IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 27, 2019, accepted February 26, 2019, date of publication March 8, 2019, date of current version March 25, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2903125

Improved Wavelet Denoising by Non-Convex Sparse Regularization Under Double Wavelet Domains

YONGJUN WU, GUANGJUN GAO, AND CAN CUI

Key Laboratory of Traffic Safety on Track of Ministry of Education, Central South University, Changsha 410075, China
Joint International Research Laboratory of Key Technology for Rail Traffic Safety, Central South University, Changsha 410075, China
National and Local Joint Engineering Research Center of Safety Technology for Rail Vehicle, Central South University, Changsha 410075, China
School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China

Corresponding author: Guangjun Gao (gjgao@csu.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 51605044, in part by the National Key Research and Development Program of China under Grant 2016YFB1200402, and in part by the Science Foundation of Hunan Province, China, under Grant 2016jj3004.

ABSTRACT This paper presents a double wavelet denoising (DWAD) method, which can preserve more details of an original signal. Although the noise removal method based on wavelet transform has been widely used, it still performs poorly for the signals with a low signal-to-noise ratio (SNR) or frequency overlap. Different from the wavelet denoising methods based on a single basis function, the DWAD considers filtering the wavelet coefficients of the noisy signal by threshold functions under two different wavelet domains, simultaneously. It considers using the difference of wavelet coefficient distribution and forcing the denoised signals under two wavelet domains to be the same to achieve more retention of details. In addition, the arctangent function is employed as a penalty function for wavelet coefficients to induce strong sparse wavelet coefficients. The DWAD is applied to one-dimensional signals and it is found that some wavelet coefficients which are smaller than the threshold could be retained during noise removal. The experiment results show that the average SNR of different noise levels is improved by at least 4.2 and 2.1 dB compared with the classical soft threshold method for the one-dimensional and image signals, respectively. Besides, the DWAD tends to obtain better performance on the details of original signals.

INDEX TERMS Wavelet transforms, signal denoising, non-convex regularization, sparse representation, double wavelet domains.

I. INTRODUCTION

Signal denoising is a common problem in signal and image processing. Wavelet transform is widely used in signal and image processing for a variety of applications. For noise reduction, it is a reasonably effective procedure to remove the noise under the wavelet representation domains. By using the threshold functions to filter the wavelet coefficients of noisy signal, the denoised signal can be obtained [1]–[3]. However, there are two problems that hinder the performance of wavelet denoising [4], [5]. 1) The basis function. For the same observed signal, the denoised signal has a difference if the employed basis function of wavelet is different. 2) The threshold function. When we remove the noise under wavelet representation domains, the threshold function would partly filter the real signal. Different threshold functions could induce different deviation.

For the denoising methods based on wavelet transform, they transform the signal to wavelet representation domains. Most of signal has the sparsity under wavelet domains. According to the sparse representation literature, wavelet denoising could be seemed as the sparse approximation under the wavelet representation domains [6]. In the process of signal reconstruction by using sparse approximation methods, the sparsity of signal has great influence on the accuracy. Strong sparsity could induce more accurate reconstructed results. For wavelet-denoising methods, the wavelet representation domain generated by specific wavelet basis function determines the sparsity of wavelet representation coefficients. Besides, the threshold function is employed to obtain the sparse representation coefficients of denoised signal, so it would also affect the sparsity of the reconstructed signal.

The associate editor coordinating the review of this manuscript and approving it for publication was Ramesh Babu N.

Therefore, it could be concluded that both basis function and threshold function have an important influence on the sparsity of denoised signal. In order to enhance the sparsity of denoised signal for wavelet denoising, the improvements for both basis function and threshold function are taken into account [7], [8].

A. DESIGN OF WAVELET REPRESENTATION DOMAINS

The wavelet representation domain depends on the wavelet basis function. Therefore, it is reasonably effective to design specific basis function for a particular signal. For image processing, some specific wavelet basis functions are constructed to obtain sparse representation coefficients [9], especially the Contourlet [10]–[12] and Curvelet [13]–[15]. Though the representation domain generated by this way is based on strict mathematic model and holds good sparsity, the mathematical complexity makes it not easy to design an appropriate basis function for the specific signal.

In addition, the redundant wavelet representation domains are considered. The noise removal methods based on undecimated wavelet transform [16], [17] and redundant wavelet domains [18], [19] are proposed. Besides, some over-complete representation domains are designed by the concatenation of two wavelet transform matrices in [20] and [21]. The concatenation transform matrices do not use the difference of wavelet coefficients distribution and are suitable for the morphological component analysis [22].

B. IMPROVEMENT OF THRESHOLD FUNCTIONS

As for threshold function, classic hard and soft threshold function are widely used in noise reduction in wavelet domains [23]. Nevertheless, both of them usually do not have good performance for noise removal. Hard threshold function does not hold good smoothness of reconstructed signal. Soft threshold function tends to filter part of original signal during filtering the noise though it keeps the reconstructed signal smooth. In order to improve the performance of wavelet threshold denoising, two classes of threshold functions are considered.

For the first class of methods, a specific threshold function is designed according to the characteristics of the distribution of wavelet coefficients. In order to balance the advantages of soft threshold function and hard threshold function, the semi-threshold [24] and garrote threshold [25] functions are proposed. Further, some nonlinear threshold functions are designed. Reference [26] proposes a new nonlinear threshold function based on non-Gaussian bivariate distributions. Instead of setting the coefficients below the threshold value to zero, [27] considers adjusting these coefficients by a polynomial function, and corresponding threshold function is designed. And the OLI-Shrink function proposed in [28] considers using the optimal linear interpolation. Besides, [29] employs the fuzzy logic to design the threshold function.

The other class of methods considers designing the threshold functions to induce sparse wavelet coefficients. The hard and soft threshold function could induce sparse representation coefficients and be employed for the sparse reconstruction with regularization by L0 norm and L1 norm, respectively [30], [31]. In order to improve the sparsity of reconstructed signal, several special penalty functions for L1 norm are designed, including the logarithmic function [32], the arctangent [33] function, the minimax-concave function [34], [35], the rational function [36] and exponential function [37], [38]. By using these penalty functions, the corresponding threshold functions would be deduced. Those threshold functions have good performance of inducing sparsity [36], [39]. However, all of them focus on the coefficients that are greater than the threshold value and ignore those smaller. Therefore, though those above-described methods are efficient, limited improvement can be achieved.

C. THE PROPOSED METHOD

The wavelet coefficients of original signal have different distribution when the signal is represented by different basis functions. When the threshold functions are used to filter noise, those coefficients that are less than threshold value would be filtered. Which makes some details of signal would be lost during the process of denoising. The part of original signal filtered by threshold function varies from wavelet basis function.

In view of the wavelet coefficients of original signal hold different distributions under different wavelet domains, some details of original signal which could be filtered in specific wavelet domain would be preserved when using other wavelet domains. Therefore, we consider employing two normal wavelet basis functions to denoise simultaneously. By using the difference of wavelet coefficient distribution and forcing the denoised signals in two wavelet domains to be the same, more detail would be preserved during filtering the noise. Besides, the penalty function is also employed to induce sparsity more strongly under two wavelet domains.

The rest of this paper is organized as fellows. In section II, we establish a denoising model based on double wavelet domains in view of the different distribution of wavelet coefficients in two wavelet domains. Section III presents some the experimental tests and the results analysis of different algorithms. In section IV, we conclude this paper.

II. SIGNAL DENOISING METHOD BASED ON DOUBLE WAVELET BASIS FUNCTIONS

A. PROBLEM STATEMENT

Generally, the real signal with noise is described as:

$$y = x + n \tag{1}$$

where the $y \in \mathbb{R}^N$ and $x \in \mathbb{R}^N$ are the observed and original signal, respectively. The n denotes the noise and it is usually assumed that obeys the Gauss distribution. For the denoised methods based on wavelet transform, they transform the signal to wavelet representation domains, and the denoising

model could be expressed as:

$$\mathbf{C} = \arg\min_{\mathbf{C}} \left\{ \frac{1}{2}\|\mathbf{C}_y - \mathbf{C}\|_2^2 + \|\mathbf{C}\|_1 \right\} \tag{2}$$

where $\mathbf{C}_y$ denotes the wavelet coefficients of observed signal y. The problem could be solved by soft threshold function and the denoised signal would be obtained by wavelet inverse transform.

