the problem becomes:

$$\begin{cases} \underset{x}{\text{minimize}} & E(x) := \beta \int_{\Omega} |\nabla x| + \sum_i ((Ax)_i - b_i \log(Ax)_i), \\ \text{subject to} & x_j \geq 0, \quad j = 1, \dots, N, \end{cases} \quad (2)$$

where $\beta > 0$ is a parameter for balancing the two terms, TV and EM. This is a convex constraint problem, and we can find the optimal solution by solving the

Karush-Kuhn-Tucker (KKT) conditions [12]:

$$-\beta \operatorname{div} \left(\frac{\nabla x}{|\nabla x|} \right)_j + \sum_i \left(a_{ij} \left(1 - \frac{b_i}{(Ax)_i} \right) \right) - y_j = 0, \quad j = 1, \dots, N,$$

$$y_j \geq 0, \quad x_j \geq 0, \quad j = 1, \dots, N, \quad y^T x = 0.$$

By positivity of $\{x_j\}$, $\{y_j\}$ and the complementary slackness condition $y^T x = 0$, we have $x_j y_j = 0$ for all $j = 1, \dots, N$. After multiplying by x_j , we obtain:

$$-\beta \frac{x_j}{\sum_i a_{ij}} \operatorname{div} \left(\frac{\nabla x}{|\nabla x|} \right)_j + x_j - \frac{\sum_i \left(a_{ij} \left(\frac{b_i}{(Ax)_i} \right) \right)}{\sum_i a_{ij}} x_j = 0, \quad j = 1, \dots, N.$$

The last term on the left hand side is an EM step (1), which we can replace as x_j^{EM} , and we finally obtain:

$$-\beta \frac{x_j}{\sum_i a_{ij}} \operatorname{div} \left(\frac{\nabla x}{|\nabla x|} \right)_j + x_j - x_j^{EM} = 0, \quad j = 1, \dots, N.$$

To solve the above equation in x , with x_j^{EM} fixed from the previous step, we use a semi-implicit iterative scheme for several steps, alternated with the EM step. The algorithm is shown below, and the convergence of the proposed algorithm is shown in next section.

```

Input:  $x^0 = 1$ ;
for  $Out=1:IterMax$  do /* IterMax: number of outer iterations */
   $x^{0,0} = x^{Out-1}$ ;
  for  $k = 1:1:K$  do /* K: number of EMupdates */
    |  $x^{k,0} = EM(x^{k-1,0})$ ; /* Including one  $Ax$  and one  $A^T y$  */
  end
   $x^{Out} = TV(x^{K,0})$ ;
end

```

Algorithm 1: Proposed EM+TV algorithm.

3 Convergence Analysis of the Proposed Algorithm

We will show, in this section, that EM+TV algorithm is equivalent to an EM algorithm with *a priori* information and provide the convergence analysis of EM+TV algorithm. The EM algorithm is a general approach for maximizing a posterior distribution when some of the data is missing [13]. It is an iterative method that alternates between expectation (E) steps and maximization (M) steps. For image reconstruction, we assume that the missing data is $\{z_{ij}\}$, describing the intensity of pixel (or voxel) j observed by detector i . Therefore the

observed data are $b_i = \sum_j z_{ij}$. We can have the assumption that z is a realization of multi-value random variable Z , and for each (i, j) pair, z_{ij} follows a Poisson distribution with expected value $a_{ij}x_j$, because the summation of two Poisson distributed random variables also follows a Poisson distribution, whose expected value is summation of the two expected values.

The original E-step is to find the expectation of the log-likelihood given the present variables x^k :

$$Q(x|x^k) = E_{z|x^k, b} \log p(x, z|b).$$

Then, the M-step is to choose x^{k+1} to maximize the expected log-likelihood $Q(x|x^k)$ found in the E-step:

$$\begin{aligned} x^{k+1} &= \operatorname{argmax}_x E_{z|x^k, b} \log p(x, z|b) = \operatorname{argmax}_x E_{z|x^k, b} \log(p(b, z|x)p(x)) \\ &= \operatorname{argmax}_x E_{z|x^k, b} \sum_{i, j} (z_{ij} \log(a_{ij}x_j) - a_{ij}x_j) - \beta J(x) \\ &= \operatorname{argmin}_x \sum_{i, j} (a_{ij}x_j - E_{z|x^k, b} z_{ij} \log(a_{ij}x_j)) + \beta J(x). \end{aligned} \quad (3)$$

From (3), what we need before solving it is just $\{E_{z|x^k, b} z_{ij}\}$. Therefore we compute the expectation of missing data $\{z_{ij}\}$ given present x^k and the condition $b_i = \sum_j z_{ij}$, denoted this as an E-step. Because for fixed i , $\{z_{ij}\}$ are Poisson variables with mean $\{a_{ij}x_j^k\}$, then the distribution of z_{ij} , is binomial distribution $(b_i, a_{ij}x_j^k/(Ax^k)_i)$, thus we can find the expectation of z_{ij} with all these conditions by the following E-step

$$z_{ij}^{k+1} \equiv E_{z|x^k, b} z_{ij} = a_{ij}x_j^k b_i / (Ax^k)_i. \quad (4)$$

After obtaining the expectation for all z_{ij} , then we can solve the M-step (3).

