BME 593

Final Project

Tommy Peng

Background

In image processing, for a typical forward problem $b = Hx$, where b is vector of the measurements of dimensions 1xM, H is the forward operator matrix with dimensions MxN and x is the vector of dimensions 1xN of parameters we want to estimate, the inverse problem can be thought of as given b, compute an estimate of x. In this project, b can be thought of as the measured sinogram and x is the image. While the inverse problem is seemingly simple to solve via $H^{-1}b = x$, the computational complexity of finding the inverse of the typically large forward operation H is makes the problem a difficult one. Furthermore, the reconstructed image may be noisy or blurred. Iterative optimization methods can be used to bypass this need to invert the forward operator, and to progressively achieve better estimates of x, in this case the image, at each iteration. In general gradient descent methods are a popular class of techniques which are used to approach optimization of convexly framed image reconstruction problems.

In this project, the goal was to design an optimization problem that includes a regularization strategy that is both reasonable and non-smooth for a X-ray computed tomography (CT) problem. In particular, the proximal gradient method (PGM) is a popular technique to solve non-smooth optimization problems within the realm of image reconstruction, and it will be used here to explore this problem in CT. PGM is a general strategy for minimization of convex optimization problems. The upside of PGM is that it allows for minimization of a convex problem which has a non-smooth component. This is an improvement from other gradient methods as it does not require the objective function to be smooth and continuously differentiable. The following is the description of the base problem:

$$\min\{F(x) = f(x) + g(x) \mid x \in R^n\}$$

$$Where\ F(x)\ is\ the\ object\ or\ loss\ function,$$

$$f(x) is\ convex\ and\ smooth, and\ g(x)\ is\ convex\ and\ non-smooth.$$

The following definitions are required:

$$Def: Prox_f(u) = \underset{x}{\operatorname{argmin}}[f(x) + \frac{1}{2}||u - x||^2]$$

$$Approximation\ of\ f(x) near\ x = y: q_t(x, y) = f(y) + \nabla f(y)^T(x - y) + \frac{1}{2t}||x - y||^2$$

The PGM can be defined as the following:

$$x_{k+1} = \underset{x}{\operatorname{argmin}}\{g(x) + \frac{1}{2t_k}||x - (x_k - t_k \nabla f(x_k)||^2$$

$$Where\ t_k\ is\ from\ the\ k^{th}\ approximation\ of\ f(x)$$

$$x_{k+1} = prox_{t_k g}(x_k - t_k \nabla f(x_k))$$

As the problem is set up, the important components that need to be determined are the data fidelity term $f(x)$ and the penalty term $g(x)$. In this case, choices for both terms are flexible, but need to be reasonable. Furthermore, $g(x)$ must me convex but non-smooth. These constraints on the problem setup will be explored in the methods section. Furthermore, the traditional PGM methods tend to be quite slow on convergence, so an accelerated proximal gradient method will be used during computation. The algorithm chosen for this project is the Fast Iterative Shrinkage Thresholding Algorithm (FISTA), which has been used and studied substantially in the fields of computed tomography and image recovery (Beck and Teboulle, 2011; Guerquin-Kern et al., 2011; Xu et al., 2016).

Methods

At the center of PGM is the objective function that we are trying to minimize, summarized below:

$$\min\{F(x) = f(x) + g(x) \mid x \in R^n\}$$

$$Where\ F(x)\ is\ the\ object\ or\ loss\ function,$$

$$f(x) is\ convex\ and\ smooth, and\ g(x)\ is\ convex\ and\ non-smooth.$$

Where $f(x)$ is typically a data fidelity term, and $g(x)$ is typically a penalty regularization term. For this project, the data fidelity term was chosen to be the squared L_2 norm between the true image and the reconstructed estimate of the image at each iteration, symbolically:

$$f(x) = ||b - Hx||^2$$

$$where\ x\ is\ the\ true\ image, H\ is\ the\ forward\ operator, and\ b\ is\ the\ measured\ data$$

In terms of this project, x is the unknown image that we want to reconstruct, H is the operator that is given to us as 'forward.m' and b is the sinogram that we decide to use. This data fidelity term was chosen as it is strictly convex, continuously differentiable, and commonly used in various subfields of image processing such as denoising and deblurring (Nikolova, 2002).

The more intricate term to be chosen is the penalty term $g(x)$. There are many possible terms to choose from the literature, many of them include an L_1 norm to ensure that the regularization term is at least non-smooth. Studies have shown that the L1-norm terms are better than other sparsity penalty terms at promoting the sharpness of the edges, an important factor to consider when reconstructing medical images for clear definition of the structures inside the body (Weller, 2012). What resides inside the L_1 norm term is also debated. The total variation term was chosen for this exercise due to its popularity and ability to improve the outlines of structures in CT images:

$$g(x) = \lambda TV(x)$$

The TV term is defined as the following in deblurring and denoising literature (Beck and Teboulle, 2009).

$$\mathbf{x} \in \mathbb{R}^{m\times n}, \quad \mathrm{TV}_{l_1}(\mathbf{x}) = \sum_{i=1}^{m-1}\sum_{j=1}^{n-1}\{|x_{i,j}-x_{i+1,j}|+|x_{i,j}-x_{i,j+1}|\} + \sum_{i=1}^{m-1}|x_{i,n}-x_{i+1,n}| + \sum_{j=1}^{n-1}|x_{m,j}-x_{m,j+1}|,$$

This term was also used in the FISTA-TV algorithm (Xu et al., 2016). λ is just a regularization constant which controls how much the regularization term contributes to the overall cost function. Since this regularization function involves a L_1 norm, it is both convex and non-smooth, therefore it satisfies both conditions required for this project. Total variation has also been identified in class as a sparsity promoting penalty term. The motivation behind this regularization is because it has been noted to be appropriate for denoising and deblurring (Beck and Teboulle, 2009). The objective function can then be summarized as:

$$F(x) = ||b - Hx||^2 + \lambda||x||_{TV}$$

In general, this project will follow the FISTA algorithm outlined by Beck and Teboulle. Beck and Teboulle's solution to the total variation regularization problem involves segments of pseudocode. The first component, fast gradient projection (FGP), is for solving the pure denoising case of the problem where $H \equiv I$. The following are definitions required for the FGP algorithm.

