Non-Local Total Variation based Low-Dose Computed Tomography Denoising

SayedMasoud Hashemi[1] Soosan Beheshti[2] Richard S.C. Cobbold[1] Narinder S. Paul [3]

Abstract— **Radiation dose of X-ray Computed Tomography (CT) imaging has raised a worldwide health concern. Therefore, low-dose CT imaging has been of a huge interest in the last decade. However, lowering the radiation dose degrades the image quality by increasing the noise level, which may reduce the diagnostic performance of the images. As a result, image denoising is one of the fundamental tasks in low-dose CT imaging. One of the state of art denoising methods, which has been successfully used in this area, is Total Variation (TV) denoising. Nevertheless, if the parameters of the TV denoising are not optimally adjusted or the algorithm is not stopped in an appropriate point, some of the small structures will be removed by this method. Here, we provide a solution to this problem by proposing a modified nonlocal TV method, called probabilistic NLTV (PNLTV). Denoising performance of PNLTV is improved by using better weights and an appropriate stopping criterion based on statistics of image wavelet coefficients. Non-locality allows the algorithm to preserve the image texture, which combined with the proposed stopping criterion enables PNLTV to keep fine details unchanged.**

I. INTRODUCTION

CT utilization has increased dramatically over the last two decades; principally due to the unsurpassed speed and detail with which cross-sectional views of all soft tissues and organs of the body can be obtained. CT results in a relatively large radiation dose to patients compared to conventional radiography and there is concern that this results in increased risk of developing cancer [1], [2]. Therefore, low-dose CT imaging is clinically desirable and has been under investigation in the last decade. However, low-dose CT images suffer from low signal-to-noise ratio and severe artifacts, which affects the diagnostic performance and the confidence of the physicians. This problem has been addressed by different techniques, which can be categorized into three major classes: projection space denoising, image space denoising, and iterative reconstruction (IR) including compressed sensing based methods [3], [4], [5], [6], [7].

IR methods consider the imaging model and the statistical properties of the CT images, which increases the image quality and decreases the noise effect. However, the true IR methods, which are usually known as model based IR [8], are computationally intensive and time consuming, which has hindered their wide clinical application to date. As a result, new denoising methods with better performance are still under investigation.

The noise properties of the projection domain is well known. This helps to design better denoising algorithms. However, projection space denoising usually degrades the sharpness of the images [9]. Therefore, in this paper we focus on the image space denoising techniques. Our goal is to propose a CT image denoising algorithm that eliminates the noise, but keeps the image details unchanged.

Modeling of the noise in the image domain is not straight forward; but it is usually assumed to be additive i.i.d Gaussian [10]. This model helps the researchers to use the state of the art denoising methods, such as dictionary learning and sparse coding based algorithms [11], [12], multi-resolution transform shrinkage/thresholding [13], and Total Variation (TV) based denoising algorithms [14], [15], [16]. TV minimization scheme, which is our focus in this paper, offers a good combination of noise removal and feature preservation. If the noisy image y is modeled by $y = \bar{y} + n$, in which the original image $\bar{y}$ is corrupted with additive zero mean i.i.d Gaussian noise n with standard deviation of σ_n, TV denoising uses the following optimization scheme to estimate the noiseless image $\hat{y}$:

$$\hat{y} = \text{argmin}_u \frac{\lambda}{2} \|y - u\|_2^2 + TV(u) \quad (1)$$

in which λ is a positive scalar, $\|x\|_2^2 = \sum_i x_i^2$ and $TV(u) = \int_\Omega |\nabla u(x)| dx$ where ∇ is the first order gradient of the image $y : \Omega \to \mathbb{R}$. A general limitation in all denoising algorithms including TV denoising approaches is losing or decreasing the contrast of the small structures [17]. One straight forward approach to achieve the best combination of noise removal and feature preservation is to tune the parameter λ. If λ is too large we may not remove enough noise. On the other hand, if λ is too small it will remove too many features and end up with a cartoon-like image [18]. Since tunning λ is a tedious task, usually it is chosen in a reasonable range, which combined with a good stopping criterion offers a good trade off between noise removal performance and feature preservation [16].

