# ESE 5932 Final Project

Kaushik Dutta
Ryan Friedman
Siqi Zhao

May 5, 2020

## 1 INTRODUCTION

Computed tomography (CT) is regarded as one of the most popular imaging modality for the diagnosis of diseases. It involves a series of x-ray projections that are taken at different angles around the patient, generating 2D slices of the 3D organs. The intensity of the CT images are expressed as *Hounsfield Unit* which is the measure of attenuation of the x-ray through different organs named after its inventor Dr. Godfrey Hounsfield. The two important parameters considered during the CT projections is the detector setting, i.e. the number of rays passing through the patient, and the number of angles at which the measurement is taken. The raw data is collected in the form of a sinogram and they are corrected to remove physical factors. Reconstruction is the mathematical process which transform this corrected projection data into tomographic images. In CT two types of reconstruction techniques are used: Analytical Reconstruction Algorithm and Optimization based Reconstruction Algorithm. For analytical reconstruction we design analytical solution of the inverse problem by understanding the physics of the scanning process. The optimization based reconstruction technique constructs the data fidelity term and minimise it by iterative regularization.

The Iterative Reconstruction algorithms are based on optimization based regularization. It starts with an initial approximation and compares the approximation to the measured data and optimizes the approximation in the direction of the measured data. The iterative method iterates until the approximation converges to the measured data or it is within the bounds of acceptable threshold of error. The two types of iterative reconstruction widely used are: Algebraic Reconstruction Techniques (e.g. ART, SART, SIRT), which reconstruct from the set of projections using linear algebra, and Statistical Reconstruction Techniques, (e.g. Expectation Maximization) which uses statistical models to reconstruct from projection data. Iterative methods do not need the formulation of the analytical solution. It is robust to minor changes in geometry as it handles missing information implicitly as compared to the analytical methods. Moreover it allows incorporation of prior knowledge to assist the reconstruction. The main challenge for the iterative reconstruction was huge computational time, but with parallel computation and use of GPUs more commercial scanner companies are shifting from the traditional analytical reconstruction to iterative based algorithms.

## 2 PROBLEM FORMULATION AND EXPERIMENTAL DETAILS

We designed our CT Reconstruction problem as an optimization problem. The data fidelity term is minimized using least square minimization using Total Variation regularization. The problem can be formulated as :

$$\underset{x}{min}\{||A(x) - d||^2 + 2\lambda \mathrm{TV}(x)\} \tag{1}$$

where $x \in \mathbb{R}^{m\times m}$ is the true image that is updated every iteration with the size $m \times m$. $d \in \mathbb{R}^{p\times q}$ is the sinogram where no of rays is denoted by $p$ and number of angles the measured is denoted by $q$. $A \in \mathbb{R}^{(m\times m)\times(p\times q)}$ is the forward operator that transforms an image to a sinogram image. $\lambda$ is the regularization parameter. $\mathrm{TV}(x) : \mathbb{R}^{m\times m} \to \mathbb{R}$ is the total variation functional that updates the image at each iteration.

For our experiment we have used the dimensions of the image as $256 \times 256$. The number of views were the measurement is taken is varied between 540, 270 and 90, while the total number of rays was fixed to 512. We have compared our designed algorithm FISTA with Total Variation with traditional Algebraic Reconstruction methods and with the reference ground truth image. For quantifying the algorithm's performance in terms of reconstruction we used Structural Similarity Index Metric (SSIM) and Mean Square Error.

## 3 METHODS

As we know from the physics of computed tomography, the measured sinogram image is convolved with several sources of noise, notably electronic noise and structural noise[4]. Hence, using an iterative method that handles noise is necessary. In formulating the CT reconstruction as an optimization problem, one of the important considerations is to find a regularization method for the ill-posed problem that preserves the edges in the true image, since identifying boundaries of objects (such as the boundary between normal and abnormal tissues) has high diagnostic value.

We decided to follow the fast gradient-based algorithm proposed by Beck and Teboulle[1], and the following section will discuss the specific implementation of their method for CT image reconstruction.