We will show that EM-Type algorithms are exactly the described EM algorithms with *a priori* information. Recalling the definition of x^{EM} , we have

$$x_j^{EM} = \sum_i z_{ij}^{k+1} / \sum_i a_{ij}.$$

Therefore, M-step is equivalent to

$$\begin{aligned} x^{k+1} &= \operatorname{argmin}_x \sum_{i, j} (a_{ij}x_j - z_{ij}^{k+1} \log(a_{ij}x_j)) + \beta \int_{\Omega} |\nabla x^k| \\ &= \operatorname{argmin}_x \sum_j (\sum_i a_{ij})(x_j - x_j^{EM} \log(x_j)) + \beta \int_{\Omega} |\nabla x^k|. \end{aligned}$$

Then we will show, in the following theorem, that the log-likelihood is increasing.

Theorem 1. *The objective functional (negative log-likelihood) $E(x)$ in (2) with x^k given by **Algorithm 1** will decrease until it attains a minimum.*

Proof. For all k and i , we always have the constraint $\sum_j z_{ij}^k = b_i$. Therefore, we have the following inequality

$$\begin{aligned}
& b_i \log((Ax^{k+1})_i) - b_i \log((Ax^k)_i) \\
&= b_i \log\left(\sum_j \frac{a_{ij}x_j^{k+1}}{(Ax^k)_i}\right) = b_i \log\left(\sum_j \frac{z_{ij}^{k+1}a_{ij}x_j^{k+1}}{b_i a_{ij}x_j^k}\right) \\
&\geq b_i \sum_j \frac{z_{ij}^{k+1}}{b_i} \log\left(\frac{a_{ij}x_j^{k+1}}{a_{ij}x_j^k}\right) \quad (\text{Jensen's inequality}) \\
&= \sum_j z_{ij}^{k+1} \log(a_{ij}x_j^{k+1}) - \sum_j z_{ij}^{k+1} \log(a_{ij}x_j^k). \tag{5}
\end{aligned}$$

This inequality gives us

$$\begin{aligned}
E^p(x^{k+1}) - E^p(x^k) &= \sum_i ((Ax^{k+1})_i - b_i \log(Ax^{k+1})_i) + \int_{\Omega} |\nabla x^{k+1}| \\
&\quad - \sum_i ((Ax^k)_i - b_i \log(Ax^k)_i) - \beta \int_{\Omega} |\nabla x^k| \\
&\leq \sum_{i,j} (a_{ij}x_j^{k+1} - z_{ij}^{k+1} \log(a_{ij}x_j^{k+1})) + \int_{\Omega} |\nabla x^{k+1}| \\
&\quad - \sum_{i,j} (a_{ij}x_j^k - z_{ij}^{k+1} \log(a_{ij}x_j^k)) - \beta \int_{\Omega} |\nabla x^k| \leq 0.
\end{aligned}$$

The first inequality comes from (5) and the second inequality comes from the M-step (3). When $E(x^{k+1}) = E(x^k)$, these two equalities have to be satisfied. The first equality is satisfied if and only if $x_j^{k+1} = \alpha x_j^k$ for all j with α being a constant, while the second one is satisfied if and only if x^k and x^{k+1} are minimizers of the M-step (3). Since the functional to be minimized in M-step (3) is strictly convex, which means that α has to be 1 and

$$\beta x_j^k \partial J(x^k)_j + \sum_i a_{ij} x_j^k - \sum_i z_{ij}^{k+1} = 0, \quad j = 1, \dots, N.$$

After plugging the E-step (4) into these equations, we have

$$\beta x_j^k \partial J(x^k)_j + (Ax^k)_i - \sum_i \frac{a_{ij} x_j^k b_i}{(Ax^k)_i} = 0, \quad j = 1, \dots, N.$$

Therefore, x^k is the minimizer of the original problem. \square

The log-likelihood function will increase for each iteration until the solution is found, and from the proof, we do not fully use the M-step. Even if the M-step is not solved exactly, it will still increase so far as $Q(x^{k+1}|x^k) > Q(x^k|x^k)$ is satisfied.

