A Mathematical Model for Adaptive Computed Tomography Sensing

Oren Barkan, Jonathan Weill, Shai Dekel, and Amir Averbuch

Abstract—**One of the main challenges in computed tomography (CT) is how to balance between the amount of radiation the patient is exposed to during scan time and the quality of the reconstructed CT image. We propose a mathematical model for adaptive CT sensing whose goal is to reduce dosage levels while maintaining high image quality at the same time. The adaptive algorithm iterates between selective limited sensing and improved reconstruction, with the goal of applying only the dose level required for sufficient image quality. The theoretical foundation of the algorithm is nonlinear Ridgelet approximation and a discrete form of Ridgelet analysis is used to compute the selective acquisition steps that best capture the image edges. We show experimental results where for the same number of line projections, the adaptive model produces higher image quality, when compared with standard limited angle, nonadaptive sensing algorithms.**

Index Terms—**Adaptive compressed sensing, adaptive CT acquisition, contour analysis, low dose CT, multiresolution analysis, nonlinear approximation, ridgelets, wavelets.**

I. Introduction

In THE last decade, several studies have shown that radiation exposure during CT scanning is a significant factor in raising the total public risk of cancer deaths [1], [20], [21]. To balance between image quality and these concerns, radiologists use the protocol: As Low as Reasonably Achievable (ALARA). It is meant to ensure that "... CT dose factors are kept to a point where risk is minimized for maximum diagnostic benefit ...", where the dose can be determined by the product of the CT tube current and the time the patient is exposed to radiation (see [2] for an overview). Currently, there are several state-of-the-art technologies that attempt to achieve dose reduction. Iterative Reconstruction (IR) methods are successful in reducing artifacts, improving resolution and lowering the noise in the reconstructed images [3], [4]. Even more recently, the Model Based Iterative Reconstruction (MBIR) [6] was introduced. It improves upon the IR methods by incorporating accurate system models coupled with statistical noise models and prior models. However, dosage levels during CT exams are still at the focus of attention and any new method that can reduce them is considered highly valuable.

This paper describes an adaptive tomography acquisition model that is superior, in the CT image quality, to existing limited angle, non-adaptive acquisition methods and in theory may allow minimal dosage levels. The method can be considered as a significant generalization of existing two-step adaptive acquisition methods [7], [8] and can potentially use the same hardware configurations that are capable of changing their geometric configuration and acquisition protocols on-the-fly [9]. For example, in [7], the authors describe an imaging C-arm system where a low-dose overview (OV) scan is used to dynamically identify an arbitrary Volume Of Interest (VOI). The OV and VOI scans are then registered and reconstructed together. In [8], the authors developed a flexible x-ray micro-CT system, named FaCT, capable of changing its geometric configuration and acquisition protocol in order to best suit an object being imaged for a particular diagnostic task. In their system, a fast, sparse-projection pre-scan is performed, the data are reconstructed, and the region of interest is identified. Next, a diagnostic-quality scan is performed where, given the region of interest, the control computer calculates an illumination window for online control of an x-ray source masking aperture to transmit radiation only through the region of interest throughout the scan trajectory.

A more advanced architecture of CT machine is the Electron Beam Tomography (EBT) [35]. In the EBT machine (illustrated in Fig. 1), instead of mechanical motion of an X-ray tube, a high speed electron beam is focused and deflected by carefully designed coils to sweep along a target ring (an anode with multiple target tracks). The entire assembly is sealed in a vacuum. Fan-shaped X-ray beams are produced and collimated to a set of detectors mounted in the detector ring. Detector ring and target ring are offset to make room for the overlapped portion. In this way a large enough X-ray tube is constructed so only the path of electrons, travelling between the cathode and the anode, is spun using a deflection coil. Since the system has no mechanical moving parts, scan time is much faster than in conventional CT system. This property enables less blurry imaging of moving structures, such as the heart and arteries.

The adaptive acquisition algorithm depict in this work can potentially use the EBT machine as a hardware acquisition device but is not limited to this type only. Standard spiral CT machines might also fit and scale with our proposed methodology with appropriate configuration.

Manuscript received December 28, 2016; revised April 12, 2017 and July 12, 2017; accepted July 24, 2017. Date of publication August 7, 2017; date of current version November 6, 2017. Parts of this manuscript appeared in a previous publication [38]. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Laura Waller. *(Corresponding author: Oren Barkan.)*

O. Barkan, J. Weill, and A. Averbuch are with the School of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel.

S. Dekel is with GE Global Research, Niskayuna, NY 12309 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCI.2017.2736788

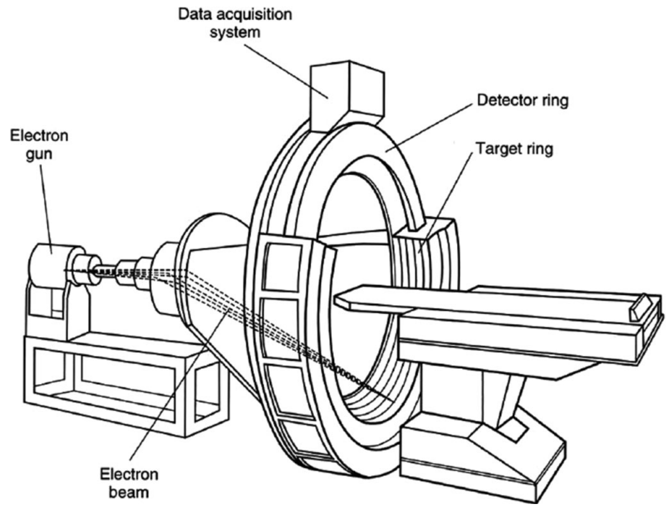


Fig. 1. Main components of the Electron Beam Tomography (EBT) machine.

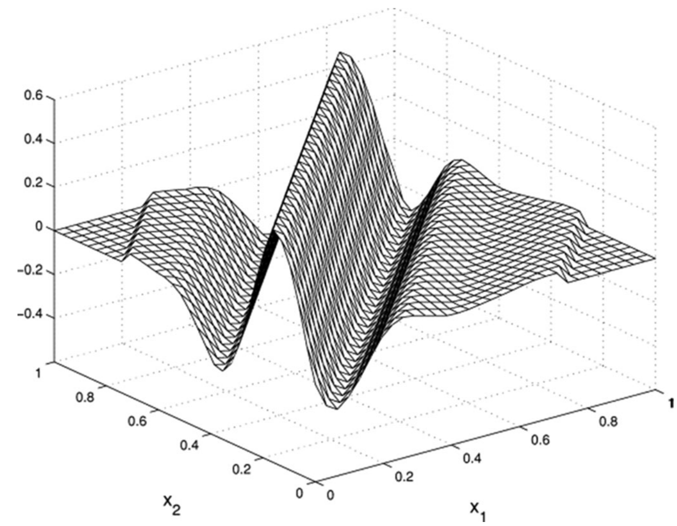


Fig. 2. A Ridgelet function $\psi_{a,b,\theta}(x_1, x_2)$.

Observe that adaptive acquisition should not be confused with adaptive reconstruction. In the latter, the acquisition model is a non-adaptive which implies uniform sampling scheme, where over a discrete set of pre-determined angles, line projections are computed at equal intervals. In this setup, the adaptive elements, if exist, are part of the post-acquisition reconstruction step.

The outline of the proposed algorithm is as follows: First, the system projects the object with an extreme low dose according to a uniform predetermined pattern and reconstructs an initial low quality image. Next, the system predicts from the reconstructed low quality image where the significant edges of the true objects are and projects along them. Then, the system iterates by incorporating the newly added line projections in order to obtain a refined approximation of the object's image. The algorithm continues to iterate between estimation of locations of significant features, adaptive acquisition and reconstruction until a convergence criterion is met. The goal is to converge to a high quality reconstruction using a minimal number of rays (line projections). The method relies on the mathematical model of Ridgelets [5], and therefore has a natural multiresolution capability, where the significance of edges is analyzed at different scales.

It is important to clarify the following fundamental assumption we make on the acquired images. To illustrate, Let I be a bi-level image, i.e., with pixels that are either '0' or '1', where the '1' values are sparse. Even on this simple image, our approach would render useless if the '1' values are scattered in random locations against the background of zeros. In such a case, as clearly explained in [24], adaptive acquisition has absolutely no advantage over non-adaptive methods. However, if I is what is called a 'cartoon' image, where the '1' values are grouped into 'nicely' connected subdomains with piecewise smooth boundaries, then the situation changes dramatically.

The mathematical theory of [5] quantifies, in the setup of CT, the geometric 'structure' of the image and how fast a Ridgelet approximation converges to the image. Our algorithm, whose goal is to acquire an unknown image, regards the adaptive Ridgelet approximation of the image as the 'optimal' benchmark and is designed to match its performance. This approach has strong ties with the waveform analysis presented in [10] that allowed the authors to classify singularities and quantify the 'stability' of limited angle tomography. Indeed, although in our work we limit the number of line projections, but do not limit the angles, the fundamental understanding of the relationship between a function's edge singularities and its Radon representation, as explained in [10], is at the core of our algorithm (see Fig. 3 and the accompanying explanation). We show in the experimental results section that for the same number of line projections, our algorithm yields higher image reconstruction quality, when compared with known limited angle, non-adaptive acquisition algorithms.

The paper is organized as follows: Sections II and III introduce necessary mathematical background. Section IV describes in detail our adaptive acquisition algorithm. Experimental results and comparisons with non-adaptive methods are given in Section V. In Section VI we draw conclusions and discuss future work.

II. Fundamentals of Ridgelets Theory

Let $\psi \in L_2(\mathbb{R})$ be a wavelet [15]. For the purpose of this paper it is sufficient that the wavelet function has two properties: compact support and vanishing moments. The latter implies that for some $r \geq 1$:

$$\int_{\mathbb{R}} \psi(x)\, x^l\, dx = 0, \quad l = 0, \ldots, r-1.$$

The classical example for a wavelet function is the Haar wavelet with one vanishing moment:

$$\psi(x) \triangleq \begin{cases} 1, & 0 \leq x \leq 1/2, \\ -1, & 1/2 < x \leq 1, \\ 0, & else. \end{cases}$$

A bivariate Ridgelet function [16] is defined by

$$\psi_{a,b,\theta}(x_1, x_2) \triangleq a^{-1/2} \psi((x_1 \cos\theta + x_2 \sin\theta - b)/a),$$

where a, b and θ are the parameters determining the scale, translation and rotation of the Ridgelet function, respectively (see Fig. 2).

