An Optimal Metric for Regularization Parameter Selection in Iterative Reconstruction for CT Image

Jiayu Duan, Xiaogang Chen, Baiqiang Shen and Xuanqin Mou

***Abstract*—Regularization parameter selection is crucial in CT iterative reconstruction because it is a balance between fidelity and penalty while there has not been so far a metric to judge if an optimal selection is made which results in the best reconstructed image quality in terms of a well balance between the apparent noise and the image fine structure. In this paper, we proposed a metric for selecting the optimal regularization parameter based on the property of natural image statistics. By using LoG operator and the pairwise products of neighboring LoG signals to extract the statistic features of the reconstructed image to account for the image quality, this proposed method evaluated all selected regularization parameters by calculating the variance of extracted statistic features and picked up the optimal regularization parameter with maximum curve of second order derivation of the calculated variance curve. Numerical and experimental results validated the efficiency of the proposed metric in terms of that the selected regularization parameter is accordance with the best visual observation. Besides, the proposed metric has low complexity of computation and only depends on features, which can be used in multiple situations.**

***Keywords*—regularization parameter, natural statistics, LoG, pairwise products, statistic features**

I. INTRODUCTION

Iterative reconstruction can cope well with incomplete data (such as low dose, limit angle, few views, interior problem, etc). Mathematically, the iterative reconstruction is a solution of an inverse problem. In the context of incomplete data**,** the inverse problem will be utmost ill-posed, which means small perturbations in observation data influences greatly the considered solutions[1]. In order to achieve a solution which approximates the noise-free reconstruction, numerous regularization strategies were proposed, such as truncated SVD, Tikhonov regularization, total variation, dictionary learning.[2] When the regularization method is settled, it is essential to select a proper regularization parameter because it compromises the data fitting with regularization.

As a matter of fact, a good solution of the iterative reconstruction relies on the selection of regularization parameter. In practice, the selection of regularization parameter can be time consuming. Hence, there are substantial strategies to select regularization parameter such as the discrepancy principle[3], Generalized cross-validation(GCV)[4], the L-curve[5,6] (introduced by Lawson and popularized by Hansen), etc.As for the discrepancy principle, it requires an estimate of the noise level, which is not always available in practice and the existence of a solution is not guaranteed for some non-smooth functionals[7]. The GCV method estimates the mean square error(MSE), but the minimization of the objective functions are nontrivial because of their flat over a broad scale. The L-curve method is totally based on data, which is sensitive to curvature estimation [8].

This work was supported by the National Key Research and Development Program of China (No. 2016YFA0202003) and National Natural Science Foundation of China (NSFC) (No. 61571359). JY Duan and XQ Mou are with the Institute of Image processing and Pattern recognition, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China; XG Chen and BQ Shen are with hangzhou power supply company of state grid zhejiang electric power company. Corresponding author: XQ Mou (e-mail:xqmou@xjtu.edu.cn).

Image quality assessment (IQA) has recently aroused a hot discussion in the area of computer tomography as it provides a hint to solve the dilemma between noise and resolution of imperfect reconstruction caused by limited quantum. In general, image quality assessment can classify into two kinds: full reference and no reference. As there is no perfect reference image in real world medical imaging, no reference IQA is recommended for accessing medical images. Recently, there are multiple methods based on IQA model to solve medical imaging problems. Woodard et al presented NR-IQA measure for structural MRI using two types of analysis of variance[11]. In [12], Nakhaie proposed a watermarking method using spread spectrum and discrete wavelet transform based on ROI processing. Dutta et al used a quantitative statistical method with closed-form analytical expressions to measure medical image quality with two metrics : covariance and resolution based on two analysis techniques: fixed point and iteration-based analysis[13]. All of the mentioned methods haven't focus on the ideal selection of regularization parameters. Inspired by nature statistics, we proposed a heuristic IQA idea based on statistical properties using sparse transformation to get rid of first and second order of statistics to help select optimal regularization parameter. With several experiment in simulated data and real data, the method performs well in selecting optimal regularization parameter.

This paper organized as follows. In section 2, we introduce the methods in details. Section 3 describes the material we used and the workflow. Section 4 presents several results of our algorithm. In the following section 5 the proposed method and its application are discussed. Our conclusion can be found in section 6.