4 GPU Implementation

In this section, we consider a fast graphics processing unit (GPU)-based implementation of the computationally challenging EM+TV algorithm. As forward projection (Ax) and backward projection ($A^T y$) are about 95% of the computational complexity of the entire algorithm, we focus on these two projections.

For forward projection, as illustrated in Fig. 1, for each source and detector pair, it is only necessary to calculate the approximate line integral, without updating the pixels. However, for backward projection, if ray tracing is used, there will be memory conflicts when it is parallelized: different threads may update the same pixel at the same time because for a given source-detector pair, and all the pixels intersecting the ray will be updated.

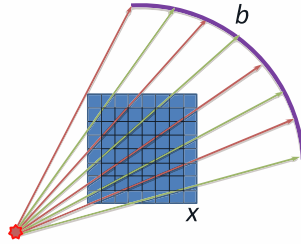


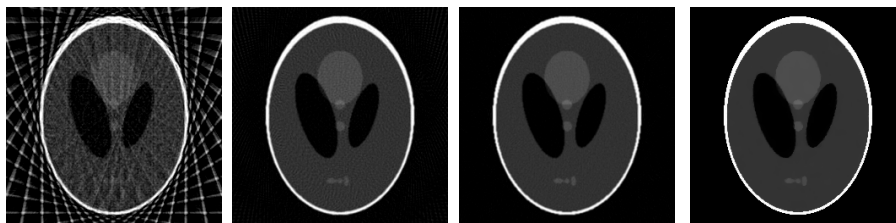
Fig. 1. Forward/Backward Projection

The forward projection can be parallelized on the GPU platform easily. A large number of threads will operate on the forward ray tracer simultaneously for different source and detector pairs. For backward projection, as there are dependencies and conflicts when two threads access one pixel, parallelization is possible, but more challenging. Compute unified device architecture (CUDA) provides atomic functions to guarantee the mutual exclusion of an address in memory, and can be used to handle such potential data conflicts. However the cost of using atomic functions is very large, therefore, in order to avoid the atomic operations and still can update all the pixels, new backward projection algorithms should be exploited. Here we propose our new backward projection algorithm in brief. For each view, we select the detectors that are far enough and set them to one group, mathematically there will be no conflicts within the group and all tracers in one group can be processed in parallel. As illustrated in Fig. 1, we can choose the tracer lines in the same color. In our case, we choose the distance between two adjoint detectors to be 6.

Finally, the EM+TV algorithm has been ported on the GPU platform. The medical image data is transferred from the host memory to device memory at the beginning of the routine; once the data is moved to device memory, the computation process is activated. And the reconstructed image is written back to the host memory when all the computation are finalized.

5 Experimental Results

In two dimensions, we compare the reconstruction results obtained by the proposed EM+TV method with those obtained by filtered back projection (FBP). For the numerical experiments, we choose the two dimensional Shepp-Logan phantom of dimension 256x256. The projections are obtained using Siddon's algorithm. The sinogram data is corrupted with Poisson noise. With the FBP method, we present results using 36 views (every 10 degrees), 180 views, and 360 views; for each view there are 301 measurements. In order to show that we can reduce the number of views by using EM+TV, we only use 36 views for the proposed method. The reconstruction results are shown in Figure 2. We notice the much improved results obtained with EM+TV using only 36 views (both visually and according to the root-mean-square-error (RMSE) between the original and reconstructed images, scaled between 0 and 255), by comparison with FBP using 36, 180 or even 360 views. Using the proposed EM+TV method, with only few samples we obtain sharp results and without artifacts.



FBP 36 (51.1003) FBP 180 (14.3698) FBP 360 (12.7039) EM+TV 36 (3.086)

Fig. 2. Reconstruction results by FBP with 36, 180, 360 views and EM+TV with 36 views (RMSE numbers are shown in parenthesis).

The EM+TV algorithm was also tested on the 3D 128x128x128 Shepp-Logan phantom. First, we obtained the projections using Siddon's algorithm [14]. Only 36 views were taken (every 10 degrees), and for each view there were 301x255 measurements. The code is implemented on the GPU platform (Tesla C1060) with a single-precision floating-point data type. The inner loop of EMupdate has three iterations and the EMupdate and TVupdate will repeat for 100 iterations. For forward projection, 512x64 blocks were used, and for each block there were 288 threads. For backward projection, 24 blocks are used and each block has 64x5x1 threads. Compared with the single-thread implementation on a CPU platform (Intel i7-920, 2.66GHz), implementation on the GPU provides more than 26x speed-up for forward projection, and 20x speed-up for backward projection, and the overall reconstruction time is about 330 seconds. The reconstructed image of the EM+TV algorithm on the GPU platform with RMSE is provided in Fig. 3, compared with the result of EM without regularization after 1000 iterations. We can see that the result of EM+TV with only 36 views delivers a very good quality compared with the EM method without TV regularization.