Specifically, we chose isotropic TV regularization[1]. The original publication of FISTA with TV[1] describes two types of TV regularization, anisotropic and isotropic TV, and the specific python package of choice, pyunloc[3], implements isotropic TV functionals. We decided to use the implemented TV function in the package and conducted some research on the differences

[1]In our presentation, we claimed that we were using the anisotropic TV because we falsely assumed that we were optimizing an 1D vector rather than a 2D image. For the 1D case, isotropic and anisotropic TV are essentially the same. The error was resulted from the fact that the majority of the group members are not in the field of imaging science as are unfamiliar with how matlab implicitly handles an image. And the error of assuming a 1D vector resulted in thinking we can use anisotropic TV as well as creating vertical artifacts in reconstructed images, we have corrected the errors in the implementation and rewritten this section. This has been a good learning opportunity for us both in terms of learning new knowledge as well as learning the process of gaining such knowledge from outside of the field. We also attempted to rewrite the method section to increase clarity, we would appreciate any feedback on this version of the description of the method.

of those two types of TV regularization functions. As being described in [2], anisotropic TV is known to favor horizontal and vertical structures for 2D images, and isotropic TV is shown to be a superior choice for regularization for a 2D image that preserves the off-axis edges.

The optimization problem can be formulated as:

$$\underset{x}{min}\{||A(x) - d||^2 + 2\lambda \mathrm{TV}(x)\} \tag{2}$$

where $x \in \mathbb{R}^{m\times n}$ is the true image with the size $\sqrt{m \times n}$. $d \in \mathbb{R}^{p\times q}$ is the sinogram image where $p : |rays| \times q : |views|$. $A \in \mathbb{R}^{m\times n}$ is the forward operator that transforms an image to a sinogram image. $\lambda$ is the regularization parameter. $\mathrm{TV}(x) : \mathbb{R}^{m\times n} \to \mathbb{R}$ is an istropic total variation functional, where

$$x \in \mathbb{R}^{m\times n}, \mathrm{TV_I}(x) = \sum_{i=1}^{m-1}\sum_{i=1}^{n-1}\sqrt{(x_{i,j} - x_{i+1,j})^2 + (x_{i,j} - x_{i,j+1})^2} \tag{3}$$

$$+ \sum_{i=1}^{m-1}|x_{i,n} - x_{i+1,n}| + \sum_{i=1}^{n-1}|x_{m,j} - x_{m,j+1}| \tag{4}$$

To solve the problem in eq.3, it requires a method that sovles non-differentiable optimization problems. We decided to use fast iterative shrinkage/thresholding algorithm(FISTA) as the algorithm to solve the optimization problem formulated for this project[1]. FISTA has two main advantages: first, it has a global convergence rate of $\mathcal{O}(\frac{1}{k^2})$ (The proof of convergence rate is detailed in [1]); second, FISTA is suitable for dealing with a non-smooth regularization function.

The general principle of the algorithm will be discussed in this paragraph and the pseudocode for the specific implementation will follow.

We discuss now how FISTA solved the optimization problem in eq.3. Let $f(x)$ denote the data fidelity term, it is clear that $f(x)$ is smooth and differentiable. The simple iterative method for solving the minimization of $f(x)$ is the gradient descent method:

$$x_k = x_{k-1} - t_k \nabla f(x_{k-1}) \tag{5}$$

As stated in [1], the optimization problem in eq.3 can be considered as a proximal regulation problem of $f(x_k)$ at $x_k$. Skipping the derivation, we have:

$$x_k = \underset{x}{\mathrm{argmin}}\{\frac{1}{2t_k}||x - (x_{k-1} - t_k\nabla f(x_{k-1}))||^2 + \lambda||x||_{tv}\} \tag{6}$$

Specifically, for FISTA, each update is:

$$x_k = p_{\mathrm{L}}(y_k) \tag{7}$$

$$t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2} \tag{8}$$

$$y_{k+1} = x_k + \frac{t_k - 1}{t_{k+1}}(x_k - x_{k-1}) \tag{9}$$

where $p_{\mathrm{L}} = \underset{x \in \mathbb{E}}{argmin}\{\frac{L}{2}||x - (y - \frac{1}{L}\nabla f(y)||^2 + ||x||_{tv}\}$

Compare to ISTA, FISTA uses the information from past 2 steps instead of 1 step to decide the new start $x_k$, This smarter choice of $x_k$ allows the proximal gradient approaches the minimum faster.