$$\boldsymbol{Operator\ L}: R^{(m-1)\times n}\times R^{m\times(n-1)} \rightarrow R^{m\times n}$$

$$\boldsymbol{L}(\boldsymbol{p},\boldsymbol{q})_{i,j} = p_{i,j} + q_{i,j} - p_{i-1,j} - p_{i,j-1}; \quad i = 1,\dots,m, j = 1,\dots,n$$

$$where\ p_{0,j} = p_{m,j} = q_{i,0} = q_{i,n} \equiv 0, for\ every\ i = 1,\dots,m\ and\ j = 1,\dots,n$$

$$\boldsymbol{Operator\ L^T}: R^{m\times n} \rightarrow R^{(m-1)\times n}\times R^{m\times(n-1)}$$

$$\boldsymbol{L^T}(\boldsymbol{x}) = (\boldsymbol{p},\boldsymbol{q}), \qquad where\ \boldsymbol{p} \in R^{(m-1)\times n}\ and\ \boldsymbol{q} \in R^{m\times(n-1)}$$

$$p_{i.j} = x_{i,j} - x_{i+1,j}; \quad i = 1,\dots,m-1, j = 1,\dots,n$$

$$q_{i.j} = x_{i,j} - x_{i,j+1}; \quad i = 1, \dots, m, j = 1, \dots, n-1$$

$$\boldsymbol{Operator\ P_c}: R^{m\times n} \to R^{m\times n}$$

$$\boldsymbol{P_c}(x)_{i,j} = \begin{cases} -10 & x_{i,j} < -10 \\ x_{i,j} & -10 \le x_{i,j} \le 400 \\ 400 & x_{i,j} > 400 \end{cases}$$

$$\boldsymbol{Operator\ P_p}: (R^{(m-1)\times n}, R^{m\times(n-1)}) \to (R^{(m-1)\times n}, R^{m\times(n-1)})$$

$$\boldsymbol{P_P}(\boldsymbol{p}, \boldsymbol{q}) = (\boldsymbol{r}, \boldsymbol{s})$$

$$r_{ij} = \frac{p_{ij}}{\max\{1, |p_{ij}|\}}\ and\ s_{ij} = \frac{q_{ij}}{\max\{1, |q_{ij}|\}}$$

The numbers were chosen for operator $\boldsymbol{P_c}$ based on the range of the high-quality reference image, as that serves as an approximate reference on what the values should be for this kind of image. This threshold can be found using any similar images of CT, and serve as a tool to restrict the values that the image data can take on so that the numbers remain appropriate.

The operators were implemented as *tp_l.m*, *tp_lt.m*, *tp_pc.m* and *tp_pp.m* respectively. With the above definitions for operators, we can finally describe the FGP algorithm (implemented as *tp_fgp.m*).

$$\boldsymbol{FGP}(\boldsymbol{b}, \boldsymbol{\lambda}, \boldsymbol{N}), b\ is\ the\ sinogram, \lambda\ is\ the\ regularization\ parameter, N\ is\ the\ iteration\ number$$

$$\boldsymbol{Step\ 0}. Take\ (\boldsymbol{r}_1, \boldsymbol{s}_1) = (\boldsymbol{p}_0, \boldsymbol{q}_0) = \left(\mathbf{0}_{(m-1)\times n}, \mathbf{0}_{m\times(n-1)}\right), t_1 = 1.$$

$$\boldsymbol{Step\ k}. (k = 1, \dots, N)\ Compute$$

$$(\boldsymbol{p}_k, \boldsymbol{q}_k) = \boldsymbol{P_p}[(\boldsymbol{r}_k, \boldsymbol{s}_k) + (\frac{1}{8\lambda}\boldsymbol{L^T}(\boldsymbol{P_c}[b - \lambda \boldsymbol{L}(\boldsymbol{r}_k, \boldsymbol{s}_k)])]$$

$$t_{k+1} = \frac{1 + \left(1 + 4t_k^2\right)^{0.5}}{2}$$

$$(\boldsymbol{r}_{k+1}, \boldsymbol{s}_{k+1}) = (\boldsymbol{p}_k, \boldsymbol{q}_k) + (\frac{t_k - 1}{t_{k+1}})(\boldsymbol{p}_k - \boldsymbol{p}_{k-1}, \boldsymbol{q}_k - \boldsymbol{q}_{k-1})$$

$$\boldsymbol{Output}: \boldsymbol{x}^* = \boldsymbol{P_c}[\boldsymbol{b} - \lambda \boldsymbol{L}(\boldsymbol{p}_n, \boldsymbol{q}_n)]$$

As the denoising problem has been defined above, the solution to the deblurring problem where the operator H is no longer the identity matrix but our forward operator can be then defined as the following FISTA problem from the linear inverse problems paper (Beck and Teboulle, 2009).

$$\boldsymbol{FISTA}$$

$$\boldsymbol{Step\ 0}. Take\ Lp_0 = 0.005, n = 2, x_0 = \mathbf{0}^{256\times256}, \boldsymbol{y}_1 = \boldsymbol{x}_0, t_1 = 1, \lambda = 0.00001.$$

$$\boldsymbol{Step\ k}.\ (\boldsymbol{k} \geq \mathbf{1})$$

$$\boldsymbol{Substep\ 1}. Find\ the\ smallest\ nonnegative\ integer\ i_k\ such\ that\ with\ Lp = n^{i_k} Lp_{k-1}$$

$$F\left(p_{Lp}(y_k)\right) \leq Q_{Lp}\left(p_{Lp}(y_k), y_k\right)$$

$$where\ p_{Lp}(y_k) = FGP\left(y_k - \left(\frac{2}{Lp}\right) H^T(Hy_k - b), \left(\frac{2\lambda}{Lp}\right), 10\right)$$

$$where\ b\ is\ the\ sinogram, and\ 10\ is\ the\ number\ of\ iterations\ FGP\ runs\ for$$

$$where\ F(x) = \left||Hx - b|\right|^2 + TV(x)$$

$$where\ Q_{Lp}(x, y) = \left||Hy - b|\right|^2 + \ dot\left(x - y, 2\mathrm{H}^{\mathrm{T}}(Hy - b)\right) + \left(\frac{L}{2}\right) * \left||x - y|\right|^2 + TV(x)$$

