X-Ray CT Image Reconstruction via Wavelet Frame Based Regularization and Radon Domain Inpainting

Bin Dong * Jia Li † Zuowei Shen ‡

December 22, 2011

Abstract

X-ray computed tomography (CT) has been playing an important role in diagnostic of cancer and radiotherapy. However, high imaging dose added to healthy organs during CT scans is a serious clinical concern. Imaging dose in CT scans can be reduced by reducing the number of X-ray projections. In this paper, we consider 2D CT reconstructions using very small number of projections. Some regularization based reconstruction methods have already been proposed in the literature for such task, like the total variation (TV) based reconstruction [1–4] and balanced approach with wavelet frame based regularization [5]. For most of the existing methods, at least 40 projections is usually needed to get a satisfactory reconstruction. In order to keep radiation dose as minimal as possible, while increase the quality of the reconstructed images, one needs to enhance the resolution of the projected image in the Radon domain without increasing the total number of projections. The goal of this paper is to propose a CT reconstruction model with wavelet frame based regularization and Radon domain inpainting. The proposed model simultaneously reconstructs a high quality image and its corresponding high resolution measurements in Radon domain. In addition, we discovered that using the isotropic wavelet frame regularization proposed in [6] is superior than using its anisotropic counterpart. Our proposed model, as well as other models presented in this paper, is solved rather efficiently by split Bregman algorithm [7,8]. Numerical simulations and comparisons will be presented at the end.

Keywords: Computed tomography, wavelet frame, total variation, split Bregman algorithm, Radon domain inpainting

*Bin Dong (dongbin@math.arizona.edu) is with the Department of Mathematics, The University of Arizona, 617 Santa Rita Ave., Tucson, AZ, 85721-0089.

†Jia Li (lijia@nus.edu.sg) is with the Department of Mathematics, National University of Singapore, Singapore, 117542.

‡Zuowei Shen (matzuows@nus.edu.sg) is with the Department of Mathematics, National University of Singapore, Singapore, 117542.

1 Introduction

X-ray computed tomography (CT), e.g. cone-beam or regular CT, has been widely used for cancer detection and radiation therapy, among many other applications. However, the major clinical concern for CT is the radiation dose imposed to the patients during imaging procedure (see e.g. [9]). The imaging dose can be controlled by reducing the number of projections or decreasing the tube current and pulse duration. However, all these methods of restricting imaging dose lead to degradations of image quality due to insufficiency of information in the data collected. The filtered back-projection (FBP) algorithm [10], which is a conventional algorithm that has been widely used, cannot produce high quality image reconstruction if the information collected by the machine is not sufficient enough (e.g. when small number of projections are used). As a result, it is necessary to design new methods to reconstruct high quality CT images from limited and noisy projection data.

One of the common CT systems is the cone-beam CT system. In 2 dimensional cases, it is known as the fan-beam CT, which will be the scanning geometry that we focus in this paper. For simplicity, we assume the the source rotates the object following a circle with a fixed radius by 360 degrees.

For a given angle θ and X-ray beamlet r, the X-ray projection operator $P^{\theta,r}$ is defined as follows:

$$P^{\theta,r}[u](t) = \int_0^{L(t)} u(\mathbf{x}_\theta + \mathbf{n}l)dl. \tag{1}$$

where u is the unknown image (X-ray attenuation coefficients) that needs to be reconstructed, $\mathbf{x}_\theta = (x_\theta, y_\theta)$ represents the coordinate of the X-ray source which is different for different projection angle θ, $\mathbf{n} = (n_x, n_y)$ is the direction vector of beamlet r, t is the coordinate on the X-ray imager which is precisely the intersection of the beamlet r with the X-ray imager, $L(t)$ is the length of the X-ray beamlet from the source to the location t on the imager. Now if $P^{\theta,r}[u](t)$ is sampled with respect to t for each angle θ, the resulting data projection can essentially be written as a vector f_θ. Now putting the vectors f_θ together for all different angles θ, we obtain an image denoted as f whose columns are formed by f_θ. We always assume that the increments of angles are equal.

We can now write the CT image reconstruction problem as a linear inverse problem

$$Pu = f, \tag{2}$$

where P is the linear operator represents the collection of discrete line integrations at different projection angles and along different beamlets. In other words, the CT image reconstruction problem is to recover image u from its partial Radon transform [11]. Notice that the matrix P only depends on the location and direction of each beamlet. Therefore, we can construct the huge sparse matrix P beforehand. In our simulations, the matrix P is generated by Siddon's algorithm [12].

In order to reduce the imaging dose, we can simply reduce the total number of projections. Since each row of the matrix P corresponds to each beamlet

used to acquire the projected image data, reducing the total number of projections can make the matrix P to be under-determined. Consequently, the linear system $Pu = f$ will have infinitely many solutions. There are some direct methods solving the linear system, e.g. the FBP algorithm [10] and algebraic reconstruction techniques (ART) [13]. However, solutions obtained from such methods are usually lack of regularity and contain artifacts, especially when the total number of projections is small.

In practice, the projected image f usually contains noise. Therefore, it is generally not a good idea to directly solve the system $Pu = f$ or the corresponding least square problem. To suppress noise and artifacts while maintaining quality of the reconstructed images, various differential operator based regularization methods, known as variational models, have been proposed in the literature, among which the total variation (TV) based regularization is one of the popular models and is proven to be effective, especially for images that are piecewise constant. TV-based image regularization model (known as the ROF model) was first proposed by [14] in the context of image denoising, and it was later extended and applied to other image processing and analysis tasks (see [15–19] and the references therein). TV-based regularization model was first applied to cone-beam CT image reconstruction in [1, 2], and later in [3] a GPU based implementation was proposed that greatly speeds up the computation efficiency of TV-based regularization. TV-based model was also applied to equally-slopped tomography in [20]. One of the standard forms for TV-based image reconstruction model can be written as follows

$$\min_u \quad \frac{1}{2}\|Pu - f\|_2^2 + \lambda\|\nabla u\|_1. \tag{3}$$

where f is the projected image, ∇ is the gradient operator and P is the projection operator.

Another regularization based image reconstruction technique is the wavelet frame based approach. The basic idea for wavelet frame based approaches is that images can be sparsely approximated by properly designed wavelet frames, and hence, the regularization used for wavelet frame based models is the ℓ_1-norm of frame coefficients. Although wavelet frame based approaches take similar forms as variational models (e.g. TV-based model (3)), they were generally considered as different approaches. Such impression was changed by the recent paper [6], where the authors established a rigorous connection between one of the wavelet frame based approaches, namely the analysis based approach, and variational models. It was shown in [6] that the analysis based approach can be regarded as a finite difference approximation of a certain type of general variational model, and such approximation will be exact when image resolution goes to infinity. Furthermore, the solutions of the analysis based approach also approximates, in some proper sense, the solutions of corresponding variational model. Such connections not only grant geometric interpretation to wavelet frame based approaches, but also lead to even wider applications of them, e.g. image segmentation [21] and 3D surface reconstruction from unorganized point sets [22]. On the other hand, the discretizations provided by wavelet frames

were shown, in e.g. [6,8,23–26], to be superior than the standard discretizations for TV-based model (3), due to the multiresolution structure and redundancy of wavelet frames which enable wavelet frame based models to adaptively choose a proper differential operators in different regions of a given image according to the order of the singularity of the underlying solutions.