Non-Local denoising methods, introduced by Buades et al. [19], propose another approach to address the feature preservation issue in denoising problem. They exploit the repetitive information present in most images and utilize a measure of similarity between nearby image patches to estimate the image structures. This allows the non-local algorithms to preserve the image texture and fine detail. This idea is used in [20] to improve the performance of TV denoising by

[1] S. Hashemi (Email: sayedmasoud.hashemiamroabadi@mail.utoronto.ca), and R.S.C Cobbold are with the Institute of Biomaterials and Biomedical Engineering, University of Toronto (UofT), Toronto, Canada.
[2] S. Beheshti is with the Department of Electrical and Computer Engineering, Ryerson University, Toronto, Canada.
[3] N.S. Paul is with University Health Network, joint Department of Medical Imaging and the Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, Canada.

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

proposing a non-local total variation (NLTV) algorithm.

In this paper, a modified NLTV, denoted by PNLM, is used for low-dose CT image denoising. We modify NLTV by utilizing better weights, introduced in [21], and improve its performance by using a good stopping criterion. The modified NLTV is describes in Section II and its performance is tested in Section III.

II. Probabilistic Non-Local Total Variation

Non-local total variation is formulated as follows [20]:

$$\hat{y} = \text{argmin}_u \frac{\lambda}{2} \|y-u\|_2^2 + TV_{NL}(u) \tag{2}$$

where $TV_{NL}(u) = \int_{\Omega\times\Omega} w(l,k)(u(l) - u(k))^2 dldk$. The weights $w(l,k)$ are adopted from NLM methods [19]:

$$w(l,k) = e^{-\int_\Omega G_a(z)|u(l+z)-u(k+z)|^2 dz/h^2} \tag{3}$$

where $\int_\Omega G_a(z)|u(l+z)-u(k+z)|^2 dz$ is the distance between patches located at l and k, G_a is a Gaussian function with standard deviation a, and h is a positive constant which acts as a scale parameter. Intuitively, the weighting function used in NLM and NLTV gives a larger weight to a pixel with a smaller patch difference. In [21] it has been shown that this weighting function could be problematic by assigning very different weights to equally probable patches. In addition, similar to classic TV denoising, the algorithm should be stopped in an appropriate point to keep the details unchanged. These two problems are addressed in the proposed method, as described in the following sections.

A. Improved Probabilistic Weighting Function

To address the weighting problem, we use the function proposed in [21] for non-local mean denoising, to modify NLTV. As the additive noise is assumed to be i.i.d and Gaussian, the difference between the patches $D_{l,k} = \sum_z d_{l+z,k+z}$ with pixel distances $d_{l,k} = (y(l) - y(k))^2/2\sigma_n^2$, can be statistically modeled by chi-square distribution. Using this assumption, the weights are defined as $w(l,k) = f(D_{l,k})$ where $f(D)$ is as follows (for $l \neq k$) [21]:

$$f(D) = \chi^2_{\eta_k}(D/\gamma_k) = \frac{(D/\gamma_k)^{\eta_k/2-1}\exp(-D/(2\gamma_k))}{2^{\eta_k/2}\Gamma(\eta_k/2)} \tag{4}$$

with $\gamma_k = var[D_{l.k}]/(2E[D_{l,k}])$ and $\eta_k = E[D_{l,k}]/\gamma_k$, which can be calculated numerically from the measured patch distances. The center pixel is weighted by $w(l,l) = \chi^2_{|\mathbb{P}|}(|\mathbb{P}|)$, in which $\mathbb{P}$ is the patch size.

B. Stopping Criterion and Parameter Adjustment

TV denoising algorithm usually removes noise as well as small structures from the images. Consequently, unless the parameter λ is carefully calibrated the denoised image is over-smoothed and the denoised image has cartoon-like features. A potential solution is to stop the algorithm in an appropriate point before the image gets very smooth. Here, we propose a stopping criterion based on the statistical properties of wavelet transform of the image. It is known that the high frequency wavelet coefficients of a noiseless image are very close to zero. In [22], this property is used by introducing a parameter $\mu(t) = 1/|J_P|\sum_{j\in J_P|\beta_j(t)|}$ which is the summation of absolute value of the wavelet coefficients β_j in high frequency subband J_P in t^{th} iteration, with $|J_P|$ representing the number of wavelet coefficients in that subband. This parameter should be very small in a noiseless image. Using a preset threshold, when μ gets smaller than that threshold the TV algorithm will be stopped. Although this criterion can be very helpful, it introduces an additional parameter that controls the stopping point and as a result, controls the result. To solve this issue, we use the idea used in Median-Absolute Deviation (MAD) [23]. MAD suggests that the median of absolute value of the high frequency wavelet coefficients of a noiseless image should be zero and the noise standard deviation of a noisy image can be estimated with this value.