$$\boldsymbol{Substep\ 2}. Set\ Lp_k = n^{i_k} Lp_{k-1}\ and\ compute$$

$$x_k = FGP\left(y_k - \left(\frac{2}{Lp_k}\right) H^T(Hy_k - b), \left(\frac{2\lambda}{Lp_k}\right), 10\right)$$

$$t_{k+1} = \frac{1 + \left(1 + 4t_k^2\right)^{0.5}}{2}$$

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

y_k is a dummy term used in the FISTA algorithm with a similar size to x_k that is used to significantly speed up the convergence. For more information about this algorithm, please consult both papers by Beck and Teboulle. An iteration number of 10 was chosen for the FGP algorithm (pure denoising component) as Beck and Teboulle suggested in their coding of the algorithm. n was chosen to be 2 in the FISTA algorithm as n has be larger than 1, and doubling the constant, Lp, seemed like a reasonable method to have relatively fast fulfillment of the inequality in substep 1. Lp_0 was chosen to be 0.005 through trial and error. Different values were chosen until the first iteration of FISTA had a $F\left(p_{Lp_0}(y_k)\right) > Q_{Lp}\left(p_{Lp_0}(y_k), y_k\right)$, which results in an increase in L per the algorithm's $Lp = n^{i_k} Lp_{k-1}$ update rule. This ensured that Lp for all iterations are appropriate in size as Lp_1 used in the first iteration was calculated from an Lp_0 that did not satisfy the inequality. λ was chosen as 0.00001 as it was a value commonly used in image reconstruction and it was used in a prior assignment.

Note that there is a dot product term in the definition of $Q_{Lp}(x, y)$. This dot product term in image reconstruction literature usually involves vectorized forms of both x and y. However, in this case, the sum of the vector output of the dot product function in MatLab with these matrices as the input, *sum(dot(x,y))*, also works, as the dot product function in MatLab calculates the inner products of the column vectors if the inputs are matrices.

Since the x_k term was calculated to find the correct Lp_k term, the *tp_findL.m* function finds both the L that satisfies the condition in substep 1 and the first line of substep 2 (the FGP calculation). The final calls for the algorithm can be found in *project_test.m*. The final FISTA algorithm was terminated after 100 iterations as it achieved sufficient clarity by eye, which was appropriate as medical images are often viewed through the eyes of the physician. The termination factor could be relative error when compared

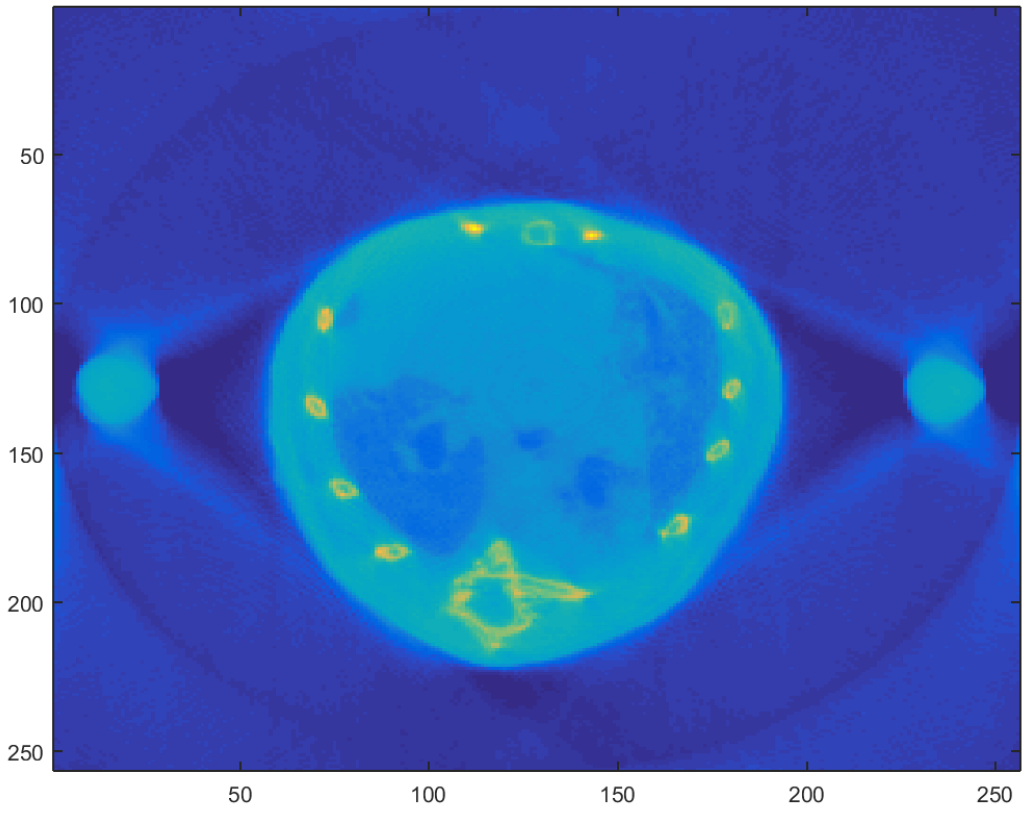
to the previous iteration's estimate of the image instead for some better results, but the reconstruction time may be too long if that relative error threshold was set too low.

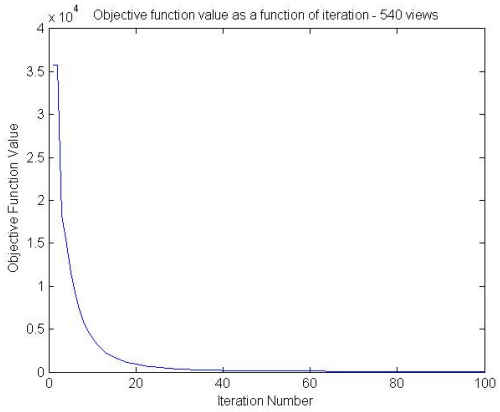
A more detailed derivation of the FGP and FISTA algorithms can be found in the papers by Beck and Teboulle.