After establishing FISTA is a fast algorithm for solving the optimization problem in eq., we discuss the dual approach to solve for isotropic total variation. The optimal solution for the inner problem is:

$$x_k = P_{\mathrm{C}}(d - \lambda\mathcal{L}(p_k, q_k)) \tag{10}$$

where linear operator $\mathcal{L} : \mathbb{R}^{((m-1)\times n)}\mathbb{R}^{(m\times(n-1))} \to \mathbb{R}^{m\times n}$ is defined by

$$\mathcal{L}(\mathbf{p}, \mathbf{q})_{i,j} = p_{i,j} + q_{i,j} - p_{i-1,j} - q_{i,j-1} \tag{11}$$

$$i = 1, \cdots, m, j = 1, \cdots, n \tag{12}$$

and $P_{\mathrm{C}}$ is an orthogonal projection operator onto a feasible set to ensure smoothness. To apply the dual approach, we need to use the mapping from $(p, q) \to (r, s)$ to ensure the smoothness.

We found an existing package in python called pyunlocbox [3] to implement the algorithm. This package has all the required components for constructing the functions for data fidelity and regularization terms, the pseudocode is in algorithm1.

Now we discuss other implementation choices in this algorithm. First, we decided to use relative tolerance as our stopping criterion since it is easy to implement and measures the change in value relative to the size of the values themselves. Second, we used L-curve criterion to decide the regularization parameter, because it gives us a principled way to decide the proper amount of regularization without relying on the reference image.

# 4 RESULTS

To select the optimal amount of regularization, we tried $\lambda$ values ranging from 10 to $10^{-5}$ and created an L-curve. Using this criteria, we found that the optimal reconstruction is with $\lambda$ is $2 \times 10^{-4}$ for 540 views, $10^{-3}$ for 270 views, and $2 \times 10^{-4}$ for 90 views (Figure 1). The reconstruction from the full sinogram for all values of $\lambda$ is shown in Supplemental Figure S1. We compared the image reconstructed with FISTA to reconstructions with SART. We chose to implement SART as our stationary method because it is less prone to overfitting than ART and converges faster than SIRT. We assessed image quality using two metrics, the structural similarity (SSIM) index and the mean squared error (MSE). We computed these metrics on the full image and on a defined region of interest (ROI). We set our ROI to cover pixel rows 175-220 and pixel columns 100-140 because the image has complex shapes in this region. The reconstructed images for 540, 270, and 90 views are shown in Figures 2, 3, and 4, respectively. The SSIM metrics are shown in Table 1 and the MSE metrics are shown

