Translation-invariant multiwavelet denoising using improved neighbouring coefficients and its application on rolling bearing fault diagnosis

SUN Hailiang¹, ZI Yanyang¹, HE Zhengjia, WANG Xiaodong and YUAN Jing
State Key Laboratory for Manufacturing and System Engineering,
Xi’an Jiaotong University, Xi’an 710049, China

Email: hailiang41@live.cn, ziyy@mail.xjtu.edu.cn.

Abstract: The deficiencies of conventional neighbouring coefficients denoising are the invariant neighbouring window size and the global threshold; therefore, it cannot accurately represent local concentrated energy of the collected signals in engineering application. The improved neighbouring coefficients named Neighbouring Coefficients Dependent on Level (NCDL) is proposed. The size of neighbouring window varies with different decomposition levels and the threshold is chosen according to the neighbourhood. Translation invariant method can effectively weaken some visual artifacts, for example Gibbs phenomena in the neighbourhood of discontinuities. Multiwavelets have two or more scaling and wavelet functions. Compared with scalar wavelet, multiwavelets offer several excellent properties such as symmetric, orthogonal, compactly support and higher order of vanishing moment. A novel denoising method - translation invariant multiwavelet denoising with improved neighbouring coefficients is presented. The simulation signal proves the validity of the presented method. This method is then applied to the fault diagnosis of a locomotive rolling bearing. The results show that the present method can effectively extract the fault characteristic frequency of a slight scrape on the outer race of the rolling bearing.

Vibration signal analysis has long been an estimated method for mechanical fault diagnosis and condition monitoring, as almost all failures can engender variations in the corresponding vibration signals. By means of appropriate signal decomposition and representation, the features embedded in the vibration signals can be extracted. Features of components with severe faults are significantly clear and can be identified easily. While for defects occur at the early stage, the noise is too heavy for feature extraction, so we have to turn to denoising method.

Wavelet is an effective tool for non-stationary signal processing and has been successfully used in many fields. It has been proved that local features in the signal can be enlarged after the scalar wavelet transformation [1]. An appropriate wavelet is crucial for the transformation result. However, scalar wavelet cannot contain orthogonality, symmetry, compact support and higher order of vanishing moment simultaneously. Therefore, Geronimo et al. develop multiwavelets [2, 3], compared with scalar wavelets, multiwavelets offer many excellent properties which are useful for denoising [4].

¹To whom any correspondence should be addressed.
Donoho [5, 6] used scalar wavelet threshold to denoising and proved its validity. The thresholding scheme compares each wavelet coefficients with the threshold and shrinks the points. V. Strela and A.T.Walden [7] utilized multiwavelets and thresholding scheme to image denoising and concluded that multiwavelets outperform scalar wavelets for most of test images. However, these methods only concerned one point and thresholded the signal term-by-term [4]. Therefore, T.T.Cai and Silverman proposed a threshold rule incorporating neighbouring coefficients and Chen introduced this method into multiwavelet denoising, which has achieved wonderful effect in image denoising and mechanical fault diagnosis [8-10]. Conventional incorporating neighbouring coefficients only concerned immediate neighbors of empirical wavelet coefficients, which did not coincide with the coefficients of multi-features. Wang has analyzed the effect of different sizes of neighboring window and presented Neighbouring Coefficients Dependent on Level (NCDL) [11]. It has variant sizes of neighbouring window with different decomposition levels and adaptive threshold selected according to its neighborhood.

For wavelet denoising, the artifacts have to do with behavior near singularities. In the neighborhood of discontinuities, wavelet denoising can exhibit pseudo-Gibbs phenomena, alternating undershoot and overshoot of a specific target level. An important observation about such artifacts: their size is connected intimately with the actual location of the discontinuity [5]. Coifman and Donoho presented translation invariant (TI) denoising, which could effectively weaken Gibbs phenomena. The average of TI had excellent denoising effect and maintained the smoothness of signals, which made TI denoising have outstanding results in medical image processing [12,13]. Multiwavelets have many advantages over scalar wavelet and TI outperforms the previous denoising methods. Regrettably, they almost have not been used to mechanical fault diagnosis.