Given $f \in L_1(\mathbb{R}^2)$, its Continuous Ridgelet Transform [16] (CRT) is defined by

$$CRT_f(a,b,\theta) \triangleq \int_{\mathbb{R}^2} \psi_{a,b,\theta}(x) f(x) dx. \tag{2.3}$$

The continuous Radon transform [17] of a bivariate function f at direction θ is defined as

$$R_f(\theta,t) \triangleq \int_{\mathbb{R}^2} f(x_1,x_2)\delta\left(x_1\cos\theta + x_2\sin\theta - t\right) dx_1 dx_2,$$

where δ is the Dirac function. The Radon and the Ridgelet transforms are related by

$$CRT_f(a,b,\theta) = \int_{\mathbb{R}} \psi_{a,b}(t) R_f\left(\theta,t\right) dt, \tag{2.4}$$

where,

$$\psi_{a,b}(x) \triangleq a^{-1/2}\psi((x-b)/a).$$

In applications, this means that the Ridgelet transform can be computed by the application of the Radon transform at a given angle, followed by 1D fast wavelet transform [15].

It is interesting to point out that Ridgelets [5] did not previously find too many applications in image processing. Their 'descendants' Curvelets [25], [26] and Shearlets [27], [28], which capture directional information as well, were found to be more useful due to their better time-frequency localization.

However, we find that Ridgelets are the right mathematical tool in the setup of CT, since the acquisition device is not able to capture, through its sampling process, well localized functionals such as Curvelet coefficients. Observe that in the context of CT reconstruction, Curvelets have been used as a regularization tool [23].

From approximation theoretical perspective, the mathematical foundation of our adaptive algorithm follows the framework of characterizing the images by the appropriate function smoothness spaces and then providing an estimate for the order of convergence.

Definition 1 [5]: For $\alpha > 0$, and $p, q > 0$, we say that $f \in \dot{R}^{\alpha}_{p,q}(\mathbb{R}^2)$, if $f \in L_1(\mathbb{R}^2)$ and

$$\|f\|_{\dot{R}^{\alpha}_{p,q}} \triangleq \left(\sum_{j=-\infty}^{\infty} 2^{j(\alpha+1/2)q}\left(\frac{1}{\pi}\int_0^{\pi}\left\|CRT_f\left(2^j,\cdot,\theta\right)\right\|_p^p d\theta\right)^{q/p}\right)^{1/q} < \infty.$$

We note that this definition requires certain conditions on the wavelet ψ, such as sufficient vanishing moments (with respect to α) and its decay, which we shall omit here. These conditions ensure that membership in the smoothness space $\dot{R}^{\alpha}_{p,q}$ does not depend on the particular wavelet used in (2.3). A typical non trivial example for a function in $\dot{R}^{\alpha}_{p,q}$ is a function with a singularity along a line such as

$$f\left(x_1,x_2\right) = 1_{\{x_1>0\}}\left(x_1,x_2\right)\left(2\pi\right)^{-1/2} e^{-\left(x_1^2+x_2^2\right)/2}$$

This function is in the Besov class [29] $B^{\alpha}_{1,1}$ only for $\alpha < 1$, which means that it almost has a first derivative in the classical sense. In contrast, this function is contained in $\dot{R}^{\alpha}_{1,1}$, for any $\alpha < 3/2$ [5], which implies that it is smoother in the scale of Ridgelet spaces than in the scale of Classical Besov spaces. In this work we assume that the functions we analyze are compactly supported in a 'standard' compact domain such as $[-1,1]^2$ and attain the value zero on its boundary. Indeed, CT images satisfy this requirement (see the examples in next sections). Therefore, by a simple zero extension argument, a function $f \in L_2([-1,1]^2)$ of this nature can also be regarded as a function in $L_1(\mathbb{R}^2) \cap L_2(\mathbb{R}^2)$. By sampling the CRT, one may obtain a discrete Ridgelet Frame system $\{\psi_\gamma\}$ with a dual system $\{\tilde{\psi}_\gamma\}$, for a countable index $\{\gamma = (a,b,\theta)\}$, such that for $f \in L_2([-1,1]^2)$,

$$f = \sum_{\gamma}\left\langle f, \tilde{\psi}_\gamma\right\rangle \psi_\gamma = \sum_{\gamma}\left\langle f, \psi_\gamma\right\rangle \tilde{\psi}_\gamma.$$

Recall, that the frame property guarantees 'stability' of the representation, in the sense that there exist constants $0 < A \leq B < \infty$, such that

$$A\|f\|_2^2 \leq \sum_{\gamma}\left|\langle f,\psi_\gamma\rangle\right|^2 \leq B\|f\|_2^2, \qquad \forall f \in L_2\left([-1,1]^2\right).$$

Let us rearrange the Ridgelet coefficients based on the size of their absolute values

$$\left|\langle f,\psi_{\gamma_1}\rangle\right| \geq \left|\langle f,\psi_{\gamma_2}\rangle\right| \geq \cdots,$$

and denote the n-term adaptive approximation to f by

$$f_n \triangleq \sum_{i=1}^{n}\left\langle f,\psi_{\gamma_i}\right\rangle \tilde{\psi}_{\gamma_i}.$$

Then, we have the Jackson-type estimate [5] for $\alpha > 1/2$ and $1/\tau = \alpha - 1/2$,

$$\|f - f_n\|_{L_2\left([-1,1]^2\right)} \leq c n^{-\alpha/2}\|f\|_{\dot{R}^{\alpha}_{\tau,\tau}}.$$

Thus, under certain assumptions on the input function, not only the convergence of the adaptive approximation is ensured, but its rate is also estimated. The outcome the theory is that the approximation rate of an adaptive Ridgelet approximation depends on the smoothness of the function in a given Ridgelet smoothness space, much in the same manner that adaptive wavelet approximation is characterized by Besov space smoothness [29].

As we shall see in Section IV, our adaptive acquisition algorithm follows the adaptive Ridgelet approximation to the image I as a model. It tries to predict from iterative approximation to I, the significant Ridgelet coefficients. Then we use these coefficients in order to select the next set of line projections that are considered as best candidates to project I with, in the subsequent iteration.

III. Efficient Total Variation Minimization Using Sparse Matrices

For a given image $I \in \mathbb{R}^{m\times m}$, with pixels values $\{I_{i,j}\}$, we define the gradient of I by

$$\nabla I_{i,j} = (I_{i,j} - I_{i-1,j}, I_{i,j} - I_{i,j-1}).$$

The Total Variation (TV) norm of the image is given by

$$|I|_{TV} \triangleq \sum_{i,j=2}^{m} \left(|I_{i,j} - I_{i-1,j}| + |I_{i,j} - I_{i,j-1}|\right).$$

Denote $N = m^2$ and let $x \in \mathbb{R}^N$ be a one-dimensional representation of I by concatenating the columns of I into a single column vector $x = (I_{1,1}, I_{2,1}, \ldots, I_{m,1}, \ldots, I_{1,m}, \ldots I_{m,m})^T$. Given an $n \times N (n << N)$ sampling matrix $A \in \mathbb{R}^{n \times N}$ and corresponding observations vector $y \in \mathbb{R}^n$, generated by $Ax = y$, the so-called TV-minimization is concerned with solving one of the following optimization problems:

$$\min_U |U|_{TV} \quad \text{s.t. } Au = y \tag{3.1}$$

$$\min_U |U|_{TV} + \mu \|Au - y\|_2^2 \tag{3.2}$$

where $u \in \mathbb{R}^N$ is the one-dimensional representation of $U \in \mathbb{R}^{m \times m}$ and μ is a given weight parameter. The minimization problem (3.2) is applied in the presence of noise in the sampling process and the weight μ depends, in part, on the expected noise level. This model is difficult to solve directly due to non-differentiability and non-linearity of the TV term. During the last few years there has been an explosion of new numeric iterative solvers (see the papers in the "Compressive Sensing Recovery Algorithms" section of [11]).

Although conceptually our method may use such solvers as black boxes, its unique features allow us to apply critical modifications that not only accelerate the iterative methods, but also make them feasible in large datasets problems when N is large. In this work, we implemented a modified version of the TVAL3 solver [12], [13]. Our modified version utilizes the fact that in our special case the matrix A is highly sparse. This is in complete contrast to the usual setup in compressed sensing, where the theory dictates a dense matrix (usually of pseudo-random nature). As we shall see in Section IV, in our case, the sparsity is due to the fact that each row of A is associated with weighted integration over a digital line in the image I and therefore a vector of weights. Each weight value corresponds to a pixel in I and determined by the amount of intersection between the analytic line and the pixel itself. As a result, only weights that are located in entries that are associated with the pixels of the digital line, have non-zero values. Thus, each row in matrix A has $\leq cm = c\sqrt{N}$(where $c < 2$) non-zero entries. We note that even if we use a more accurate model based interpolation, where the line is given more significant width, the matrix A would remain sparse. This structure allows us to reduce memory consumption, to adaptively update a sparse data structure for A and to implement fast linear algebra operations. This idea is not new to the CT community. Moreover, for practical clinical data sizes in 3D helical uniform acquisition, the matrix A can be too large to hold in memory and must be computed on the fly. Also, its form is carefully determined from the geometry of the focal spots and detectors [14]. In this work we focus on the 2D model and in future work we plan to investigate whether in the 3D case our smaller adaptive sampling set can be stored in memory or computed on-the-fly.

We now explain, for the sake of completeness, our modification of the TVAL3 algorithm. For the constrained optimization problem such as (3.1), there are a number of methods that approach the original constrained problem by a sequence of unconstrained subproblems. One of them is the Quadratic Penalty Method [30]. This method puts a quadratic penalty term instead of the constraint in the objective function where each penalty term is a square of the constraint violation with multiplier. Though its wide usage, it requires to increase the multipliers to infinity so as to guarantee the convergence, which may cause the ill-conditioning problem, numerically. Another method concerning the constrained optimization problem is the Augmented Lagrangian method [31] (an augmented Lagrange method has been already used in CT reconstruction [32]). According to this method, the corresponding Augmented Lagrangian of the left-hand side minimization in (3.1), is given by

$$\begin{aligned} &L_A(w, u, v, \lambda, \mu, \beta) \triangleq \\ &\sum_{s=1}^{N} \Big(\|w_s\|_1 - \langle v_s, (Du)_s - w_s \rangle \\ &+ \frac{\beta_s}{2} \|(Du)_s - w_s\|_2^2 \Big) \\ &- \left\langle \lambda^t, Au - y \right\rangle + \frac{\mu}{2} \|Au - y\|_2^2 , \end{aligned}$$

where $w_s, v_s \in \mathbb{R}^2$, $\|w_s\|_1 \triangleq |w_s(1)| + |w_s(2)|$, $(Du)_s \triangleq \nabla U_{i(s),j(s)}$, $1 \leq s \leq N$, and the two vectors λ and v are the Lagrangian multipliers. To solve (3.1), the following Alternating Direction scheme is used: Denote the approximate minimizers of (3.1) at the kth inner iteration by $w^{(k)}$ and $u^{(k)}$. Then $w^{(k+1)}$ and $u^{(k+1)}$ can be attained by solving two separated subproblems. The first is the 'w-subproblem':

$$\begin{aligned} w^{(k+1)} &= \arg\min_w L_A(w, u^{(k)}) \\ &= \sum_{s=1}^{N} \left(\|w_s\|_1 - \left\langle v_s, \left(Du^{(k)}\right)_s - w_s \right\rangle \right. \\ &\qquad \left. + \frac{\beta_s}{2} \left\| \left(Du^{(k)}\right)_s - w_s \right\|_2^2 \right). \end{aligned}$$