II. METHODS

Typically, an inverse problem always seeks an appropriate solution $x \in R^N$ to $AX = b$. To alleviate the ill-posed problem, the solution can be achieved by using Lagrange Multiplier as objective function:

$$x_\beta = \arg\min_x \{ \left\| Ax - b^\delta \right\|_2^2 + \beta\psi(x) \} \quad (1)$$

A fidelity which quantifies the deviation between measured data and predicted data is the main constituent of objective

DOI: 10.12059/Fully3D.2017-11-3102002

function. In iterative reconstruction, b^{δ} represents acquired projection data consisting of exact projection data ($b^{\dagger}$) and inevitable acquisition error ε during the acquisition physics, including the noise induced by limited quantum, i.e., $b^{\delta} = b^{\dagger} + \varepsilon$. We denote A: $R^N \rightarrow R^M$ as a bound and linear operator which is called system matrix in CT. The second term denoted as $\psi(x)$ represents the regularization term solely based on prior knowledge rather than the data. β as the regularization parameter compromises the fidelity term and penalty term.

In fact, it is proved that the original solution x could be sparsely represented as follows:

$$x^{\dagger} = \phi\alpha \quad (2)$$

Where α represents the sparse coefficients of $x^{\dagger}$ in the domain ϕ. During our previous study, we found that after processing images with sparse operator, the coefficients of the images follow the distribution of Laplacian. Elements in α have characteristic statistical properties which can be molded as Laplacian distribution with zero mean. Here is the histogram of multiple medical images in Fig.1.

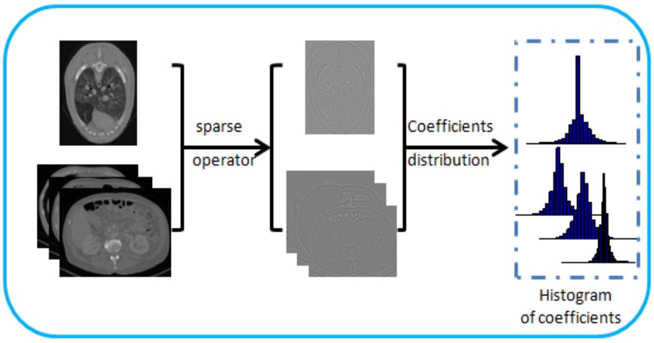


Fig. 1. The sparse coefficients of medical images

In addition, the noise component in the photon measurement in X-ray CT mainly consists of the noise caused by detection system electronics and the noise in X-ray CT is characterized by Gaussian random variable with zero mean[14]. In our hypothesis, when proper sparse operator is used the coefficients of noise also follow the Gaussian distribution. We separate the ε as follow:

$$\varepsilon = A\phi\gamma \quad (3)$$

In our previous work[8], it is explained in detail that the perfect solution relies on the selection of β. We use a simple model to validate the proposed model. In the assumption, we decompose the $\hat{x}_{\beta}$ as follow:

$$x_{\beta} = \phi\eta_{\beta} = \phi\alpha_{\beta} + \phi\gamma_{\beta} \quad (4)$$

α_{β} and γ_{β} denote the coefficients of signal and noise respectively.

In simulation, we selected a simple reconstruction with forward projection and adding noise to the phantom. Reconstructed with multiple regularization parameters, a series of recon images were generated. After de-convolution operation, we extracted the noise signal coefficients in different scale of regularization parameters and used kernel function to model the distribution of coefficients. The workflow is describe in Fig.2 and the curves of noise in different state of regularization parameters showed in Fig.3.

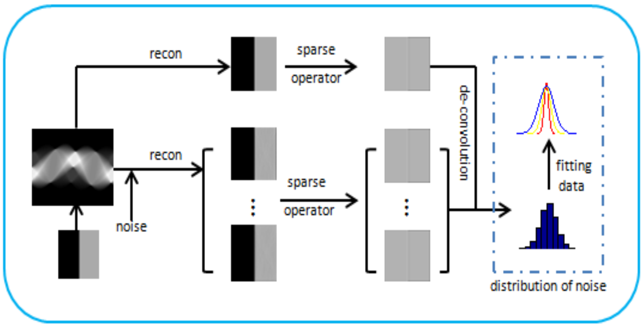


Fig. 2. The workflow of the validation experiment

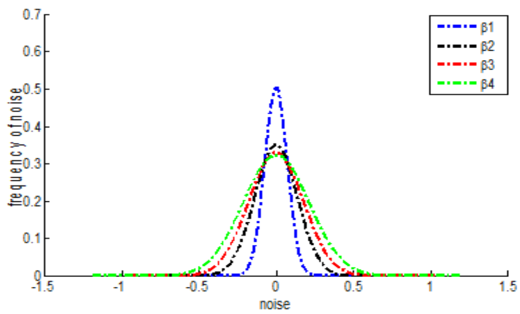


Fig. 3. The distribution of noise coefficients. From β_1 to β_4 the regularization parameter decreases gradually .