Therefore, translation invariant is introduced into multiwavelet denoising with improved neighbouring coefficients. Conventional methods including Hilbert envelope spectrum, cepstrum and scalar wavelet thresholding used to analyze the simulation signal are compared comprehensively. The simulation result shows that the presented method performs better in signal denoising. Finally, this method is applied to fault diagnosis of a locomotive rolling bearing. Compared with other methods, the presented method accurately extracts characteristic frequency of a slight scrape on the outer race of a rolling bearing.

1. Summary of multiwavelet transformation theory

For multiwavelet bases, a vector of $r$ scaling functions $\Phi(t) = (\phi(t), \phi_2(t), \ldots, \phi_r(t))^T$ and a vector of $r$ wavelet functions $\Psi(t) = (\psi_1(t), \psi_2(t), \ldots, \psi_r(t))^T$ satisfy the following two-scale matrix refinement equations:

$$\Phi(t) = \sum_k H_k \Phi(2t-k) \quad k \in \mathbb{Z}$$
$$\Psi(t) = \sum_k G_k \Phi(2t-k) \quad k \in \mathbb{Z}$$  \hspace{1cm} (1)

where $\{H_k\}$ and $\{G_k\}$ are respectively $r \times r$ matrix low-pass and high-pass filters.

Here we consider the multiple scaling functions and multiwavelet functions which are developed by Geronimo, Hardin and Massopust (GHM) [2], as shown in figure 1.

GHM multiwavelets are very popular with several advantages over scalar wavelets. They are symmetric or anti-symmetric, orthogonal, supported on the interval [0,1] and [1,2] respectively and with approximation order 2. Similar as the scalar wavelet, the decomposition and reconstruction of multiwavelets are as follows

$$s_{j-1,n} = \sum_k H_{k-2n} s_{j,k} \quad d_{j-1,n} = \sum_k G_{k-2n} s_{j,k}$$
$$s_{j,k} = \sum_n H_{k-2n} s_{j-1,n} + G_{k-2n} d_{j,k}$$  \hspace{1cm} (2)

The two-scale matrix refinement equations have been used to analyze the simulation signal.
where, $s_{j-1,n}$ is the $r$ dimension low frequency components, $d_{j-1,n}$ is the $r$ dimension high frequency components. * is the conjugate transpose operator.

The decomposition of signal features can get more concentrated energy distribution and higher SNR because of the excellent properties of multiwavelets. The multi-filters require $r$ -dimensions vectors input, so the pre-processing of one dimension signal $f[n]$ is needed. Downie and Silverman [14] pointed out that oversampling in the pre-processing has the best effect for signal denoising, while the post-processing was the inverse transform of the pre-processing. For GHM multiwavelets in this paper, the corresponding pre-processing is shown in equation (4), $c = \sqrt{2}$ is to make sure that the vanishing moment order is 1.

$$s_{0,n} = \begin{pmatrix} f[n] \\ c f[n] \end{pmatrix}$$

(4)

where, $s_{0,n}$ is a 2-dimension signal after the pre-processing.

2. Translation invariant multiwavelet denoising using improved neighbouring coefficients

Translation invariant (TI) denoising suppresses noise by averaging over thresholded signals of all circular shifts. TI multiwavelet denoising [13] combines the advantages of multiwavelets and TI, it obtains new signals which have phase difference with the original ones by time-domain-shift and changes the position of singular points of the signals, so this method can weaken or even avoid the Gibbs phenomena caused by singular points. In order to find out the optimal translation, this method averages the signals which are over thresholded with some circular shifts in a certain range.

The conventional neighboring coefficients procedure chose constant size of neighboring window $l = 3$ at every level after wavelet decomposition. However, this method is not accurate enough. Because the size of neighboring window is invariant but the dependence of coefficients is variant at different levels. Wang presented the improved neighboring coefficients denoising after studying the regularity of wavelet coefficients dependency [11], to resolve the problem of constant neighborhood. The formula of incorporating neighboring coefficients is as equation (5)

$$S_{j,k} = \sum_{n=-N}^{N} d_{j,k+n}^2 \quad N = N_0 - j$$

(5)

where, $j$ is the level of wavelet decomposition, the size of neighboring window is $2N+1$, $N_0$ is a constant, it should be selected according to the signal duration of features and the support of wavelet filters. The combination of improved neighboring coefficients and TI is as in figure 2. It is obvious that the presented method incorporates more coefficients at low level while fewer at high level.