Note that the 'w-subproblem' is separable with respect to each w_s, $1 \leq s \leq N$, and has a closed form solution [12]. The second subproblem, also known as the 'u-subproblem' is:

$$\begin{aligned} u^{(k+1)} &= \arg\min_u L_A(w^{(k+1)}, u) \\ &= \sum_{s=1}^{N} \left(\left\| w_s^{(k+1)} \right\|_1 - \left\langle v_s, (Du)_s - w_s^{(k+1)} \right\rangle \right. \\ &\quad \left. + \frac{\beta_s}{2} \left\| (Du)_s - w_s^{(k+1)} \right\|_2^2 \right) \\ &\quad - \left\langle \lambda^t, Au - y \right\rangle + \frac{\mu}{2} \|Au - y\|_2^2 . \end{aligned}$$

The 'u-subproblem' can be solved using a steepest decent method, but since this might be too costly for large scale problems, an aggressive 'one-step' of the steepest decent can be

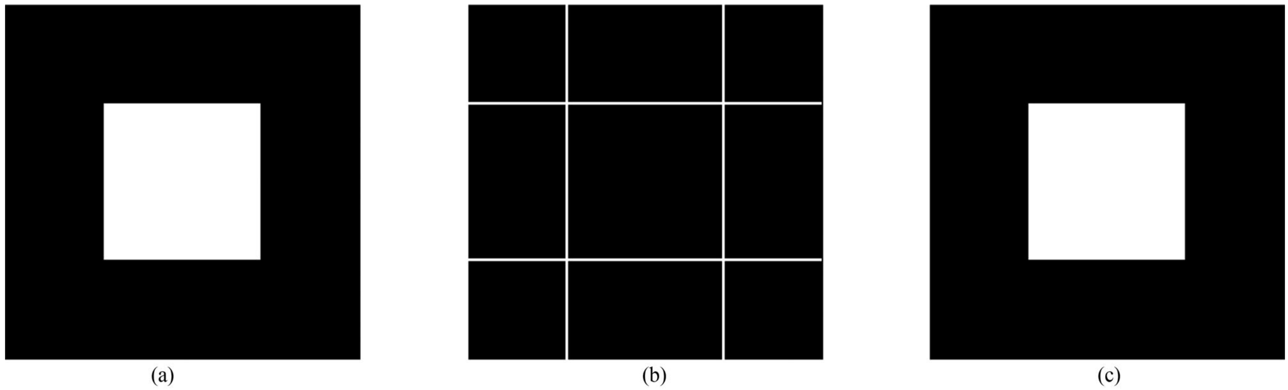


Fig. 3. 256×256 Square example. (a) The original image. (b) The sampling pattern of 8 line projections. (c) Reconstruction using the sampling pattern in (b).

computed as an iteration (see the details in [12]). After attaining $w^{(k+1)}$ and $u^{(k+1)}$, the multiplier updating is performed based on the analysis of [33], [34],

$$v_s^{(k+1)} = v_s^{(k)} - \beta_s \left(\left(Du^{(k+1)} \right)_s - w_s^{(k+1)} \right), \ 1 \le s \le N,$$
$$\lambda^{(k+1)} = \lambda^{(k)} - \mu (Au^{(k+1)} - y).$$

This second update step is exactly an example of where our modification accelerates significantly the TV minimization, by either storing and applying the matrix A in a sparse form or by computing and applying the sparse rows of A on the fly. Finally, choose new penalty parameters $\beta_s^{(k+1)} \ge \beta_s^{(k)}$ and $\mu^{(k+1)} \ge \mu^{(k)}$. The stopping criteria are one of the following: The quantities $|\nabla L_A(w^{(k)}, u^{(k)}, v^{(k)}, \lambda^{(k)}, \mu^{(k)}, \beta^{(k)})|$, $\sum_{s=1}^{N} \|(Du^{(k)})_s - w_s^{(k)}\|_2$ and $\|Au^{(k)} - y\|_2$ are sufficiently small, or the relative change $\|u^{(k+1)} - u^{(k)}\|_2$ is sufficiently small.

Inside the main loop of the Alternating Direction scheme, the number of rows in A is increased by a predetermined fixed constant M at each iteration (see Section IV), where the rows are line projections determined from the Ridgelet analysis of the approximation image.

IV. Adaptive Tomography Acquisition

Before presenting the details of the Adaptive Tomography Acquisition (ATA) algorithm, we first provide an instructive and useful example: Assume we had an access to an optimal 'oracle'. We then ask, how many line projections are needed as rows in the matrix A, such that the image of Fig. 3(a) can be reconstructed with high precision, by solving (3.1) ? The surprising result is that, equipped with an 'oracle', this image can be reconstructed with perfect precision, where the matrix A in (3.1) contains only 8 rows associated with 8 line projections. Thus, the numbers of samples satisfies $n = 0.000122N$, which is a tiny fraction of the size of the image, $N = 256 \times 256$.

This is achieved by selecting the unique four pairs of line projections that are the immediate neighbors of each of the four lines associated with the edges of the white square. Fig. 3(b) and (c) show the locations of the line projections and the reconstructed image, respectively.

The moral of this example, which correlates well with the theory in [5], is that during the acquisition process, we should try to adaptively sample the line projections that are aligned and centered around the edges of the image. Obviously, the image I is unknown and we do not have access to an 'oracle'. As we shall see, this is exactly where the multiresolution nature of the Ridgelet model is useful.

In the next sections, a detailed description of the main steps of the Adaptive Tomography Acquisition (ATA) algorithm is given (see also flowchart in Fig. 4).

A. Initialization

We create an initial sampling matrix A using a relatively small uniform set of line projections and sample the image I to obtain an initial observations vector y. The number of the initial line projections is determined relative to the image size. For example, for an image of size 256×256, we measured 8 equally spaced line integrals at eight uniformly spaced angles, which generates a total of 64 initial measurements that are about 0.1% of the image size. Fig. 7(a) illustrates this non-adaptive uniform sampling pattern for an image of size 256×256. We also initialize the current iteration counter to $k = 0$ and the initial approximation to I to $U^{(-1)} = 0$.

B. TV Minimization

The inputs for this step are: an updated sampling matrix A (with new additional rows that correspond to the newly acquired line projections), an observations vector y and the previous approximation $U^{(k-1)}$ as the initial guess. Then, we apply TV minimization step (3.1) or (3.2) (depend on expected noise levels) to compute $U^{(k)}$. Recall, that in our setup, the sparse nature of A enables to process large-scale images.

We have an option to select a tradeoff between reconstruction quality and performance. We do not necessarily need to completely solve the TV minimization problem by iterating the TV solver until it converges as in [12]. Instead, we apply only a fixed and limited number of iterations of the TV solver or alternatively, terminate the iterations using a less demanding stopping criterion and proceed to the next step. This speeds up this step in the algorithm, but in some cases, its effect on the next

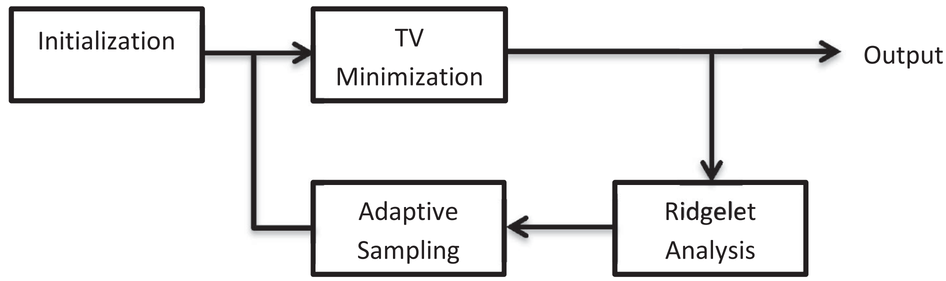


Fig. 4. Flowchart of ATA algorithm.

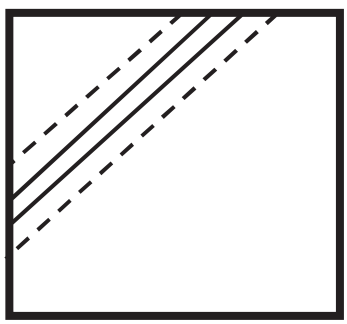


Fig. 5. Line integrals that were acquired per a significant Ridgelet coefficient: The external dashed lines correspond to the support of the Ridgelet and the inner lines are the sampled line projections.

ATA $(A, I, M, L, \varepsilon)$

Input: *A - Initial sampling matrix , I - Input image, M -Number of the Ridgelet coefficients subset considered in each iteration. L- Total number of line projections. ε - Stopping threshold.*

Output: *U - Reconstructed image.*

Notations: *u - 1D vector representation of U , A(I) – Sampled image I using the sampling matrix A.*

1. $U^{(-1)} \leftarrow 0.\ k \leftarrow 0.\ \delta \leftarrow 2\varepsilon$

2. **While** $\delta > \varepsilon$ **and** number of rows in $A < L$

- **2.1.** $y \leftarrow A(I)$
- **2.2.** *Obtain $u^{(k)}$ by solving problem (3.1) or (3.2), using A, y and $U^{(k-1)}$ as the initial guess .*
- **2.3.** *Compute the discrete Ridgelet coefficients of $U^{(k)}$.*
- **2.4.** *Find the M Ridgelet coefficients that have the largest absolute values that have not been sampled yet.*
- **2.5.** *For each of the newly found Ridgelet coefficients: add new rows to A associated with line projections, whose sampling approximates the value of the Ridgelet coefficient on the image I .*
- **2.6.** $\delta \leftarrow \left| u^{(k)} - u^{(k-1)} \right|$. $k \leftarrow k+1$.

3. Return $U^{(k-1)}$.

Fig. 6. The ATA Algorithm.

analysis step implies that more line projections are needed to be acquired in order to achieve the same reconstruction quality.

In any case, our adaptive acquisition process terminates if one of the following conditions: $\|U^{(k)} - U^{(k-1)}\|_2 \leq \varepsilon$ or number of rows in $A \geq L$, holds, where ε is a predetermined threshold and L is a limit on the total number of line projections that is permitted to be acquired.

