Ground Penetrating Radar Weak Signals Denoising via Semi-soft Threshold Empirical Wavelet Transform

Xu Qiao\(^1\), Feng Yang\(^1\), Jing Zheng\(^2\)

\(^1\) China University of Mining and Technology (Beijing), School of Mechanical Electronic & Information Engineering, Beijing 100083, China
\(^2\) China University of Mining and Technology (Beijing), College of Geophysics and Surveying Engineering, Beijing 100083, China

Corresponding Author Email: iamqiaoxu@163.com

https://doi.org/10.18280/isi.240213

Received: 8 January 2019
Accepted: 28 March 2019

Keywords: road security, ground penetrating radar, empirical wavelet transform, signal denoising, threshold function

ABSTRACT

Ground penetrating radar (GPR) weak signals have the characteristics of low signal-to-noise ratio (SNR) and high frequency, which is a major challenge to noise attenuation. In this paper, we propose a GPR denoising approach based on empirical wavelet transform (EWT) combined with semi-soft thresholding. According to the frequency characteristics of signal, a spectrum segmentation strategy is designed. It can adaptively decompose signal and noise into different modes. The mode which contains more valid signals is processed by hard thresholding to reserve amplitude; the other modes which contain useless signals are processed by soft threshold functions to maintain the continuity of the signal. After weak signal denoising by our proposed method, we compared its performance on synthetic and field data using complete ensemble empirical mode decomposition (CEEMD) and synchrosqueezed wavelet transform (SWT). The proposed method denoising performance is better than other two methods.

1. INTRODUCTION

The underground damages are challenging the road security. These underground damages are usually caused by underground construction, rainfall and so on. Underground damages may cause road slurry, cracking and collapse.

![Figure 1. Road collapse](image)

The underground damages have become a serious problem a few years ago. Several big voids were found under the road near the construction site of high-rise buildings in Beijing. In addition, lots of underground damages have been reported. For example, Figure 1 shows a road collapse, 10 m long, 5 m wide, and 3 m deep, caused by underground damages.

To prevent road collapse, some people decided to find underground damages. Among various non-destructive tests, a ground penetrating radar (GPR) test was selected because GPR is the most convenient and accurate tool [1]. The GPR test can be carried out while driving on the road, so it is suitable for urban area because it does not need traffic control and can quickly survey large area. Ground penetrating radar signals become weak when investigation depth increases. Ground penetrating radar weak signals have the characteristics of low signal-to-noise ratio and high frequency, which needs to be reduced.

Many methods have been developed for random noise reduction and reflected events recovery. Recently, Time-frequency peak filtering method [2], singular value decomposition method [3], local signal-and-noise orthogonalization [4], and dictionary method [5], have been exclusively used for GPR data denoising. Because of the high SNR of GPR signals, it is easier to denoise. But these methods are not ideal for GPR weak signals denoising. Some algorithms are applied to GPR weak signals denoising, such as wavelet transform (WT) and empirical mode decomposition (EMD). In order to increase the amplitude of the GPR weak signals, time-reversal-based processing techniques are designed [6]. However, these methods perform poorly in weak reflection events, particularly when the SNR is below a given threshold. The two improved methods of EMD [7], EEMD [8] and CEEMD [9], are effective signal analysis methods. But they have no adequate theoretical basis. Some algorithms are applied to GPR weak signals denoising, like low-rank filtering [10], synchrosqueezed wavelet transform (SWT) [11] and hidden Markov model [12-13], are not suitable for downhole GPR weak signal denoising. Pattern recognition and incremental nonnegative matrix factorization [14] and neural network [15] are methods for GPR event detection and SNR enhancement. Whereas, the main work of these methods focus on signal detection instead of SNR enhancement. Multi-resolution analysis algorithms [16] has good performance for downhole GPR weak signals denoising, but its time complexity is high.
Recently, empirical wavelet transform (EWT) has been proven to be an useful signal analysis method [17]. It applies Littlewood-Paley and Meyer’s wavelets to extract the amplitude modulated-frequency modulated (AM-FM) components, which are defined as Intrinsic Mode Function (IMF). An equivalent of the Harmonic wavelet, when the basis function is real, is Littlewood–Paley wavelet [18]. Meyer wavelet is a classic wavelet. It has many good properties, for example derivation infinitely, smoothness, attenuates fast [19]. EWT has been applied in many fields. When analyzing the time-frequency characteristics of GPR signals, EWT can provide a high resolution [20]. Meanwhile, it effectively remove the powerline interference and baseline wander in the electrocardiogram signal [21]. In research and application, EWT is a new and effective algorithm.