For these reasons, as well as the fact that images data are always discrete, we use wavelet frames as the tool for CT image reconstruction. We note that wavelet frame based regularization was first applied for cone-beam CT image reconstruction in [5]. The model in that paper is essentially balanced approach (see [23,27]). In this paper, we propose a new model that uses the analysis based approach of wavelet frame method instead, because it is more effective in terms of removing artifacts and keeping key features. More importantly, the model proposed here automatically builds in a Radon domain inpainting mechanism which enables us to reconstruct high quality images with very small number of projections.

Notice that all the above-mentioned work treats the projected image f as given and try to recover a desirable u from the given f. The challenge here is when only few number of projections are used, the information contained in f is far from being enough. All regularization based methods assume certain prior knowledge on the desired recovery u which, to some extent, overcomes the lack of information in f. However, the prior knowledge on the projected image f has yet not been utilized. We observe that when measurements are sufficient, the projected image f should be piecewise smooth and hence can be sparsely approximated by wavelet frames. Therefore, motivated by wavelet frame based image super-resolution models [23, 27–29], we will propose a model that simultaneously reconstruct the CT image and its corresponding higher resolution projected image by penalizing ℓ_1-norm of their wavelet frame coefficients. Note that the attempt of increasing image resolution can be regarded as an image inpainting problem [30, 31]. Our numerical simulations show that the recovered images u from our proposed model are of higher quality than the images recovered by TV-based model and analysis based approach without Radon domain inpainting. These experiments indicate that utilizing the prior knowledge of the projected image f, in addition to the prior knowledge of u, can further improve the quality of the reconstructed CT image.

The rest of the paper is organized as follows. In Section 2 we will present our wavelet frame based CT image reconstruction model, together with an efficient algorithm. Numerical simulations will be provided in Section 3 and concluding remarks will be given in Section 4.

2 Models and Algorithms

2.1 Tight Wavelet Frames

We now briefly introduce the concept of tight frames and framelets. Interesting readers should consult [32–34] for theories of frames and framelets, [35]

for a short survey on theory and applications of frames, and [26] for a more detailed survey.

A countable set $X \subset L_2(\mathbb{R})$ is called a tight frame of $L_2(\mathbb{R})$ if

$$f = \sum_{h \in X} \langle f, h \rangle h \quad \forall f \in L_2(\mathbb{R}),$$

where $\langle \cdot, \cdot \rangle$ is the inner product of $L_2(\mathbb{R})$. The tight frame X is called a tight wavelet frame if the elements of X is generated by dilations and translations of finitely many functions called framelets. The construction of framelets can be obtained by the unitary extension principle (UEP) of [33]. In our implementations, we will use the piecewise linear B-spline framelets constructed by [33]. Given a 1-dimensional framelet system for $L_2(\mathbb{R})$, the s-dimensional tight wavelet frame system for $L_2(\mathbb{R}^s)$ can be easily constructed by using tensor products of 1-dimensional framelets (see e.g. [26, 32]).

In the discrete setting, a discrete image u is an s-dimensional array. We will use W to denote fast tensor product framelet decomposition and use $W^\top$ to denote the fast reconstruction. Then by the unitary extension principle [33], we have $W^\top W = I$, i.e. $u = W^\top W u$ for any image u. We will further denote an L-level framelet decomposition of u as

$$Wu = \{W_{l,i,j}u : 1 \leq l \leq L, (i,j) \in I\},$$

where I denotes the index set of all framelet bands. More details on discrete algorithms of framelet transforms can be found in [26].

2.2 Wavelet Frame Based CT Image Reconstruction Models

We start with some simple notations. We denote f_0 as the observed projected image defined on grid Λ of size $N_m \times N_p$ where N_m is the number of grid points on the X-ray imager and N_p is the number of projections used to acquire f_0. Denote the higher resolution projected image as f defined on the grid $\Omega \supset \Lambda$. From a practical concern, we will only consider inpainting with respect to projection angles (See Figure 1). Therefore, as a typical configuration, Ω is an $N_m \times 2N_p$ grid whose even columns are the columns of Λ. Hence, the number of projections for the operator P that corresponds to f is $2N_p$. Let R_Λ be the restriction operator defined as $(R_\Lambda v)[i,j] := v[i,j]$ for $(i,j) \in \Lambda$ and zero elsewhere. Then the constraint f should satisfy is $R_\Lambda f = f_0$, meaning the recovered high resolution projected image should be consistent with the observed image f_0. In our model, however, we will not enforce such constraint to be exactly satisfied since f_0 always contains noise. We note that one can also replace R_Λ by some other operators that enforce data consistency (see e.g. [36, 37]).

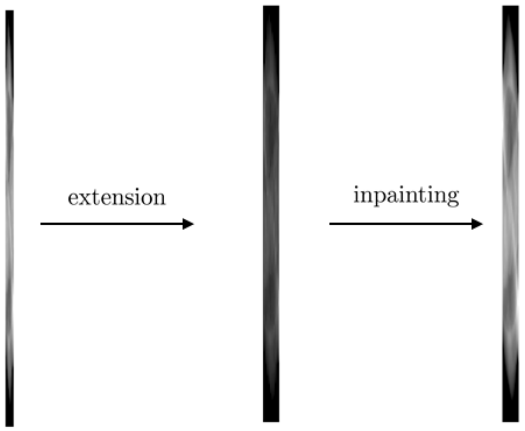

Figure 1: The strategy of inpainting in Radon domain.

Now, we propose our CT image reconstruction model as follows

$$\begin{aligned}\min_{f,u}\ \frac{1}{2}\|R_{(\Omega\setminus\Lambda)}(Pu-f)\|_2^2+\lambda_1\|W_1 f\|_{1,p}+\lambda_2\|W_2 u\|_{1,p}\\ +\frac{\kappa}{2}\|R_\Lambda f-f_0\|_2^2+\frac{1}{2}\|R_\Lambda(Pu)-f_0\|_2^2,\end{aligned}\tag{4}$$

where the norm $\|\cdot\|_{1,p}$ is defined as

$$\|Wu\|_{1,p}=\left\|\sum_{l=1}^{L}\Big(\sum_{(i,j)\neq(0,0)}|W_{l,i,j}u|^p\Big)^{\frac{1}{p}}\right\|_1,\quad p=1,2.\tag{5}$$

When $p=1$, we shall refer to the norm $\|\cdot\|_{1,1}$ as the *anisotropic* ℓ_1-norm of the frame coefficients, which is the standard ℓ_1-norm used for frame based image restoration problems. When $p=2$, we shall refer to the norm $\|\cdot\|_{1,2}$ as the *isotropic* ℓ_1-norm of the frame coefficients, which was proposed in [6]. It was shown [6] that for image restoration problems, isotropic ℓ_1-norm outperforms anisotropic ℓ_1-norm for analysis based approach in terms of both quality of the restoration and efficiency of the corresponding numerical algorithm. In this paper, we will show that for CT image reconstruction, isotropic ℓ_1-norm is also superior than anisotropic ℓ_1-norm.

In model (4), W_1 and W_2 denote two different tight wavelet frame transform. In our simulations, we will use cubic B-spline framelet system for W_1 with 3 levels of decomposition, and linear B-spline framelet system for W_2 with 1 level of decomposition.