C. Ridgelet Analysis

The inputs to this step is an improved approximation $U^{(k)}$ to I. We compute a subset of $U^{(k)}$ Ridgelet coefficients by the application of Radon transform followed by the application of wavelet transform, as shown in (2.4). Since in our application we only use Ridgelets for analysis, we do not need to use the inverse Ridgelet transform as in [16] and that simplifies the implementation. In practice, we realize that if we choose the number of angles to be a quarter of the image length, then our sampling scheme is sufficiently dense for high quality reconstruction, but not too much as to lead to subsequent unnecessary acquisition. Thus, for an images of size 256×256, we compute the Ridgelet coefficients for only 64 uniformly spaced angles, $\theta \in \{0, \pi/64, \ldots, 63\pi/64\}$, with 256 line projections per angle. For our experimental results, we applied the univariate discrete Haar wavelet [15] transform at each of the 64 angles to the 256 computed line projections. Then we subsampled the coefficients to avoid unnecessary subsequent acquisition. Specifically, we compute the Ridgelet coefficients $\alpha^{(k)}_{a,b,\theta}(x) \triangleq \langle U^{(k)}, \psi_{a,b,\theta} \rangle$ using the Haar wavelet function $\psi_{a,b}(x)$, in 4 different resolutions, where $a = 2^j,\ j = 0, \ldots, J_\theta$ and $J_\theta \in \{0, 1, 2, 3\}$ depends on the angle θ. In this case, the discrete sampling of Ridgelet coefficients is controlled by pairs (θ, J_θ) according to the patterns that appear inTable I.

D. Adaptive Sampling

The analysis of the Ridgelet coefficients $\alpha^{(k)}_{a,b,\theta}$, computed in the previous step, enables us to decide who are the new line projections that are added to A as new rows. Specifically, we choose these line projections to be associated with the M largest Ridgelet coefficients that have not yet been marked as sampled by the algorithm. The goal of the selected line projections is to approximate (2.3) where ψ is the Haar wavelet. In our experiments, we select $M = 0.1L$, where L is the total number of line projection the algorithm is allowed to acquire. Then, we sample the image I by the updated matrix A to obtain an updated

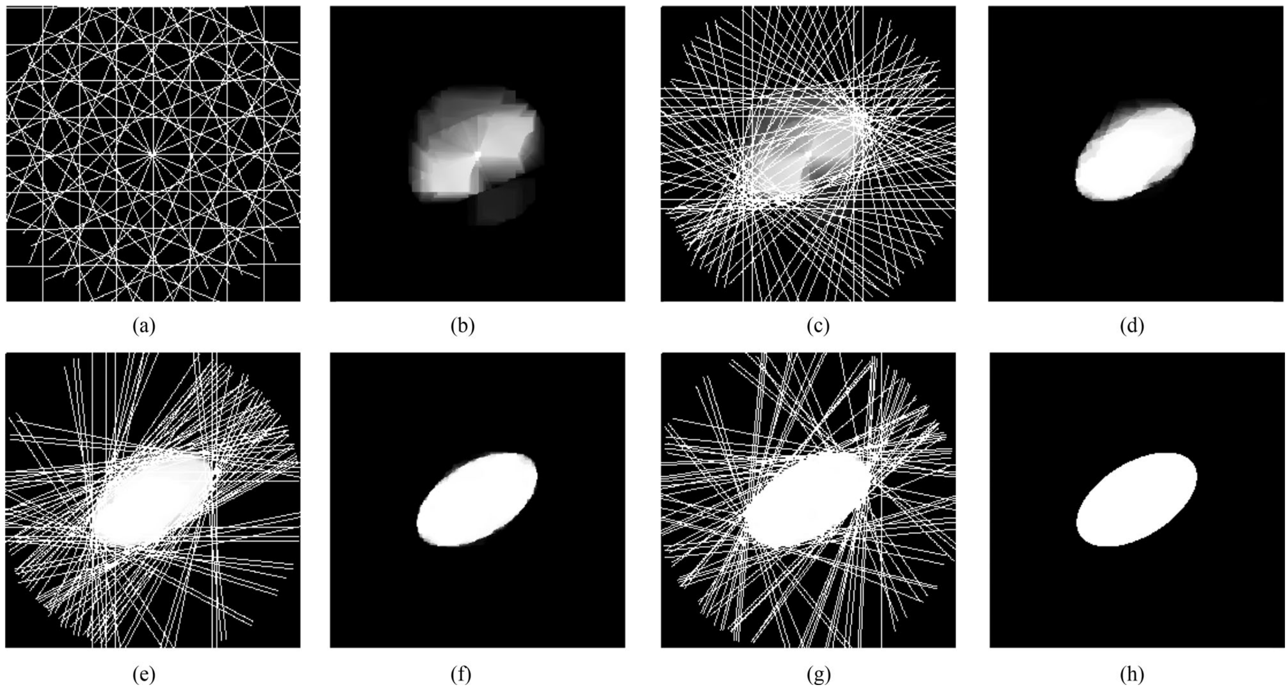


Fig. 7. Output from 4 iterations of the ATA algorithm on the 'Ellipse' image (h): (a): Initial non-adaptive uniform acquisition pattern. (b): Initial approximation $U^{(0)}$ from the non-adaptive acquisition. (c, e, g): Newly added line projections in iteration number 0,1 and 2 respectively. (d, f, h): The corresponding approximations $U^{(k+1)}, k = 0, 1, 2$.

TABLE I
SAMPLING PATTERN FOR 256×256 IMAGE.

Angle $(0 \le l < 8)$	$\frac{8l\pi}{64}$	$\frac{(8l+1)\pi}{64}$	$\frac{(8l+1)\pi}{64}$	$\frac{(8l+1)\pi}{64}$	$\frac{(8l+1)\pi}{64}$	$\frac{(8l+1)\pi}{64}$	$\frac{(8l+1)\pi}{64}$	$\frac{(8l+1)\pi}{64}$
J_θ	3	0	1	0	2	0	1	0

See Section IV-C for Explanation

observations vector y. In Fig. 5, we see an illustration of the support of the Haar Ridgelet function (dashed lines) and the associated line projections (the two inner lines) within its support.

Now, we look closer at the implication of using only two line projections to approximate the value of the Haar Ridgelet. Assume that the Ridgelet coefficient $\alpha_{a,b,\theta}^{(k)}(x)$ has not been marked as sampled yet but it is significant enough to be sampled at the current iteration. Let $R_I(\theta, \cdot)$ be the Radon transform of the unknown image I at a fixed angle θ. In this case, the two values of the line projections that we acquire are $R_I(\theta, b + a/4)$ and $R_I(\theta, b + 3a/4)$. These values are considered as the approximation: $a^{-1/2}(R_I(\theta, b + a/4) - R_I(\theta, b + 3a/4)) \approx CRT_I(a, b, \theta)$.

The ATA algorithm is described in Fig. 6. As already discussed in Section I, the ATA algorithm can be potential use the EBT machine hardware as an acquisition device, however it is not limited to this type of architechture. A general fan-beam or parallel machine can be adapted to support the proposed method: at each iteration, the algorithm determines who are the optimal rays that should be acquired in the next series of the gantry's rotations. It is important to clarify that in the presence of noise, it would be plausible to sample the same Ridgelet coefficient more than once (prescribed number of times).

E. Example and Analysis

In Fig. 7, we see the output from 4 iterations of the ATA algorithm on the 'Ellipse' image [shown in Fig. 7(h)]. Fig. 7(a) and (b) show the uniform acquisition pattern described in the initialization step and the resulted first approximation image $U^{(0)}$, respectively. Fig. 7(c), (e) and (g) show the newly sampled line projections associated with the next M largest unsampled Ridgelet coefficients in iterations 0,1 and 2. Fig. 7(d), (f) and (h) show the resulted approximation images $U^{(k+1)}$ $(k = 0, 1, 2.)$ produced by solving (2.1) after updating A with the new corresponding rows. We see that the algorithm quickly identifies the edges of the ellipse and only samples around them with more samples along the longer axis first. Moreover, initially, when the approximation $U^{(k+1)}$ is still blurred [see Fig. 7(b)], the algorithm finds from the Ridgelet analysis that it should first acquire line projections associated with Ridgelet coefficients from coarse resolution as seen in Fig. 7(c) and (e). Only after the

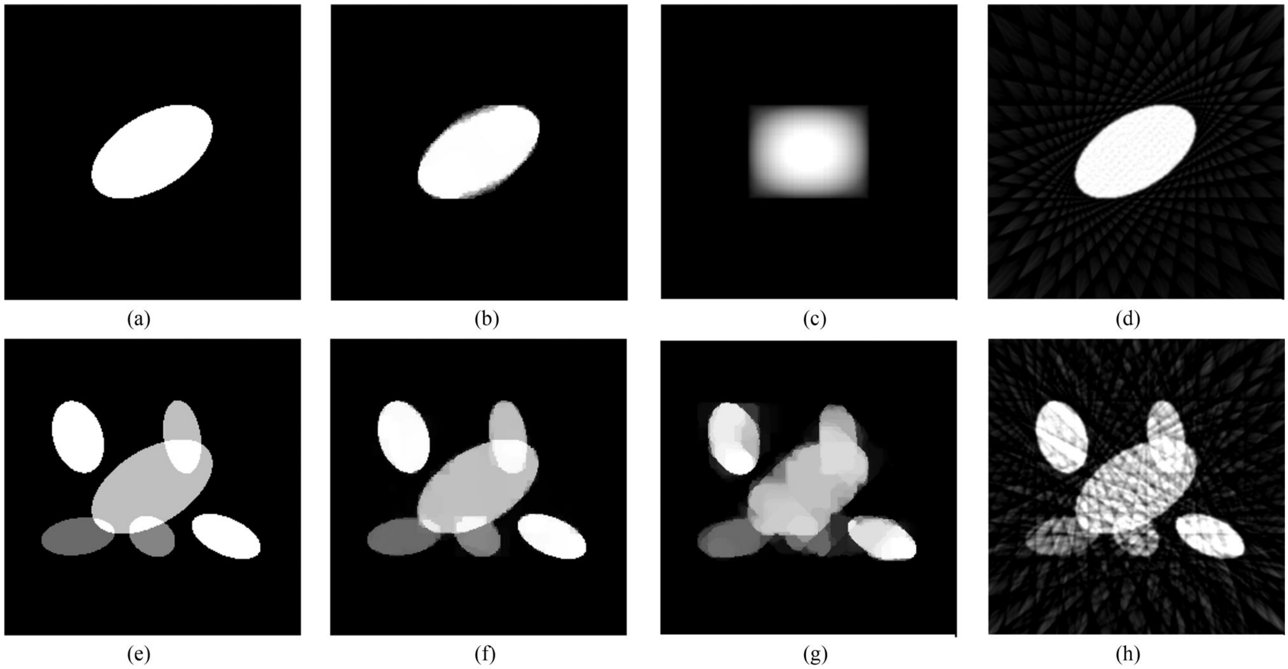


Fig. 8. PSNR comparison between ATA and non-adaptive acquisition methods. (a-d) Reconstruction results for the 'Ellipse' image. (a) ATA. 338 line projections. Perfect reconstruction. (b) NAS. 512 line projections. PSNR = 27.98 dB. (c) NAF. 512 line projections. PSNR = 17.40 dB. (d) FBP. 5120 line projections. PSNR = 23.88 dB. (e-h) Reconstruction results for the '6-ellipses' image. (e) ATA. 971 line projections. Perfect reconstruction. (f) NAS. 1024 line projections. PSNR = 29.73 dB. (g) NAF. 1024 line projections. PSNR = 21.92 dB. (h) FBP. 5120 line projections. PSNR = 19.64 dB.