In this paper, we research on noise attenuation for GPR weak signal via EWT algorithm. EWT effectively decomposes GPR weak signal. Meanwhile, the combination of hard threshold function and soft threshold function, improved threshold function can better preserve signal. As a consequence, underground events become clearer after denoising. This paper is structured as follows. Section II describes the theory of EWT. Section III introduces spectrum segmentation schemes and threshold functions. Section IV presents the results and the comparative analysis. Eventually, Section V concludes this paper.

2. THEORY OF EWT

EWT is a completely adaptive and data driven method with solid mathematical foundations. The steps for this approach are as follows.

1. IMF needs to meet the following rules:

   • To acquire signal Fourier spectrum \( F(\omega) \), we compute the fast Fourier transform (FFT) of signal \( f(t) \).

   • Segment the Fourier spectrum correctly. The set of boundaries is \( \omega_k \) for \( k=0,1, \ldots, K \). Separate the Fourier spectrum \([0,\pi]\) into \( K \)-1 contiguous parts and \( \omega^0 = 0 \), \( \omega^K = \pi \). In Section III, the segmentation method will be introduced in detail.

   • Littlewood-Paley and Meyer wavelets can be used to design low-pass and band-pass filter.

   The expression of Littlewood-Paley wavelet is as follows:

   \[
   \varphi(t) = \frac{\sin(4\pi t) - \sin(2\pi t)}{2\pi}
   \]  

   The expression of Meyer wavelet is as follows:

   \[
   \varphi(\omega) = \begin{cases}
   1, & |\omega| < \frac{2\pi}{3} \\
   \cos\left(\frac{\pi}{2} \beta \left(1 - \frac{|\omega|}{\frac{2\pi}{3}}\right)\right), & \frac{2\pi}{3} \leq |\omega| < \frac{4\pi}{3} \\
   0, & \text{otherwise}
   \end{cases}
   \]  

   \[
   \psi(\omega) = \begin{cases}
   e^{-i\omega \tau} \varphi(\omega), & \frac{2\pi}{3} \leq |\omega| < \frac{4\pi}{3} \\
   e^{i\omega \tau} \varphi(\omega), & \frac{4\pi}{3} \leq |\omega| \leq \frac{8\pi}{3} \\
   0, & \text{otherwise}
   \end{cases}
   \]  

   where, \( \varphi(\omega) \) is scaling function and \( \psi(\omega) \) is Meyer’s wavelet function.

   The empirical scaling function \( \varphi_k(\omega) \) and the empirical wavelets \( \psi_k(\omega) \) are defined by expressions of equation (4) and (5).

\[
\varphi_k(\omega) = \begin{cases}
1, & |\omega| \leq 1 - \gamma \\
\cos\left(\frac{\pi}{2} \beta \left(1 - \frac{|\omega|}{1 - \gamma}\right)\right), & (1 - \gamma) \omega^k \leq |\omega| \leq (1 + \gamma) \omega^k \\
0, & \text{ otherwise}
\end{cases}
\]  

\[
\psi_k(\omega) = \begin{cases}
1, & (1 + \gamma) \omega^k \leq |\omega| \leq 1 + \gamma \\
e^{-i\omega \tau} \varphi(\omega), & (1 - \gamma) \omega^k \leq |\omega| \leq (1 + \gamma) \omega^k \\
e^{i\omega \tau} \varphi(\omega), & (1 - \gamma) \omega^k \leq |\omega| \leq (1 + \gamma) \omega^k \\
0, & \text{ otherwise}
\end{cases}
\]  

where, \( \omega \) is the frequency, \( \omega^k \) is the kth boundary frequency, \( \gamma \) is a parameter that ensures no overlap between the two consecutive transitions, and \( \beta(x) \) is an arbitrary function that satisfies the following conditions such as

\[
\beta(x) = \begin{cases}
0, & x \leq 0 \\
1, & x \geq 1 \\
\beta(x) + \beta(1-x) = 1, & 0 < x < 1
\end{cases}
\]  