By using wavelet transform, signal could be seemed as linear combine of wavelet basis under different scales. Both of the observed and original signals can be expressed by two wavelet basis functions:

$$\begin{cases} \mathbf{x} = \mathbf{W}_1\mathbf{C}_1 = \mathbf{W}_2\mathbf{C}_2 \\ \mathbf{y} = \mathbf{W}_1\mathbf{C}_{1y} = \mathbf{W}_2\mathbf{C}_{2y} \end{cases} \tag{3}$$

where $\mathbf{C}_1$, $\mathbf{C}_2$ and $\mathbf{C}_{1y}$, $\mathbf{C}_{2y}$ are the wavelet coefficients of original signal and observed signal under two wavelet domains, respectively. $\mathbf{W}_1$ and $\mathbf{W}_2$ denote the wavelet transform matrices based on two basis functions. In this paper, we consider to employ the translational-invariant (i.e., undecimated) wavelet transform [40], [41], which satisfies the Parseval frame condition:

$$\mathbf{W}_1^T\mathbf{W}_1 = \mathbf{W}_2^T\mathbf{W}_2 = \mathbf{I} \tag{4}$$

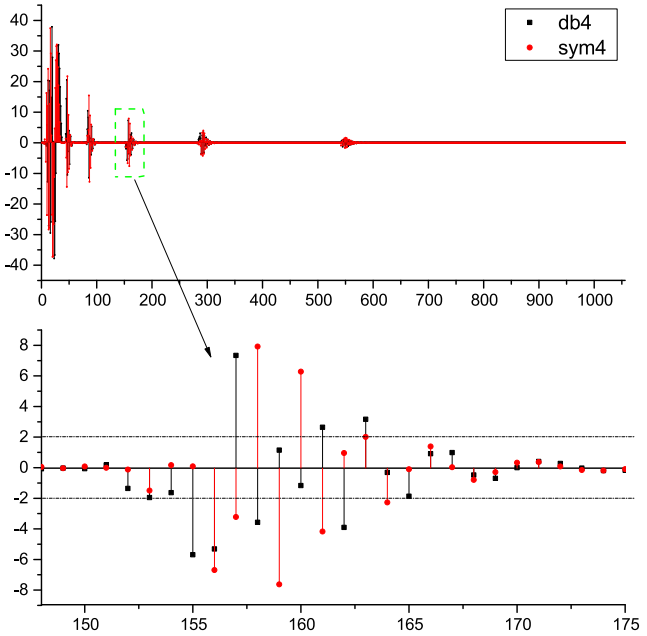


FIGURE 1. Wavelet coefficients (sorting by decomposition scales) of Doppler signal under db4 and sym4 wavelet domains.

Fig. 1 shows that the wavelet coefficients of classical Doppler signal under Daubechies (*db*4) and the Symlets (*sym*4) wavelet basis function with 4 vanishing moments are with different distribution. Comparing the amplitude of coefficients under two wavelet domains, it could be found that some wavelet coefficients in *db*4 domain are significant while they are small in *sym*4 domain. Which means some details of signal will be retained under other wavelet domains though they are filtered by threshold function under specific wavelet domains. Therefore, in order to preserve the details of signal as far as possible and improve the performance of denoising, we propose a unified model based on double wavelet basis functions (DWAD) to denoise the signal with non-convex optimization in this paper. The DWAD is descried by the constrained optimization problem as follow:

$$\begin{aligned} &\arg\min_{\mathbf{C}_1,\mathbf{C}_2} \frac{1}{2}\|\mathbf{C}_{1y} - \mathbf{C}_1\|_2^2 + \lambda_1 P_1(\mathbf{C}_1) + \lambda_2 P_2(\mathbf{C}_2) \\ &\textit{subject to}\ \ \mathbf{W}_1\mathbf{C}_1 = \mathbf{W}_2\mathbf{C}_2 \end{aligned} \tag{5}$$

where λ_1 and λ_2 denote the regularization parameters. The P_1 and P_2 are the penalty function terms for representation coefficients under two representation domains, respectively. In this paper, we consider using the same penalty function for the wavelet coefficients under two wavelet domains. Therefore, P_1 and P_2 have the same form and could be expressed as:

$$\begin{aligned} P_1(\mathbf{C}_1) &= \sum_{j=1}^{N} \phi((\mathbf{C}_1)_j, a) \\ P_2(\mathbf{C}_2) &= \sum_{j=1}^{N} \phi((\mathbf{C}_2)_j, a) \end{aligned} \tag{6}$$

where the ϕ is the penalty function, and N denotes the number of coefficients. The parameter a controls the non-convexity of the penalty function and it should be in the range of 0 to $1/\lambda$ [39]. Fig. 2 (a) and (b) illustrate several penalty functions and corresponding threshold functions, respectively. Those penalty functions are defined as [42]:

$$\phi(\mathbf{C}_j, a) = \begin{cases} \dfrac{1}{a} log(1 + a\,|x|), & log \\ \dfrac{|x|}{1 + a\,|x|\,/2}, & rat \\ \dfrac{2}{a\sqrt{3}}(tan^{-1}(\dfrac{1 + 2a\,|x|}{\sqrt{3}} - \dfrac{\pi}{6})), & atan \end{cases} \tag{7}$$

Specially, if both of wavelet basis functions is the same, which means $\mathbf{W}_1 = \mathbf{W}_2$, then the proposed model could be rewritten as:

$$\arg\min_{\mathbf{C}_1} \frac{1}{2}\|\mathbf{C}_{1y} - \mathbf{C}_1\|_2^2 + (\lambda_1 + \lambda_2)P_1(\mathbf{C}_1) \tag{8}$$

In this case, the proposed model would degenerate into the classical denoising model.

Note that the proposed model in 5 is different the over-complete wavelet transforms model which constructed the representation dictionary by the concatenation of two wavelet transform matrices [20], [21]. The over-complete wavelet transform matrix $D = [\mathbf{W}_1\ \mathbf{W}_2]$ does not satisfy the Parseval frame condition:

$$D^T D = \begin{bmatrix} \mathbf{I} & \mathbf{W}_1^T\mathbf{W}_2 \\ \mathbf{W}_2^T\mathbf{W}_1 & I \end{bmatrix} \neq \mathbf{I} \tag{9}$$

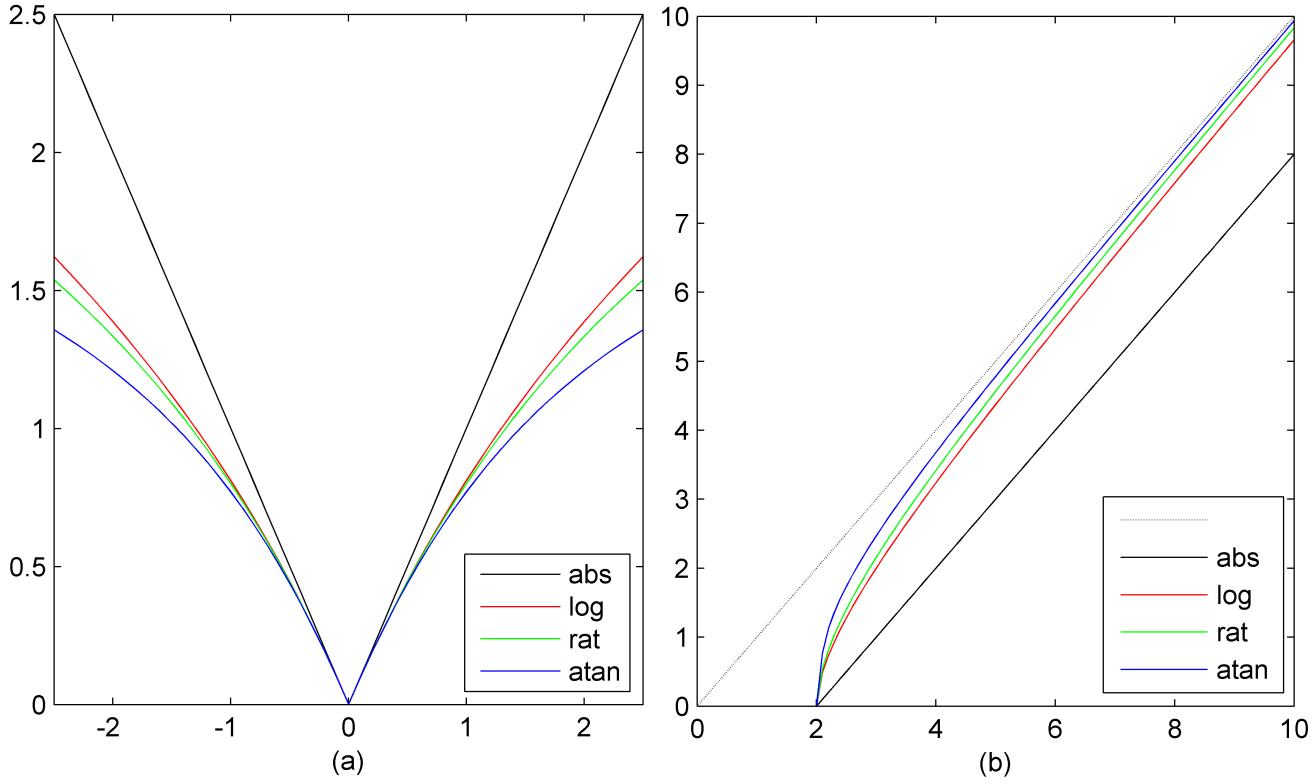


FIGURE 2. (a) Penalty functions, (b) threshold functions.