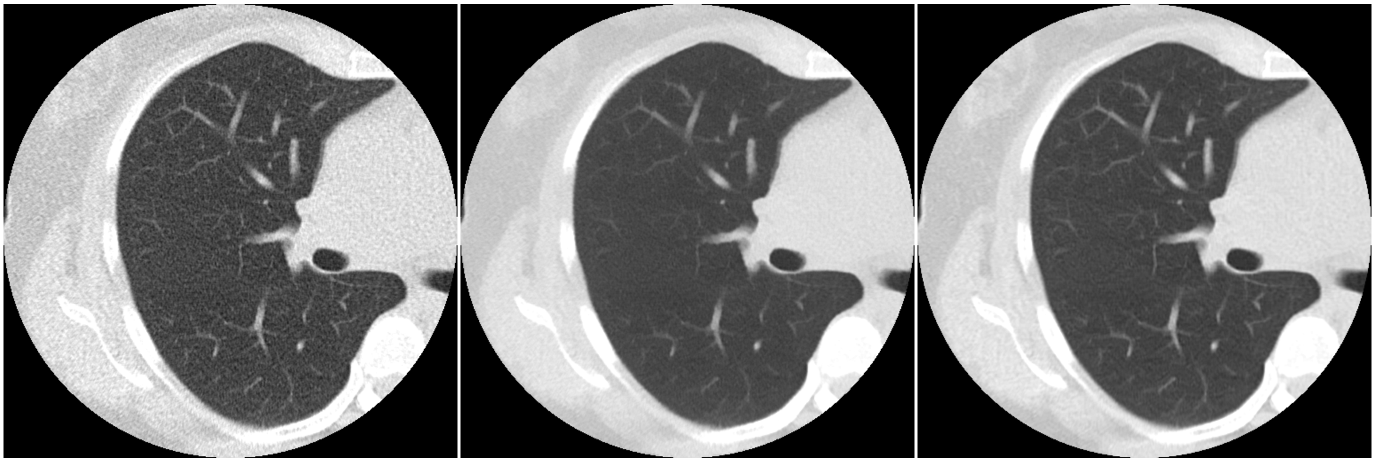
We calculate MAD at each iteration of the proposed scheme shown in Algorithm 1. In this algorithm we start with a relatively small λ, which over-smooths the image. In each iteration the image is denoised with the modified NLTV, the noisy image is partially added to this denoised image to recover the removed structures, and λ is increased for the next iteration. Consequently, MAD should be decreased in consecutive iterations until an optimum point, in which the MAD values start increasing. Therefore, PNLTV should be stopped in this point. The split Bregman method proposed in [20] is used to solve PNLV optimization scheme at each iteration.

Algorithm 1 Proposed iterative scheme for CT denoising.

Initialize: $\alpha \geq 1$, $0 \leq \beta < 1$, small λ, $maxiter$
$u^0 \leftarrow$ original image, $u \leftarrow$ original image
while $iter < maxiter$ **do**,
  $[cA, cH, cV, cD] = dwt2(u)$
  $MAD^{iter} = median(|cD|)$
  if $MAD^{iter-1} < MAD^{iter}$ **then**
    Break;
  end if
  $u^{iter} = Split\_Bregman(u, \lambda)$
  $\lambda \leftarrow \lambda \times \alpha$
  $u = \beta \times u^{iter} + (1-\beta)u^0$
end while