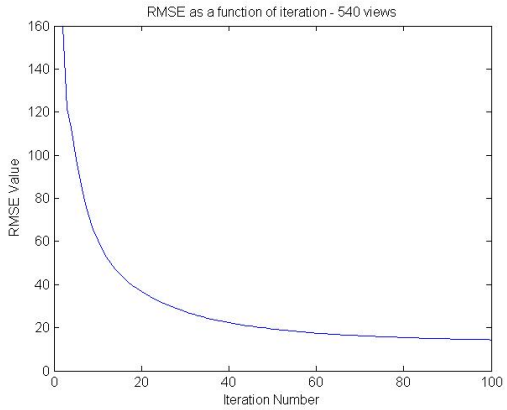
Results

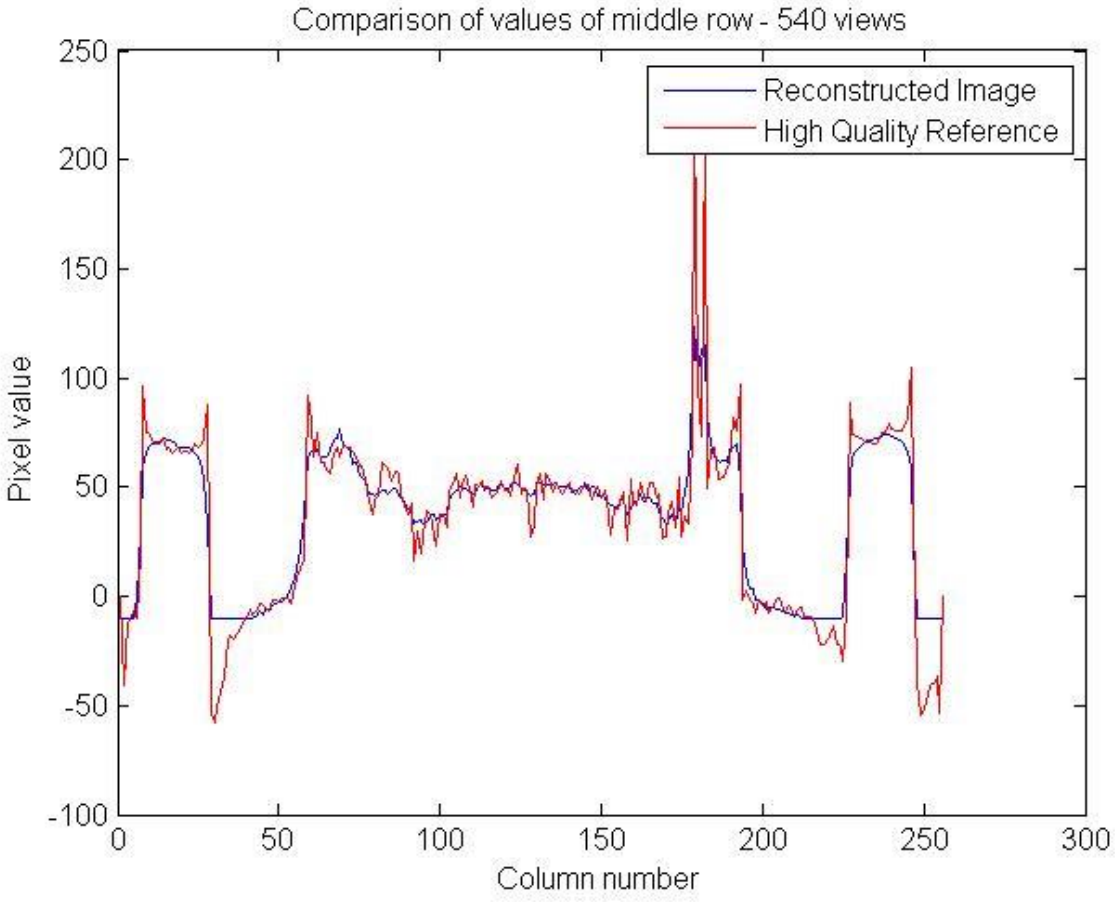
The results here are for three different input sinograms, which were created using the method given in the readme file attached to the project instructions. The three sinograms were of 540 (all available), 270 and 90 uniformly spaced tomographic views. The mean squared error between the reconstructed and high-quality provided image will be plotted as a function of iterations for all 3 sinograms. The reconstructed images at 100 iterations will also be shown for all 3 sinograms. Finally, the central row of both the reconstructed image and the high-quality provided image will both be extracted, and the values will be plotted against each other.

