Noise Removal for Medical X-ray images in Multiwavelet Domain

Ling Wang, Jianming Lu, Yeqiu Li, Takashi Yahagi

Abstract—When a signal is embedded in an additive Gaussian noise, its estimation is often done by finding a wavelet basis that concentrates the signal energy in few coefficients and then thresholding the noisy coefficients. However, in many practical problems such as medical X-ray image, astronomical and low-light images, the recorded data is not modeled by Gaussian noise but as the realization of a Poisson process. Multiwavelet is a new development to the body of wavelet theory. Multiwavelet simultaneously offers orthogonality, symmetry and short support which are not possible in scalar 2-channel wavelet systems. After reviewing this recently developed theory, a new theory and algorithm for denoising medical X-ray images using multiwavelet multiple resolution analysis (MRA) are presented and investigated in this paper. The proposed covariance shrink (CS) method is used to threshold wavelet coefficients. The form of thresholds is carefully formulated which is the key to more excellent results obtained in the extensive numerical simulations of medical image denoising compared to conventional methods.

I. INTRODUCTION

Noise in medical X-ray images [1] is primarily categorized into quantum mottle, which is related to the number of incident X-rays, and due to artificial noise which is due to the grid, etc. The effect of quantum mottle is manifested as an increase in the graininess of an X-ray image as the dose is reduced. Therefore, noise reduction is of great significance in medical X-ray images. The noise of X-ray images obeys a Poisson law and, hence, is highly dependent on the underlying light intensity pattern being imaged [2].

Over the past decade, wavelet thresholding has received a lot of attention from researchers in image denoising. Scalar orthogonal wavelets with a single mother-wavelet function have played an important role in denoising [4]. The denoising is done on the detail coefficients of wavelet transform, it has been shown that this algorithm offers the advantages of smoothness and adaptation. However, it is known that there is a limitation on the time-frequency localization of scalar wavelet functions.

Multiwavelet has been developed recently by using translates and dilates of more than one mother wavelet function [9]. They are known to have several advantages over scalar wavelets. In addition, Strela et al. [9] claimed that multiwavelet hard thresholding offers better results than the traditional scalar wavelet hard thresholding.

In this paper, a new shrink theory and denoising algorithm for medical X-ray images using multiwavelet multiple resolution analysis (MRA) is presented and investigated. We propose a new shrinkage method—covariance shrink (CS) method. The CS method is subband adaptive—a separate threshold is computed for each detail subband of multiwavelet transform. We propose near-optimal thresholds and a more suitable image denoising by extensive numerical simulations. We show in this paper that multiwavelet denoising gives better results than the conventional scalar wavelet denoising and conventional multiwavelet denoising.

II. SCALAR WAVELET AND MULTIWAVELET DENOISING

Scalar wavelet and multiwavelet denoising attempt to remove the noise present in the signal while preserving the signal characteristics, regardless of its frequency content. They involve three steps:

1) A linear forward wavelet transform.
2) Nonlinear shrinkage denoising.
3) A linear inverse wavelet transform.

Experimental results show that small wavelet coefficients are dominated by noise, while coefficients with a large absolute value carry more signal information than noise. Replacing noisy coefficients (small coefficients below a certain threshold value) by zeros and inverse wavelet transformation may lead to a reconstruction that has less noise. This method is indeed effective and thresholding is a simple and efficient method for noise reduction.

The most well-known thresholding methods include VisuShrink [14] and SureShrink [15]. These threshold choices enjoy asymptotic minimax optimality over function spaces such as Besov spaces. BayesShrink for denoising determines the threshold for each subband assuming a Generalized Gaussian Distribution (GGD). The results obtained for image denoising using BayesShrink are much better than VisuShrink and SureShrink.

In this paper, a new multiwavelet thresholding method based on the two types of multivariate thresholding approach is proposed. We call the proposed method covariance shrink (CS) method. BayesShrink and traditional multiwavelet vector thresholding approach will be the main comparison to the proposed method here, and as will be seen later in this paper, the proposed threshold often yields better results.
III. THE PROPOSED METHOD

In this section, we are going to introduce the proposed system structure and the CS method.

A. Proposed system structure

We propose a new denoising method in this paper for medical X-ray images in the multiwavelet domain. Fig. 1 shows the proposed system structure. The algorithm effectively performs the process of “transform,” “denoise,” and “inverse transform.” The proposed algorithm can be characterized as consisting of five main components.

1) Prefilter the shifted image into multiple streams by a specific prefilter.
2) Transform the preprocessed image into the multiwavelet domain using an orthogonal periodic multiwavelet transform.
3) Calculate the thresholds by using the covariance shrink (CS) method and apply the bivariate soft thresholding to the wavelet coefficients (leave the scaling coefficients alone).
4) Postprocess the image.
5) Inverse transform the result.