From the different curves of distribution, we found that the distribution of noise is Gaussian-like. When the regularization parameter is increasing, the tail of the distribution becomes higher and higher because its contains the information of reconstruction image coefficients whose distribution is Laplace. During the process, there exist the statement where the image quality is preserved with smallest noise. With these phenomena, we assume that a generalized Gaussian distribution (GGD) of the sparse coefficients can be used to capture the changes of the images with different β. In general, the distribution of the α_{β} and γ_{β} can be modeled as Gaussian distribution with zero mean.

$$g(x; a, \sigma^2) = \frac{a}{2\eta\gamma(\frac{1}{a})} e^{-(\frac{|x|}{\eta})^a} \quad (5)$$

where $\eta = \delta\sqrt{\frac{\gamma(\frac{1}{a})}{\gamma(\frac{3}{a})}}$ and b $\gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$, $x > 0$

In the formula, $a = 2$ yields the Gaussian distribution and $a = 1$ yields the Laplacian distribution.

During the selecting process, for example, we approach the β^* by decreasing the β. The reconstructed images with different β appear to be more and more clearer and the noise is smaller and smaller until without the noise. The distribution of sparse coefficients of $\hat{x}_\beta$ in this process can be explained by this formula:

$$f(\eta_\beta) = g(\eta_\beta; a_1, \delta_1^2) \quad (6)$$

where a_1 represents the shape parameter and δ_1^2 denotes the variance.

As the process proceeding, the shape parameter decreases so does the variance and $a_1 \approx 1$ when $\beta = \beta^*$. While as the β becomes smaller, the reconstructed image becomes noisy. The current x_β which is decomposed as formula 4. The distribution of sparse coefficient η_β can be modeled as follows:

$$f(\eta_\beta) = t(1-t) g(\eta_\beta; a_1, \delta_1^2) * g(\gamma_\beta; a_2, \delta_2^2) \quad (7)$$

In this formula, t represents the harmonic coefficients which compromise the signal with noise. Commonly, the optimal reconstruction appears in the circumstance which the details are shown and the noise isn't. In order to get the optimal reconstructed image, we exploited a novel method. In our method, we select the variance of the image as a media to help us select optimal β^*. Same as the former process, when $\beta > \beta^*$, the variance increases as the β decreases and when β continues to decrease, the variance increases much quicker because the noise is denser than signal. So the maximum curvature of the variance curve can be used to set the optimal point of β^*.We can use the moment-matching based approach to gain the parameters like $(t, a_1, a_2, \delta_1^2, \delta_2^2)$, and by using these parameters, the desirable reconstruction is expected.

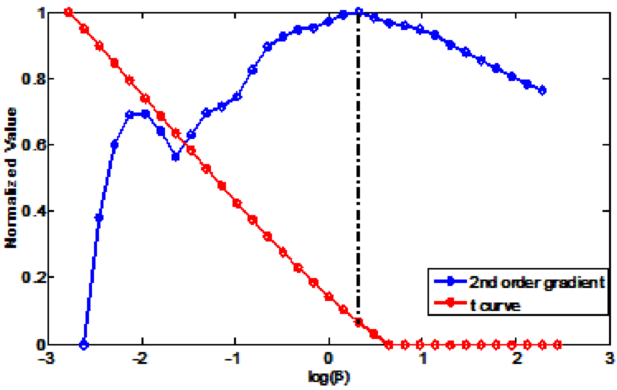


Fig. 4. T-curve vs the 2nd order gradient of sparse coefficients in residual images.

In the simulation, we validated our method in Shepp-Logan phantom image. First, by using forward projection, the sinogram of phantom was generated. After adding noise in the sinogram, we used multiple of regularization parameters to reconstruct the phantom. Then, wiener filter was applied to estimate and eliminate noise. With the known ground-truth, several residual images were produced. Using sparse operator, we can get sparse coefficients in reconstruction images with different regularization parameters. Calculating all the standard deviation of noise and residual coefficients, the t in formula 7 can be solved. We set too small value as 0 in t curve. We found that the optimal regularization parameter selected by proposed method occurs in the neighbor where t appears to rise.