**Algorithm 1** TV-FISTA

**Input** $d$, $A$, $\lambda$, $maxiter$, $tol$

$(r_0, s_0) = (p_0, q_0) = (0_{(m-1)\times n}, 0_{m\times(n-1)})$, $t_1 = 1$

**if** $k \leq maxiter$ **then**

*Forward*

  **Compute** $x_f = \frac{2}{\lambda} A^T(A(x_{k-1}) - d)$

*Backward*

  $x_k = x_f - \frac{1}{\lambda} \boldsymbol{\nabla} \cdot (r_k, s_k)$

  $obj_k = \frac{1}{2}||x_f - x_k||^2 + \lambda||\nabla x_k||^2$

  $tol_k = \frac{|obj_{k-1} - obj_k|}{obj_k}$

*Update projector*

  $r_k = r_{k-1} - \frac{1}{8\lambda}\frac{\nabla x_k}{\partial p_k}$

  $s_k = r_{k-1} - \frac{1}{8\lambda}\frac{\nabla x_k}{\partial q_k}$

*FISTA*

  $t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2}$

  $p_k = \frac{r_k}{\max\{1, \sqrt{r_k^2 + s_k^2}\}}$

  $q_k = \frac{s_k}{\max\{1, \sqrt{r_k^2 + s_k^2}\}}$

  **if** $tol_k \leq tol$ **then**

    **Stop**

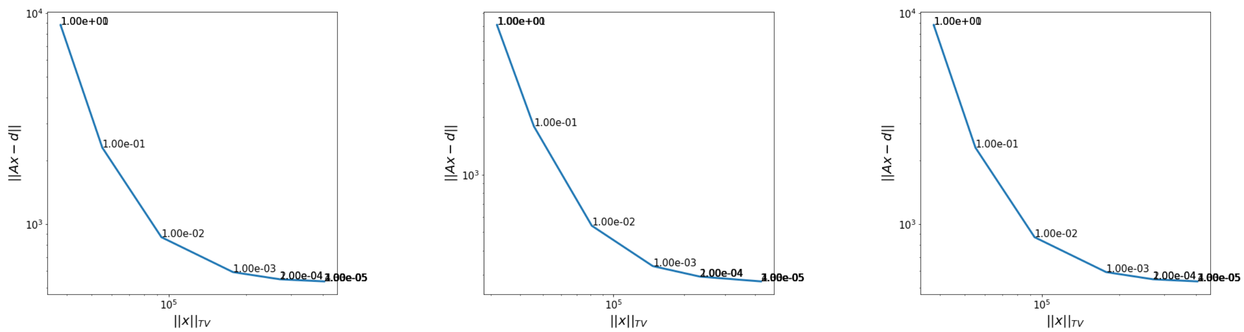
  $k = k + 1$

in Table 2. In virtually every case, FISTA outperforms SART. However, SART slightly outperforms FISTA in the following cases:

- The SSIM on the full image is higher with SART than with FISTA, although this is not the case for MSE.
- The reconstruction of the ROI from 270 views is slightly better with SART than with FISTA.

| Number of views | FISTA | FISTA ROI | SART | SART ROI |
|---|---|---|---|---|
| 540 | 0.860 | 0.968 | 0.885 | 0.855 |
| 270 | 0.888 | 0.837 | 0.877 | 0.849 |
| 90 | 0.889 | 0.843 | 0.792 | 0.741 |

Table 1: SSIM of image reconstructions.




(a) Full sinogram (b) 270 views (c) 90 views

Figure 1: L-curves for FISTA reconstruction of image.

| **Number of views** | **FISTA** | **FISTA ROI** | **SART** | **SART ROI** |
|---|---|---|---|---|
| 540 | 33.18 | 40.02 | 40.91 | 244.75 |
| 270 | 38.44 | 254.20 | 43.62 | 251.27 |
| 90 | 39.08 | 251.92 | 80.66 | 397.23 |

Table 2: MSE of image reconstructions.

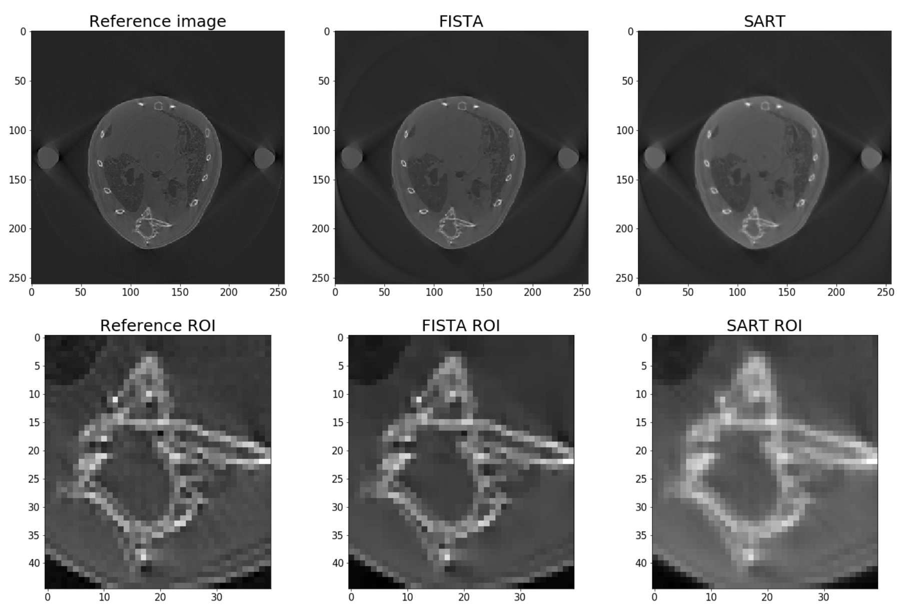



Figure 2: Reconstructed image using FISTA and SART from the full sinogram.