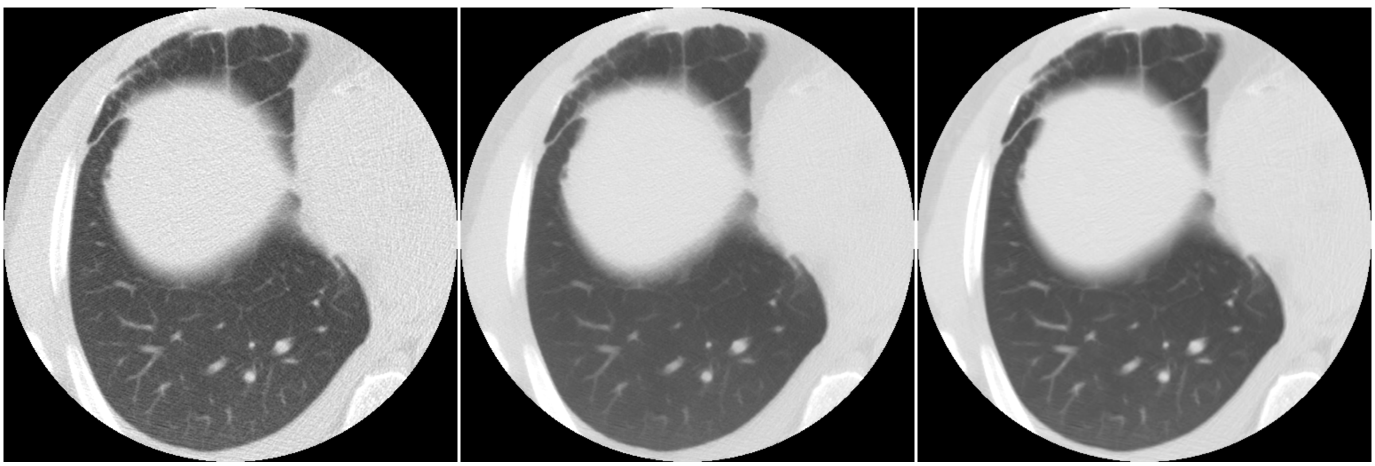
III. Results

To evaluate the proposed algorithm, 200 low and ultra-low dose chest CT images are denoised with the proposed method (PNLTV) and a regular non-local TV denoising method. Three axial slices of these patients are shown in Figures 1, 2, and 3. Since we do not have access to high quality images, the denoised images are evaluated qualitatively by comparing the fine details which are removed or kept unchanged. In addition, to have a quantitative measure of the denoising performance, noise standard deviation (STD) of the low and ultra-low dose images are compared with the corresponding denoised images. Noise STD is measured from a smooth



(A) (B) (C)

Fig. 1: Denoising of a low dose axial chest CT: (A) original image with noise STD of 64HU, (B) image denoised by NLTV, noise STD is 20HU, and (C) image denoised by PNLTV, noise STD is 24HU.



(A) (B) (C)

Fig. 2: Denoising of a low dose axial chest CT: (A) original image with noise STD of 54HU, (B) image denoised by NLTV, noise STD is 15HU, and (C) image denoised with PNLTV, noise STD is 15HU.

region in the images.

The low-dose protocol uses tube current $\times$ rotation time of 50mAs and tube voltage of 120kVp; and the ultra-low dose protocol uses 25mAs and 120kVp. All the images are shown with window-level/window-width of -550/1600.

Figures 1, 2, and 3 compare the CT images denoised with PNLTV and the non-local TV denoising method proposed in [20], denoted by NLTV. As it can be seen in these figures, although none of the fine details and small textures are removed by NLTV and there is no cartoon feeling in any of the images processed by NLTV, in some cases the contrast of very small structures are decreased. However, PNLTV keeps all the details and the textures unchanged. The noise standard deviation measurements show a 60$\pm$10% noise reduction by PNLTV and a 65%$\pm$15% noise reduction with NLTV. This shows that NLTV removes slightly more noise, but PNLTV preserves the fine details more effectively. In addition, based on the qualitative assessment of the images in our group, the images denoised by PNLTV are clinically preferred.

In this section, the parameters in Algorithm 1 are as follows: $\beta = 0.6$, $\alpha = 1.5$, and $maxiter = 50$.