The first term $\frac{1}{2}\|R_{(\Omega\setminus\Lambda)}(Pu-f)\|_2^2$ together with the last term $\frac{1}{2}\|R_\Lambda(Pu)-f_0\|_2^2$ serve as the data fidelity terms for our model. The reason that we are

not using the simpler fidelity term $\frac{1}{2}\|(Pu-f)\|_2^2$ is because f is the estimated projection data which is generally not as reliable as f_0. Therefore, in the domain Λ where the actual projected image f_0 is available, we should make sure that $Pu \approx f_0$ on Λ. The term $\frac{1}{2}\|R_\Lambda f - f_0\|_2^2$ makes sure that the recovered higher resolution projected image f is consistent with f_0 on Λ. The terms $\lambda_1\|W_1 u\|_{1,p}$ and $\lambda_2\|W_2 f\|_{1,p}$ are regularization terms which guarantee that the reconstructed images u and f are piecewise smooth.

We observe that the model (4) is general and includes the analysis based approach for CT reconstruction without Radon domain inpainting as a special case. Indeed, if we assume $\Omega = \Lambda$ and $\kappa = \infty$, then we can rewrite (4) equivalently as

$$\min_u \quad \frac{1}{2}\|Pu - f\|_2^2 + \lambda\|Wu\|_{1,p}, \tag{6}$$

where we simply use W to denote W_2 in (4). When $p = 1$ in (6), model (6) is the analysis based approach proposed for general image restoration problems [8, 38, 39]. Throughout the rest of this paper, we shall refer to model (6) with $p = 1$ as anisotropic wavelet frame based model; and refer to model (6) with $p = 2$ as isotropic wavelet frame based model.

2.3 Alternative Optimization Algorithms

To solve the model (4), we propose an alternative optimization algorithm summarized in Algorithm 1.

Algorithm 1 Wavelet Frame Based CT Image Reconstruction

Step 0. Solve model (6) to obtain an initial reconstruction u^0.
while stopping criteria is not met **do**
Step 1. Solve

$$f^{k+1} := \arg\min_f \ \frac{1}{2}\|R_{(\Omega\setminus\Lambda)}(Pu^k - f)\|_2^2 + \lambda_1\|W_1 f\|_1 + \frac{\kappa}{2}\|R_\Lambda f - f_0\|_2^2.$$

Step 2. Solve

$$u^{k+1} := \arg\min_u \ \frac{1}{2}\|R_{(\Omega\setminus\Lambda)}(Pu - f^{k+1})\|_2^2 + \lambda_2\|W_2 u\|_1 + \frac{1}{2}\|R_\Lambda(Pu) - f_0\|_2^2.$$

end while

To solve model (6) as well as step 1 and 2 in Algorithm 1, we shall use the split Bregman algorithm. The split Bregman algorithm was first proposed in [7] which was shown to be convergent and powerful in [7, 40] when it is applied to various variational models used for image restoration, e.g., ROF [14] and nonlocal variational models [41]. Convergence analysis of the split Bregman was given in [8]. Here we briefly review the basic ideas of split Bregman algorithm. Interested readers can refer to [7, 8] for more details.

Consider the following minimization problem

$$\min_{u} \quad E(u) + \lambda \|Wu\|_{1,p}, \tag{7}$$

where $E(u)$ is a smooth convex functional. Letting $d = Wu$, then (7) can be rewritten as

$$\min_{u, d=Wu} \quad E(u) + \lambda \|d\|_{1,p}. \tag{8}$$

Note that both u and d are variables now. The derivation of splitting Bregman iteration for solving (8) is based on Bregman distance ([7, 8]). It was recently shown (see e.g. [42,43]) that the split Bregman algorithm can also be derived by applying augmented Lagrangian method (see e.g. [44]) to (8). The connection between split Bregman algorithm and Douglas-Rachford splitting was addressed by [45]. We shall skip the detailed derivations and directly describe the split Bregman algorithm that solves (7) through (8) as follows,

$$\begin{cases} u^{k+1} = \arg\min_u E(u) + \frac{\mu}{2}\|Wu - d^k + b^k\|_2^2, \\ d^{k+1} = \arg\min_d \lambda\|d\|_{1,p} + \frac{\mu}{2}\|d - Wu^{k+1} - b^k\|_2^2, \\ b^{k+1} = b^k + Wu^{k+1} - d^{k+1}. \end{cases} \tag{9}$$

By [46–48], the second subproblem has a simple analytical solution based on soft-thresholding operator. Therefore, (9) can be written equivalently as

$$\begin{cases} u^{k+1} = \arg\min_u E(u) + \frac{\mu}{2}\|Wu - d^k + b^k\|_2^2, \\ d^{k+1} = \mathcal{T}^p_{\lambda/\mu}(Wu^{k+1} + b^k), \\ b^{k+1} = b^k + (Wu^{k+1} - d^{k+1}), \end{cases} \tag{10}$$

where $\mathcal{T}^p_t$ is the soft-thresholding operator defined by

$$(\mathcal{T}^p_t(v))_{l,i,j} = \begin{cases} v_{l,i,j}, & \text{if } (i,j) = (0,0) \\ \frac{v_{l,i,j}}{|v_{l,i,j}|}\max(|v_{l,i,j}| - t, 0), & \text{if } (i,j) \neq (0,0), p = 1 \\ \frac{v_{l,i,j}}{R_l}\max(R_l - t, 0), & \text{if } (i,j) \neq (0,0), p = 2 \end{cases} \tag{11}$$

with $R_l = (\sum_{(i,j)\neq(0,0)} |v_{l,i,j}|^2)^{\frac{1}{2}}$.

Note that the last two steps of (10) are simple and computationally efficient, while computation costs for the first step is usually more expensive as it involves solving some linear system. In our simulations, we use conjugate gradient method to solve such linear system.

By the split Bregman algorithm (9), we solve Step 1 of Algorithm 1 as follows:

$$\begin{cases} f^{k+1} := (R_{(\Omega\setminus\Lambda)} + \mu_1 I + \kappa R_\Lambda)^{-1}(R_{(\Omega\setminus\Lambda)}(Pu^k) + \mu_1 W_1^\top(d_1^k - b_1^k) + \kappa R_\Lambda f_0), \\ d_1^{k+1} := \mathcal{T}^p_{\lambda_1/\mu_1}(W_1 f^{k+1} + b_1^k), \\ b_1^{k+1} := b_1^k + (W_1 f^{k+1} - d_1^{k+1}). \end{cases} \tag{12}$$

Note that if $\kappa = 1$, in the first equation we have $R_{(\Omega\setminus\Lambda)} + R_\Lambda = I$. The stopping criteria is posed as $\|d_1^k - W_1 u^k\| \leq \epsilon_f$ with ϵ_f being a given tolerance. Conjugate gradient method is used to solve the linear system in the first step.

The algorithm for Step 0 and Step 2 of Algorithm 1 is given as follows: (The algorithms for Step 0 and Step 2 are almost the same except the operator P corresponds to different number of projections and we have $\Omega = \Lambda$ for Step 0.)

$$\begin{cases} u^{k+1} := (P^\top P + \mu_2 I)^{-1}(P^\top (R_{(\Omega\setminus\Lambda)f} + f_0) + \mu_2 W_2^\top (d_2^k - b_2^k)), \\ d_2^{k+1} := \mathcal{T}^p_{\lambda_2/\mu_2}(W_2 u^{k+1} + b_2^k), \\ b_2^{k+1} := b_2^k + (W_2 u^{k+1} - d_2^{k+1}). \end{cases} \tag{13}$$

The stopping criteria is posed as $\|d_2^k - W_2 u^k\| \leq \epsilon_u$ with ϵ_u being a given tolerance. Conjugate gradient method is used to solve the linear system in the first step.