Therefore, the representation coefficients that are employed to represent $\mathbf{x} = D\mathbf{C}_d$ under over-complete wavelet dictionary could not be obtained by the transposed matrix:

$$\mathbf{C}_d \neq D^T\mathbf{x} = \begin{bmatrix} \mathbf{W}_1^T \\ \mathbf{W}_2^T \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \end{bmatrix} \tag{10}$$

B. THE DWAD ALGORITHM

The problem 5 is a convex optimization problem with constraints when parameter a in penalty function is with $0 \leq a \leq 1/\lambda$. it is not easy to resolve this problem directly. But the optimization method based on convex convergence theory could be employed to solve the problem. The alternating direction method of multipliers (ADMM) is well suited to distributed convex optimization [43]. According to the augmented Lagrangian method [43], [44], the constrained problem 5 could be rewritten as unconstrained optimization problem by a quadratic penalty approach. We form the augmented Lagrangian:

$$L(\mathbf{C}_1, \mathbf{C}_2, \mu) = \frac{1}{2}\|\mathbf{C}_{1y} - \mathbf{C}_1\|_2^2 + \lambda_1 P_1(\mathbf{C}_1) + \lambda_2 P_2(\mathbf{C}_2) + \frac{\mu}{2}\|\mathbf{W}_1\mathbf{C}_1 - \mathbf{W}_2\mathbf{C}_2 - \mathbf{d}\|_2^2 \tag{11}$$

where $\mu > 0$. According to [43] and [45], the problem 5 can be solved by iteratively minimizing $\mathbf{C}_1$ and $\mathbf{C}_2$ in 11. For each iteration in ADMM, there are three steps and the optimal solution of the sub-problem is obtained alternately.

$$\mathbf{C}_1^{k+1} = \arg\min_{\mathbf{C}_1} \frac{1}{2}\|\mathbf{C}_{1y} - \mathbf{C}_1\|_2^2 + \lambda_1 P_1(\mathbf{C}_1) + \frac{\mu}{2}\|\mathbf{W}_1\mathbf{C}_1 - \mathbf{W}_2\mathbf{C}_2^k - \mathbf{d}^k\|_2^2 \tag{12}$$

$$\mathbf{C}_2^{k+1} = \arg\min_{\mathbf{C}_2} \frac{\mu}{2}\|\mathbf{W}_1\mathbf{C}_1^{k+1} - \mathbf{W}_2\mathbf{C}_2 - \mathbf{d}^k\|_2^2 + \lambda_2 P_2(\mathbf{C}_2) \tag{13}$$

$$\mathbf{d}^{k+1} = \mathbf{d}^k - (\mathbf{W}_1\mathbf{C}_1^{k+1} - \mathbf{W}_2\mathbf{C}_2^{k+1}) \tag{14}$$

where we suggest initializing $\mathbf{d} = 0$ and $\mathbf{C}_2^0 = \mathbf{C}_{2y}$. The sub-problem 12 and 13 are strictly convex and could be solved under two wavelet domains, respectively.

For sub-problem 12, the transform matrix $\mathbf{W}_1$ is coupled with the optimal variable $\mathbf{C}_1$, and it could not be able to solve the sub-problem easily. Therefore, we consider using the Majorization-Minimization (MM) approach to minimize the sub-problem and the iteration algorithm is employed to approximate the optimal solution. The surrogate sub-objective function for sub-problem 12 is describes as:

$$G_{1n}(\mathbf{C}_1, \mathbf{C}_{1n}) = \frac{1}{2}g_{1n}(\mathbf{C}_1, \mathbf{C}_{1n}) + \lambda_1 P_1(\mathbf{C}_1) \tag{15}$$

where $\mathbf{C}_{1n}$ is the optimal solution for the previous iteration. The surrogate function $g_{n1}(\mathbf{C}_1, \mathbf{C}_{1n})$ is designed by MM

framework [46], [47], and it is defined as:

$$\begin{aligned} g_{n1}(\mathbf{C}_1, \mathbf{C}_{1n}) &= \|\mathbf{C}_{1y} - \mathbf{C}_1\|_2^2 + \mu\|\mathbf{W}_1\mathbf{C}_1 - \mathbf{W}_{2d}\|_2^2 \\ &\quad + \mu(\mathbf{C}_1 - \mathbf{C}_{1n})^T(\alpha\mathbf{I} - \mathbf{W}_1^T\mathbf{W}_1)(\mathbf{C}_1 - \mathbf{C}_{1n}) \\ &= (1 + \mu\alpha)\|\mathbf{b}_1 - \mathbf{C}_1\|_2^2 + K_1 \end{aligned} \tag{16}$$

where

$$\begin{aligned} \mathbf{W}_{2d} &= \mathbf{W}_2\mathbf{C}_2 + \mathbf{d} \\ \mathbf{b}_1 &= \frac{1}{1+\mu\alpha}(\mathbf{C}_{1y} + \mu(\mathbf{W}_1^T\mathbf{W}_{2d} \\ &\quad + \mathbf{W}_1^T\mathbf{W}_1\mathbf{C}_{1n} - \alpha\mathbf{C}_{1n})) \end{aligned} \tag{17}$$

In 16, the K_1 does not depend on $\mathbf{C}_1$ and the parameter α should be satisfied:

$$\alpha \geq maxeig(\mathbf{W}_1^T\mathbf{W}_1) \tag{18}$$

Specially, we consider that the wavelet transform matrices satisfy the Parseval frame condition in this paper. Therefore, we could combine the 9 and 16. Then the sub-problem 12 could be equivalent to the following problem:

$$\arg\min_{\mathbf{C}_1} \frac{1+\mu}{2}\|\frac{\mathbf{C}_{1y} + \mu\mathbf{W}_1^T\mathbf{W}_{2d}}{1+\mu} - \mathbf{C}_1\|_2^2 + \lambda_1 P_1(\mathbf{C}_1) \tag{19}$$

Finally, the optimal solution of the sub-problem 12 would be directly obtained by the threshold function.

$$\mathbf{C}_1^{k+1} = Th(\frac{\mathbf{C}_{1y} + \mu\mathbf{W}_1^T(\mathbf{W}_2\mathbf{C}_2^k + \mathbf{d}^k)}{1+\mu}; \frac{\lambda_1}{1+\mu}) \tag{20}$$

where the threshold function $Th()$ is defined by the corresponding penalty function. The parameter λ_1 denotes the regularization parameter under the first wavelet domain.

For the sub-problem 13 in which the optimal variable $\mathbf{C}_1$ is coupled with the transform matrix $\mathbf{W}_2$, it is similar to the sub-problem 12. Therefore, we can also employ the Majorization-Minimization (MM) approach to solve the sub-problem 13. The corresponding surrogate sub-objective function is describes as:

$$G_{2n}(\mathbf{C}_2, \mathbf{C}_{2n}) = \frac{\mu\beta}{2}\|\mathbf{b}_2 - \mathbf{C}_2\|_2^2 + K_2 + \lambda_2 P_2(\mathbf{C}_2) \tag{21}$$

where

$$\begin{aligned} \mathbf{W}_{1d} &= \mathbf{W}_1\mathbf{C}_1 - \mathbf{d} \\ \mathbf{b}_2 &= \frac{1}{\beta}(\mathbf{W}_2^T\mathbf{W}_{1d} + \beta\mathbf{C}_{2n} - \mathbf{W}_2^T\mathbf{W}_2\mathbf{C}_{2n}) \\ \beta &\geq maxeig(\mathbf{W}_2^T\mathbf{W}_2) \end{aligned} \tag{22}$$

Considering the situation where the wavelet transform matrices satisfy the Parseval frame condition, the sub-problem 13 could be simplified to the following problem:

$$\arg\min_{\mathbf{C}_2} \frac{\mu}{2}\|\mathbf{W}_2^T\mathbf{W}_{1d} - \mathbf{C}_2\|_2^2 + \lambda_2 P_2(\mathbf{C}_2) \tag{23}$$

Therefore, we could also employ the threshold function that is designated by specific penalty function to generate the optimal solution:

$$\mathbf{C}_2^{k+1} = Th(\mathbf{W}_2^T(\mathbf{W}_1\mathbf{C}_1^{k+1} - \mathbf{d}^k); \frac{\lambda_2}{\mu}) \tag{24}$$

where the λ_2 is the regularization parameter under the second wavelet domain. Note that, the representation coefficients of the original signal under different wavelet domains hold different sparsity.

For the wavelet domain in which the representation coefficients are with strong sparsity, we consider setting small regularization parameter. Therefore, we suggest that the regularization parameters for two wavelet domains are set as:

$$\begin{aligned} \lambda_1 &= \gamma\lambda \\ \lambda_2 &= (1-\gamma)\lambda \end{aligned} \tag{25}$$

where γ denotes proportion parameter and updates for each iteration. It is defined as:

$$\begin{aligned} \gamma^1 &= \frac{\mathbf{P}_1(\mathbf{C}_{1y})}{\mathbf{P}_1(\mathbf{C}_{1y}) + \mathbf{P}_2(\mathbf{C}_{2y})} \\ \gamma^{k+1} &= \frac{\mathbf{P}_1(\mathbf{C}_1^k)}{\mathbf{P}_1(\mathbf{C}_1^k) + \mathbf{P}_2(\mathbf{C}_2^k)} \end{aligned} \tag{26}$$

Through alternately solving two sub-problems by using the threshold function deduced by the penalty function, the filtered wavelet coefficients would be obtained under two wavelet domains, respectively. Different details of original signal would be preserved under different wavelet domains. The $\mathbf{d}$ in 14 is updated by the difference of denoised signal based on two wavelet basis functions, which imposes that the solutions of two sub-objective functions deduce the same denoised signal. Therefore, more details of original signal could be preserved during filtering the noise. The denoised signal can be obtained by two wavelet transform matrices, respectively.

$$\hat{\mathbf{x}} = \mathbf{W}_1\mathbf{C}_1 \approx \mathbf{W}_2\mathbf{C}_2 \tag{27}$$

C. THE DWAD PROCEDURE AND PARAMETERS SETTING

According to 12 - 27, we could conclude the procedure of proposed DWAD. Finally, the iterative denoising algorithm based on double wavelet basis functions could be described as Algorithm 1.