After introducing our fundamental theory and method based on, we now talk about the sparse operator employed in our method. LoG (Laplacian of Gaussian) filter has a center surrounded profile that is symmetrically sensitive to intensity changes across all orientations[9]. And it can easily get rid of first and second order statistical redundancy. Besides, it is proved that the LoG models the earliest stage output of human visual neural system and performs well in image quality assessment (IQA) model design. The LoG operator can be described as follows:

$$h_{LoG} = \frac{\partial^2}{\partial x^2} g(x, y|\sigma) + \frac{\partial^2}{\partial y^2} g(x, y|\sigma) = \frac{1}{2\pi\sigma^2} \frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (8)$$

where $g(x, y|\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$ is the isotropic function with scale parameter σ. By using pairwise products of neighboring LoG signals along main diagonal direction, we found that the signs of the adjacent coefficients can get rid of more redundancy compared with single LoG in Table 1.

Table 1. The mutual information entropy

Coefficients	LoG	Pairwise of LoG
mutual information entropy	0.2652	0.2115
mutual information entropy (Gaussian white noise)	0.0912	0.1861

Calculating the mutual information entropy of coefficients extracted by LoG and pairwise of LoG, it is cleared that with the pairwise operation the redundancy is further removed compared with boundary line.

After pairwise products of adjacent LoG, we calculate the similarity of adjacent LoG coefficients with this formula:

$$S(i, j) = \frac{2L(i+1, j+1)L(i, j)}{L^2(i+1, j+1) + L^2(i, j) + c} \quad (9)$$

where c is a small positive constant to avoid numerical instability when denominator is small. In experiment c=1. We calculate variance of similarity coefficients and select the optimal regularization parameter by the maximum curve of 2^{nd} order gradient.

Considering the removing efficiency of this multiplication, it can be used to extract the sparse coefficients and to help analyze the distributions of sparse coefficients. Hence, in our experiment, we use the LoG operator with σ =0.5 and use diagonal direction multiplication to reduce redundancy.

III. EXPERIMENTS

We use several experiments to demonstrate the efficiency of our method. In general, we use two different regularization methods: total variation (TV) and dictionary learning. Futhermore, the experiment data varies from sheep lung perfusion to real data of human abdomen. During this process, it proves the independency of the method which can be used without limits. The sheep was scanned on the SIEMENS Somatom Sensation 64-slice CT scanner to acquire a 697 × 1160 sinogram. The human abdomen raw projection derived from Mayo Clinic and use Single-slice rebinning method[10] to obtain 736X2304 sinogram.

The order subsets methods were used to accelerate the algorithm and choose TV regularization parameter ranging from 0.28×1.1^{-7} to 0.28×1.1^{7} with scale 1.1 in experiment. By using dictionary regularization, regularization parameter ranging from 0.1×1.1^{-15} / 0.2×1.1^{-7} to 0.1×1.1^{30} / 0.2×1.1^{7} . With the different regularization parameters, a series of reconstructed images were created, 15 TV images and 46/15 dictionary images respectively, which the quality of the reconstructed images vary from noisy ones to over smooth ones. We filtered every reconstructed image with LoG operator to remove redundancy and got the sparse coefficients. With adjacent multiplication, the further redundancy was removed. Then the similarity coefficients was calculated using equation(9). The variance is estimated using the moment-matching based approach. The curves of the similarity variance vs β , along with the second order gradient of variance vs β, were plotted. The result demonstrates that the maximum curve of the second order of variance shows the optimal regularization parameter.

IV. RESULTS

The results for optimal regularization parameter selecting are displayed as follows. The quality of the reconstructed images is from noisy ones to over-smooth ones which guarantees the suggested optimal regularization parameter by the proposed method is in the range.

The result of selecting TV regularization parameter is depicted in Fig. 3.(a) showing noisy image with small regularization parameter, and (b) shows the optimal reconstructed image with selected regularization parameter by proposed method. (c) is the over smooth image with large regularization parameter. The result indicates that our method performs well in selecting TV regularization parameter. The reconstructed image selected by proposed method shows a good trade-off between noise and resolution. In second experiment, by using the dictionary as penalty, same as first experiment, the result shows in Fig. 4 and Fig .5. In the experiment, the sparsity and precision parameter is fixed, we only take regularization parameter in consideration. The reconstructed images (a_1), (a_2) are noisy due to the small regularization parameters. And in (b_1),(b_2), the images exhibit the optimal trade-off between resolution and noise. Additionally, the (c_1),(c_2) are the result with large regularization parameters which are over-smooth. The perpendicular line on the 2^{nd} order of the variance curve points out the optimal regularization parameter by the proposed method. The result indicates the proposed method also performs well in selecting dictionary learning regularization parameter. Overall, the suggested method can be used in different kinds of regularization parameter and different reconstructed target.