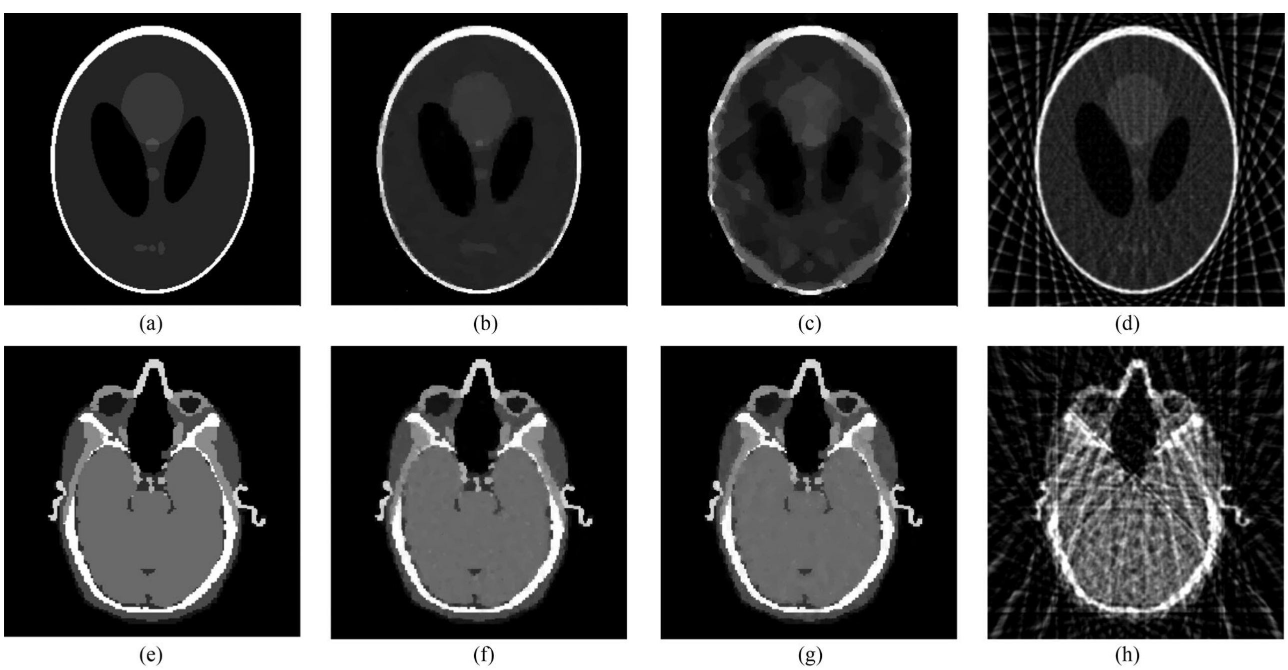


Fig. 9. PSNR comparison between ATA and non-adaptive acquisition methods. (a-d) are reconstruction results for the 'Shepp-Logan' image. (a) ATA. 1630 line projections. Perfect reconstruction. (b) NAS. 1792 line projections. PSNR = 26.44 dB. (c) NAF. 1792 line projections. PSNR = 19.43 dB. (d) FBP. 5120 line projections. PSNR = 18.04 dB. (e-h) are reconstruction results for the 'Zubal-Head' image. (e) ATA. 3834 line projections. Perfect reconstruction. (f) NAS. 4096 line projections. PSNR = 31.16 dB. (g) NAF. 4096 line projections. PSNR = 33.31 dB. (h) FBP. 5120 line projections. PSNR = 17.67 dB.

approximation contains sufficiently sharp edges [see Fig. 7(f)], Ridgelet coefficients from finer resolution become significant and the line projections associated with them are acquired as seen in Fig. 7(g). In summary, the ATA algorithm attempts to acquire only line projections that are around and aligned with edge singularities that are ordered by resolution.

Let us assume that the ATA algorithm manages to almost accurately identify the most significant Ridgelets coefficients of

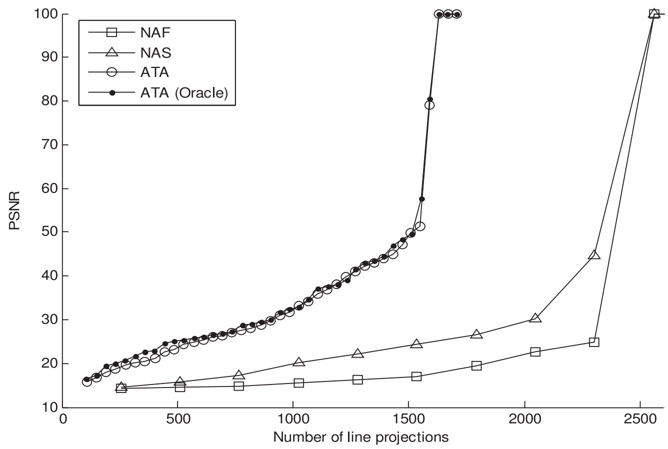


Fig. 10. PSNR comparison between the ATA algorithm, ATA with oracle and non-adaptive acquisition methods of NAS and NAF after their applications for the 'Shepp-Logan' image.

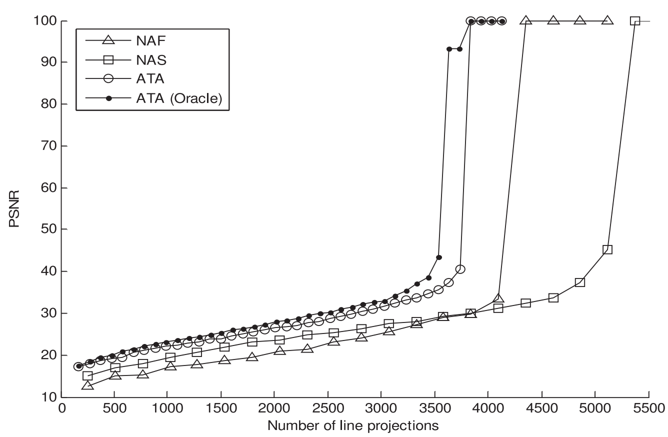


Fig. 11. PSNR comparison between the ATA algorithm, ATA with oracle and non-adaptive acquisition methods of NAS and NAF after their applications for the 'Zubal-Head' image.

the true image, I using only the approximate iterated image, $U^{(k)}$ (we validate this assumption empirecally in Section V). Under this assumption, we may analyze the rate of convergence using the theory discussed in Section II. Observe that a function belongs to the Radon 'smoothness' space of Definition 1, with a high value of 'smoothness' index α, if it has 'sparse' directional information, which decreases sufficiently fast as $a = 2^j$ is smaller. As an example, consider 'cartoon' functions which are piecewise constants over polygonal domains. At a fixed angle θ, the Radon transform $R_f(t, \theta)$ is a simple piecewise linear function of the variable t, with compact support and a bounded number of discontinuities, depending of the number of segments in the polygonal boundary. Assume the wavelet ψ has at least two vanishing moments, then for small values of a in (2.4), the Ridgelet transform $CRT_f(a, b, \theta)$ will be non-zero only in segments of total length $\leq ca$, where c is an absolute constant depending on the choice of ψ and the geometry of the polygonal boundary. Therefore, in the p-norm, for $a = 2^j, j < 0$, we get an estimate $\|CRT_f(2^j, \cdot, \theta)\|_p^p \leq c(f, \psi, p)2^{j/2}$. This implies that $f \in \dot{R}^{\alpha}_{p,q}(\mathbb{R}^2)$, for any 'smoothness' $\alpha > 0$. Therefore, from the Jackson estimate we may conclude that the adaptive algorithm will converge for these simple prototype functions at the rate $n^{-\alpha/2}$ for any $\alpha > 0$, which matches the perfect reconstruction results we obtain for these functions in this work. For more complex images, the convergence analysis is similar, but more involved. One needs to estimate the Radon smoothness of the function in order to understand the rate of convergence. This type of analysis of adaptive methods has been carried out for wavelet image compression by characterizing images as functions in Besov spaces [29].

V. Experimental Setup and Results

In this section we compare the ATA algorithm with known limited angle (non-adaptive) methods and also examine the quality of the estimate for the significant Ridgelet coefficients of the image I produced by our algorithm. We provide extensive experiment for phatom images and real clinical data. In both cases, we show that for a given number of line projections measured on the image I, ATA produces a significantly better approximation to I. We use the standard Peak Signal to Noise Ratio (PSNR),

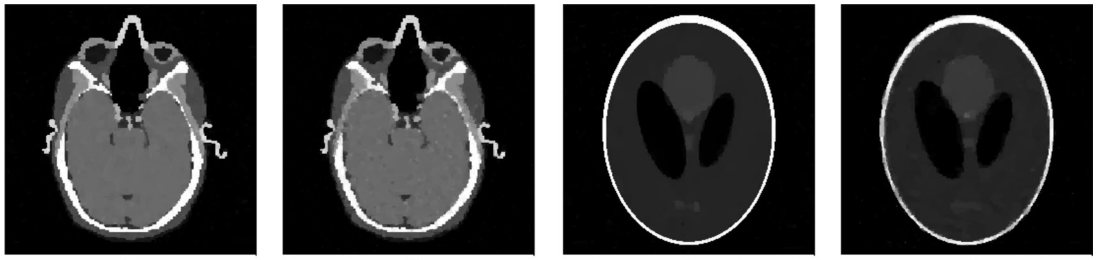
$$PSNR(I, U) \triangleq 10\log_{10}\left(\left(N^{-1}\sum_{i,j}|I_{i,j} - U_{i,j}|^2\right)^{-1}\right)$$

measured in dB, to quantify an approximation U to the image I where the image pixels take values in [0,1],.