It is very important to select appropriate wavelet function for wavelet denoising. In the DWAD method, two criteria are used to determine the wavelet basis function. The first criterion is the sparsity of wavelet coefficients. For sparse approximation methods, the sparsity of representation coefficients has great influence on the accuracy of signal reconstruction. Therefore, the wavelet basis functions, which induce strong sparse representation coefficients, should be preferred. Specifically, we choose the wavelet basis functions with smaller L1 norm for the observed signal. Secondly, we consider the correlation coefficient of two wavelet functions. In the DWAD method, the difference of wavelet coefficients for signal under two representation domains is utilized. Some details of original signal would be preserved in one wavelet domain while they are filtered under another wavelet domain. The difference of wavelet coefficients in two representation domains is beneficial to the preservation

Algorithm 1 DWAD Algorithm

Input:

Wavelet transform matrix $\mathbf{W}_1 \in \mathbb{R}^{N\times N}$, $\mathbf{W}_2 \in \mathbb{R}^{N\times N}$.
Observed vector $y \in \mathbb{R}^N$
Number of iterations *Iter*

Output:

An estimate of $\hat{\mathbf{x}} \in \mathbb{R}^N$ the original signal $\mathbf{x}$

Initialize regularization parameters λ, μ
$\mathbf{C}_{1y} \leftarrow \mathbf{W}_1^T\mathbf{y}$
$\mathbf{C}_{2y} \leftarrow W_2^T\mathbf{y}$
$\mathbf{C}_2 \leftarrow \mathbf{C}_{2y}$
$\mathbf{d}_0 \leftarrow \mathbf{0}$
$\gamma \leftarrow \frac{P_1(\mathbf{C}_{1y})}{P_1(\mathbf{C}_{1y})+P_2(\mathbf{C}_{2y})}$
for $i = 0 \rightarrow Iter$ **do**
  $\mathbf{W}_{2d} \leftarrow \mathbf{W}_2\mathbf{C}_2 + \mathbf{d}$
  $\mathbf{b}_1 \leftarrow \frac{\mathbf{C}_{1y}+\mu\mathbf{W}_1^T\mathbf{W}_{2d}}{1+\mu}$
  $\lambda_1 \leftarrow \gamma\lambda$
  $\mathbf{C}_1 \leftarrow Th(\mathbf{b}_1, \frac{\lambda_1}{1+\mu})$
  $\mathbf{W}_{1d} \leftarrow \mathbf{W}_1\mathbf{C}_1 - \mathbf{d}$
  $\mathbf{b}_2 \leftarrow \mathbf{W}_2^T\mathbf{W}_{1d}$
  $\lambda_2 \leftarrow (1-\gamma)\lambda$
  $\mathbf{C}_2 \leftarrow Th(\mathbf{b}_2, \frac{\lambda_2}{\mu})$
  $\mathbf{d} \leftarrow \mathbf{d} - (\mathbf{W}_1\mathbf{C}_1 - \mathbf{W}_2\mathbf{C}_2)$
  $\gamma \leftarrow \frac{P_1(\mathbf{C}_1)}{P_1(\mathbf{C}_1)+P_2(\mathbf{C}_2)}$
end for
return $\hat{\mathbf{x}} \leftarrow \mathbf{W}_1\mathbf{C}_1$

of details. Therefore, we trend to choose two wavelet basis functions which hold low correlation coefficient.

For ADMM, which is employed to solve the DWAD model, the parameter μ affects the convergence speed of the algorithm. In order to improving the convergence in practice, a simple scheme is proposed and works well according to [43]. In this scheme, μ is updated for each iteration:

$$\mu^{k+1} = \begin{cases} s_1\mu^k, & \|r^k\|_2 > s\|e^k\|_2 \\ \mu^k/s_2, & \|e^k\|_2 > s\|r^k\|_2 \\ \mu^k, & otherwise \end{cases} \tag{28}$$

where the r^k is the residual of k step and is defined $r^k = \mathbf{W}_1\mathbf{C}_1^k - \mathbf{W}_2\mathbf{C}_2^k$. The e denotes the dual residual and it is defined $e^k = \mu(\mathbf{W}_2\mathbf{C}_2^{k+1} - \mathbf{W}_2\mathbf{C}_2^k)$. The parameters s_1 and s_2 are usually set 2, and it is a typical choice to be $s = 10$.

III. NUMERICAL EXAMPLES

The proposed denoising method is an improvement of wavelet transform and tries to preserve more details of original signal during noise removal. Therefore, the proposed method is suitable for the signals with good sparsity under wavelet domains. In order to measure the performance of the proposed method, some numerical examples are carried out. In view of the noise removal methods based on wavelet transform are not suitable for signals with low signal-to-noise ratio, it is further suggested that the SNR should be higher than 10dB and peak signal to noise ratio (PSNR) be higher than 20dB for the one dimension and image signals, respectively. The decimated and undecimated wavelets transforms, both of them satisfy the Parseval frame condition, are employed to test the proposed method, respectively. Besides, the influence of regularization parameter is analyzed.

A. THE DWAD BASED ON DECIMATED WAVELET TRANSFORMS

Firstly, in order to test the performance of reserving the wavelet coefficients of original signals, especially for the coefficients which are smaller than the threshold and usually represents the details of original signals, two different transform matrices are generated by the decimated wavelet transforms.

We consider employing the classical Doppler signal length of 1024 as the original signal and the corrupted signal by Additive White Gaussian Noise (AWGN) as observed signal. We set the signal-to-noise ratio (SNR) of the noisy signal to be 16. The observed signal holds smaller L1 norm under *sym*4 and *db*4 wavelet domains. Therefore, the two decimated wavelet transform matrices for the proposed DWAD are generated by the *sym*4 and *db*4 wavelet basis functions. Two 5-scale decimated wavelet transform are generated by two wavelet basis functions, respectively. For one dimensional signal, the convergence speed of DWAD algorithm is very fast. We suggest $Iter = 20$ iterations with the parameter $\mu = 0.8$.

In this numerical example, we focus on the performance of detail preservation of the original for different methods. Therefore, we consider using the universal threshold under different wavelet scales for noise filtering. Under *db*4 and *sym*4 wavelet domains, the classical soft method and the penalty functions methods (with the logarithmic function (Log), rational function (Rat) and arctangent function (Atan)) are employed to compare with the proposed DWAD method. For soft threshold method, the threshold is set to be 2.5σ. As for the penalty methods and proposed DWAD, we suggest that the threshold and parameter are $\lambda = 3.0$ and $a = 1/\lambda$, respectively.

Fig. 3 shows the results of denoised signal by soft method and the DWAD method without penalty function. According to the results, it could be found that the proposed DWAD holds better performance of noise removal and smoothness. Further, in order to evaluate the performance of different methods, we consider employing the SNR to measure the noise removal, and employing the root-mean-square-error (RMSE) and structural similarity index (SSIM) to measure the detail preservation ability of the original signal. Table 1 shows each index of compared methods.

As illustrated in Table 1, by using the penalty functions, the effect of noise removal and detail preserving are greatly enhanced. The Atan penalty function usually obtains the best performance. The results of denoised signal by Atan penalty

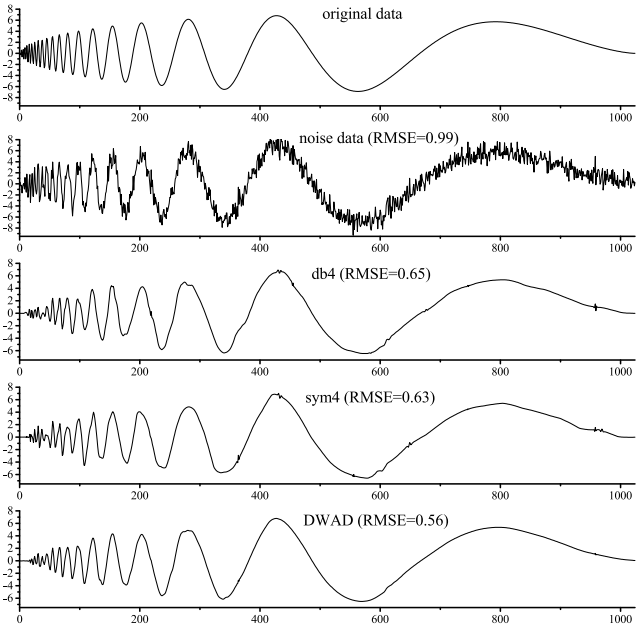


FIGURE 3. Denoised signal by the soft method and DWAD without penalty function.

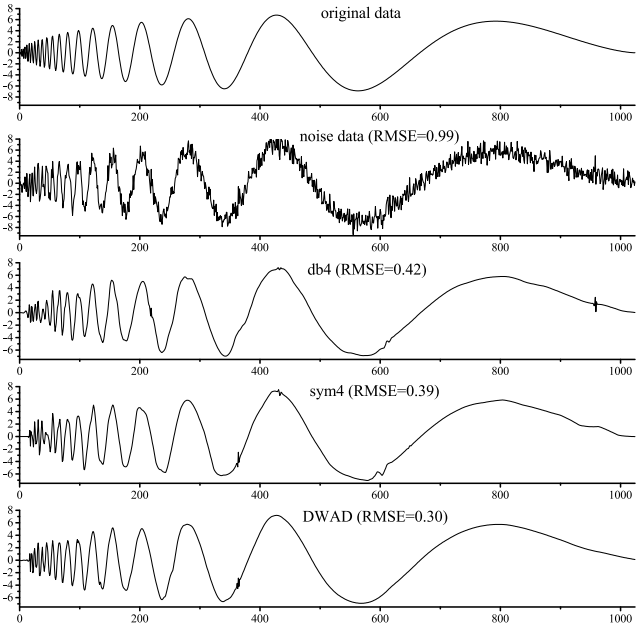


FIGURE 4. Denoised signal by the Atan penalty method and DWAD with Atan.

method and the DWAD method with Atan penalty function are illustrated in Fig. 4. Besides, when using the same penalty function, the proposed DWAD method has better performance than the method which based on only one wavelet domain in various indicators. By using the DWAD method, the SNR has an increase of about 2.3 dB.