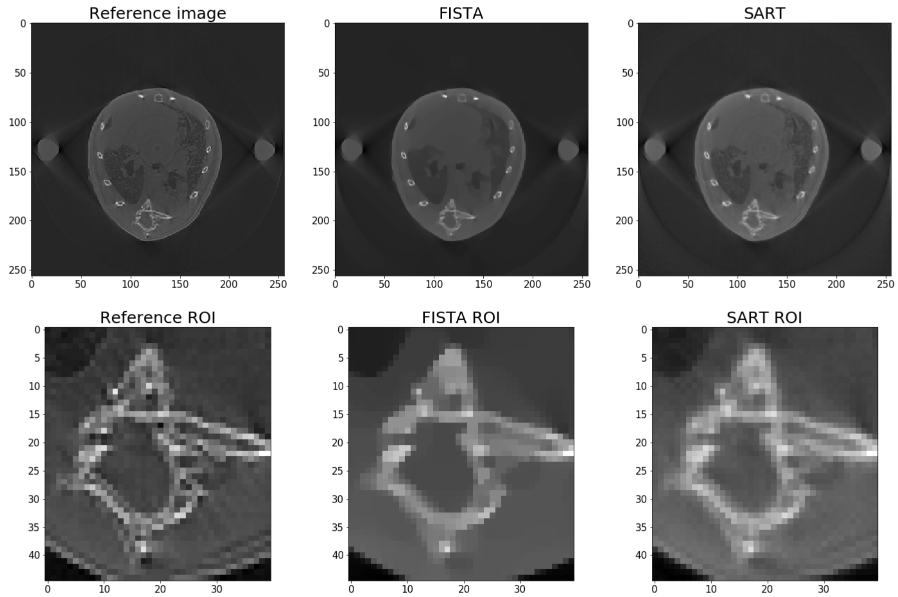



Figure 3: Reconstructed image using FISTA and SART from the 270 views.

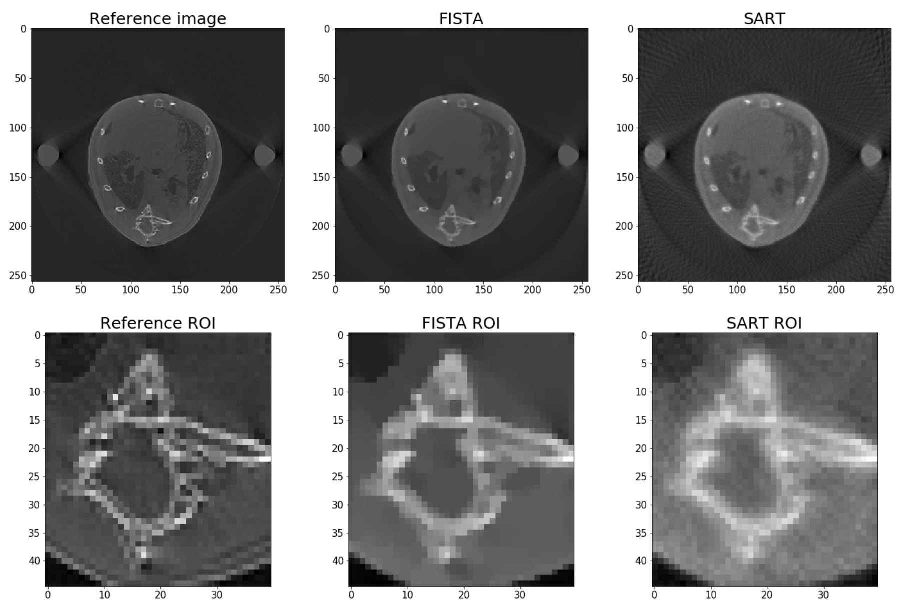



Figure 4: Reconstructed image using FISTA and SART from the 90 views.

# 5 DISCUSSION AND CONCLUSIONS

In general, FISTA outperforms SART at reconstructing the image. By comparing reconstructions on the three different sinograms, we see that the SSIM and MSE is substantially more stable with FISTA (Tables 1, 2). Therefore, FISTA creates more robust reconstructions with less data. Although the global SSIM with the full sinogram is slightly higher for SART, this is because FISTA creates a sharper halo effect around the actual image (Figure 2) likely due to the total variation regularization. Indeed, the SSIM over the ROI is better with FISTA than with SART and we can see that the colors and shading more closely resemble the reference image. The FISTA reconstruction also has sharper edges compared to the SART reconstruction. This is especially apparently as fewer views are used (Figure 3). Finally, FISTA is much more tolerant to noise. In the 90 view reconstruction, the SART reconstruction contains wavy patterns whereas the FISTA reconstruction does not contain any noise artifacts (Figure 4). As a result, the FISTA reconstruction over the ROI does a much better job preserving the complex structures. In addition, the SART reconstruction appears to have a much noisier, non-uniform background in the ROI.