540 views

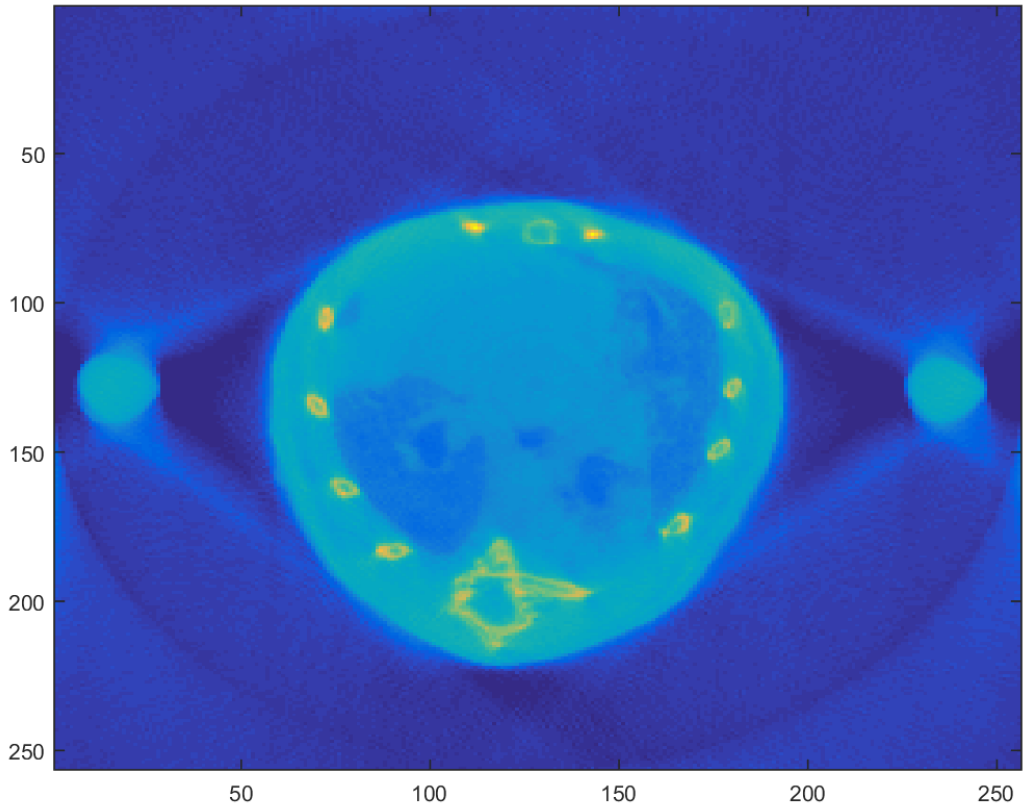


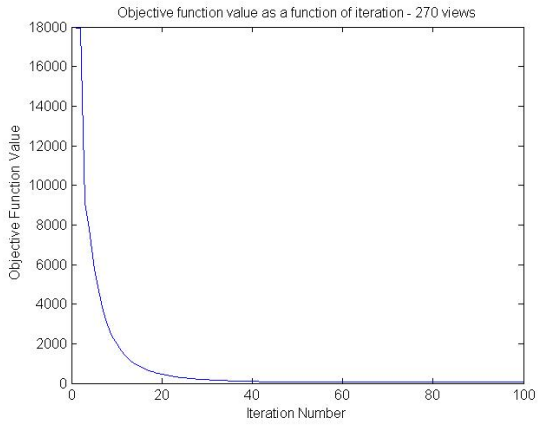

x 10^4
Objective function value as a function of iteration - 540 views
Objective Function Value
Iteration Number

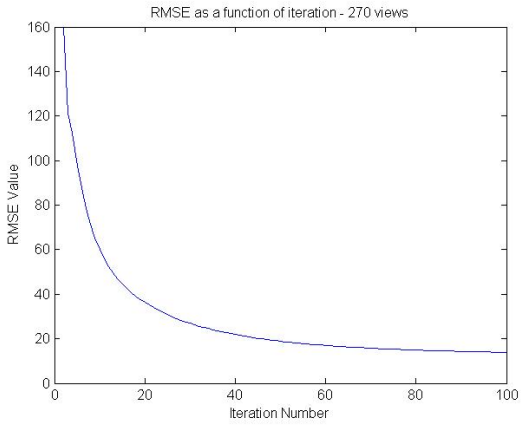

RMSE as a function of iteration - 540 views
RMSE Value
Iteration Number

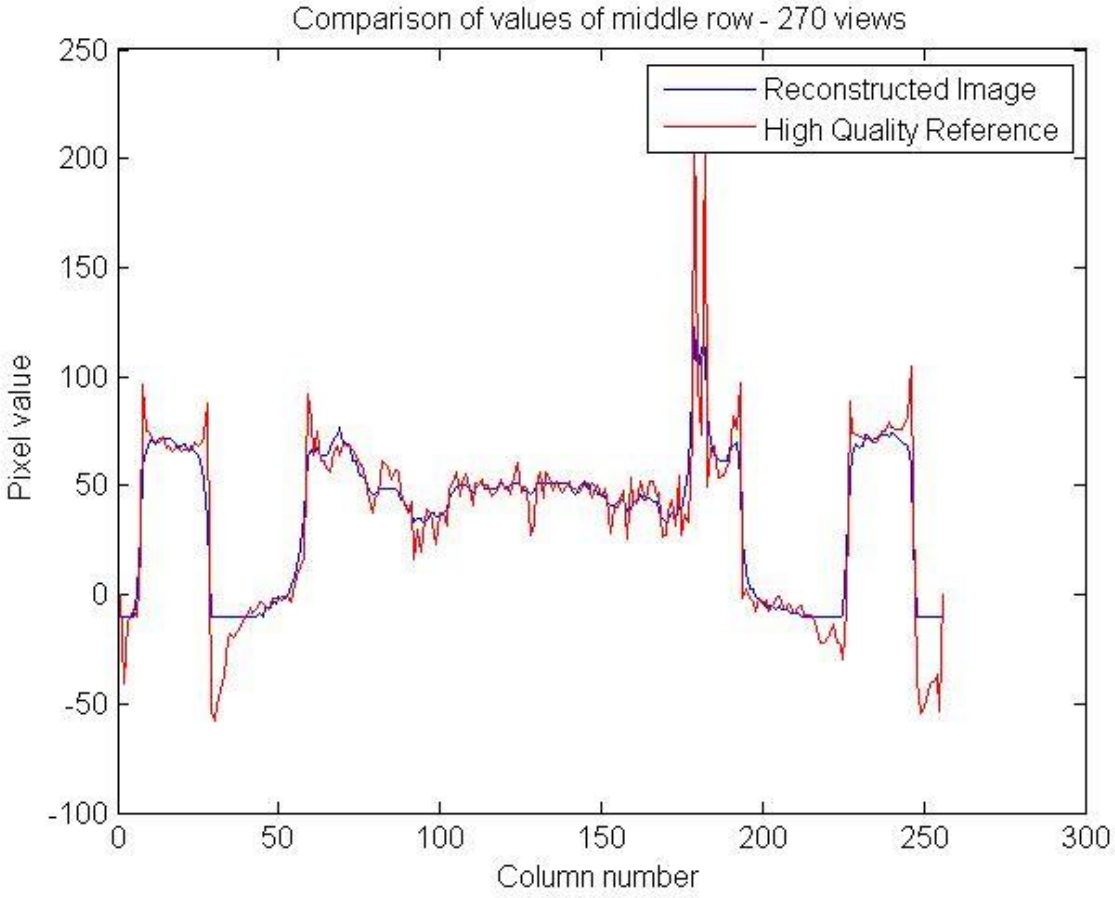

Comparison of values of middle row - 540 views
Reconstructed Image
High Quality Reference
Pixel value
Column number