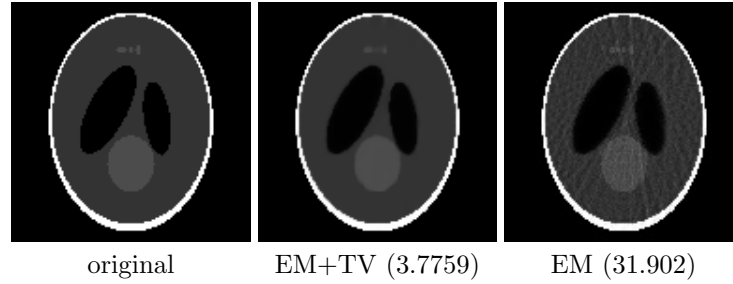


Fig. 3. Reconstructed images (RMSE numbers are shown in parenthesis).

Additionally, EM+TV was applied to a larger $256 \times 256 \times 256$ phantom, which has smaller features. Different numbers of views and choices of parameter β are chosen, and the results with RMSE are in Fig. 4. β is the parameter used to balance the EM and TV steps, if the parameter is too large, the result will be blurred, otherwise if the parameter is too small, the result will have artifacts. The problem of choosing the best β is under investigation.

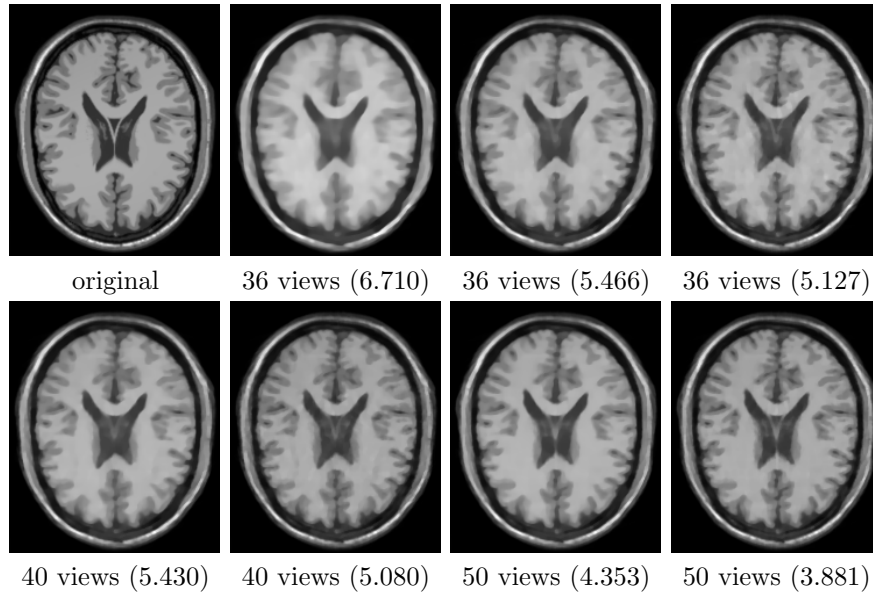


Fig. 4. Reconstructed data. From left to right, top to bottom: The middle slice of original image, result by 36 views with $1/\beta = 2.5, 10, 30$; 40 views with $1/\beta = 10, 30$; 50 views with $1/\beta = 10, 30$ (RMSE numbers are shown in parenthesis).

The numerical experiments above show the efficiency of EM+TV for CT reconstruction using fewer views. However these chosen views are equally spaced. The next experiment is to show that we can choose some special views and further reduce the number of views. This test is done on 2D phantom data, but can also be applied to 3D case. Instead of choosing 24 equally spaced views, we

can add six (or fewer) views to 18 equally spaced views (every 20 degrees). The six views are chosen at 70, 90, 110, 250, 270 and 290 degrees, and the results with corresponding RMSE are shown in Fig 5. From the results, we can see that, if the six chosen views are equally spaced, the result is worse than that of 24 equally spaced views. From the phantom, we can see that there are many small objects (features) along the middle vertical line. Therefore, we choose additional views from the top and bottom, which gives better results. The results with 22 views (three views have sources from up and one view has source from down (3U1D) and 1U3D) are even better than those with 24 equal views, both visually and from the RMSE calculation.