To summarize, we conclude that FISTA creates more robust reconstructions with less data, creates sharper edges, and is more tolerant to noise than stationary iterative methods such as SART.

# 6 AUTHOR CONTRIBUTIONS

All authors contributed equally to this project. KD generated the forward operators and sinograms, chose image quality metrics, and wrote the Introduction and Problem Formuation. RF implemented SART, wrote the final code, and wrote the Resuts and Discussion sections. SZ wrote the pesudocode of the FISTA algorithm, developed prototype FISTA code, and wrote the Methods section.

# 7 CODE AVAILABILITY

All code used for this project, with instructions for use, is available at `https://github.com/rfriedman22/ese5932_final_project/`.

## REFERENCES


[1] Amir Beck and Marc Teboulle. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Trans. Image Process.*, 18(11):2419–2434, November 2009.

[2] Laurent Condat. Discrete total variation: New definition and minimization. *SIAM J. Imaging Sci.*, 10(3):1258–1290, January 2017.

[3] Michaël Defferrard, Rodrigo Pena, and Nathanaël Perraudin. Pyunlocbox: Optimization by proximal splitting.

[4] Xinhui Duan, Jia Wang, Shuai Leng, Bernhard Schmidt, Thomas Allmendinger, Katharine Grant, Thomas Flohr, and Cynthia H McCollough. Electronic noise in CT detectors: Impact on image noise and artifacts. *AJR Am. J. Roentgenol.*, 201(4):W626–32, October 2013.

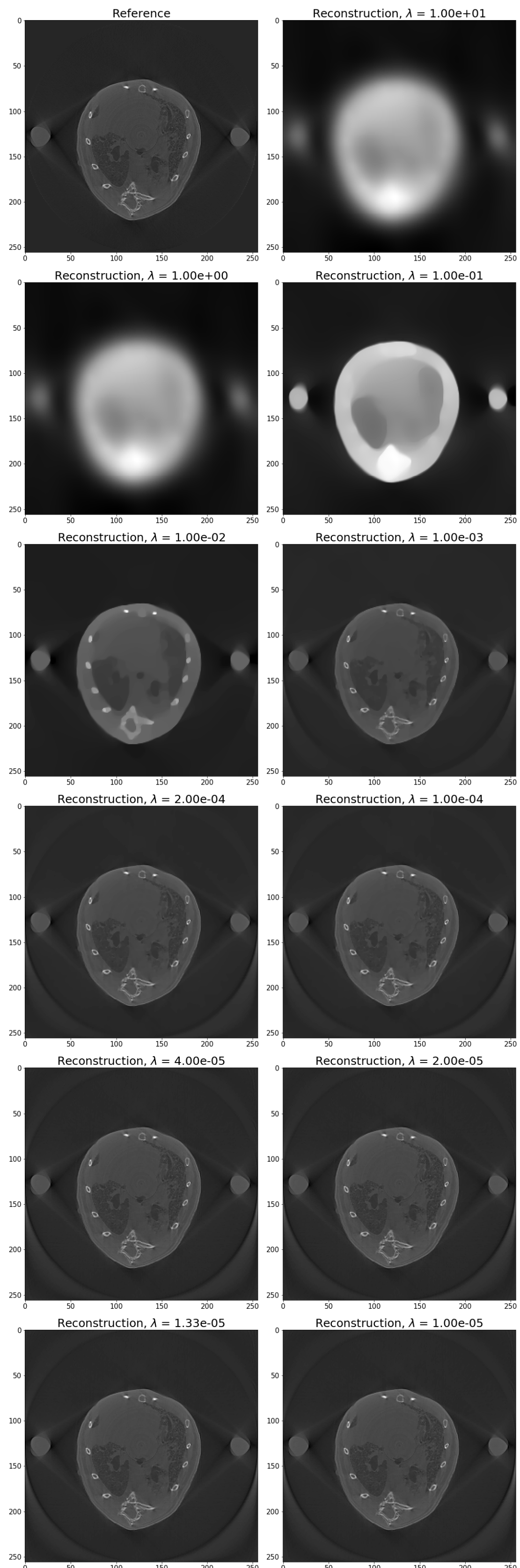



Figure S1: FISTA reconstruction from full sinogram with different values of $\lambda$.