As shown in Fig. 4, by using the proposed DWAD, the details of original signal are well preserved while the noise is suppressed well. Fig. 5 shows the wavelet coefficients of

TABLE 1. RMSE, SNR and SSIM of compared algorithms based on *db*4 and *sym*4 basis functions.

Methods	threshold	RMSE	SNR	SSIM
Soft_db4	2.5	0.65	16.0	0.9927
Soft_sym4	2.5	0.63	16.2	0.9928
Soft_DWAD	2.5	0.56	17.2	0.9943
Log_db4	3.0	0.45	19.1	0.9973
Log_sym4	3.0	0.41	19.8	0.9975
Log_DWAD	3.0	0.32	22.2	0.9987
Rat_db4	3.0	0.44	19.4	0.9976
Rat_sym4	3.0	0.40	20.1	0.9978
Rat_DWAD	3.0	0.31	22.4	0.9988
Atan_db4	3.0	0.42	19.7	0.9979
Atan_sym4	3.0	0.39	20.4	0.9981
Atan_DWAD	3.0	**0.30**	**22.7**	**0.9989**

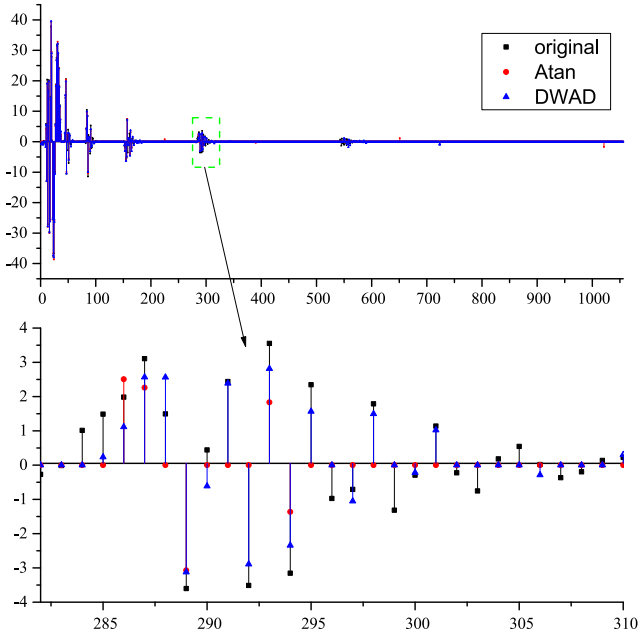


FIGURE 5. Wavelet coefficients (sorting by decomposition scales) of the denoised signals obtained by the penalty function and DWAD method under *db*4 wavelet domain.

the original signal, denoised signal by the Atan method and denoised signal by the DWAD method under *db*4 wavelet domain. According to Fig. 5, we could found that the DWAD method partially retains some wavelet coefficients which are filtered to be zero by using the Atan method.

Further, to more clearly compare the relative bias of the reconstructed wavelet coefficients by the Atan penalty function and the DWAD with Atan penalty function, these two sets of reconstructed wavelet coefficients are illustrated in Fig. 6 (a) and (b), respectively. As it is shown in Fig. 6, for both of two methods, the reconstructed wavelet coefficients $\mathbf{C}_r$ hold strong linear correlation with the wavelet coefficients $\mathbf{C}$ of the original signal in general. However, for those reconstructed wavelet coefficients that are less than the threshold, the reconstructed wavelet coefficients by the DWAD with Atan penalty function lie closer to the identity than the reconstructed wavelet coefficients by the Atan

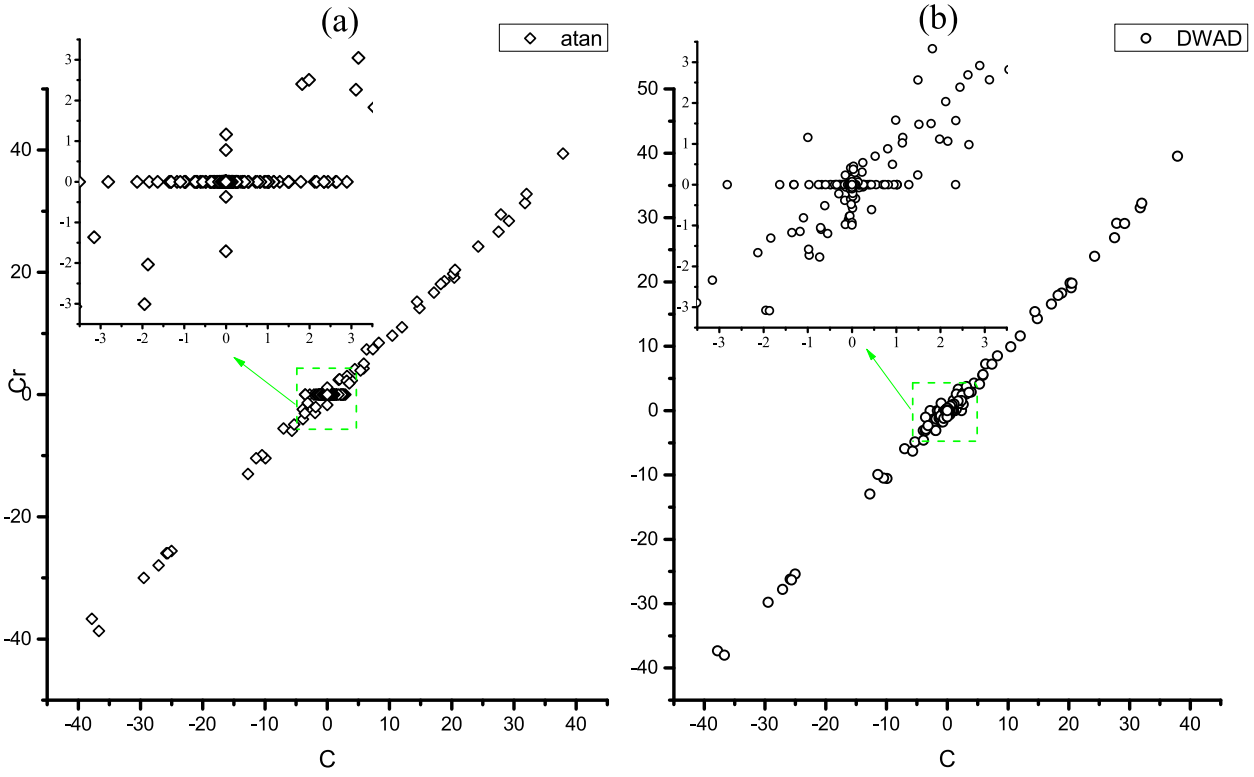


FIGURE 6. Relative bias of reconstructed wavelet coefficients by two methods under *db*4 wavelet domain.

penalty function. This is why the proposed DWAD could preserve details of original signal during noise removal.

Therefore, the proposed DWAD holds better performance of noise filtering and details preservation than those methods which are based on only one wavelet domain.

B. THE DWAD BASED ON UNDECIMATED WAVELET TRANSFORMS

For the threshold denoising methods under wavelet domains, it often introduces some artifacts such as spurious noise spikes and pseudo-Gibbs oscillations, especially for the signal with discontinuities. This hinders the further improvement of the wavelet denoising effect. The reason for pseudo-Gibbs artifacts is due to non-zero coefficients being erroneously set to zero [41]. In order to solve this problem, [41] consider employing the total variation (TV) as regularization in wavelet denoising model. Considering that the proposed DWAD model could partly preserve the wavelet coefficients which are set to zero by classical wavelet denoising model, we would test the artifacts reduction performance of DWAD in this numerical example.

To further improve the performance of wavelet denoising, we consider using the undecimated wavelet transforms in the proposed DWAD method. The observed signal holds small L1 norm under *haar* wavelet domain, because the original signal contains staircase. Besides, the *haar* and *sym*4 wavelet basis functions holds low correlation coefficient. Therefore, two 5-scale undecimated wavelet transform are generated by the *haar* and *sym*4 wavelet basis functions, respectively. Similarly, we suggest that the parameter μ is 0.8 and the number of iterations *Iter* is 20.

The hard threshold, soft threshold, penalty functions and the WATV proposed in [41] methods are employed to compare with the proposed DWAD method. According to [41], the regularization parameter λ is set to vary with wavelet scale and it could be expressed as:

$$\lambda_j = \frac{\lambda\sigma}{2^{j/2}} \tag{29}$$

where σ denotes the noise level and λ is a constant. For the WATV method, the λ is suggested to be 2.5 in [41]. As for the hard threshold, soft threshold, penalty function and DWAD methods, we recommend setting it to be 4.0, 3.0, 4.0 and 4.5, respectively. The parameter a for each wavelet scale is $1/\lambda_j$.