270 views

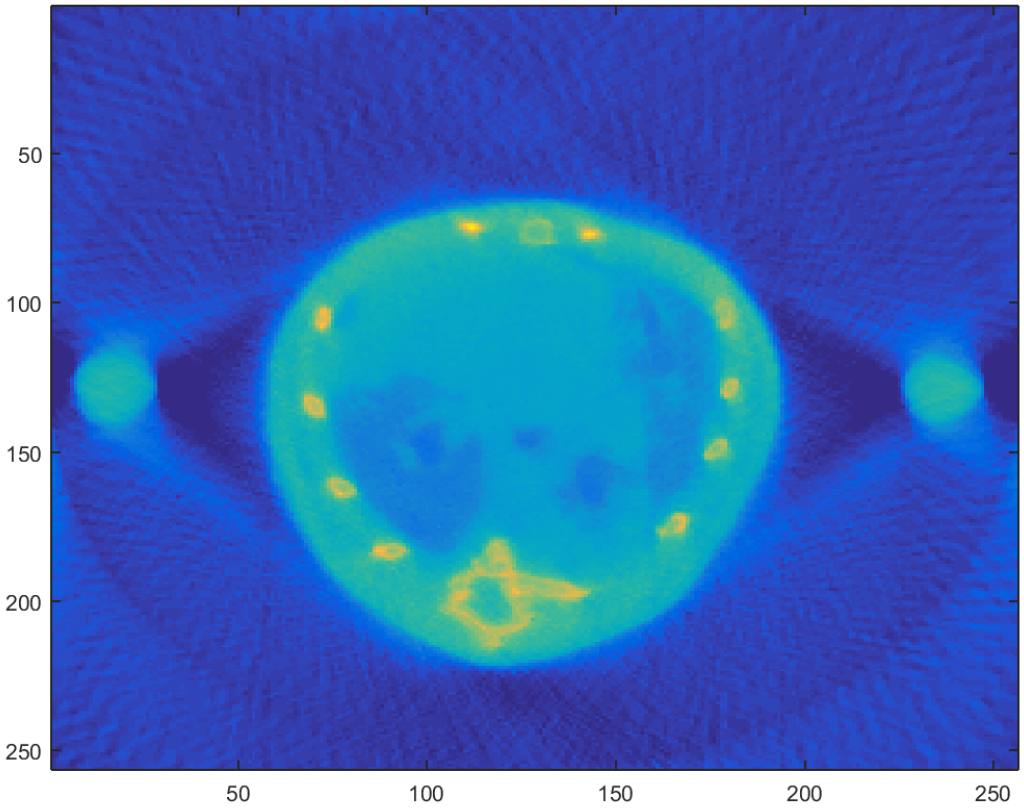

50
100
150
200
250
50
100
150
200
250

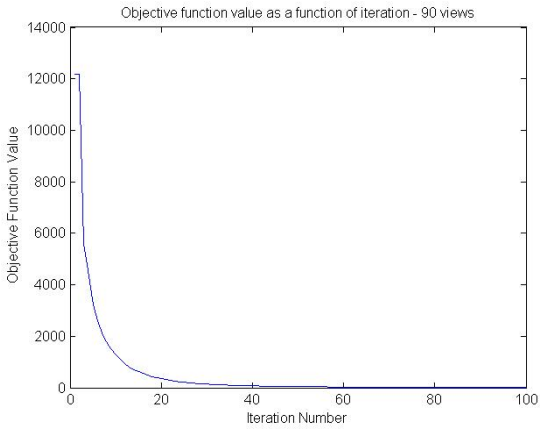

Objective function value as a function of iteration - 270 views
Objective Function Value
Iteration Number
0
2000
4000
6000
8000
10000
12000
14000
16000
18000
0
20
40
60
80
100

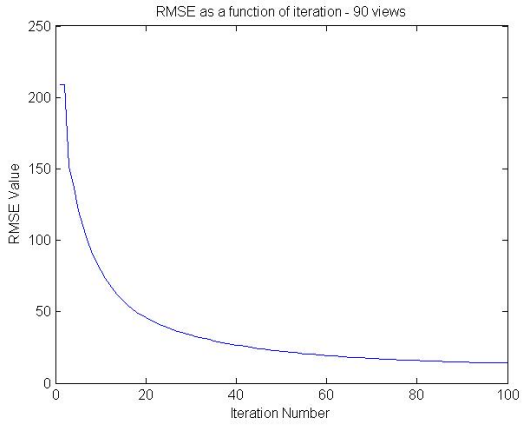

RMSE as a function of iteration - 270 views
RMSE Value
Iteration Number
0
20
40
60
80
100
120
140
160
0
20
40
60
80
100

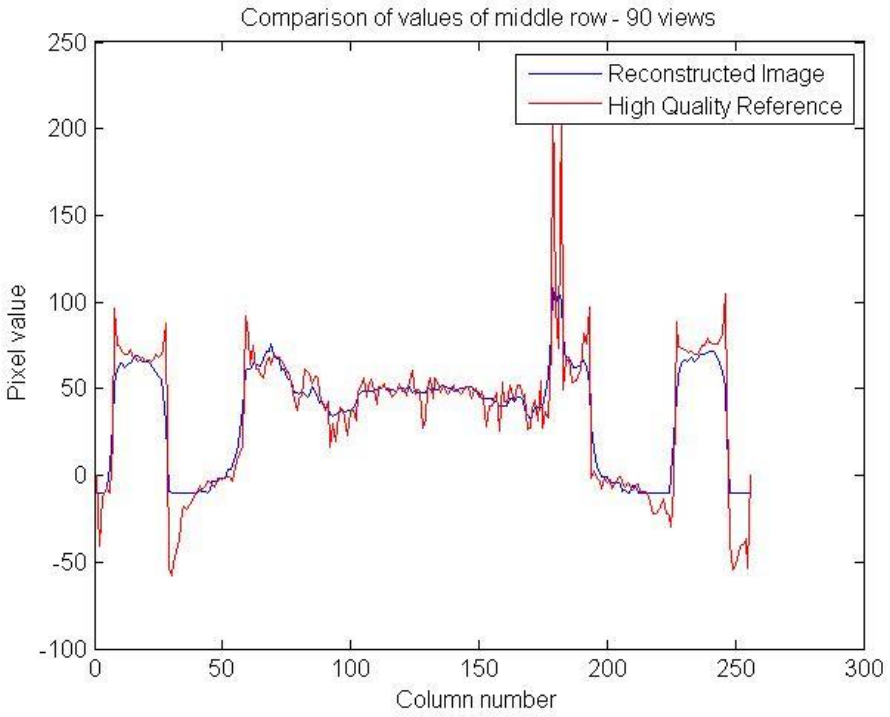

Comparison of values of middle row - 270 views
Reconstructed Image
High Quality Reference
Pixel value
Column number
250
200
150
100
50
0
-50
-100
0
50
100
150
200
250
300