IV. Conclusion

A modified non-local total variation denoising method was proposed in this paper to improve the contrast to noise ratio of the low and ultra low dose CT images. Total variation based methods often remove the image details, unless its parameters are adjusted precisely. Two approaches were combined here to overcome this problem: (1) using non-local TV method, which improves the preservation of the fine details by exploiting the similar patches in the image, and (2) using an efficient stopping criterion, which uses statistical properties of image wavelet coefficients, and stops the algorithm before it is over smoothed by TV denoising. In addition, non-local TV is improved by utilizing a probabilistic weighting function, which is based on the statistical

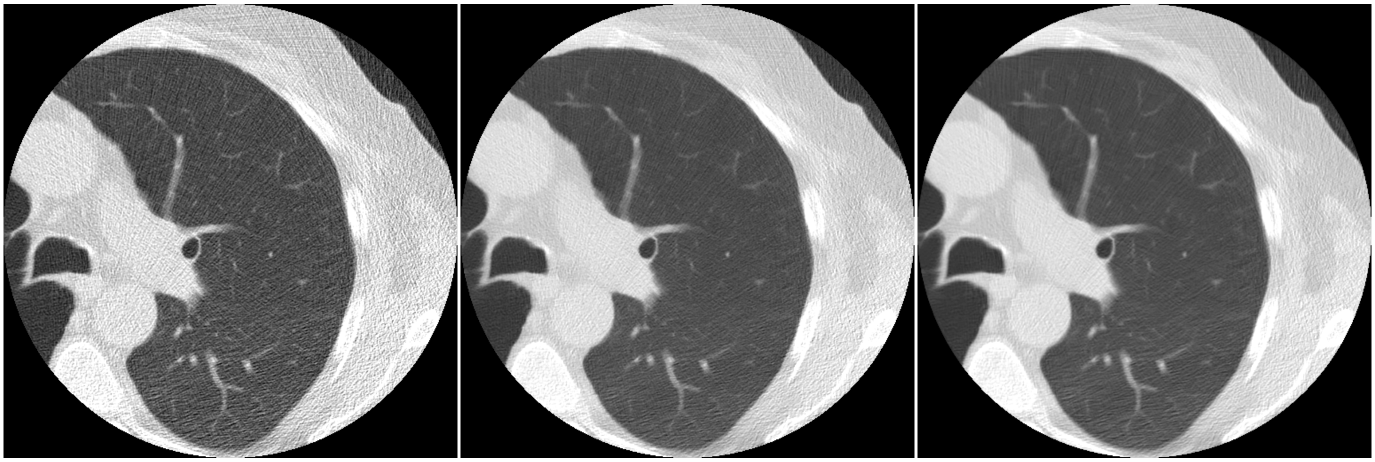


Fig. 3: Denoising of an ultra low dose axial chest CT: (A) original image with noise STD of 120HU, (B) image denoised by NLTV, noise STD is 45HU, and (C) image denoised by PNLTV, noise STD is 47HU.

model of the patch distances. The simulation results show that the proposed method, denoted by PNLTV, decreases the noise by 60% without losing the image details.

REFERENCES

[1] D.J. Brenner and E.J. Hall, "Computed tomography - an increasing source of radiation exposure," *The New England Journal of Medicine*, vol. 57, pp. 2277–2284, Nov. 2007.

[2] M.S. Pearce, J.A. Salott, M.P. Little, K. McHugh, C. Lee, K.P. Kim, N.L. Howe, C.M. Ronckers, P. Rajaraman, L. Parker A.W. Craft, and A.B. de Gonzálezl, "Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study," *The Lancet*, vol. 380, no. 9840, pp. 499–505, Aug. 2012.

[3] H. Lee, L. Xing, R. Davidi, R. Li, J. Qian, and R. Lee, "Improved compressed sensing-based cone-beam CT reconstruction using adaptive prior image constraints," *Physics in Medicine and Biology*, vol. 57, no. 8, pp. 2287–2307, Feb. 2012.

[4] B.P. Fahimian, Y. Zhao, Z. Huang, R. Fung, Y. Mao, C. Zhu, M. Khatonabadi, J.J. DeMarco, S.J. Osher, M.F. McNitt-Gray, and J. Miao, "Radiation dose reduction in medical x-ray ct via fourier-based iterative reconstruction," *Medical Physics*, vol. 40, no. 3, pp. 031914–1–031914–10, Mar. 2013.