B. The proposed covariance shrink method

Firstly, we analysis the coefficients of multiwavelet transform. Fig. 2 shows the decomposition of the multiwavelet in a multiwavelet pyramid [18] obtained after L=2 steps, in the particular case of r=2.

We can conclude that if we use a different threshold for each subband, we can get better result than using the same threshold for all subbands because of the characteristic of coefficients in the multiwavelet domain.

Secondly, we introduce the bivariate thresholding method. The success of wavelet thresholding lies in that, usually, the signal will be compressed into a few large coefficients, whereas the noise components will be in the small coefficients only. Univariate thresholding is developed by Donoho and Johnstone, and two kinds of thresholding methods (soft and hard) are discussed. Univariate thresholding is successfully used in multiwavelet denoising by Strela et al. [10]. Even though univariate thresholding works in multiwavelet denoising, it does not give enough noise reduction. This is because multiwavelet transform normally produces correlated coefficients. Downie and Silverman developed the bivariate thresholding method by setting the universal threshold to $\lambda = 2\log n$ [16]. Suppose we apply the DMWT with an appropriate prefilter to a noisy function, then we obtain M stream coefficients, $D_{j,m,n}^* = D_{j,m,n}^s + E_{j,m,n}$, where $D_{j,m,n}^s$ denotes the signal coefficient, and $E_{j,m,n}$ indicates a multivariate normal distribution $N(0, V_j)$. The matrix $V_j$ is the covariance matrix for the error term that depends on the resolution scale $j$. Suppose we use a threshold of $\lambda$, then, the hard thresholding rule in bivariate thresholding can be written as

$$\hat{D}_{j,m,n} = \begin{cases} D_{j,m,n} & \text{if } \theta_{j,m,n} \geq \lambda \\ 0 & \text{otherwise} \end{cases}$$

(1)

where $\theta_{j,m,n} = \sqrt{D_{j,m,n}^T V_j D_{j,m,n}}$. The bivariate soft thresholding can be formulated as

$$\hat{D}_{j,m,n} = \begin{cases} D_{j,m,n} \left(1 - \frac{\lambda}{\theta_{j,m,n}}\right) & \text{if } \theta_{j,m,n} \geq \lambda \\ 0 & \text{otherwise} \end{cases}$$

(2)

To compute $\theta_{j,k}$, we need $V_j$, which can be obtained from $V_j = Var(D_{j,m,n})$, and we can explicitly obtain the covariance structure of the wavelet coefficients at each level. Here, we can use a robust covariance estimation method to estimate $V_j$ directly from the observed coefficients [17].

Finally, we introduce the proposed Covariance Shrink method. The thresholds of the CS method depend on the amount of each subband coefficients, coefficient’s values of

---

**Fig. 1.** Flow chart of the proposed method.

**Fig. 2.** Multiwavelet image decomposition up to two levels, with r=2.
each subband, multiwavelet base, and preprocessor. We calculate the thresholds by using the following equations.

\[ \lambda_{j,h} = \beta_{j,h} \sqrt{2\sigma^2(m \times n)} \]  
\[ \lambda_{j,v} = \beta_{j,v} \sqrt{2\sigma^2(m \times n)} \]  
\[ \lambda_{j,d} = \beta_{j,d} \sqrt{2\sigma^2(m \times n)} \]

where \( j \) is the decomposition level; \( h, v, d \) are the horizontal direction, the vertical direction, and the diagonal direction, respectively; \( c \) is the average variance, which depends on the selection of the multiwavelet base and preprocessor [7]; \( \sigma^2 \) is the variance of Gaussian noise; \( m \times n \) is the size of the subband of interest. To calculate the thresholds, we need to calculate the following equations.

1) Calculate the average value of coefficient magnitudes of each subband

\[ H_{j\text{mean}} = \frac{1}{m \times n} \sum_{p=1}^{m} \sum_{q=1}^{n} |H_{j,m,n}| \]  
\[ V_{j\text{mean}} = \frac{1}{m \times n} \sum_{p=1}^{m} \sum_{q=1}^{n} |V_{j,m,n}| \]  
\[ D_{j\text{mean}} = \frac{1}{m \times n} \sum_{p=1}^{m} \sum_{q=1}^{n} |D_{j,m,n}| \]

2) Calculate the average value of coefficient’s magnitudes of each level from 1)

\[ \text{Average}_j = \frac{H_{j\text{mean}} + V_{j\text{mean}} + D_{j\text{mean}}}{3} \]  
3) Calculate \( \beta_{j,h} \), \( \beta_{j,v} \), and \( \beta_{j,d} \).