A. Evaluated Methods

Given an $m \times m$ image, we prescribe a target of n samples. Denote $d = n/m$ (assuming $n \bmod m = 0$). We compare five acquisition and reconstruction methods:

1. *Filtered Back Projection (FBP):* For the FBP method we sampled $60 \times m$ line projections (regardless of the target limit), which are m equally spaced line integrals over the angles $0, \pi/60, \ldots, 59\pi/60$. We then used the MATLAB implementation ('iradon') to obtain a reconstructed image.
2. *Non Adaptive Equally Spaced (NAS):* We used equally spaced rotations and a fixed number of line projections at each angle such that the total number of line projections matched the prescribed budget. We then applied TV minimization to this sampled data. Specifically, $m/2$ (equally spaced) line projections were acquired over the angles $0, \pi/2d, 2\pi/2d, \ldots, (2d-1)\pi/2d$.
3. *Non Adaptive Uniform Fourier (NAF):* This method is used in [22]. In this mode, we uniformly select lines in the Fourier domain of the image and use Fourier coefficients on these lines as the entries of the sampling matrix A. Specifically, m Fourier coefficients were taken on the lines associated with the angles $0, \pi/d, 2\pi/d, \ldots, (d-1)\pi/d$.
4. *Adaptive Tomography Acquisition (ATA):* Our proposed adaptive algorithm (see Section IV).
5. *ATA using an oracle (ATA oracle):* The same algorithm used in 4, but now we allow ATA to use the Ridgelet analysis of I instead of using the Ridgelet analysis performed on the iterated approximated image.



(a) (b) (c) (d)

Fig. 12. PSNR comparison between ATA and NAS for the reconstruction of the ‘Zubal-head’ image at simulated incident photon count $\gamma_I = 1000000$. (a–b) Reconstruction results for the ‘Shepp-Logan’ (c–d) Reconstruction results for the ‘Zubal-Head’. (a) ATA. 4096 line projections. PSNR = 33.44 dB. (b) NAS. 4096 line projections. PSNR = 29.26 dB. (c) ATA. 4096 line projections. PSNR = 37.58 dB. (d) NAS. 4096 line projections. PSNR = 25.30 dB.

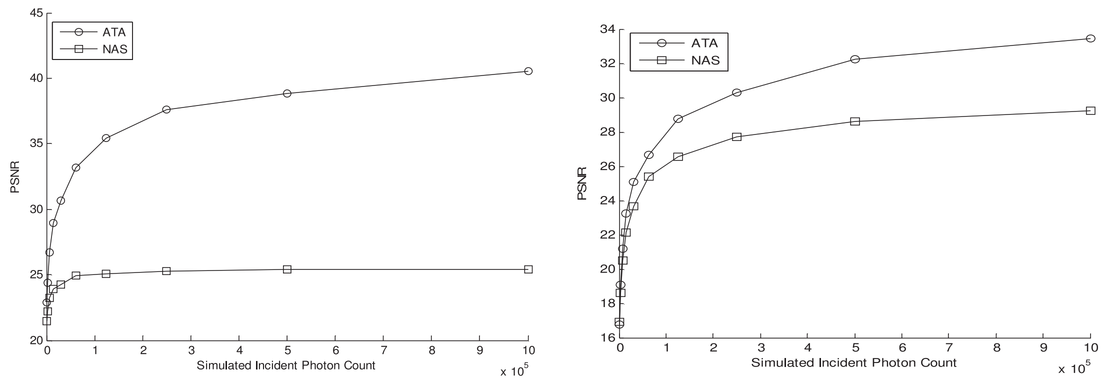


Fig. 13. PSNR comparison between ATA and NAS after their application to the ‘Shepp-Logan’ (left) and ‘Zubal-Head’ (right) for various simulated incident photon counts.

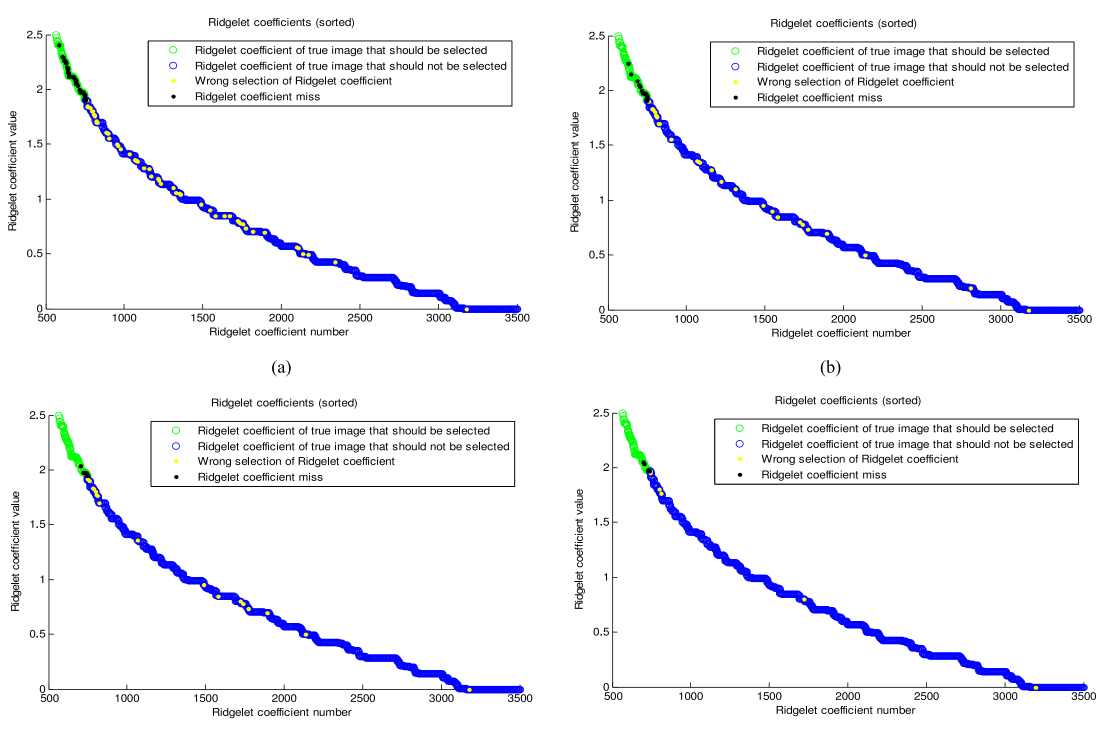


Fig. 14. Analysis of Ridgelet coefficients for the 128 × 128 ‘Shepp-Logan’ image: Plot of the Ridgelet coefficients sorted by their absolute value. (a–d) ROI of miss and wrong selection of the Ridgelet coefficients by ATA with a total number of 4, 8, 16 and 32 iterations, respectively.

TABLE II
MISSED RIDGELET COEFFICIENTS

Number of total iterations	Number of missed coefficients	Fraction of missed coefficients	Relative weight of missed coefficients
4	58	0.077	**0.022**
8	25	0.033	**0.009**
16	17	0.022	**0.006**
32	6	0.008	**0.002**

B. Phantom Data Experiments

We start by providing experiments on phantom images. The experiement conducted on 256×256 well known phantom test images: 'Shepp Logan' and 'Zubal head' [18] and additional two phantom images generated at our lab, the 'Ellipse' and '6-ellipses' images.The first tests are noise free. In these tests we solve the problem (2.1). In Figs. 8 and 9 we see that the ATA algorithm achieves perfect reconstruction using the smallest number of line projections, while the uniform limited angle, non-adaptive acquisition algorithms, NAS and NAF (that equipped with the same TV minimization solver) and FBP, achieve significantly lower image quality. Figs. 10 and 11 show a comparison betweePSNR values of methods 2-5 for different numbers of line projections for the 'Shepp-Logan' and 'Zubal Head' images, respectively. We see, that despite of not having the image I available at the time of acquisition, our algorithm manages to perform almost as good as an algorithm equipped with an 'oracle' that uses the Ridgelet analysis of the image I.

Next, we show results with simulated low dose as in [3]. For a selected parameter of incident photon count γ_I, the simulated detected photon counts $\tilde{\gamma}$, were chosen as Poisson distributed random variables with mean equal to $\gamma_I e^{-p}$, where p is a noiseless line projection. The simulated noisy projection,$\tilde{p}$, is then determined by $\tilde{p} = -\log(\tilde{\gamma}/\gamma_I)$. This time, in our iterations, we solve the problem (2.2), which provides better regularity for noisy data. In Fig. 12 we see a comparison of ATA and NAS using dose simulation for the 'Shepp-Logan' and 'Zubal-Head' images. We see that the image quality produced by ATA for the 'Shepp-Logan' is higher for a smaller number of line projections. We can also see a clear advantage of ATA over NAS on the 'Zubal Head' image under a dose simulation. In Fig. 13(a) and (b), we see plots of PSNR reconstruction values at various simulated dose levels for the Shepp Logan and Zubal Head images, respectively.

As described in Section IV, the ATA algorithm tries to estimate the Ridgelet coefficients of I by analyzing the approximation, at each iteration, of the image $U^{(k)}$. The next experiment is designed to evaluate the estimation error of the Ridgelet coefficients regarding I. Given a fix limited number of line projections, denoted by L, we examine the correctness of the Ridgelet coefficients that were selected by the ATA algorithm, as a function of the total number of iterations. Fig. 14(a) shows a plot of the Ridgelet coefficients computed on the 128×128 'Shepp-Logan' image, sorted by their absolute values. The graph is divided into two separated zones. The green zone represents the optimal selection of the Ridgelet coefficients, given a predetermined budget, from which the algorithm determines the L rays to be projected. The blue zone represents the residue (or tail) of the Ridgelet coefficients that should not be selected by the algorithm in the optimal case, i.e., if the ATA algorithm picked a subset of k coefficients during its run, green zone will sprawl over the first k Ridgelet coefficients of image I and the rest will be marked as blue. Fig. 14(b)–(e) show the same plot (ROI) for 4,8,16 and 32 total number of iterations, respectively. Black markers represent coefficients in the green zone that were missed by ATA. Yellow markers represent the coefficients that were wrongly selected by ATA instead of the correct ones in black. As we can see, the correctness of the selection process gets improved as the total number of iterations increases. Table II summarizes the results according to the following criterions: number of missed Ridgelet coefficients, relative fraction of missed Ridgelet coefficients (this is the number of missed coefficients / total number of the Ridgelet coefficients of the original image that should be selected) and relative weight of missed Ridgelet coefficients (which is the summation over the absolute values of the missed Ridgelet coefficients / the summation over the absolute values of the total number of the Ridgelet coefficients of the original image that should be selected).

C. Clinical Data Experiments

In this section we report results for real clinical data. Since we found the methods ATA and NAS to significantly outperform the method NAF and FBP, we provide results for ATA and NAS, only. The images in the experiments were randomly selected from two different datasets: one of brain images and the other of abdomen images [42]. Figs. 15 and 16 show the results for brain and abdomen images, respectively. Here, we include an additional reconstruction measure we name Relative Error (RErr):

$$RErr(I,U) \triangleq \frac{100\|U - I\|_F}{\|I\|_F}$$

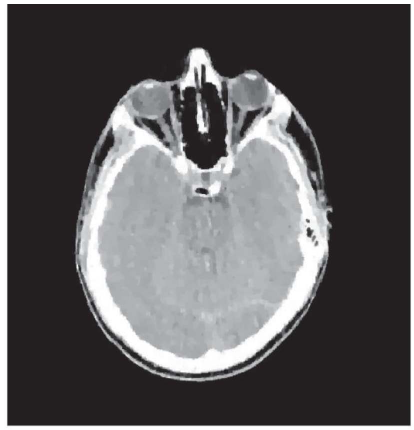
where $\|\centerdot\|_F$ stands for the Frobenius norm.