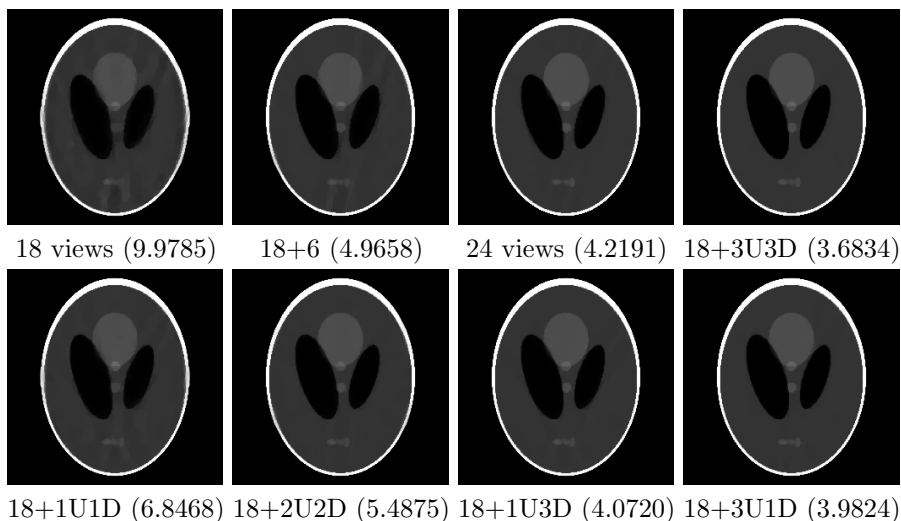


Fig. 5. Adaptive study of the views. From left to right, top to bottom: reconstruction results from 18 equal views, 6 equal views added onto 18 equal views, 24 equal views, 6 special views added onto 18 views, 2 views (90, 270) added, 4 views (70, 110, 250, 290) added, 4 views (90, 250, 270, 290) added, 4 views (70, 90, 110, 270) added (RMSE numbers are shown in parenthesis).

6 Conclusion

In this paper, a method that combines EM and TV for CT image reconstruction is proposed. This method can provide very good results using much fewer number of views. It requires fewer measurements to obtain a comparable image, which results in significant decrease of the radiation dose. The method is extended to three dimensions and can be used for real data. One of the challenges in EM+TV is long computation time. We have demonstrated that by suitable parallel algorithm design and efficient implementation EM+TV on a GPU platform, execution time can be reduced by well over an order of magnitude. In addition, we believe there are opportunities for further optimizations in

areas such as memory access, instruction flow, and parallelization of the backward algorithm that can further improve execution time. We believe that, as demonstrated, the combination of algorithms and optimized implementation on appropriate platforms has the potential to enable high-quality image reconstruction with reduced radiation exposure, while also enabling relatively fast image reconstruction times.

Acknowledge

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

References

1. Shepp, L., Logan, B.: The Fourier reconstruction of a head section. *IEEE Transaction on Nuclear Science* **21** (1974) 21–34
2. Pan, X., Sidky, E., Vannier, M.: Why do commercial CT scanners still employ traditional filtered back-projection for image reconstruction? *Inverse Problems* **25** (2009) 123009
3. Kak, A., Slaney, M.: *Principles of Computerized Tomographic Imaging*. Society of Industrial and Applied Mathematics (2001)
4. Shepp, L., Vardi, Y.: Maximum likelihood reconstruction for emission tomography. *IEEE Transaction on Medical Imaging* **1** (1982) 113–122
5. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Phys D*, **60** (1992) 259–268
6. Sidky, E., Pan, X.: Image reconstruction in circular cone-beam computed tomography by total variation minimization. *Physics in Medicine and Biology* **53** (2008) 4777–4807
7. 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)
8. Jia, X., Lou, Y., Li, R., Song, W., Jiang, S.: GPU-based fast cone beam CT reconstruction from undersampled and noisy projection data via total variation. *Medical Physics* **37** (2010) 1757–1760
9. 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
10. Chen, J., Yan, M., Vese, L.A., Villasenor, J., Bui, A., Cong, J.: EM+TV for reconstruction of cone-beam CT with curved detectors using GPU. In: *Proceedings of International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*. (2011) 363–366
11. *Compressive Sensing Resources*: (<http://dsp.rice.edu/cs>)
12. Boyd, S., Vandenberghe, L.: *Convex Optimization*. Cambridge University Press (2004)
13. 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
14. Siddon, R.: Fast calculation of the exact radiological path for a three-dimensional CT array. *Medical Physics* **12** (1986) 252–255