90 views

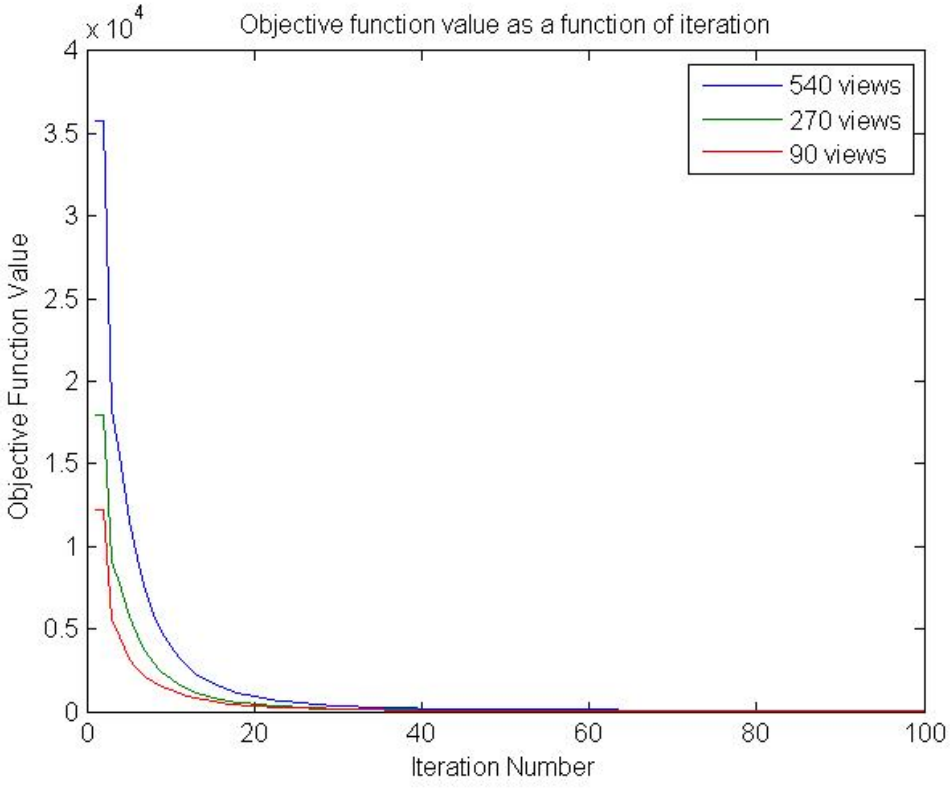

50
100
150
200
250
50
100
150
200
250

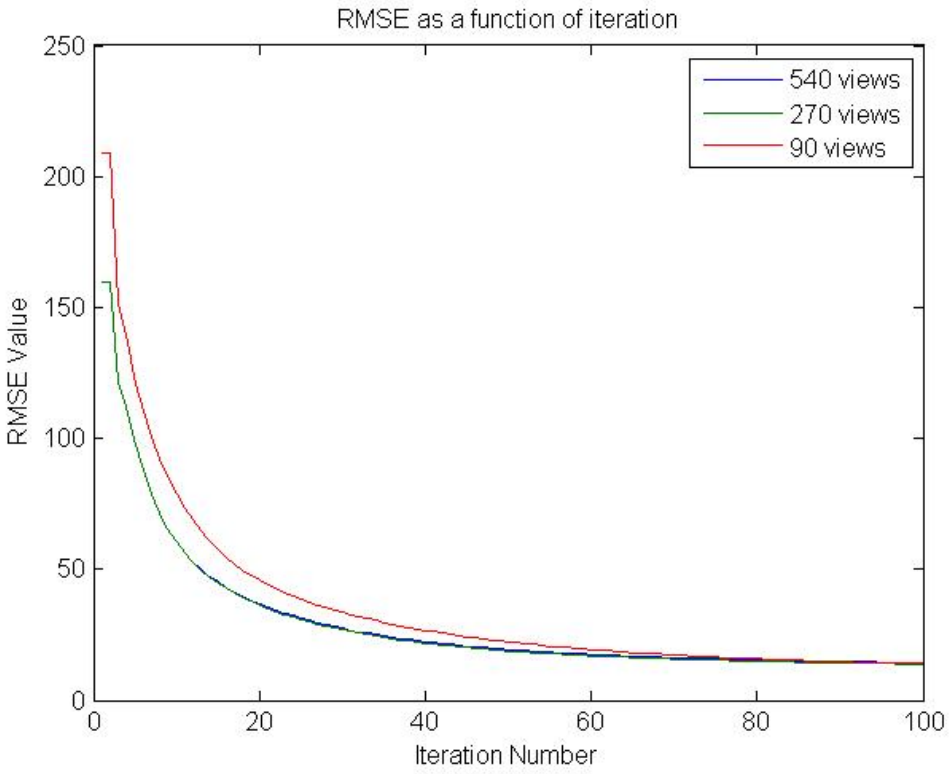

Objective function value as a function of iteration - 90 views
Objective Function Value
Iteration Number
0
2000
4000
6000
8000
10000
12000
14000
0
20
40
60
80
100


RMSE as a function of iteration - 90 views
RMSE Value
Iteration Number
0
50
100
150
200
250
0
20
40
60
80
100