The original signal and observed signal with noise level of $\sigma = 4$ are illustrated respectively in Fig. 7 (a) and Fig. 8 (a). Fig. 7 (b) - (e) show the denoised signals by different methods under *haar* wavelet domain, whereas Fig. 8 (b) - (e) illustrate the corresponding results under *sym*4 wavelet domain. Both of Fig. 7 (f) and Fig. 8 (f) denote the denoised result by using the proposed DWAD method. The result of DWAD achieves the best performance of RMSE. For those methods those are employed to compare with DWAD, the results under *haar* wavelet domain have better RMSE than under *sym*4 wavelet domain.

TABLE 2. Average SNR of several algorithms by different representation domains with 20 trials (RMSE,PSNR).

σ	Hard threshold		Soft threshold		Penalty (Atan)		WATV		DWAD
	haar	*sym4*	*haar*	*sym4*	*haar*	*sym4*	*haar*	*sym4*	
1	(0.48, 31.5)	(0.42, 32.6)	(0.66, 28.6)	(0.75, 27.5)	(0.44, 32.2)	(0.42, 32.5)	(0.50, 31.1)	(0.38, 33.4)	**(0.37, 33.6)**
2	(0.72, 27.9)	(0.74, 27.7)	(1.19, 23.6)	(1.36, 22.4)	(0.73, 27.8)	(0.83, 26.7)	(0.80, 27.0)	(0.69, 28.3)	**(0.64, 28.9)**
4	(1.37, 22.3)	(1.61, 21.0)	(2.10, 18.6)	(2.34, 17.7)	(1.50, 21.6)	(1.83, 19.8)	(1.38, 22.3)	(1.35, 22.5)	**(1.24, 23.3)**
6	(2.05, 18.9)	(2.37, 17.6)	(2.80, 16.1)	(3.03, 15.4)	(2.26, 18.0)	(2.64, 16.7)	(1.90, 19.5)	(1.99, 19.1)	**(1.87, 19.7)**
8	(2.66, 16.6)	(2.98, 15.6)	(3.36, 14.5)	(3.54, 14.1)	(2.92, 15.8)	(3.25, 14.8)	**(2.41, 17.4)**	(2.55, 17.0)	(2.54, 17.0)
Avg.	(1.46, 23.4)	(1.62, 22.9)	(2.02, 20.3)	(2.20, 19.4)	(1.57, 23.1)	(1.79, 22.1)	(1.40, 23.5)	(1.39, 24.1)	**(1.33, 24.5)**

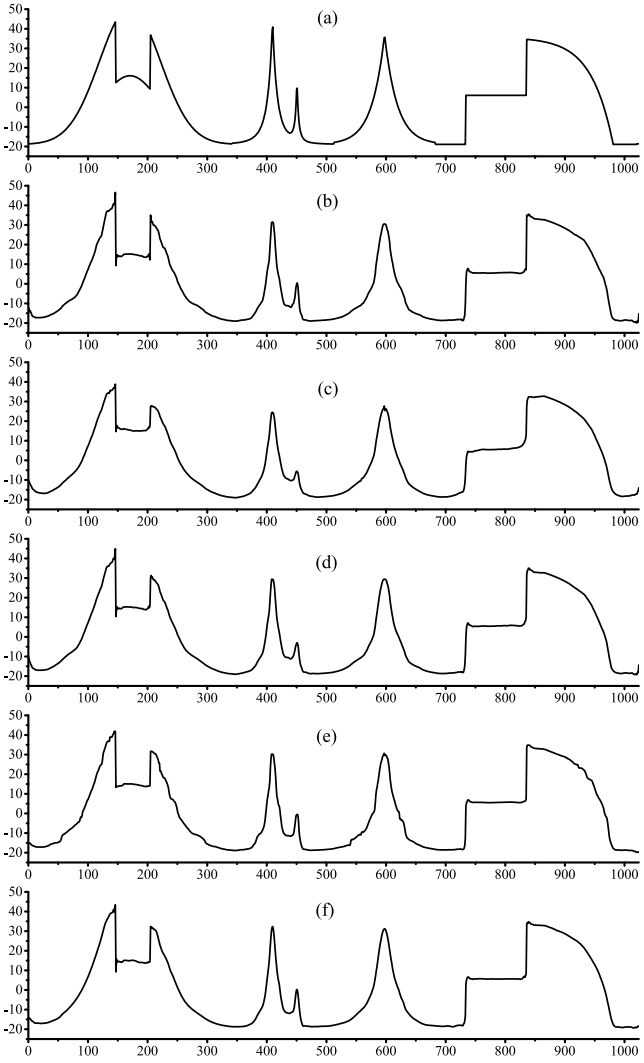


FIGURE 7. Denoised signal by different methods under *haar* wavelet domain. (a) Sy. (b) Hard (RMSE = 1.38). (c) Soft (RMSE = 2.08). (d) Atan (RMSE = 1.56). (e) WATV (RMSE = 1.40). (f) DWAD (RMSE = 1.22).

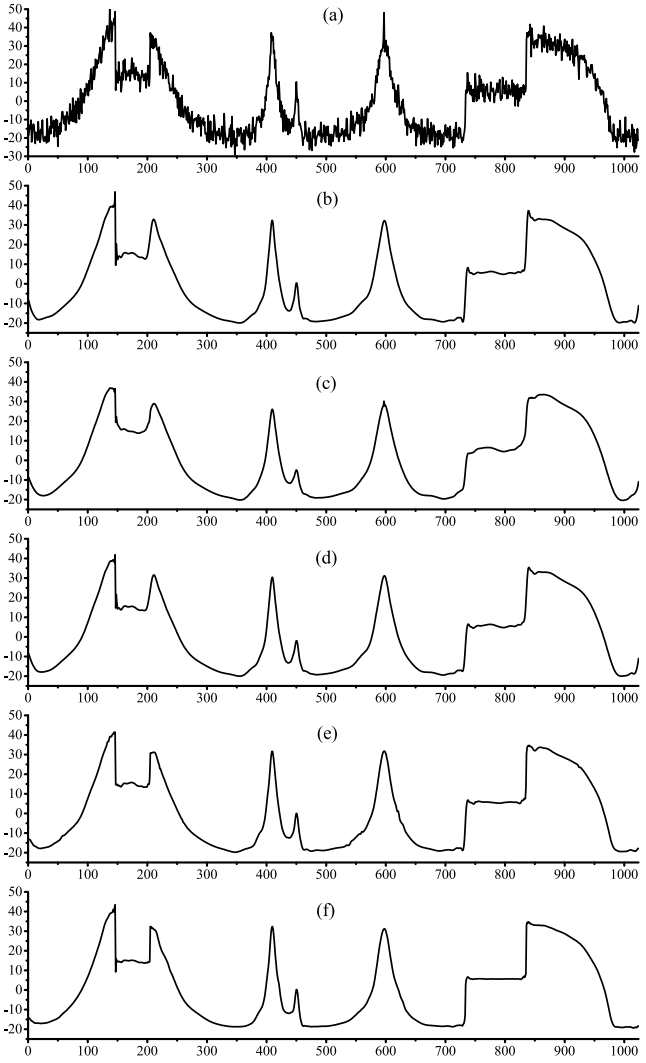


FIGURE 8. Denoised signal by different methods under *sym4* wavelet domain. (a) Sy (σ = 4). (b) Hard (RMSE = 1.70). (c) Soft (RMSE = 2.34). (d) Atan (RMSE = 1.89). (e) WATV (RMSE = 1.40). (f) DWAD (RMSE = 1.22).

Besides, the impact of the threshold values for each method is illustrated Fig. 9 and Fig. 10 for two wavelet domains, respectively. These results indicate that the proposed DWAD method has good performance of preserving details and achieves lower RMSE than other methods. As the results show in Fig. 7, Fig. 8, Fig. 9 and Fig. 10, the proposed method has good performance in noise filtering, smoothness preservation and artifact suppression.

Moreover, the corrupted signals by AWGN with different standard deviation are employed to test the performance of different methods with different noise levels. Table 2 shows the average signal-to-noise ratio (SNR) of denoised signal for 20 trials. The results indicate that the proposed DWAD

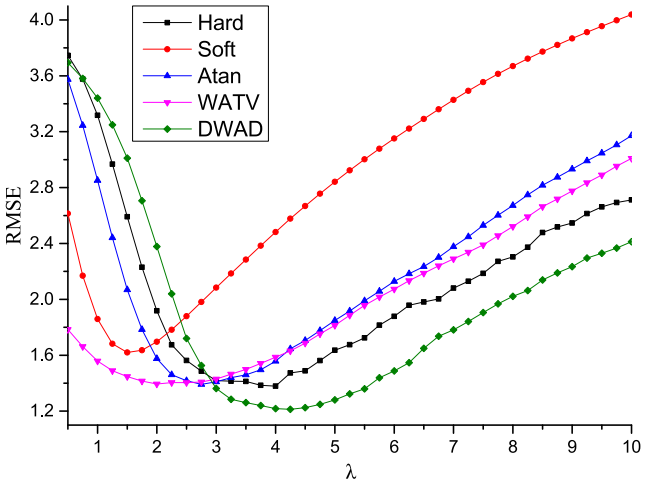


FIGURE 9. The RMSE with different thresholds under *haar* wavelet domain.

