1 Methods

Given the several sources of noise for computed tomography[5], such as electronic noise and structural noise, the reconstruction of the CT image is indeed an ill-posed problem. To solve this inverse problem, we formulate the CT image reconstruction as an optimization problem. One of the important considerations for formulating this optimization problem is to find a regularization method 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 the anisotropic TV regularization functional. While anisotropic TV is known to favor horizontal and vertical structures for 2D images[3], it simplifies some of the implementation of the algorithm. So the anisotropic TV is implemented as the first path, and the reconstruction quality is good. A decision is made not to implement the isotropic TV functional.

Hence the optimization problem can be formulated as:

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

where $x \in \mathbb{R}^{m\times 1}$ is the flattened true image with the size $\sqrt{m\times 1}$. $d \in \mathbb{R}^n$ is the flattened sinogram image with the size $\sqrt{n\times 1}$. $A \in \mathbb{R}^{m\times n}$ is the forward operator that transforms an image to a sinogram image. λ is the regularization parameter. $\mathrm{TV}(x) : \mathbb{R}^{m\times 1} \to \mathbb{R}$ is an anistropic total variation functional, where

$$x \in \mathbb{R}^{m\times 1}, \mathrm{TV}(x) = \sum_{i=1}^{m-1} |x_i - x_{i+1}| \tag{2}$$

To solve the problem in eq.1, it requires a method that sovles non-differentiable optimization problems. As discussed in [1], the gradient projection method is particularly suited for this task because of the faster convergence rate compare to other methods such as primal-dual Newton-based methods[2]. The specific innovation of [1] is a new proximal gradient method called fast iterative shrinkage/thresholding algorithm (FISTA), which has a global convergence rate of $\mathcal{O}(\frac{1}{k^2})$ (The proof of convergence rate is detailed in [1]).

The general principle of the algorithm will be discussed in this paragraph and the pseudocode for the specific implementation will follow. The first key insight is to apply a orthogonal projection operator on P_C, which is a box projection that maps x onto a feasible set (which can change in each iteration). As discussed in [1], the solution is

$$x_k = \mathrm{P}_C(x_{k-1} - t_k \nabla f(x_{k-1})) \tag{3}$$

where f denotes the data fidelity term.

The next step is to construct the proximal map. Let $\mathcal{P}_1$ denote the set of all pairs of matrices $(\mathbf{p}, \mathbf{q})$. Note that for our special case where $x \in \mathbb{R}^{m\times 1}$, $\mathcal{P}_1$ becomes matrices $\mathbf{p}$ satisfying

$$|p_i| \leq 1, i = 1, \cdots, m-1 \tag{4}$$

Following *Proposition 4.1* in [1], we can write out the objective function as

$$\underset{\mathbf{p}\in\mathcal{P}}{max}\underset{x\in C}{min}\{||x-d|||_{\mathrm{F}}^2 + 2\lambda \mathrm{Tr}(\mathcal{L}(\mathbf{p}^{\mathrm{T}}x))\} \tag{5}$$

and the optimal solution of the inner minimization problem is

$$x = \mathrm{P}_C(d - \lambda\mathcal{L}(\mathbf{p})) \tag{6}$$

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

$$\mathcal{L}(\mathbf{p})_i = p_i - p_{i-1}, i = 1, \cdots, m-1 \tag{7}$$

Plugging in the objective function, gradient and step size (which is an upper bound on the Lipschitz constant, and in term depends on the regularization parameter), we can write out the pseudocode for the algorithm, see algorithm.1.

We found an existing package in python called pyunlocbox [4] to implement the algorithm. This package has all the required components for constructing the functions for data fidelity and regularization terms, and has implemented the FISTA version.

Algorithm 1 TV-FISTA

Input d, A, λ, $maxiter$, tol
$r_0 = p_0 = 0_{m-1}$, $t_1 = 1$
if $i \leq maxiter$ **then**
Forward
  Compute $x_f = \frac{2}{\lambda}A^T(A(x) - d)$
Backward
  $x_i = x_f - \frac{1}{\lambda}\boldsymbol{\nabla}\cdot x_f$
  $obj = \frac{1}{2}||x_f - x_i||^2 + \lambda||\nabla x_i||^2$
  $tol_i = \frac{|obj_{i-1} - obj_i|}{obj_i}$
Update projector
  $r_i = r - \frac{1}{4\lambda}\nabla x_i$
FISTA
  $t_{i+1} = \frac{1+\sqrt{1+4t_i^2}}{2}$
  $p_i = \frac{r_1}{\max\{1, p_i\}}$
  $r_i = p_i + \frac{t_i - 1}{t_i * (p_i - p_{i-1})}$
  if $tol_i \leq tol$ **then**
    Stop
  $i = i + 1$

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.

In general, we made an iterative plan for implementing this algorithm. We started from a series of simple choices, such as optimizing in the image domain with image represented as a row vector (in contrast to wavelet domain), the anisotropic TV, and relative tolerance. While those choices are unsophisticated compare to many advanced tools, we reasoned that we could adjust our implementation choices if the reconstruction quality is poor. To our surprise and delight, the first try of the reconstruction quality is good. We hence adhered to our implementation choices in regularization function and stopping criterion. The detailed reconstruction result will be presented in the result section.

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] Tony F Chan, Gene H Golub, and Pep Mulet. A nonlinear primal-dual method for total variation-based image restoration, 1996.

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

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

[5] 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.