We want to remark that at each exterior iteration of Algorithm 1, we initialize Step 1 using the estimated projection image f, as well as d_1 and b_1 calculated from a previous exterior iteration, except for the very first iteration, we choose $f^0 = 0$ and $d_1^0 = b_1^0 = 0$. Similarly, we initialize Step 2 using the recovered CT image u, as well as d_2 and b_2 calculated from a previous exterior iteration, except for the very first iteration, we choose u^0, d^0 and b^0 from Step 0 as initial guesses.

2.4 Convergence Analysis

Convergence of alternative optimization methods, also called coordinate descent methods, have been well-studied in the literature [49–53]. In particular, [52, Theorem 4.1] can be directly applied to prove the convergence of Algorithm 1. On the other hand, each of the subproblem of Algorithm 1 is solved by the split Bregman algorithm and its convergence is studied by [8, 54, 55]. However, since the split Bregman algorithm is an iterative algorithm, the subproblems of Algorithm 1 cannot be solved exactly in practice. In that case, [52, Theorem 4.1] is no longer applicable to ensure convergence. In this section, we provide a convergence analysis of Algorithm 1 when each of the subproblem is solved inexactly with some error. We will show that if the errors decay fast enough with respect to k, then the sequence $\{(u^k, f^k)\}_k$ is bounded and any cluster point is a solution to (4).

Let $F(u, f) : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$ be the objective function of (4), i.e.

$$\begin{aligned} F(u,f) = \frac{1}{2}\|R_{(\Omega\setminus\Lambda)}(Pu - f)\|_2^2 + \lambda_1 \|W_1 f\|_{1,p} + \lambda_2 \|W_2 u\|_{1,p} \\ + \frac{\kappa}{2}\|R_\Lambda f - f_0\|_2^2 + \frac{1}{2}\|R_\Lambda(Pu) - f_0\|_2^2. \end{aligned} \tag{14}$$

All the parameters λ_1, λ_2 and κ are positive constants. It is clear that $F(u, f)$ is convex and continuous. We say that $(\bar{u}, \bar{f})$ is a coordinatewise minimizer of

$F(u, f)$ if

$$F(\bar{u}, \bar{f}) \leq F(\bar{u} + v_u, \bar{f}) \quad \text{and} \quad F(\bar{u}, \bar{f}) \leq F(\bar{u}, \bar{f} + v_f) \quad \forall v_u \in \mathbb{R}^n, \forall v_f \in \mathbb{R}^m.$$

In general, a coordinatewise minimizer is not necessarily a global minimizer. However, we will show in the following lemma that for our $F(u, f)$ defined by (14), any coordinatewise minimizer is also a global minimizer. Note that this lemma can also be derived using [52, Lemma 3.1]. For completeness of this paper, we include a direct proof of this lemma. We first recall that the subgradient of a convex subdifferentiable function $G : \mathbb{R}^N \mapsto \mathbb{R}$ at $x \in \mathbb{R}^N$, denoted as $(\partial G)(x)$, is defined by the following inequality

$$G(x + y) \geq G(x) + \langle y, (\partial G)(x) \rangle.$$

In general, ∂G is a set belonging to $\mathbb{R}^N$. By Fermat's rule (see e.g. [56, Theorem 10.1]), $\bar{x}$ is a global minimizer of the convex subdifferentiable function G if and only if $0 \in (\partial G)(\bar{x})$.

Lemma 2.1. *Any coordinatewise minimizer of the objective function $F(u, f)$ defined by* (14) *is also a global minimizer.*

Proof. First, we denote $F(u, f) := F_1(u, f) + F_2(u) + F_3(f)$ with

$$F_1(u, f) := \frac{1}{2}\|R_{(\Omega\setminus\Lambda)}(Pu - f)\|_2^2, \quad F_2(u) := \lambda_2\|W_2 u\|_{1,p} + \frac{1}{2}\|R_\Lambda(Pu) - f_0\|_2^2$$

and

$$F_3(f) := \lambda_1\|W_1 f\|_{1,p} + \frac{\kappa}{2}\|R_\Lambda f - f_0\|_2^2.$$

It is clear that F_1 is differentiable, and F_2 and F_3 are subdifferentiable. Suppose $(\bar{u}, \bar{f})$ is a coordinatewise minimizer of F. Then, by Fermat's rule, we have

$$0 \in (\partial_u F)(\bar{u}, \bar{f}) \quad \text{and} \quad 0 \in (\partial_f F)(\bar{u}, \bar{f}). \tag{15}$$

Since F_1, the only term of F where u and f are not separable, is differentiable, F_2 only depends on u and F_3 only depends on f, we thus have

$$\partial F = \partial F_1 + \partial F_2 + \partial F_3 = (\partial_u F_1, \partial_f F_1) + (\partial_u F_2, 0) + (0, \partial_f F_3) = (\partial_u F, \partial_f F).$$

In other words, ∂F is set formed by the tensor of $\partial_u F$ and $\partial_f F$. Therefore, it is obvious from (15) that $0 \in (\partial F)(\bar{u}, \bar{f})$. □

Suppose each subproblem of Algorithm 1 is not solved exactly and at each iteration we introduce errors $\varepsilon_k > 0$ and $\delta_k > 0$ defined as

$$\begin{aligned} F(u^{k+1}, f^k) &\leq F(\tilde{u}^{k+1}, f^k) + \varepsilon_k \\ F(u^{k+1}, f^{k+1}) &\leq F(u^{k+1}, \tilde{f}^{k+1}) + \delta_k, \end{aligned} \tag{16}$$

where $\tilde{u}^{k+1} = \arg\min_u F(u, f^k)$ and $\tilde{f}^{k+1} = \arg\min_f F(u^{k+1}, f)$. Now, we analyze the convergence of the sequence $\{(u^k, f^k)\}_k$ defined by (16). We have the following convergence theorem.

Theorem 2.2. *Assume that* $\sum_{j=0}^{\infty}(\varepsilon_k+\delta_k)<\infty$. *Then the sequence* $\{(u^k,f^k)\}_k$ *defined by* (16) *is bounded and any of its cluster point is a global minimizer of* $F(u,f)$.