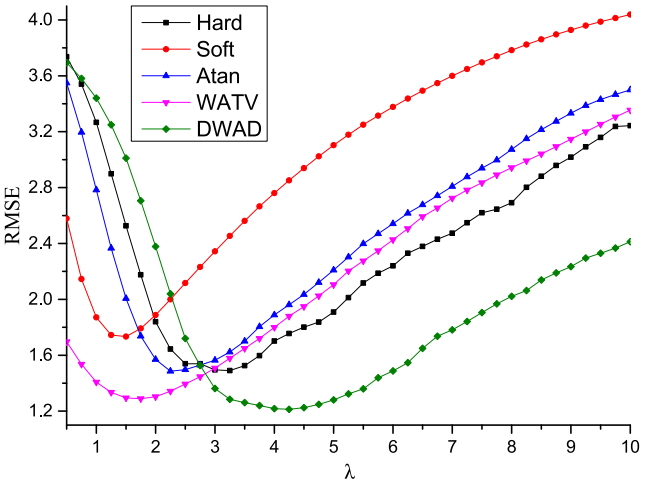


FIGURE 10. The RMSE with different thresholds under *sym*4 wavelet domain.

has better performance for the signal with lower noise level. Compared with classical soft threshold method, the average SNR of different noise levels is improved by at least 4.2dB by using the proposed DWAD.

C. APPLICATION IN IMAGE DENOISING

In order to test the performance of proposed DWAD for the process of 2 dimensional signals, we consider the problem of image denoising. The classical image of Lena with additive Gaussian noise is employed to test. The original image and noisy image with noise level $\sigma = 25$ are showed in Fig. 11 (a) and (b), respectively. We compared the proposed DWAD method with the penalty function methods under two wavelet domains, the hard threshold method based on the discrete Curvelet transformation [13], the non-local means (NLM) method proposed in [48], and the block-matching 3D filtering (BM3D) proposed as a state-of-art filtering algorithm in [49].

For the penalty function methods with the Atan function, the threshold for each wavelet scale is set according to 29, and the parameter λ is set to be 4.25. The denoised images under the *sym*4 and *db*2 wavelet domains are showed in Fig. 11 (c) and (d), respectively. As for the discrete Curvelet transformation, NLM and BM3D methods, we use the parameters recommended in the corresponding literature. The results of noise removal by these methods are illustrated in Fig. 11 (e), (f) and (g), respectively.

For the DWAD method, the original and noisy images have smaller L1 norm under *haar*, *sym*2 and *db*2 wavelet domains than the *db*4 and *sym*4 wavelet domains. However, the *haar*, *sym*2 and *db*2 wavelet basis functions hold large correlation coefficients with each other. Therefore, we consider using *sym*2 and *db*4 wavelet basis functions to generate two decimated wavelet transform matrices for image denoising. The threshold and parameter for penalty function are $\lambda = 3.25$ and $a = 1/\lambda$, respectively.

In order to analyze the convergence and complexity of the DAWD, we set the number of iterations *Iter* being 50, and consider denoising the noisy image by the DWAD with different parameter μ. The residual ratio p between reconstructed images in two wavelet domains is employed to estimate the convergence of DAWD, and it could be defined as:

$$p = \frac{2\|(\mathbf{W}_1\mathbf{C}_1 - \mathbf{W}_2\mathbf{C}_2)\|_F^2}{\|\mathbf{W}_1\mathbf{C}_1\|_F^2 + \|\mathbf{W}_2\mathbf{C}_2\|_F^2} \times 100\% \tag{30}$$

For different parameter μ, Fig. 12 illustrates the variation of residual ratio p with the number of iterations. It could be found that residual ratio p converges rapidly with the increase of iteration times. When the μ is set 1.0, p converges to about 0.8%. In the case of μ is set 1.5, 2.0, 5.0 and 10.0, residual ratio converges to about 0. Besides, larger parameter μ induces faster the convergence rate at the first few iterations. However, the DWAD trends to converge at the same rate when the number of iterations is greater than 15. If the value of parameter μ is too large, such as 5 or 10, it would be automatically adjusted according to 28 during the iteration process. Therefore, when μ is greater than or equal to 1.5, it could guide the fast convergence of DWAD algorithm, and has little influence on the results. In view of this, we consider setting $\mu = 1.5$ and $Iter = 25$ in image denoising. The corresponding result of denoising are shown in the Fig. 11 (h), and the residual ratio p converges to about 0.1%. As the results in Fig. 11 shows that the DWAD method could obtain better performance of noise removal and details preservation than the penalty function and Curvelet transform methods. Compared with the result of NLM method, more details of original image would be preserved in the result of the DWAD method.

Further, noisy images with different noise levels are employed to test the performance of several methods. The PSNR and the structural similarity (SSIM), which are defined in [50], are employed to evaluate the performance of noise reduction and preserving details and features, respectively.

FIGURE 11. Denoising an experimental image with noise level 25. (a) Original image. (b) Noisy image. (c) Result of *sym*4. (d) Result of *db*2. (e) Result of Curvelet. (f) Result of NLM. (g) Result of BM3D. (h) Result of DWAD.

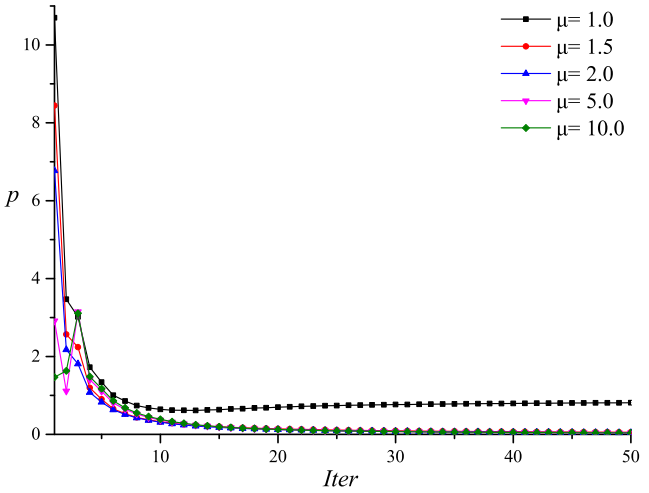


FIGURE 12. The variation of residual ratio *p* with the number of iterations.

TABLE 3. Performance of image denoising by several algorithms for different noise levels (PSNR, SSIM, time(s)).

σ	*sym*4	*db*2	*Curvelet*	NLM	BM3D	DWAD
5	36.12	35.77	35.75	36.95	38.71	37.64
	0.9741	0.9727	0.9758	0.9776	0.9840	0.9812
10	33.06	32.59	33.33	34.75	35.91	34.63
	0.9430	0.9402	0.9540	0.9582	0.9691	0.9631
15	31.22	30.79	31.96	33.02	34.26	32.91
	0.9085	0.9052	0.9336	0.9385	0.9547	0.9465
20	29.87	29.48	30.92	31.66	33.04	31.70
	0.8718	0.8685	0.9143	0.9186	0.9406	0.9307
25	28.77	28.41	30.05	30.57	32.08	30.75
	0.8346	0.8305	0.8956	0.8987	0.9265	0.9153
30	27.83	27.49	29.32	29.65	31.28	29.96
	0.7981	0.7931	0.8777	0.8789	0.9128	0.9002
35	27.01	26.68	28.69	28.88	30.58	29.30
	0.7621	0.7571	0.8604	0.8590	0.8999	0.8857
40	26.27	25.98	28.14	28.20	29.88	28.73
	0.7275	0.7230	0.8447	0.8391	0.8862	0.8716
45	25.58	25.34	27.66	27.59	29.33	28.24
	0.6942	0.6901	0.8290	0.8193	0.8779	0.8581
50	24.96	24.73	27.21	27.04	28.85	27.82
	0.6629	0.6587	0.8139	0.7996	0.8664	0.8456
Avg.	29.07	28.73	30.30	30.83	32.39	31.17
	0.8177	0.8139	0.8899	0.8888	0.9218	0.9098
time	0.05	0.05	0.69	1.60	1.75	4.15

Table 3 shows the corresponding PSNR and SSIM of the denoised images by different methods. Besides, in order to test the computational complexity, we consider run these algorithms on 64-bit computer (with 3.20 GHz CPU). The average operating time is shown in the bottom line of Table 3.

According to the results in Table 3, the BM3D method always achieves the best performance. The proposed DWAD achieves great improvement in PSNR and SSIM by comparing with the methods based on single representation domain, though it does not obtain the best performance and performs poorest in computational complexity. Note that, the NLM method obtains higher PSNR than the DWAD at the noise level 10 and 15, but it holds smaller SSIM. According to the average of PSNR and SSIM, the proposed DWAD holds better performance of noise removal and details preservation than the Curvelet transform and NLM methods. Compared with classical soft threshold method, the average PSNR of different noise levels is improved by at least 2.1dB by using the proposed DWAD.

IV. CONCLUSION

This paper proposes a novel denoising method based on double wavelet basis functions. In view of the wavelet

coefficients of original signal hold different distributions under different wavelet domains, some details of original signal which could be filtered in specific wavelet domain would be preserved when using other wavelet domains. We consider utilizing difference of the representation coefficients of signal under two different wavelet domains to weaken the influence of threshold function for those coefficients which are less than the threshold. Meanwhile, the non-convex penalty function is employed to induce strong sparsity for those wavelet coefficients which are greater than the threshold.