The threshold of improved neighboring coefficients is as shown in equation (6)
Where, \( \lambda_j = 2 \log n_j \), \( \alpha \) is an adjusting coefficient of the threshold which is determined by the length of the neighborhood, \( n \) is the length of collected signals. It can be seen that the threshold became term-by-term when \( N = 0 \), the same as soft threshold; while it became conventional neighboring coefficients when \( N = 1 \).

\[
d_{j,k} = \begin{cases} 
d_{j,k} \left( 1 - \frac{\alpha \lambda^2_{j,k}}{S_{j,k}^2} \right), & S_{j,k}^2 \geq \alpha \lambda^2 \\
0, & S_{j,k}^2 < \alpha \lambda^2 \end{cases}
\]  

\( (6) \)

Figure 2. The algorithm of TI and improved neighbouring coefficients.

The block diagram of translation invariant multiwavelet denoising using improved neighboring coefficients is shown in figure 3. The steps of the presented algorithm are as follows:
1) Prefilter the original noisy signal into multiple streams by a specific prefilter.
2) Shift the data with circular-shift-operator and apply multiwavelet transform to obtain scale coefficients and wavelet coefficients.
3) Apply Neighbouring Coefficients Dependent on Levels (NCDL) to shrink the empirical wavelet coefficients.
4) Calculate the denoised multiple streams from TI by reversing the processes in step 2.
5) Postfilter the denoised multiple streams to get the denoised signal.

Figure 3. The block diagram of the presented procedure.

3. The simulation
In order to verify the validity of the proposed method, a simulation experiment is designed as follows: Gaussian white noise of large magnitude is added to the periodic impulse signal that is shown in figure 4(c), the sample frequency is 1000 Hz and the impulse signal is

\[ h(t) = \text{sgn}(t)e^{-100|t|}\sin(2\pi \cdot 50 \cdot t) \]

\[ f(t) = \sum_{k} A_k h(t - kT - \tau) + n(t) \]  \hspace{1cm} (7)

where, \( T \) is the period of shock impulse signals, \( \tau \) is the phase of the impulse and \( A_k \) is the amplitude of the impulse. \( \tau \) and \( A_k \) are used to simulate the randomness in the bearing signal. The impulse signal is displayed in figure 4(a). In this case \( T = 0.2s \), the characteristic frequency is 5Hz. The periodic impulse signal is considered as characteristics of fault. \( n(t) \) represents the signal caused by other reasons. It is considered to be noise here. The variance of added Gaussian white noise is 0.3 and signal-noise-ratio (SNR) is 0.5.

Frequency spectrum, envelope spectrum and cepstrum are most common methods in the frequency domain analysis. Figure 4(b) and 4(d) show frequency spectrum of signals in figure 4(a) and 4(c) respectively. We can see 50Hz in both figure 4(b) and figure 4(d), which represents the characteristic frequency of these oscillation attenuations. In figure 4(b), 5Hz is a little weak compared with its harmonics, which represents the characteristic frequency of shock impulses. In figure 4(d), there are no clear symbols of 5Hz and its harmonics because of the interference of noise. In figure 5(a), 5Hz and its harmonics are clearer than that of figure 4(d), and we can see 0.22s in figure 5(b), which approximately equals to the period of shock pulses in time domain.

**Figure 4.** (a) The periodic shock impulses (b) Spectrum of the periodic shock impulses (c) The simulation noisy signal (d) Spectrum of the simulation noisy signal.