when \( \sigma \leq 10 \), calculate \( \beta \) by using the following equations.

\[ \beta_{j,h} = \frac{\text{Average}_j}{H_{j\text{mean}}}, \beta_{j,v} = \frac{\text{Average}_j}{V_{j\text{mean}}}, \beta_{j,d} = \frac{\text{Average}_j}{D_{j\text{mean}}} \]  
otherwise, when \( \sigma > 10 \)

\[ \beta_{j,h} = \frac{\text{Average}_j}{H_{j\text{mean}}}, \beta_{j,v} = \frac{\text{Average}_j}{V_{j\text{mean}}}, \beta_{j,d} = \frac{\text{Average}_j}{D_{j\text{mean}}} \]  

IV. EXPERIMENTAL RESULTS

In this section, we show the experimental results and analyze the performance of the proposed algorithm. We test the proposed method on different medical X-ray images with different doses. To evaluate the medical X-ray image quality, we compute the mean-to-standard-deviation ratio (MSR) [13] in the desired region of interest (DROI). DROI is the region that the doctor is most interested in.

\[ \text{MSR} = \frac{\mu_d}{\sigma_d} \]

where \( \mu_d \) and \( \sigma_d \) are the mean and the standard deviation computed in the DROI.

In order to evaluate the efficiency of the proposed method more precisely, we also compute the Peak Signal to Noise Ratio (PSNR) which itself remains largely unsurpassed as a method of image quality measurement. PSNR we used in medical X-ray images is defined as:

\[ \text{PSNR} = 20 \log_{10}(\frac{1}{MSE}) \]  
\[ MSE = \frac{1}{H \times L} \sum_{x,y} (d(x,y)/\max(D) - o(x,y)/\max(O))^2 \]  

\( (d(x,y) \in D, o(x,y) \in O) \)

where the high-dose X-ray image is \( D \) (\( d(x,y) \in D \)) that contains \( H \) by \( L \) pixels and the reconstructed image is \( O \) (\( o(x,y) \in O \)).

We evaluate the proposed method by comparing to the following methods.

- The traditional multiwavelet method (MWV). The multiwavelet base is \( CL \), the preprocessing is row repeated preprocessing, the thresholding is scalar hard thresholding [7] and the threshold is calculated by the following equation.

\[ T = \sqrt{2 \times c \sigma^2 \log M \times N} \]

- \( c \) is the average variance.

- \( \sigma^2 \) is the variance of Gaussian noise.

- \( M \times N \) is the amount of coefficients in the multiwavelet transform.

- BayesShrink in wavelet domain (BS). The wavelet base is \( db5 \) (Daubechies wavelet with window is 5), the thresholding is univariate soft thresholding.

The PSNRs are given in Tables I. The MSRs are given in Tables II. Output results are given in Figs. 3-8. From the previously method tables, we can see that the proposed method gives better values for PSNR, MSR, and CNR than the traditional methods. From Figs. 3-8, we can see that we can get the best denoised image by using the proposed method.

From the figures and tables shown above, we know that the presented algorithm achieves the highest performance.
denoising works better than univariate denoising. Multiwavelet results than scalar wavelet. In addition, multiwavelet bivariate good general method. Multiwavelet denoising gives better combined with repeated row preprocessing appears to be a very impressive. Chui-Lina scaling functions and wavelets groups of noisy 2D test images, and the results are visually generally outperform scalar wavelet in image denoising for six main is proposed. Experimental results show that multiwavlets ical X-ray images using CS thresholding in multiwavelet do-
dose = 4.17691E-08c/kg.

Fig. 3. The results using different methods applied to Spine1, for dose = 4.17691E-08c/kg.

Fig. 4. The amplified images using different methods applied to Spine1, for dose = 4.17691E-08c/kg.

V. CONCLUSION

In this paper, a new method for removing noise from medical X-ray images using CS thresholding in multiwavelet domain is proposed. Experimental results show that multiwavelets generally outperform scalar wavelet in image denoising for six groups of noisy 2D test images, and the results are visually very impressive. Chui-Lina scaling functions and wavelets combined with repeated row preprocessing appears to be a good general method. Multiwavelet denoising gives better results than scalar wavelet. In addition, multiwavelet bivariate denoising works better than univariate denoising. Multiwavelet can obtain better result than scalar wavelet when noise the level is high and can preserve edges better than scalar wavelet. All experimental results show that the proposed method gives better results than conventional methods. Our future research will focus on finding more accommodating convergent values for wavelet coefficients. We also want to generalize it to all wavelets and employ it for other kinds of noise.

REFERENCES