Many functions satisfy these conditions, the most commonly used in the literature is

\[
\beta(x) = x^4(35 - 84x + 70x^2 - 20x^3)
\]  

4. Extract different modes according to the scale function and the empirical wavelet function. The scale coefficients are given by the inner products with the scale function:

\[
W(1,t) = \langle f, \varphi \rangle = \int f(\tau) \varphi(\tau - t) d\tau
\]  

The detailed coefficients are obtained by the inner product with empirical wavelets:

\[
W(k,t) = \langle f, \psi \rangle = \int f(\tau) \psi(\tau - t) d\tau
\]  

where, \( \tau \) is the time point, \( \langle \cdot, \cdot \rangle \) is the inner product, over-bar represents the conjugation of a function and \( W(k,t) \) denotes the coefficients for the kth filter bank at the \( t \)th time point. The noise coefficients are less than the effective signal coefficients. Then by the mechanism of the scale difference of empirical wavelet coefficients for signal and noise, the empirical wavelet coefficients of noisy are threshold processed.

3. GPR WEAK SIGNAL DENOISING VIA EWT

3.1 Segmentation of the Fourier spectrum

The segmentation mechanism of Fourier spectrum is a vital
step in separating signals and noise, which provides a theoretical basis for the adaptive analysis of signals. Therefore, even in the case of low SNR, EWT can analyze the signal effectively. There are four segmentation methods in EWT, namely localmax, localmaxmin, localmaxminf, and adaptive, respectively. Because adaptive method shows better frequency division ability, it is used in this paper.

The main idea of adaptive method is that starts from an initial set of boundaries and then adapts them as the smallest window around considered boundary. The minimum are computed from the original spectrum. Since 0 and π are always used in the first boundary and the last boundary, we need to find extra k − 1 boundaries. Assuming that initial boundary is set to a_k , k = 0,1,...,K , a_k ∈ [0,ω_max] and a_0 = 0 , a_K = ω_max (ω_max is the maximum frequency of the signal), the window can be calculated by equation (10).

\[
\begin{align*}
    L_k &= a_k - \text{round}\left(\frac{a_k - a_{k+1}}{2}\right) \\
    U_k &= a_k + \text{round}\left(\frac{a_{k+1} - a_k}{2}\right)
\end{align*}
\]

(10)

where, round(*) means rounded function. The lower and upper are the lower and upper limits of the window, respectively.

The expression of the boundary is as follows:

\[
\omega_k = \text{arg min}_\omega F_1(\omega), \quad \omega \in [L_k,U_k]
\]

(11)

where, \(F_1(\omega)\) is minimum value within a window.

To demonstrate the feasibility of adaptive method, we use the sinusoidal wave with SNR equaling to -50 dB to simulate signal. The expression of sinusoidal wave \(x(t)\) is as follows:

\[
x(t) = a_1 \sin(\omega_1 t) + a_2 \sin(\omega_2 t)
\]

(12)

where, \(a_1,a_2\) are amplitudes \(\omega_1,\omega_2\) are angular frequencies and \(t\) is time. The spectrum of the pure signal, noisy and mixed signal are shown in Figure 2, we can see that the dominant frequency of the mixed signal ranges from 0 to 500 MHz. In order to make full use of the spectral characteristics, we set \(k = 3\). Boundaries and filtering bank of the adaptive are shown in Figure 3.

**Figure 2.** The spectrum of the signal. (a) Pure signal. (b) Noisy. (c) Mixed signal

**Figure 3.** Boundaries and filtering bank of the adaptive. (a) Boundaries. (b) Filtering bank

**Figure 4.** Empirical wavelet coefficients from three subbases. (a) Band1. (b) Band2. (c) Band3

### 3.2 Threshold functions

Hard threshold function has been widely used, since it is good on keeping signal amplitude. But there are some shortcomings. For example, the reconstructed data will be discontinuous at threshold point; the reconstructed signal is prone to Pseudo-Gibbs phenomenon. A new threshold approach overcomes the disadvantage which is not continuous, but advantage of hard threshold is not preserved.

We assume that \(W_{ki}, i = 1,2,...,K, t = 1,2,...,N\) is empirical wavelet coefficient matrix of the \(i\)th mode. \(\hat{W}_{ki}\) is coefficient matrix after threshold. \(\lambda_i\) is threshold which is related to the variance of the noise. \(sgn(*)\) is the sign function. The expressions of hard threshold function, soft threshold function and the improved threshold function are as follows.