Several methods such as translation invariant Daubechies8 (Db8) scalar wavelet denoising with soft threshold, translation invariant multiwavelet denoising with soft threshold, multiwavelet denoising with improved neighboring coefficients (NCDL) and the presented procedure are utilized to analyze the simulation signal. The signal is decomposed into four levels, the pre-processing and post-processing adopt over-sampling mode. In order to find out the optimal circular-shift of TI, the article compares mean square errors (MSE) of denoising results with different circular-shifts. MSE of them can be seen in table 1. We can see that when the circular-shift is 32, it has the minimum MSE. The decomposition level is set to 4.

| circular-shift | 4  | 8  | 16 | 32 | 64 |
|----------------|----|----|----|----|----|
| MSE            | 0.0214 | 0.0196 | 0.0191 | 0.0173 | 0.0181 |
As shown in figure 6, these methods all can suppress much noise. In figure 6(a), translation invariant Db8 scalar wavelet denoising with soft threshold shows its limitation in extracting periodic impulses, because this method ignores most impulses of the simulation signal. In figure 6(b), translation invariant multiwavelet denoising with soft threshold performs a little better than that of figure (a), it could extract almost all periodic pulses, while at 0.7~1.5s, the impulses are influenced by noise. In figure 6(c), multiwavelet denoising with improved neighboring coefficients can extract all periodic impulses, shown as these triangles; while at some specific intervals, shown in the rectangles, several singular points appear. The result of present procedure is shown in figure 6(d), equidistant pulse shocks appear at every 0.2 seconds, even at 0.7~1.5s where other methods lose their capabilities, the characteristic frequency is 5Hz. It agrees with the simulation signal and has a highest SNR. So the present procedure outperforms other methods in signal denoising.

In order to contrast the denoising capability of different methods quantitatively, MSE of denoising results with these methods are calculated. The results of MSE are in table 2, it can be seen that the proposed method has the minimum MSE, which means this method has the best denoising effect.

| Wavelets | Db8+TI | TI+GHM | NCDL+GHM | NCDL+TI+GHM |
|----------|--------|--------|----------|-------------|
| MSE      | 0.0219 | 0.0205 | 0.0192   | 0.0173      |

Figure 5. (a) Hilbert envelope spectrum of the simulation signal (b) cepstrum of the simulation signal.

Figure 6. Denoising results for the simulation signal (a) TI+Db8 (b) TI+GHM (c) NCDL+GHM (d) the presented method.
4. The Engineering application

When a defect occurred on the outer race or interior race of a rolling bearing, impulses are generated when the rolling elements pass over the defect, and the period is the interval of two balls passing over the defect respectively. Due to the damping of the bearing, the impulse has signal duration, in other words, the energy of these impulses is distributed on several points. Therefore, the characteristic signals of a bearing with defect can represent as the local concentrated energy.

The proposed method is applied to fault diagnosis of a locomotive rolling bearing with defect. As shown in figure 7, there is a slight scrape on its outer race. The specification of the rolling bearing is shown in table 3. The sampling frequency is 12.8 KHz and the rotational speed is 650 rpm. Then the characteristic frequency of the defect on the outer race can be calculated by equation (8). In this equation, \( f \) is the rotational frequency, \( d \) is the roller diameter, \( D \) is the pitch diameter, \( \alpha \) is the contact angle and \( z \) is the number of rollers. From equation (8), the characteristic frequency is \( f_o = 78.17 \text{ Hz} \), and the period corresponding with the characteristic frequency is 12.8ms.

\[
f_o = \frac{f}{2} \left(1 - \frac{d}{E} \cos \alpha \right) z
\]

Table 3. The specification of the locomotive rolling bearing.

| model     | interior diameter | outer diameter | roller diameter | roller number | pressure angle (\( \alpha \)) |
|-----------|-------------------|----------------|-----------------|---------------|-------------------------------|
| 552732QT  | 160(mm)           | 290(mm)        | 34(mm)          | 17            | 0\(^\circ\)                   |

Figure 7. The locomotive rolling bearing with a slight scrape on the outer race.

Figure 8 depicts the time-domain waveform and its spectrum of the vibration signal collected from the bearing with a slight scrape on its outer race, the noise is so heavy that it is difficult to find out periodic impulse in both time-domain waveform and its spectrum. In figure 9(a), the characteristic frequency of the scrape defect cannot be seen, while its double frequency 158Hz appears. In figure 9(b), 12.5ms can be seen, which is very close to the period of the slight scrape defect, however, it is interfered by strong background noise.