Proof. By (16), we have

$$\begin{aligned}F(u^{k+1},f^{k+1}) &\le F(u^{k+1},\tilde{f}^{k+1})+\delta_k \le F(u^{k+1},f^k)+\delta_k \le F(\tilde{u}^{k+1},f^k)+\varepsilon_k+\delta_k\\ &\le F(u^k,f^k)+\varepsilon_k+\delta_k.\end{aligned}$$

Thus, we have

$$F(u^{k+1},f^{k+1}) \le F(u^0,f^0)+\sum_{j=0}^{k}(\varepsilon_j+\delta_j).$$

Since $\sum_{j=0}^{\infty}(\varepsilon_k+\delta_k)<\infty$ and all level sets of F, i.e. $\{(u,f)\in\mathbb{R}^n\times\mathbb{R}^m : F(u,f)\le C\}$, is compact, it is clear that the sequence $\{(u^k,f^k)\}_k$ is indeed bounded and hence has convergent subsequence. Without loss of generality, we assume that the sequence itself converges and has limit $(\bar{u},\bar{f})$. Using (16) again, we have

$$\begin{aligned}F(u^{k+1},f^k) &\le F(u^{k+1}+v_u,f^k)+\varepsilon_k \quad \forall v_u\in\mathbb{R}^n\\ F(u^{k+1},f^{k+1}) &\le F(u^{k+1},f^{k+1}+v_f)+\delta_k \quad \forall v_f\in\mathbb{R}^m.\end{aligned}$$

Letting $k\to\infty$, and using the continuity of F and the fact that $\varepsilon\to 0$ and $\delta_k\to 0$, we have

$$\begin{aligned}F(\bar{u},\bar{f}) &\le F(\bar{u}+v_u,\bar{f}) \quad \forall v_u\in\mathbb{R}^n\\ F(\bar{u},\bar{f}) &\le F(\bar{u},\bar{f}+v_f) \quad \forall v_f\in\mathbb{R}^m.\end{aligned}$$

We have thus shown that $(\bar{u},\bar{f})$ is a coordinatewise minimizer of F and the rest of the proof follows from Lemma 2.1. □

3 Numerical Simulations

In this section, we shall compare our proposed model (4) (using $p=2$) with TV-based model (3), anisotropic wavelet frame based model ((6) with $p=1$) and isotropic wavelet frame based model ((6) with $p=2$). In all experiments, for model (4) we always set $\kappa=1,\lambda_2=0.01,\mu_2=0.00002$. The parameter μ_1 is set up as 200 for Catphan phantom and 1000 for other phantoms. Only the parameter λ_1 varies due to the strength of the noise and number of projections. Stronger noise or less projections will make the parameter λ_1 larger, which is the same as the parameter λ in model (3) and (6) with both $p=1$ and $p=2$. We test these models using a digital NURBS-based cardiac-torso (NCAT) phantom [57–59].

In our experiments, all the projection data is created by set $f=P(\tilde{u})+\epsilon$, where ϵ is a noisy signal corresponding to an X-ray tube current of certain

mA used in a typical scanning protocol [60]. Roughly speaking, at each pixel location the noise is generated from a Guassian distribution with zero mean and a variance depending on the value of projected image at that location. We will choose two different noise levels: mild and strong noise which can be seen in Figure 2. To get some ideas of the actual strength of noise comparing to signal, we remark that the pixel values of the projected image with 20 projections for example range within $[0, 5.7613]$ with mean 2.3967. For each model, we calculate the relative error, correlation and the total computation time. The relative error and the correlation is defined in (17) and (18) respectively as follows:

$$err(u) = \frac{\|u - \tilde{u}\|_2}{\|\tilde{u}\|_2} \tag{17}$$

$$corr(u) = \frac{(u - \bar{u})(\tilde{u} - \bar{\tilde{u}})}{\|u - u\|_2 \|\tilde{u} - \bar{\tilde{u}}\|_2} \tag{18}$$

where $\tilde{u}$ denotes the ground truth, $\bar{u}$ and $\bar{\tilde{u}}$ denote the mean values of u and $\tilde{u}$ respectively.

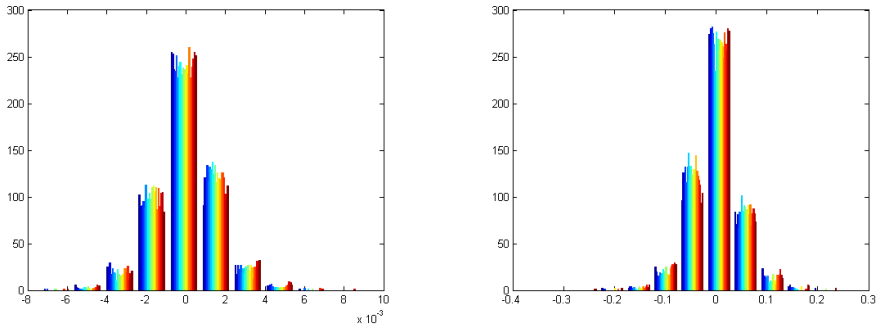

Figure 2: The distribution of the noise adding in the Radon domain with 20 projections. Images from left to right represent the mild and strong noise, respectively.

Table 1 and 2 show that our proposed model performs best among all tested models. The trend of the relative error can be seen in Figure 4 and 6, which gives us the conclusion that our proposed model (4) has the fastest decreasing speed of relative error. Also, both the anisotropic wavelet frame based model and isotropic wavelet frame based model generate better results than TV-based model (3). It is worth noticing that the isotropic wavelet frame based model, the relative error is lower and the correlation is higher comparing to anisotropic wavelet frame based model, since the isotropic ℓ_1-norm can protect edges in all directions.

The reconstructed images for all the tested models and cases are shown in Figure 3 and Figure 5. We observe that the reconstructed images have almost no artifacts using either isotropic wavelet frame based model or our proposed

model (4). In particular, for $N_p = 15$, some of the key structures in the lung region are lost for TV-based model, the anisotropic wavelet frame based model recovers more structure than TV-based model, while our proposed model (4) recovers even more structure than the anisotropic wavelet frame based model.

In addition to simulated CT data, we also test our algorithm to real CT data. We apply all algorithms to reconstruct a transverse slice on the source rotation plane of the physical Catphan phantom (The Phantom Laboratory, Inc., Salem, NY) using GE CT scanner at mA level 100 with 984 projections being acquired by the scanner. We select 30 projections among the 984 projections (i.e. $N_p = 30$) for reconstruction. Reconstructed images are shown in Figure 7. Judging from the spatial resolution of the reconstructed CT images, it is clear that all wavelet frame based models produce better results than the TV-based model, among which our inpainting model (4) performs the best.

Finally we note that if we use model (4) and do inpainting in Radon domain twice, the relative error can be further reduced and the correlation will be enhanced at the cost of more computation time (see Table 3).

Table 1: Comparison of relative error (in percentage), correlation (in percentage) and the running time (in seconds) of the algorithm with mild real noise.

N_p	TV-based Model			Anisotropic			Isotropic			Inpainting Model (4)		
	error	corr	Time	error	corr	Time	error	corr	Time	error	corr	Time
10	19.3	96.8	107	15.2	98.0	100	13.6	98.4	113	12.4	98.7	285
15	12.4	98.7	124	9.9	99.1	121	8.4	99.4	138	7.2	99.6	369
20	8.8	99.4	137	7.7	99.5	128	6.2	99.7	140	5.2	99.8	396
30	6.3	99.7	172	5.8	99.7	151	4.7	99.8	173	4.1	99.8	523
40	5.1	99.8	204	4.5	99.8	188	3.4	99.9	203	2.9	99.9	640
60	3.8	99.9	265	3.5	99.9	427	2.7	99.9	370	2.2	100.0	807

Table 2: Comparison of relative error (in percentage), correlation (in percentage) and the running time (in seconds) of the algorithm with strong real noise.

N_p	TV-Based Model			Anisotropic			Isotropic			Inpainting Model (4)		
	error	corr	Time	error	corr	Time	error	corr	Time	error	corr	Time
15	15.3	98.1	146	11.1	99.0	130	10.2	99.1	126	9.5	99.2	338
20	12.6	98.6	295	8.9	99.3	142	8.2	99.4	159	7.9	99.4	430
30	11.2	98.9	354	7.5	99.5	186	7.1	99.6	177	7.0	99.6	526
40	10.4	99.1	398	7.1	99.6	218	6.6	99.6	207	6.5	99.6	654

Table 3: Comparison of relative error (in percentage), correlation (in percentage) and the running time (in seconds) of the multiple inpainting in Radon domain with the regularization of wavelet frame for mild real noise.