The results of proposed method based on the decimated wavelet transforms show that the wavelet coefficients of original signal could be preserved better during filtering the noise, especially for the coefficients which are smaller than the threshold. Besides, the experiment results for show that the average SNR of different noise levels is improved by at least 4.2dB and 2.1dB compared with classical soft threshold method for the one dimensional and image signals, respectively. In addition, the RMSE and SSIM of denoised signals indicate that the proposed method tends to obtain better performance on the details of original signals.

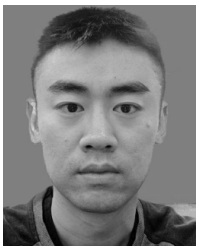
REFERENCES

[1] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Trans. Image Process.*, vol. 9, no. 9, pp. 1532–1546, Sep. 2000.
[2] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
[3] M. Srivastava, C. L. Anderson, and J. H. Freed, "A new wavelet denoising method for selecting decomposition levels and noise thresholds," *IEEE Access*, vol. 4, pp. 3862–3877, 2016.
[4] L. Zhang, P. Bao, and Q. Pan, "Threshold analysis in wavelet-based denoising," *Electron. Lett.*, vol. 37, no. 24, pp. 1485–1486, Nov. 2001.
[5] A. E. Cetin and M. Tofighi, "Projection-based wavelet denoising [lecture notes]," *IEEE Signal Process. Mag.*, vol. 32, no. 5, pp. 120–124, Sep. 2015.
[6] R. Zhao, X. Liu, C.-C. Li, M. Sun, and R. J. Sclabassi, "A new denoising method based on wavelet transform and sparse representation," in *Proc. Int. Conf. Signal Process.*, 2008, pp. 171–174.
[7] L. Jing-Yi, L. Hong, Y. Dong, and Z. Yan-Sheng, "A new wavelet threshold function and denoising application," *Math. Problems Eng.*, vol. 2016, no. 3, 2016, Art. no. 3195492.
[8] B. N. Singh and A. K. Tiwari, "Optimal selection of wavelet basis function applied to ECG signal denoising," *Digit. Signal Process.*, vol. 16, no. 3, pp. 275–287, 2006.
[9] V. Delouille, M. Jansen, and R. von Sachs, "Second-generation wavelet denoising methods for irregularly spaced data in two dimensions," *Signal Process.*, vol. 86, no. 7, pp. 1435–1450, 2006.
[10] A. L. da Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, Oct. 2006.
[11] R. Eslami and H. Radha, "Translation-invariant contourlet transform and its application to image denoising," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3362–3374, Nov. 2006.
[12] H. Sadreazami, M. O. Ahmad, and M. N. S. Swamy, "A study on image denoising in contourlet domain using the alpha-stable family of distributions," *Signal Process.*, vol. 128, pp. 459–473, Nov. 2016.
[13] J.-L. Starck, E. J. Candes, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Trans. Image Process.*, vol. 11, no. 6, pp. 670–684, Jun. 2002.
[14] V. M. Kamble, P. Parlewar, A. G. Keskar, and K. M. Bhurchandi, "Performance evaluation of wavelet, ridgelet, curvelet and contourlet transforms based techniques for digital image denoising," *Artif. Intell. Rev.*, vol. 45, no. 4, pp. 509–533, 2016.
[15] T. Qiao *et al.*, "Effective denoising and classification of hyperspectral images using curvelet transform and singular spectrum analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 1, pp. 119–133, Jan. 2017.
[16] J.-L. Starck, J. Fadili, and F. Murtagh, "The undecimated wavelet decomposition and its reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 297–309, Feb. 2007.
[17] X. Wang, X. Liu, A. Zhang, and F. U. Bo, "Undecimated wavelet Bayesian image denoising method with its threshold determined by curve fitting," *Pattern Recognit. Artif. Intell.*, vol. 29, no. 4, pp. 322–331, Apr. 2016.
[18] H. Li and F. Liu, "Image denoising via sparse and redundant representations over learned dictionaries in wavelet domain," in *Proc. Int. Conf. Image Graph.*, 2009, pp. 754–758.
[19] J. Chen, Y. Zi, Z. He, and X. Wang, "Adaptive redundant multiwavelet denoising with improved neighboring coefficients for gearbox fault detection," *Mech. Syst. Signal Process.*, vol. 38, no. 2, pp. 549–568, 2013.
[20] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," *Proc. Nat. Acad. Sci. USA*, vol. 100, no. 5, pp. 2197–2202, 2003.
[21] I. W. Selesnick and M. A. T. Figueiredo, "Signal restoration with overcomplete wavelet transforms: Comparison of analysis and synthesis priors," *Proc. SPIE*, vol. 7446, Sep. 2009, Art. no. 74460D.
[22] M. Elad, M. A. T. Figueiredo, and Y. Ma, "On the role of sparse and redundant representations in image processing," *Proc. IEEE*, vol. 98, no. 6, pp. 972–982, Jun. 2010.
[23] D. L. Donoho and J. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
[24] H.-Y. Gao and A. G. Bruce, "Waveshrink with firm shrinkage," *Statist. Sinica*, vol. 7, no. 4, pp. 855–874, 1997.
[25] H.-Y. Gao, "Wavelet shrinkage denoising using the non-negative garrote," *J. Comput. Graph. Statist.*, vol. 7, no. 4, pp. 469–488, 1998.
[26] L. Sendur and I. W. Selesnick, "Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2744–2756, Nov. 2002.
[27] M. Nasri and H. Nezamabadi-Pour, "Image denoising in the wavelet domain using a new adaptive thresholding function," *Neurocomputing*, vol. 72, no. 4, pp. 1012–1025, 2009.
[28] A. Fathi and A. R. Naghsh-Nilchi, "Efficient image denoising method based on a new adaptive wavelet packet thresholding function," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3981–3990, Sep. 2012.
[29] J. Saeedi, M. H. Moradi, and K. Faez, "A new wavelet-based fuzzy single and multi-channel image denoising," *Image Vis. Comput.*, vol. 28, no. 12, pp. 1611–1623, 2010.
[30] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, Nov. 2009.
[31] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001.
[32] R. Charnigo, J. Sun, and R. Muzic, "A semi-local paradigm for wavelet denoising," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 666–677, Mar. 2006.
[33] S. Rajan, S. Wang, R. Inkol, and A. Joyal, "Efficient approximations for the arctangent function," *IEEE Signal Process. Mag.*, vol. 23, no. 3, pp. 108–111, May 2006.
[34] İ. Bayram, "On the convergence of the iterative shrinkage/thresholding algorithm with a weakly convex penalty," *IEEE Trans. Signal Process.*, vol. 64, no. 6, pp. 1597–1608, Mar. 2016.
[35] C.-H. Zhang, "Nearly unbiased variable selection under minimax concave penalty," *Ann. Statist.*, vol. 38, no. 2, pp. 894–942, 2010.
[36] I. Selesnick and M. Farshchian, "Sparse signal approximation via non-separable regularization," *IEEE Trans. Signal Process.*, vol. 65, no. 10, pp. 2561–2575, May 2017.
[37] A. Lanza, S. Morigi, and F. Sgallari, "Convex image denoising via non-convex regularization with parameter selection," *J. Math. Imag. Vis.*, vol. 56, no. 2, pp. 195–220, 2015.
[38] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed ℓ^0 norm," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 289–301, Jan. 2009.
[39] I. W. Selesnick and I. Bayram, "Sparse signal estimation by maximally sparse convex optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1078–1092, Mar. 2013.
[40] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, "Noise reduction using an undecimated discrete wavelet transform," *IEEE Signal Process. Lett.*, vol. 3, no. 1, pp. 10–12, Jan. 1996.
[41] Y. Ding and I. W. Selesnick, "Artifact-free wavelet denoising: Non-convex sparse regularization, convex optimization," *IEEE Signal Process. Lett.*, vol. 22, no. 9, pp. 1364–1368, Sep. 2015.

[42] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479–2493, Jul. 2009.
[43] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
[44] M. V. Afonso, J.-M. Bioucas-Dias, and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345–2356, Sep. 2010.
[45] J. Eckstein and D. Bertsekas, "On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Program.*, vol. 55, no. 3, pp. 293–318, Jun. 1992.
[46] K. Lange, D. R. Hunter, and I. Yang, "Optimization transfer using surrogate objective functions," *J. Comput. Graph. Statist.*, vol. 9, no. 1, pp. 1–20, Mar. 2000.
[47] M. A. T. Figueiredo, J. M. Bioucas-Dias, and R. D. Nowak, "Majorization–minimization algorithms for wavelet-based image restoration," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 2980–2991, Dec. 2007.
[48] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2, Jun. 2005, pp. 60–65.
[49] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
[50] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

YONGJUN WU received the B.S. and M.S. degrees from Central South University, China, in 2013 and 2016, respectively, where he is currently pursuing the Ph.D. degree. His research interests include intelligent traffic monitoring, signal processing, image processing, and machine learning.

GUANGJUN GAO received the M.S. and Ph.D. degrees from Central South University, China, in 2001 and 2008, respectively, where he is currently a Professor with the School of Traffic and Transportation Engineering. His research interests include impact safety protection technology, aerodynamics, and traffic safety of rail transit.



CAN CUI received the B.S. and M.S. degrees from Central South University, China, in 2013 and 2016, respectively, where he is currently pursuing the Ph.D. degree. His research interests include intelligent traffic monitoring and image processing.