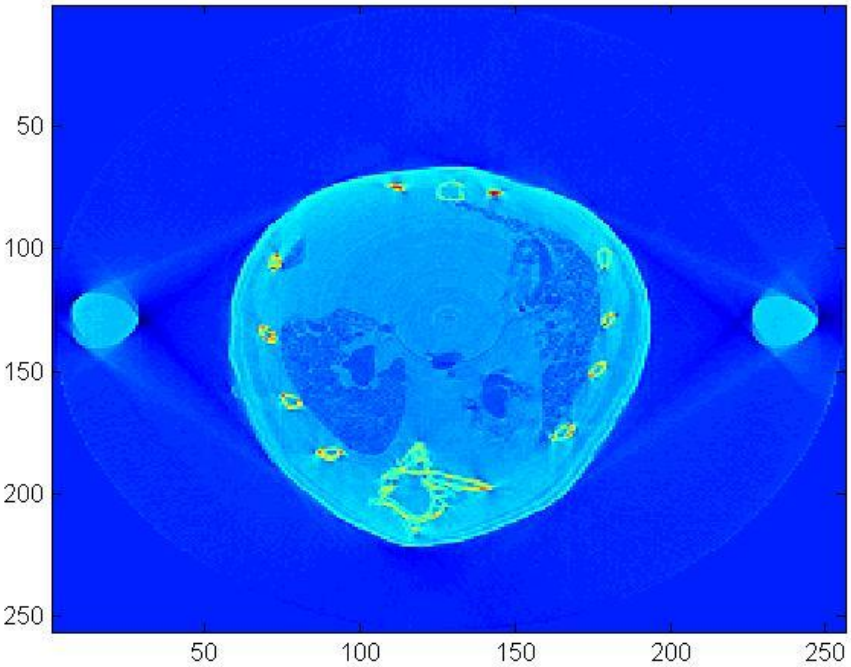

Comparison of values of middle row - 90 views
Reconstructed Image
High Quality Reference
Pixel value
Column number
250
200
150
100
50
0
-50
-100
0
50
100
150
200
250
300

All 3 views


x 10^4
Objective function value as a function of iteration
540 views
270 views
90 views
Objective Function Value
Iteration Number
4
3.5
3
2.5
2
1.5
1
0.5
0
0
20
40
60
80
100


RMSE as a function of iteration
540 views
270 views
90 views
RMSE Value
Iteration Number
0
50
100
150
200
250
0
20
40
60
80
100

Discussion


50
100
150
200
250
50
100
150
200
250

Judging from the high-quality image provided above, the reconstructed images were of acceptable quality. The smaller topological structures within the image was appropriately reconstructed, as seen by the similarities between the yellow and dark blue sectors within the circular component (presumably torso) of both the reconstructed images and the high-quality image. This similarity can be confirmed by the plots of the middle row values across the reference and reconstructed images. In general, the reconstructed images as smoother than the high-quality provided image, this is can be thought of as the direct effect of the total variation term inside the objective function as it penalizes the reconstruction if the variation within the image values is too large, so the reconstructed images are more likely to be smoother. This smoothness is also seen through the image plots. The image plots are comparable as they were all plotted with the same color scale. The high-quality image has red spots which corresponds to the high peak values, while the reconstructed images lack these high red spots and has more mellow yellow or smoother color transitions, therefore suggesting that it is smoother.

The objective function value for all 3 different sinograms decreased dramatically within the first 40 iterations, eventually reaching levels close to 0. In terms of objective function value, the sinograms with smaller number of views tended to have smaller objective function values. This makes sense as the number of elements in the b and Hx arrays were less with less views, so the norm value out of the data fidelity term should also be smaller.

An interesting observation is that as the number of views decreased from 270 to 90, the RMSE across all iterations seemed to increase. This suggests that as the number of views decreases the ability of the algorithm to converge to desired output decreases, this is to be expected as with fewer views, the algorithm has less information to reconstruct images of the same size, so the error in reconstruction should be higher. However, it is also interesting to note that there was not a big difference between the RMSE values between 270 and 540 views, this suggests that additional views in the sinogram past a

certain threshold number of views contribute less and less to improving the convergence ability of the algorithm.

Another interesting observation was that the running time for the algorithm was shorter with a smaller number of views in the sinogram (this is to be expected as there are fewer calculations required during each forward operation). This data was not collected empirically, but was noticeable as I ran the algorithm. This suggests that there is a trade-off between image reconstruction quality and speed. With higher numbers of views (up to a threshold), the image reconstruction quality increase, but speed decreases. Therefore, it would be prudent to find the balance the diminishing returns of increased views on quality with the increase in speed in future studies.

Further improvements can be made to the algorithm. Perhaps one of the easiest to make is the starting image and projection operators of the algorithm. In this case, if the image is known to be of a cross section of the human torso, the algorithm could potentially speed up by simply implementing some prior knowledge terms into the projection operators and starting images. For example, we could apply a warm start where the starting image included some defined fine structures that included the rib cage and the spinal cord cross-sections as they are generally found in all cross sections of the human torso. While this algorithm produced images that were appropriate for general use, the images produced were still not as clearly defined in terms of structures as the high-quality reference image. One way to improve it is to increase the number of iterations. Another way to improve it is to choose other regularization terms, which will change the proximal projection operation that we use.

Conclusion

An objective function with a least squares data fidelity term and a total variation regularization term was applied in the setting of the computed tomography problem. The FISTA algorithm used in the project was a subclass of gradient projection algorithms, which can be described as a gradient descent step and

a proximal projection step after the initialization of variables. A backtracking method was used to calculate the optimal step size in the gradient descent step, and the proximal projection operator was calculated through the FGP algorithm using proofs shown by Beck and Teboulle. This project serves a validation that the FISTA algorithm suggested by Beck and Teboulle can be applied to computed tomography problems.

References

Beck, Amir, and Marc Teboulle. Fast Gradient-Based Algorithms for Constrained Total Variation Image Denoising and Deblurring Problems - IEEE Xplore Document. N.p., 11 June 2009. Web. 20 Apr. 2017.

Beck, Amir, and Marc Teboulle. "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems." SIAM Journal on Imaging Sciences 2.1 (2009): 183-202. Web.

Guerquin-Kern, M., J.-C. Baritaux, and M. Unser. "Efficient Image Reconstruction under Sparsity Constraints with Application to MRI and Bioluminescence Tomography." 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2011): n. pag. Web.

Nikolova, Mila. "Minimizers of Cost-Functions Involving Nonsmooth Data-Fidelity Terms. Application to the Processing of Outliers." SIAM Journal on Numerical Analysis 40.3 (2002): 965-94. Web.

Weller, Daniel S. "Accelerating Magnetic Resonance Imaging by Unifying Sparse Models and Multiple Receivers." Thesis. Massachusetts Institute of Technology, 2012. Print.

Xu, Qiaofeng, Deshan Yang, Jun Tan, Alex Sawatzky, and Mark A. Anastasio. "Accelerated Fast Iterative Shrinkage Thresholding Algorithms for Sparsity-regularized Cone-beam CT Image Reconstruction." Medical Physics 43.4 (2016): 1849-872. Web.