In all of the examples, ATA obtained better PSNR and RErr values than the NAS method. Yet, in Fig. 16 it is unclear whether the reconstructed pulmonary vessels in b.2 are better visualized than in b.3.

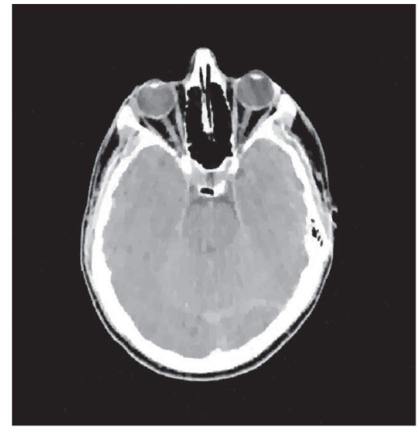
D. Computational Complexity

The running times of the ATA algorithm, simulated on Matlab, are about 5–10 slower than the non-adaptive methods (NAS, NAF) for the same number of line projections. This relates to the choice of M, the number of new number of line projections L,

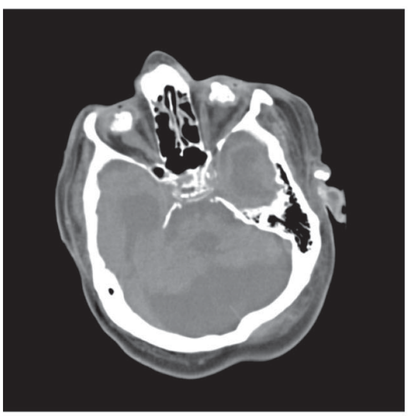
(a.1) Original Image – Brain 1 (512x512).



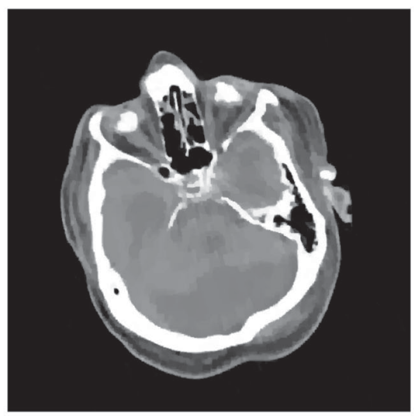
(b.1) NAS. 10485 line projections. PSNR: 27.97dB. RErr: 9.71%



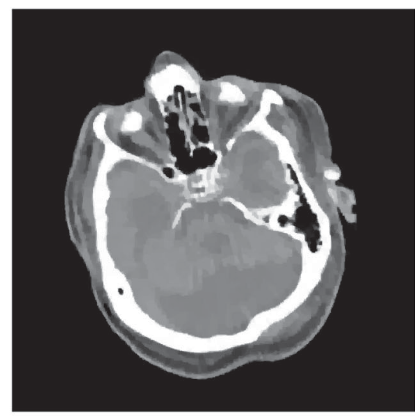
(c.1) ATA. 10485 line projections. PSNR: 31.12dB. RErr: 6.75%



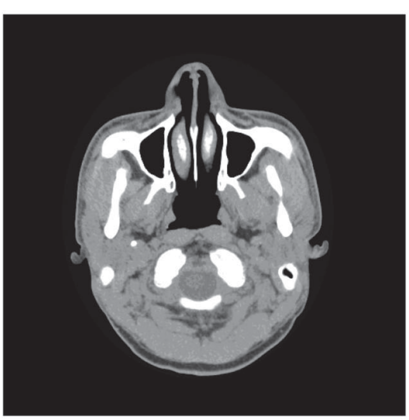
(a.2) Original Image – Brain 2 (512x512).



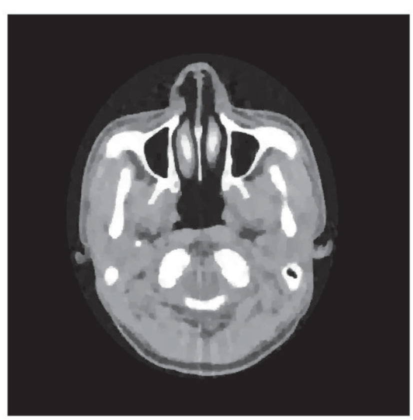
(b.2) NAS. 10485 line projections. PSNR: 29.86dB. RErr: 8.6%



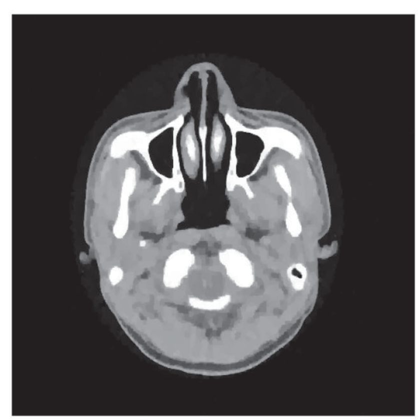
(c.2) ATA. 10485 line projections. PSNR: 31.60dB. RErr: 7.04%



(a.3) Original Image – Brain 3 (512x512).

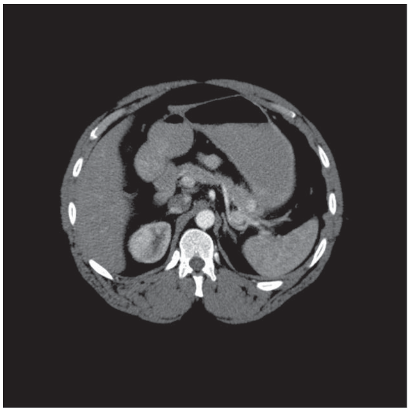


(b.3) NAS. 10485 line projections. PSNR: 30.18dB. RErr: 9.08%

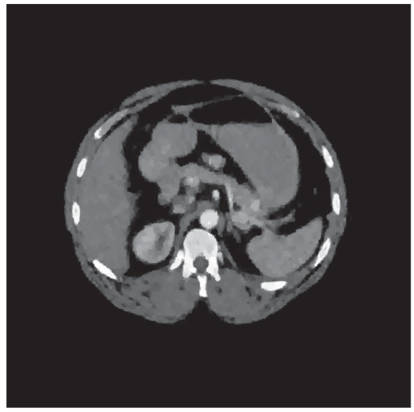


(c.3) ATA. 10485 line projections. PSNR: 32.86dB. RErr: 6.67%

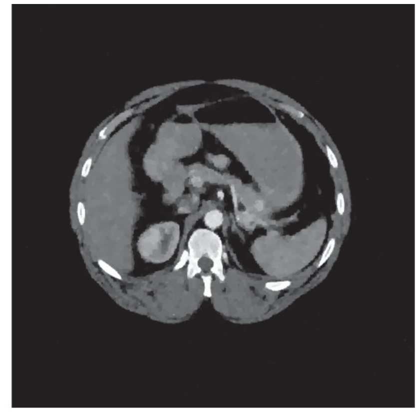
Fig. 15. Three examples for clinical data (brain images). (a.1-3) Original image. (a.1) Original Image – Brain 1 (512×512). (a.2) Original Image – Brain 2 (512×512). (a.3) Original Image – Brain 3 (512×512). (b.1-3) Reconstruction by NAS. (b.1) NAS. 10485 line projections. PSNR: 27.97 dB. RErr: 9.71%. (b.2) NAS. 10485 line projections. PSNR: 29.86 dB. RErr: 8.6%. (b.3) NAS. 10485 line projections. PSNR: 30.18 dB. RErr: 9.08%. (c.1-3) Reconstruction by ATA. (c.1) ATA. 10485 line projections. PSNR: 31.12 dB. RErr: 6.75%. (c.2) ATA. 10485 line projections. PSNR: 31.60 dB. RErr: 7.04%. (c.3) ATA. 10485 line projections. PSNR: 32.86 dB. RErr: 6.67%.



(a.1) Original Image – Abdomen 1 (512x512).



(b.1) NAS. 10485 line projections. PSNR: 30.23dB. RErr: 14.84%

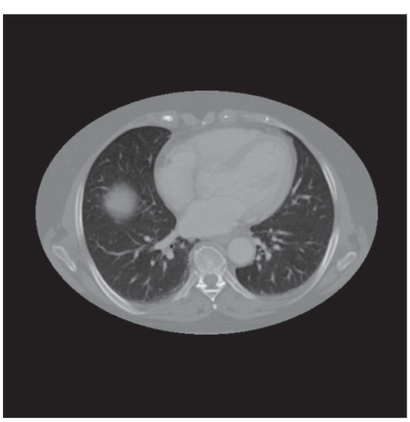


(c.1) ATA. 10485 line projections. PSNR: 31.49dB. RErr: 12.83%

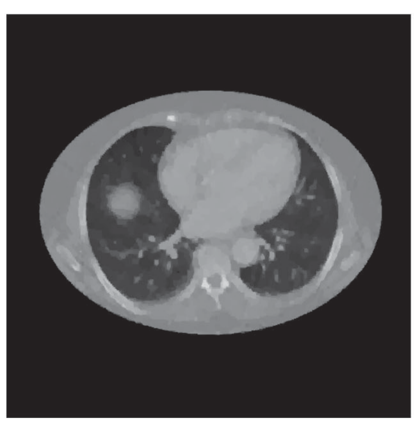
(a.2) Original Image – Abdomen 2 (512x512).

(b.2) NAS. 20971 line projections. PSNR: 33.22 dB. RErr: 7.75%

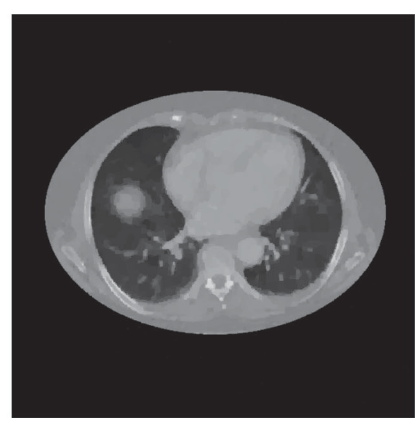
(c.2) ATA. 20971 line projections. PSNR: 34.53 dB. RErr: 6.62%



(a.3) Original Image - Abdomen 3 (512x512).