N_p	Isotropic			Inpainting Once			Inpainting Twice		
	error	correlation	Time	error	correlation	Time	error	correlation	Time
10	13.6	98.4	113	12.4	98.7	285	12.3	98.7	411
15	8.4	99.4	138	7.2	99.6	369	7.0	99.6	550
20	6.2	99.7	140	5.2	99.8	396	5.0	99.8	818

4 Conclusions

In this paper, we proposed a simultaneous CT image reconstruction and Radon domain inpainting model using wavelet frame based regularization. The

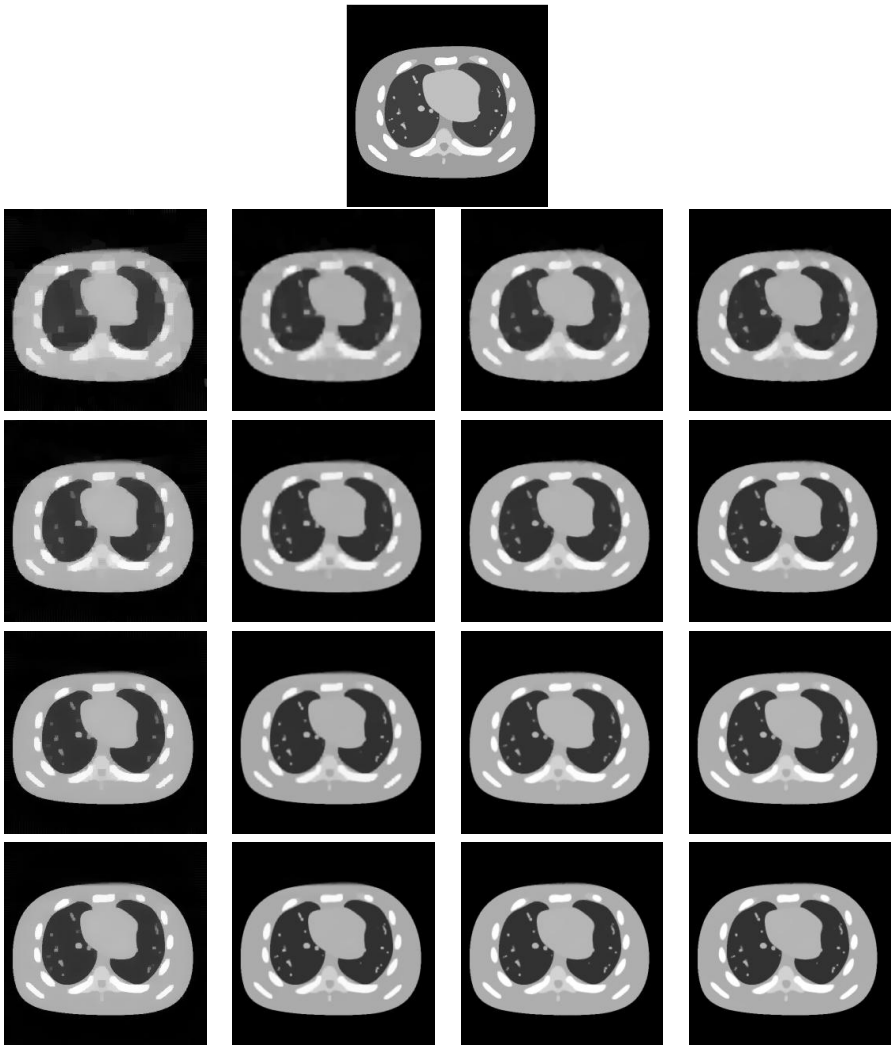

Figure 3: The tomographic result with mild real noise. The image on top is the true data $\tilde{u}$. The following rows represent the results using 15, 20, 30 and 40 projections, respectively. Images from left to right in each row are the results obtained by TV-based model, anisotropic wavelet frame based model, our proposed isotropic wavelet frame based model and our proposed model (4) with inpainting in Radon domain.

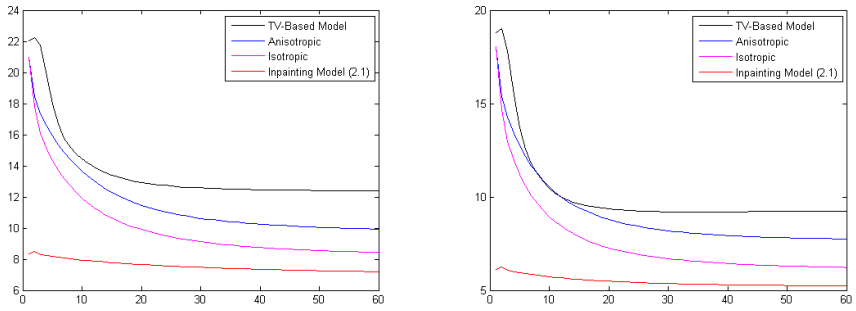

Figure 4: The change of relative error during the iteration for the cases with mild real noise. The two graphs represent the results using 15 and 20 projections, respectively.

proposed model reconstructs high quality CT images as well as high resolution projected images based on the observed low resolution projected images. Fast numerical algorithms were also introduced base on the split Bregman algorithm. Our numerical simulations show that the proposed model outperforms the TV-based model, as well as analysis based approach without Radon domain inpainting.

Acknowledgements

We would like to thank Dr. Xun Jia (Department of Radiation Oncology, University of California, San Diego) for providing us with the MATLAB program of Siddon's algorithm, the real CT data, and many valuable discussions on the subject.

References

[1] E. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Physics in medicine and biology*, vol. 53, p. 4777, 2008.

[2] E. Sidky, C. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *Journal of X-Ray Science and Technology*, vol. 14, no. 2, pp. 119–139, 2006.

[3] X. Jia, Y. Lou, R. Li, W. Song, and S. Jiang, "GPU-based fast cone beam CT reconstruction from undersampled and noisy projection data via total variation," *Medical physics*, vol. 37, p. 1757, 2010.