For the optimal choice of circular-shift, Coifman suggested the circular-shifts should be accomplished in order \( n \log_2(n) \). The article selects several numbers for this application, while 32 is the optimal one which can obtain a better denoising effect and has high computation efficiency. The decomposition level is set to 4. Translation invariant Db8 scalar wavelet denoising with soft threshold is applied and the result is as shown in figure 10(a). Unfortunately, this method can hardly extract the periodic pulses. In figure10(b), TI multiwavelet denoising with soft threshold is then applied , this method performs a little better than that of figure (a) and extracts a few impulses, however, it shows its limitation at some specific intervals. In figure 10(c), multiwavelet denoising with improved neighboring coefficients is applied and it can extract all periodic impulses, shown as these triangles; while at some specific intervals, shown in the rectangles, several singular points appear. As shown in figure 10(d), the proposed method gives a significantly distinct result. The period of the impulses is about 12.8ms that verifies the calculation of bearing defect. It is obvious that this method is effective for extracting the features corresponding to the slight scrape on the outer race and the noise is suppressed properly.
5. Conclusion

From the results of the simulation and fault diagnosis of a bearing defect on its outer race, it can be seen that translation invariant multiwavelet denoising with improved neighboring coefficients is able to weaken visual artifacts such Gibbs phenomena and the presented method is superior to previous wavelet denoising methods and classic signal processing methods. Thus the proposed method is a powerful tool for bearing fault diagnosis in the early stage. It is also a promising diagnosis method for other faults in rotating machinery.

Although the present method achieves the obvious improvement compared with conventional neighboring coefficients. There are several other issues, such the optimal choice of sizes of neighboring window, the optimal threshold and the adaptive selection of range of circular-shift still remaining in wavelet denoising. Actually, as some experts said, every method has its shortcomings. For the advanced signal processing techniques presently used in fault diagnosis, wavelet transform, spectral kurtosis, cycle-stationary analysis, EMD, etc., it is difficult to find one technique is superior to
the others in any cases. Generally, aiming at the specific diagnosis problem, the best one may be
selected by comparison.

As mentioned above, the present method has overcome some deficiencies of the conventional
neighboring coefficients denoising and provided a powerful tool for nonlinear and non-stationary
signal analysis. However, some issues still need to be improved urgently. It is an interesting and
significant research topic to develop these based on the presented method. The authors would like to
investigate this topic in future.

Acknowledgments
This research is supported financially by the project of National Natural Science Foundation of China
(No.50975220), National Basic Research Program of China (No.2009CB724405) and Important
National Science and Technology Specific Projects (No. 2009ZX04014-015).

Reference
[1] Lin J. 2001 J. Jour of NDT&E, 34 25-30.
[2] GERONIMO JS, HARDIN DP and MASSOPUST PR. 1994 J. JAT, 78 373-401.
[3] GERONIMO JS, HARDIN DP and MASSOPUST PR. 1996 J. SIAM JMA, 27 1158-92.
[4] Wang Xiandong. 2009 D(in Chinese). Xi’an Jiaotong University.
[5] Donoho DL. 1995 J. IEEE Trans. on Inf Theory, 41 613-27.
[6] Donoho DL and Johnstone IM. 1994 J. Biometrika, 81 425-55.
[7] StrrelaV and Walden AT. 1998 J. Imperial College, statistics section, TR-98-01.
[8] Cai T T and Silverman BW. 2001 J. Ind J Stat B, 63 127-48.
[9] Chen GY and Bui TD. 2003 J. IEEE Signal Processing Letters, 10 211-4.
[10] Yuan J, He ZJ and Wang XD. 2009 J (in Chinese). J Mech Eng, 45 155-60.
[11] Wang XD, Zi YY and He ZJ. 2011 J. MSSP, 25 285-304.
[12] Coifman RR and Donoho DL. 1995 J. Lecture Notes in Stat, 103 125-50.
[13] Bui TD and Chen GY. 1998 J. IEEE Trans. on Signal Processing, 46 414-20.
[14] Downie TR and Silverman BW. 1998 J. IEEE Trans. on Signal Processing, 46 2558-61.