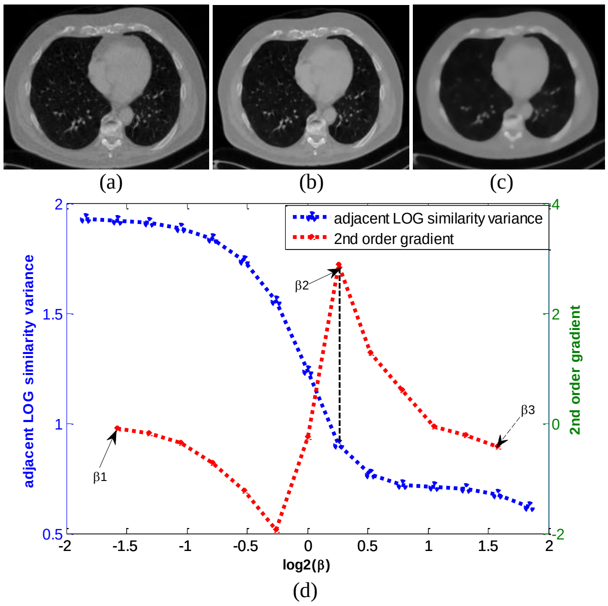
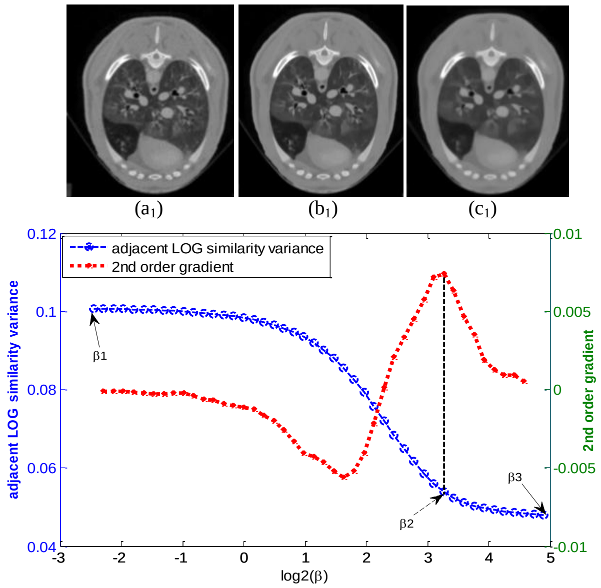


Fig. 5. (a), (b) and (c) are human abdomen reconstructions with different regularization parameters. Reconstructions correspond to different β as indicated by the black arrows in (d), respectively. Curves of the variance vs regularization parameter β and along with 2^{nd} order gradient of the variance vs β , respectively. The perpendicular line which is in black points out the optimal regularization parameter by the proposed method. Note that the abscissa is normalized.

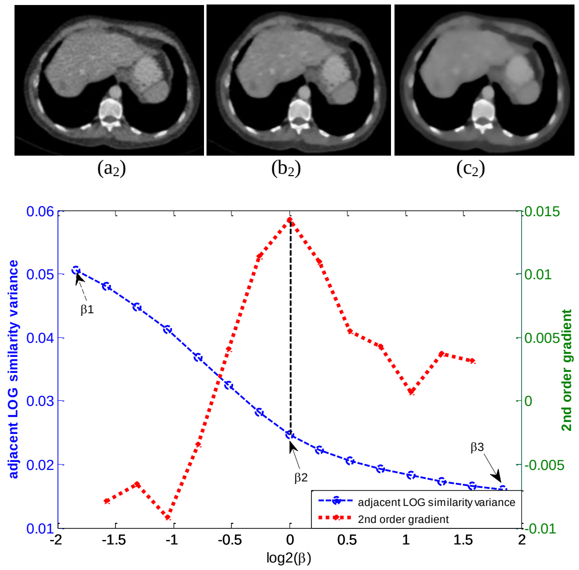
V. DISCUSSION

As a handy tool, the proposed method using the statistic characteristic of nature image can be used to select the optimal regularization parameter. This process is very meaningful because the quality of reconstructed image depends on the selection of regularization parameter. Via this method, it can be a metric to judge if an optimal selection is made which results in the best reconstructed image quality in terms of a well balance between the noise and the image fine structure. In the following work, setting up CT database, this method can be used as a standard to select optimal reconstructions. Now it has proved its ability in choosing optimal regularization parameter in CT. So it is expected that this method can also participate in various fields, such as MRI/PET/SPECT iterative reconstruction, natural image processing, etc. Besides, this proposed idea is not restricted to specific penalty term because the selection is based on statistic features so it can be widely used in total variation, low rank, and dictionary learning, etc.