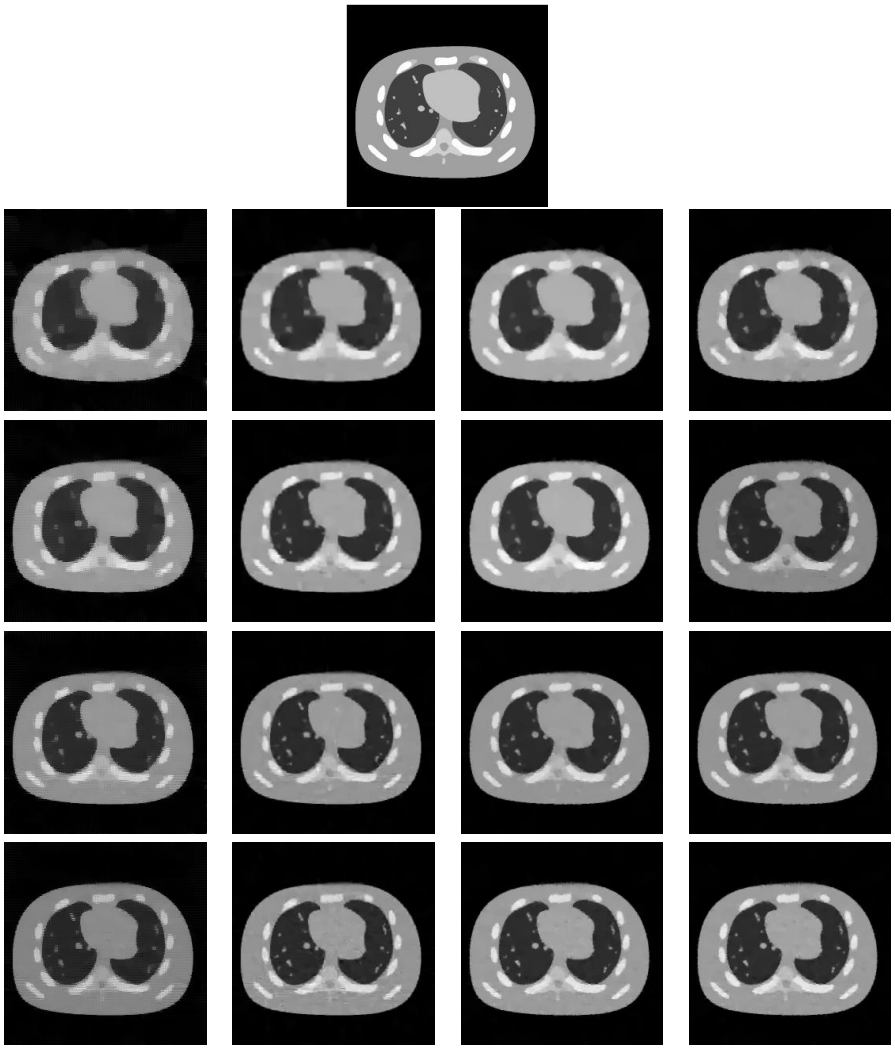

Figure 5: The tomographic result with strong real noise. The image on top is the true data $\tilde{u}$. The following rows represent the results using 15, 20, 30 and 40 projections, respectively. Images from left to right in each row are the results obtained by TV-based model, anisotropic wavelet frame based model and our proposed isotropic wavelet frame based model and our proposed model (4) with inpainting in Radon domain.

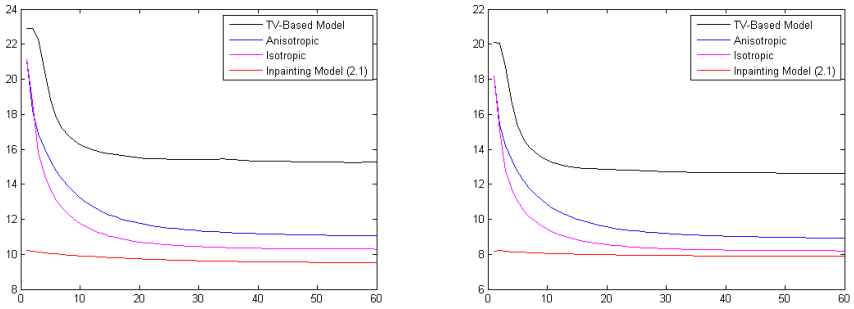

Figure 6: The change of relative error during the iteration for the cases with strong real noise. The two graphs represent the results using 15 and 20 projections, respectively.

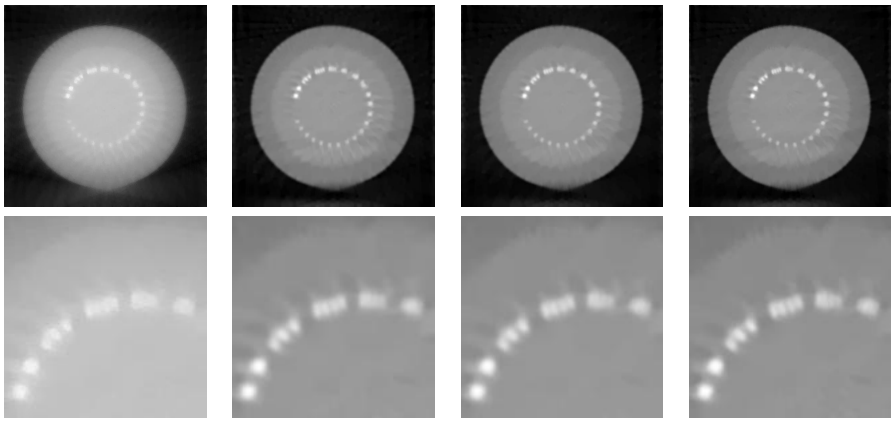

Figure 7: The reconstruction results for Catphan phantom. The top row from left to right presents the reconstruction results by TV-based model, anisotropic wavelet frame based model, isotropic wavelet frame based model and model (4) with inpainting in Radon domain. The bottom row presents the corresponding zoom-in views of the reconstruction results.

[4] K. Choi, J. Wang, L. Zhu, T. Suh, S. Boyd, and L. Xing, "Compressed sensing based cone-beam computed tomography reconstruction with a first-order method," *Medical physics*, vol. 37, p. 5113, 2010.

[5] X. Jia, B. Dong, Y. Lou, and S. Jiang, "GPU-based iterative cone-beam CT reconstruction using tight frame regularization," *Physics in Medicine and Biology*, vol. 56, pp. 3787–3807, 2011.

[6] J. Cai, B. Dong, S. Osher, and Z. Shen, "Image restorations: total variation, wavelet frames and beyond," *preprint*, 2011.

[7] T. Goldstein and S. Osher, "The split Bregman algorithm for L1 regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[8] J. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration," *Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal*, vol. 8, no. 2, pp. 337–369, 2009.

[9] M. Islam, T. Purdie, B. Norrlinger, H. Alasti, D. Moseley, M. Sharpe, J. Siewerdsen, and D. Jaffray, "Patient dose from kilovoltage cone beam computed tomography imaging in radiation therapy," *Medical physics*, vol. 33, p. 1573, 2006.

[10] L. Feldkamp, L. Davis, and J. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Am. A*, vol. 1, no. 6, pp. 612–619, 1984.

[11] J. Radon, "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten," *Berichte Sächsische Akademie der Wissenschaften*, vol. 69, pp. 262–267, 1917.

[12] R. L. Siddon, "Fast calculation of the exact radiological path for a 3-dimensional ct array," *Medical Physics*, vol. 12, pp. 252–5, 1985.

[13] R. Gordon, R. Bender, and G. Herman, "Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography* 1," *Journal of theoretical Biology*, vol. 29, no. 3, pp. 471–481, 1970.

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

[15] Y. Meyer, *Oscillating patterns in image processing and nonlinear evolution equations: the fifteenth Dean Jacqueline B. Lewis memorial lectures*. Amer Mathematical Society, 2001.

[16] G. Sapiro, *Geometric partial differential equations and image analysis*. Cambridge Univ Pr, 2001.

[17] S. Osher and R. Fedkiw, *Level set methods and dynamic implicit surfaces*. Springer, 2003.

[18] G. Aubert and P. Kornprobst, *Mathematical problems in image processing: partial differential equations and the calculus of variations*. Springer, 2006.

[19] T. Chan and J. Shen, *Image processing and analysis: variational, PDE, wavelet, and stochastic methods*. Society for Industrial Mathematics, 2005.

[20] Y. Mao, B. Fahimian, S. Osher, and J. Miao, "Development and optimization of regularized tomographic reconstruction algorithms utilizing equally-sloped tomography," *Image Processing, IEEE Transactions on*, vol. 19, no. 5, pp. 1259–1268, 2010.