[5] M. Hashemi, S. Beheshti, P.R. Gill, N.S. Paul, and R.S.C. Cobbold, "Fast fan/parallel beam CS-based low-dose CT reconstruction," in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, May 2013.

[6] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 182–202, 2009.

[7] J. Browne and A.B. de Pierro, "A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography," *IEEE Transaction on Medical Imaging*, vol. 15, no. 5, pp. 687–699, Oct. 1996.

[8] M. Katsura, I. Matsuda, M. Akahane, K. Yasaka, S. Hanaoka, H. Akai, J. Sato, A. Kunimatsu, and K. Ohtomo, "Sparse mri: the application of compressed sensing for rapid mr imaging," *Investigation Radiology*, vol. 48, no. 4, pp. 206–212, Apr. 2013.

[9] T. Li, X. Li, J. Wang, J. Wen, H. Lu, J. Hsieh, and Z. Liang, "Nonlinear sinogram smoothing for low-dose x-ray CT," *IEEE Transaction on Nuclear Science*, vol. 51, no. 5, pp. 2505–2513, Oct. 2004.

[10] T. Lei and W. Sewchand, "Statistical approach to x-ray CT imaging and its applications in image analysis. i. statistical analysis of x-ray CT imaging," *IEEE Transactions on Medical Imaging*, vol. 11, no. 1, pp. 53–61, Mar 1992.

[11] F. Yu, Y. Chen, and L. Luo, "Ct image denoising based on sparse representation using global dictionary," in *Complex Medical Engineering (CME), 2013 ICME International Conference on*, May 2013, pp. 408–411.

[12] K. Abhari, M. Marsousi, J. Alirezaie, and P. Babyn, "Computed tomography image denoising utilizing an efficient sparse coding algorithm," in *Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on*, July 2012, pp. 259–263.

[13] A. Borsdorf, R. Raupach, T. Flohr, and J. Hornegger, "Wavelet based noise reduction in CT-images using correlation analysis," *IEEE Transactions on Medical Imaging*, vol. 27, no. 12, pp. 1685–1703, Dec 2008.

[14] X.Q. Zhang and J. Froment, "Constrained total variation minimization and application in computerized tomography," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, Anand Rangarajan, Baba Vemuri, and AlanL. Yuille, Eds., vol. 3757 of *Lecture Notes in Computer Science*, pp. 456–472. Springer Berlin Heidelberg, 2005.

[15] E. Sideky, Y. Duchin, X. Pan, and C. Ullberg, "A constrained, total-variation minimization algorithm for low-intensity x-ray CT," *Medical Physics*, vol. 38, no. 7, pp. S117–S125, July 2011.

[16] Y. Wang and H. Zhou, "Total variation wavelet-based medical image denoising," *International Journal of Biomedical Imaging*, vol. 2006, 2006.

[17] J.M. Pauly M. Lustig, D. Donoho, "Sparse MRI: the application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, pp. 1182–1195, 2007.

[18] S. Osher, A. Sole, and L. Vese, "Image decomposition and restoration using total variation minimization and the H^1," *Multiscale Modeling Simulation*, vol. 1, no. 3, pp. 349–370, 2003.

[19] A. Buades, B. Coll, and J.M. Morel, "A non-local algorithm for image denoising," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005.*, June 2005, vol. 2, pp. 60–65.

[20] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *SIAM Journal on Imaging Sciences*, vol. 3, no. 3, pp. 253–276, 2010.

[21] Y. Wu, B. Tracey, P. Natarajan, and J.P. Noonan, "Probabilistic non-local means," *IEEE Signal Processing Letters*, vol. 20, no. 8, pp. 763–766, Aug. 2013.

[22] Y. Wang and H.M. Zhou, "Total variation wavelet based medical image denoising," *International Journal of Biomedical Imaging*, vol. 2006, 2006.

[23] D.L. Donohoa and I.M. Johnstonea, "Adapting to unknown smoothness via wavelet shrinkage," *Journal of the American Statistical Association*, vol. 90, no. 432, pp. 1200–1224, 1995.