(d1)

Fig. 6. First row are low dose sheep lung reconstructions with different regularization parameters. Reconstructions correspond to different β as indicated by the black arrows in (d_1). Curves of the variance vs regularization parameter β and along with 2nd order gradient of the variance vs β ,respectively. The perpendicular line which is in black points out the optimal regularization parameter by the proposed method.



(d2)

Fig. 7.(a_2), (b_2), (c_2) are quarter-dose human abdomen reconstructions with different regularization parameters. Reconstructions correspond to different β as indicated by the black arrows in (d_2). Curves of the variance vs regularization parameter β and along with 2nd order gradient of the variance vs β ,respectively.

Compared with single LoG operator, the adjacent multiplication can get rid of more redundancy. And in real data, it is more robust than single LoG operator. Now this proposed method require a series of reconstructed images but our ultimate goal is to design an algorithm to select regularization parameter adaptively.

VI. Conclusion

This method has been tested that shows encouraging results on several examples. It is based on the early human vision system which is a heuristic idea for parameter selecting. What's more, the adjacent multiplication ensures the sparsity and robustness of the method. This method can be used as a standard to help select regularization parameter.

Acknowlegement

The real data of the experiment derives from Mayo Clinics, so the authors would like to say thanks here to Dr. Cynthia McCollough, the Mayo Clinic, the American Association of Physicists in Medicine, and grants EB017095 and EB017185 from the national Institute of Biomedical Imaging Bioengineering from the USA.

References

[1] Kunisch, Karl, and J. Zou. "Iterative choices of regularization parameters in linear inverse problems." Inverse Problems14.5(1998):34-40.

[2] Kilmer, Misha E., and D. P. O'Leary. "Choosing Regularization Parameters in Iterative Methods for Ill-Posed Problems." Siam Journal on Matrix Analysis & Applications 22.4(2000):1204-1221.

[3] Morozov, V. A. "On the solution of functional equations by the method of regularization." Soviet Math Dokl 7.3(1966):510-512.

[4] G. Golub, M .Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter." Technometrics, vol.21, no. 2,pp. 215-223, 1979

[5] Hansen, By P C. "Analysis of Discrete Ill-Posed Problem by means of L-Curve." Soc. Industr. Appl. Mathem. Rev. 1992 2010.

[6] Lawson C L, Richard J. Hanson "Solving Least Squares Problems[J]. Mathematics of Computation, 1976, 30(135).

[7] Ito, Kazufumi, B. Jin, and T. Takeuchi. "A Regularization Parameter for Nonsmooth Tikhonov Regularization." Siam Journal on Scientific Computing 33.3(2011):1415-1438.

[8] Xuanqin Mou, Ti Bai, Xi Chen, Hengyong Yu, Qingsong Yang and Ge Wang. "Optimal Selection for Regularization Parameter in Iterative CT Reconstruction Based on the Property of Natural Image Statistics."Fully 3D,2015

[9] Xue, W., et al. "Blind Image Quality Assessment Using Joint Statistics of Gradient Magnitude and Laplacian Features." IEEE Transactions on Image Processing 23.11(2014):4850-62.

[10] Noo, F, and M. R. Defrise. "Single-slice rebinning method for helical cone-beam CT. " Physics in Medicine & Biology 44.2(1999):561-57.

[11] Li S C, Paramesran R. Review of medical image quality assessment[J]. Biomedical Signal Processing & Control, 2016, 27:145-154.

[12] Nakhaie A A, Shokouhi S B. No reference medical image quality measurement based on spread spectrum and discrete wavelet transform using ROI processing[J]. 2011, 8069(5):000121-000125.

[13] Dutta J, Ahn S, Li Q. Quantitative Statistical Methods for Image Quality Assessment[J]. Theranostics, 2013, 3(10):741-56.

[14] Lei,Tianhu, Statistics of medical imaging , Boca Raton, FL, 2011. p125-129