[21] B. Dong, A. Chien, and Z. Shen, "Frame based segmentation for medical images," *Communications in Mathematical Sciences*, vol. 9, no. 2, pp. 551–559, 2011.

[22] B. Dong and Z. Shen, "Frame based surface reconstruction from unorganized points," *accepted by Journal of Computational Physics*, 2011.

[23] R. Chan, T. Chan, L. Shen, and Z. Shen, "Wavelet algorithms for high-resolution image reconstruction," *SIAM Journal on Scientific Computing*, vol. 24, no. 4, pp. 1408–1432, 2003.

[24] R. Chan, L. Shen, and Z. Shen, "A framelet-based approach for image inpainting," *Research Report*, vol. 4, p. 325, 2005.

[25] J. Cai, S. Osher, and Z. Shen, "Linearized Bregman iterations for frame-based image deblurring," *SIAM J. Imaging Sci*, vol. 2, no. 1, pp. 226–252, 2009.

[26] B. Dong and Z. Shen, "MRA Based Wavelet Frames and Applications," *IAS Lecture Notes Series, Summer Program on "The Mathematics of Image Processing", Park City Mathematics Institute*, 2010.

[27] R. Chan, S. Riemenschneider, L. Shen, and Z. Shen, "Tight frame: an efficient way for high-resolution image reconstruction," *Applied and Computational Harmonic Analysis*, vol. 17, no. 1, pp. 91–115, 2004.

[28] R. Chan, S. Riemenschneider, L. Shen, and Z. Shen, "High-resolution image reconstruction with displacement errors: A framelet approach," *International Journal of Imaging Systems and Technology*, vol. 14, no. 3, pp. 91–104, 2004.

[29] J. Cai, R. Chan, L. Shen, and Z. Shen, "Tight Frame Based Method for High-Resolution Image Reconstruction," *Contemporary Applied Mathematics, A. Damlamian and S. Jaffard eds, Higher Education Press (Beijing)*, pp. 1–36, 2010.

[30] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pp. 417–424, ACM Press/Addison-Wesley Publishing Co., 2000.

[31] J. Cai, R. Chan, and Z. Shen, "A framelet-based image inpainting algorithm," *Applied and Computational Harmonic Analysis*, vol. 24, no. 2, pp. 131–149, 2008.

[32] I. Daubechies, *Ten lectures on wavelets*, vol. CBMS-NSF Lecture Notes, SIAM, nr. 61. Society for Industrial Mathematics, 1992.

[33] A. Ron and Z. Shen, "Affine Systems in $L_2(\mathbb{R}^d)$: The Analysis of the Analysis Operator," *Journal of Functional Analysis*, vol. 148, no. 2, pp. 408–447, 1997.

[34] I. Daubechies, B. Han, A. Ron, and Z. Shen, "Framelets: Mra-based constructions of wavelet frames," *Applied and Computational Harmonic Analysis*, vol. 14, pp. 1–46, Jan 2003.

[35] Z. Shen, "Wavelet frames and image restorations," *Proceedings of the International Congress of Mathematicians, Hyderabad, India*, 2010.

[36] S. K. Patch, "Computation of unmeasured third-generation vct views from measured views," *IEEE Transactions on Medical Imaging*, vol. 21-7, pp. 801–13, 2002.

[37] H. Yu, Y. Wei, J. Hsieh, and G. Wang, "Data consistency based translational motion artifact reduction in fan-beam CT," *Medical Imaging, IEEE Transactions on*, vol. 25, no. 6, pp. 792–803, 2006.

[38] M. Elad, J. Starck, P. Querre, and D. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Applied and Computational Harmonic Analysis*, vol. 19, no. 3, pp. 340–358, 2005.

[39] J. Starck, M. Elad, and D. Donoho, "Image decomposition via the combination of sparse representations and a variational approach," *IEEE transactions on image processing*, vol. 14, no. 10, pp. 1570–1582, 2005.

[40] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *SIAM Journal on Imaging Sciences*, 2010. accepted.

[41] G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing," *Multiscale Model Sim*, vol. 7, pp. 1005–1028, Jan 2008.

[42] E. Esser, "Applications of Lagrangian-based alternating direction methods and connections to split Bregman," *CAM Report*, vol. 9, p. 31, 2009.

[43] X. Tai and C. Wu, "Augmented Lagrangian method, dual methods and split Bregman iteration for ROF model," *Scale Space and Variational Methods in Computer Vision*, pp. 502–513, 2009.

[44] R. Glowinski and P. Le Tallec, *Augmented Lagrangian and operator-splitting methods in nonlinear mechanics*. Society for Industrial Mathematics, 1989.

[45] S. Setzer, "Split Bregman algorithm, Douglas-Rachford splitting and frame shrinkage," *Scale Space and Variational Methods in Computer Vision*, pp. 464–476, 2009.

[46] D. Donoho, "De-noising by soft-thresholding," *IEEE transactions on information theory*, vol. 41, no. 3, pp. 613–627, 1995.

[47] P. Combettes and V. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Modeling and Simulation*, vol. 4, no. 4, pp. 1168–1200, 2006.

[48] Y. Wang, W. Yin, and Y. Zhang, "A fast algorithm for image deblurring with total variation regularization," *Rice University CAAM Technical Report TR07-10*, 2007.

[49] J. Bezdek and R. Hathaway, "Convergence of alternating optimization," *Neural, Parallel & Scientific Computations*, vol. 11, no. 4, pp. 351–368, 2003.

[50] J. Friedman, T. Hastie, H. Höfling, and R. Tibshirani, "Pathwise coordinate optimization," *The Annals of Applied Statistics*, vol. 1, no. 2, pp. 302–332, 2007.

[51] Z. Luo and P. Tseng, "On the convergence of the coordinate descent method for convex differentiable minimization," *Journal of Optimization Theory and Applications*, vol. 72, no. 1, pp. 7–35, 1992.

[52] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *Journal of optimization theory and applications*, vol. 109, no. 3, pp. 475–494, 2001.

[53] P. Tseng and S. Yun, "A coordinate gradient descent method for nonsmooth separable minimization," *Mathematical Programming*, vol. 117, no. 1, pp. 387–423, 2009.

[54] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.

[55] J. Eckstein and D. Bertsekas, "On the douglasrachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 1, pp. 293–318, 1992.

[56] R. Rockafellar and J. Roger, *Variational analysis*. Springer Verlag, 1997.

[57] W. Segars, B. Tsui, D. Lalush, E. Frey, M. King, and D. Manocha, "Development and application of the new dynamic Nurbs-based Cardiac-Torso (NCAT) phantom.," *Biomedical 5 Engineering, (Chapel Hill, NC: University of North Carolina)*, 2001.

[58] W. Segars, D. Lalush, and B. Tsui, "A realistic spline-based dynamic heart phantom," *Nuclear Science, IEEE Transactions on*, vol. 46, no. 3, pp. 503–506, 1999.

[59] W. Segars, D. Lalush, and B. Tsui, "Modeling respiratory mechanics in the MCAT and spline-based MCAT phantoms," *Nuclear Science, IEEE Transactions on*, vol. 48, no. 1, pp. 89–97, 2001.

[60] J. Bushberg, J. Seibert, E. Leidholdt Jr, and J. Boone, *The essential physics of medical imaging*. Williams & Wilkins, 2002.