(b.3) NAS. 7864 line projections. PSNR: 34.10dB. RErr: 7.34%



(c.3) ATA. 7864 line projections. PSNR: 35.11dB. RErr: 6.52%

Fig. 16. Three examples for clinical data (Abdomen images). (a.1-3) Original image. (a.1) Original Image – Abdomen 1 (512×512). (a.2) Original Image – Abdomen 2 (512×512). (a.3) Original Image - Abdomen 3 (512×512). (b.1-3) Reconstruction by NAS. (b.1) NAS. 10485 line projections. PSNR: 30.23 dB. RErr: 14.84%. (b.2) NAS. 20971 line projections. PSNR: 33.22 dB. RErr: 7.75%. (b.3) NAS. 7864 line projections. PSNR: 34.10 dB. RErr: 7.34%. (c.1-3) Reconstruction by ATA. (c.1) ATA. 10485 line projections. PSNR: 31.49 dB. RErr: 12.83%. (c.2) ATA. 20971 line projections. PSNR: 34.53 dB. RErr: 6.62%. (c.3) ATA. 7864 line projections. PSNR: 35.11 dB. RErr: 6.52%.

the choice $M = 0.1L$ yields about 10 iterations, where the matrix A, in the kth iteration, contains $\sim 0.1kL$ rows. Solving these iterations is about 5.5 slower than solving the TV-minimization of order n only once. The rest of the running time of ATA is spent on the Ridgelet analysis computations that are performed at each iteration. Nevertheless, it is important to clarify that since the aim of this work is to provide a proof of concept, the experiments were conducted on Matlab. A possible real time implementation will be in C++, while utilizing parallel computation and GPUs (e.g., Ridgelet analysis step can be easily parallelized as well as the frequent matrix-vector multiplications that are executed during the TV minimization step). In this case, the overhead induced by our method may be negligible. Another possible improvement is the investigation of an efficient updating scheme that will enable the adaptive algorithm, at each iteration, to reconstruct a new approximation image based on the newly acquired line projections and the approximation from the previous iteration (without solving the whole TV minimization problem from the beginning).

VI. Conclusion and Future Work

In this paper we proposed a mathematical model for adaptive CT acquisition whose theoretical goal is to radically reduce dosage levels, while maintaining high quality reconstruction. We presented numerical simulations that demonstrate the potential of the mathematical model of adaptive acquisition and compared our results to known limited angle, non adaptive acquisition methods.

Our future research will focus on enhancement of the ATA algorithm. We believe this can be done by improving the contour analysis step which is currently based only on Ridgelet analysis. We are exploring the adaptation of edge detection methods e.g., Canny edge detector for the ATA algorithm. The author also developed a multiresolution version of the Gradient-based Hough Transform and already experienced promising results when incorporating it into the ATA algorithm.

In this work, our measure for dose reduction by ATA, which is the number of projected rays, does not simulate accurately the radiation dose level in real world. Dose in a CT scan depends on the machine's flux intensity, with lower flux intensity implies lower dose, but higher Poisson-type noise in the detected measurements. Therefore we wish to examine more realistic models beyond the simplistic model of the total number of line projections. By doing so, we also expect to utlize additional dimension of adaptation – the dose dimension which enable to control the intensity of each ray individually. In this way, the total measure of dose will be determined from the summation over the intensities of the selected rays and not just as the total number of projected rays. Furthermore, we plan to simulate true 3D scanning and add motion correction.

As discussed in Section V, running time is a drawback of the ATA algorithm. At each iteration, the ATA algorithm updates the sampling matrix A and solves the whole optimization problem again, which is an expensive computational task. We plan to explore an efficient updating scheme that allows the current approximated image to be modified solely by the newly acquired line projections and by the previous approximated image. This efficient scheme will eliminate the need to solve the whole problem from the beginning.

Lastly, it should be interesting to test other TV solvers such as [19], [39], [40], [41] and see if they (or modified versions of them) are better suited to the adaptive scheme proposed in this paper.

References

[1] D. Brenner, C. Ellison, E. Hall, and W. Berdon, "Estimated risks of radiation-induced fatal cancer from pediatric CT," *Amer. J. Roentgenology*, vol. 176, pp. 289–296, 2001.

[2] C. McCollough, M. Bruesewitz, and J. Kofler, "CT dose reduction and dose management tools: Overview of available options," *Radiographics*, vol. 26, pp. 503–512, 2006.

[3] A. Delaney and Y. Bresler, "Globally convergent edge-preserving regularized reconstruction: An application to limited-angle tomography," *IEEE Trans. Image Proc.*, vol. 7, no. 2, pp. 204–221, Feb. 1998.

[4] J. Thibault, K. Sauer, C. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multislice helical CT," *Med. Phys.*, vol. 34, pp. 4526–4544, 2007.

[5] E. Candès, "Ridgelets: Theory and applications," Ph.D. dissertation, Stanford Univ., Stanford, CA, USA, 1998.

[6] "MBIR aims to outshine ASIR for sharpness, CT dose reduction," AuntMinnie, May 18, 2010.

[7] D. Kolditz, Y. Kyriakou, and W. Kalender, "Volume-of-interest (VOI) imaging in C-arm flat-detector CT for high image quality at reduced dose," *Med. Phys.*, vol. 37, pp. 2719–2730, 2010.

[8] J. Moore, H. Barret, and L. Furenlid, "Adaptive CT for high-resolution, controlled-dose, region-of-interest imaging," in *Proc. IEEE Nucl. Sci. Symp. Conf.*, 2009, pp. 4154–4157.

[9] B. Rutt, "Scanned projection radiography using high speed computed tomographic scanning system," US Patent 4 573 179, Feb. 25, 1986.

[10] T. Quinto, "Singularities of the X-ray transform and limited data tomography in $\mathbb{R}^2$ and $\mathbb{R}^3$," *SIAM J. Math. Anal.*, vol. 24, pp. 1215–1225, 1993.

[11] "Compressive sensing resources." [Online]. Available: http://dsp.rice.edu/cs.

[12] C. Li, "Compressive sensing for 3D data processing tasks: Applications, models and algorithms," Ph.D. dissertation, , Rice Univ., Houston, TX, USA, Apr. 2011.

[13] TVAL3. [Online]. Available: http://www.caam.rice.edu/∼optimization/L1/TVAL3/.

[14] B. De Man and S. Basu, "Distance-driven projection and backprojection in three dimensions," *Phys. Med. Biol.*, vol. 49, pp. 2463–2475, 2004.

[15] S. Mallat, *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*, 3rd ed. New York, NY, USA: Academic, 2009.

[16] M. N. Do and M. Vetterli, "The finite ridgelet transform for image representation," *IEEE Trans. Image Proc.*, vol. 12, no. 1, pp. 16–28, Jan. 2003.

[17] J. Hsieh, *Computed Tomography, Principles, Design, Artifacts and Recent Advances*, 2nd ed. Hoboken, NJ, USA: Wiley, 2009.

[18] I. Zubal, C. Harrell, E. Smith, Z. Rattner, G. Gindi, and P. Hoffer, "Computerized three-dimensional segmented human anatomy," *Med. Phys.*, vol. 21, pp. 299–302, 1994.

[19] Y. Mao, B. P. Fahimian, S. J. Osher, and J. Miao, "Development and optimization of regularized tomographic reconstruction algorithms," *IEEE Trans. Image Proc.*, vol. 19, no. 5, pp. 1259–1268, May 2010.

[20] R. Nelson, "Thousands of new cancers predicted due to increased use of CT," Medscape news, Dec. 17, 2009.

[21] I. Shuryak, R. Sachs, and D. Brenner, "Cancer risks after radiation exposure in middle age," *J. Natl. Cancer Inst.*, vol. 102, pp. 1606–1609, 2010.

[22] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Dec. 2006.

[23] J. Frikel, "A new framework for sparse regularization in limited angle x-ray tomography," in *Proc. IEEE Conf. Biomed. Imag.*, 2010, pp. 824–827.

[24] E. Arias-Castro, E. J. Candès, and M. Davenport, "On the fundamental limits of adaptive sensing," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 472–481, Jan. 2013.

[25] E. Candès, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," *Multiscale Model. Simul.*, vol. 5, pp. 861–899, 2006.

[26] E. Candès and D. Donoho, "New tight frames of curvelets and optimal representations of objects with piecewise singularities," *Commun. Pure Appl. Math.*, vol. 57, pp. 219–266, 2003.
[27] G. Kutyniok and D. Labate, "Construction of regular and irregular shearlet frames," *J. Wavelet Theory Appl.*, vol. 1, pp. 1–10, 2007.
[28] D. Labate, W-Q. Lim, G. Kutyniok, and G. Weiss, "Sparse multidimensional representation using shearlets," *SPIE Proc. Wavelets XI*, vol. 5914, pp. 254–262, 2005.
[29] R. DeVore, "Nonlinear approximation," *Acta Numer.*, vol. 7, pp. 50–150, 1998.
[30] R. Courant, "Variational methods for the solution of problems with equilibrium and vibration," *Bull. Amer. Math. Soc.*, vol. 49, pp. 1–23, 1943.
[31] M. Fortin and R. Glowinski, *Augmented Lagrangian Methods: Application to the Numerical Solution of Boundary Value Problems*. Amsterdam, The Netherlands: Elsevier, 1983.
[32] S. Ramani and J. Fessler, "Parallel MR image reconstruction using augmented Lagrangian methods," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 694–706, Mar. 2011.
[33] M. Hestenes, "Multiplier and gradient methods," *J. Optim. Theory Appl.*, vol. 4, pp. 303–320, 1969.
[34] M. Powell, "A method for nonlinear constraints in minimization problems," in *Optimization*, R. Fletcher, Ed. New York, NY, USA: Academic, 1969, pp. 283–298.
[35] C. H. McCollough, "Principles and performance of electron beam CT," in *Medical CT and Ultrasound: Current Technology and Applications*, L. W. Goldman and J. B. Fowlkes, Eds. Madison,WI, USA: Adv. Med. Publishing, 1995, pp. 487–518.
[36] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Commun. ACM*, vol. 15, pp. 11–15 Jan. 1972.
[37] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.
[38] O. Barkan, J. Weill, S. Dekel, and A. Averbuch, "Adaptive compressed tomography sensing," in *Proc. IEEE Comput. Vis. Pattern Recog.*, 2013, pp. 2195–2202.
[39] J. A. M. Lorenzo, J. H. Juesas, and W. Blackwell, "A single-transceiver compressive reflector antenna for high-sensing-capacity imaging," *IEEE Antennas Wireless Propag. Lett.*, vol. 15, pp. 968–971, 2016.
[40] J. H. Juesas *et al.*, "Consensus-based imaging using ADMM for a compressive reflector antenna," in *Proc. 2015 IEEE Int. Symp. Antennas Propag. USNC/URSI Natl. Radio Sci. Meeting*, 2015, pp. 1304–1305.
[41] A. Molaei *et al.*, "Interferometric sounding using a compressive reflector antenna," in *Proc. 2016 10th Eur. Conf. IEEE Antennas Propag.*, 2016, pp. 1–4.
[42] [Online]. Available: http://www.osirix-viewer.com/resources/dicom-image-library/

Oren Barkan, photograph and biography not available at the time of publication.

Jonathan Weill, photograph and biography not available at the time of publication.

Shai Dekel, photograph and biography not available at the time of publication.

Amir Averbuch, photograph and biography not available at the time of publication.