Hard threshold function is

\[
\hat{W}_{ij} = \begin{cases} 
W_{ij}, & |W_{ij}| \geq \lambda_i \\
0, & |W_{ij}| < \lambda_i
\end{cases}
\]

(13)

Soft threshold function is

\[
\hat{W}_{ij} = \begin{cases} 
\text{sgn}(W_{ij})(|W_{ij}| - \lambda_i), & |W_{ij}| \geq \lambda_i \\
0, & |W_{ij}| < \lambda_i
\end{cases}
\]

(14)

The semi-soft threshold function is

\[
\hat{W}_{ij} = \begin{cases} 
\mu W_{ij} + (1 - \mu)\text{sgn}(W_{ij})(|W_{ij}| - \lambda_i), & |W_{ij}| \geq \lambda_i \\
0, & |W_{ij}| < \lambda_i
\end{cases}
\]

(15)
where
\[ \mu = 1 - \alpha (|\hat{y} - y|)^2, 0 \leq \alpha \leq 1, 0 \leq \mu \leq 1 \]

(16)

When \( \alpha \to 0 \) is the hard threshold function, \( \alpha \to 1 \) is the soft threshold function. In general, \( \alpha = 0.05 \). Due to subtraction, the coefficients by threshold processed are less than the original coefficients. The amplitude of the denoised signals will be less than the pure signals.

To check the feasibility and validity of our threshold algorithm, we use the four threshold methods to process the synthetic signals. We not only analyze the denoised waveform, but also analyze the denoising results using SNR, peak signal to noise ratio (PSNR), mean square error (MSE) and mean absolute error (MAE). Higher SNR, PSNR and lower MSE, MAE mean better denoising results. The expression are as follows.

\[
\text{SNR} = 10 \log \left( \frac{\sum_{t=1}^{N} F_{t}^2}{\sum_{t=1}^{N} (\hat{W}_{t} - F_{t})^2} \right)
\]

(17)

\[
\text{PSNR} = 10 \log \left( \frac{(2^n - 1)^2}{\frac{1}{N} \sum_{t=1}^{N} (\hat{W}_{t} - F_{t})^2} \right)
\]

(18)

\[
\text{MSE} = \frac{\sum_{t=1}^{N} (\hat{W}_{t} - F_{t})^2}{N}
\]

(19)

\[
\text{MAE} = \frac{\sum_{t=1}^{N} |\hat{W}_{t} - F_{t}|}{N}
\]

(20)

where \( F_{t} \) is the pure signal, \( \hat{W}_{t} \) is the denoised signal, and \( N \) is the length of signal.

After signal denoising via CEEMD, SWT, results are shown Figure 5 and Figure 6.

![Figure 5. IMFs via CEEMD algorithm. (a) IMF5 + IMF4 + IMF3. (b) IMF2. (c) IMF1](image)

We analyze quantitatively the denoising results of the three methods with the help of SNR, PSNR, MSE and MAE.

**Table 1. Denoising results of different methods**

<table>
<thead>
<tr>
<th>Method</th>
<th>SNR(dB)</th>
<th>PSNR(dB)</th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed signal</td>
<td>19.843</td>
<td>33.987</td>
<td>2.276</td>
<td>0.0815</td>
</tr>
<tr>
<td>CEEMD</td>
<td>31.122</td>
<td>39.626</td>
<td>0.621</td>
<td>0.0218</td>
</tr>
<tr>
<td>SWT</td>
<td>31.796</td>
<td>39.963</td>
<td>0.575</td>
<td>0.0206</td>
</tr>
<tr>
<td>EWT</td>
<td>32.427</td>
<td>40.278</td>
<td>0.534</td>
<td>0.0192</td>
</tr>
</tbody>
</table>

The results show that the proposed algorithm has better performance than CEEMD and SWT methods. Under the proposed algorithm, different threshold functions are compared with the help of SNR, PSNR, MSE and MAE. From Table 2, we can find that semi-soft threshold method obtains the largest SNR and the smallest MSE.

**Table 2. Denoising results of different threshold methods**

<table>
<thead>
<tr>
<th>Method</th>
<th>SNR(dB)</th>
<th>PSNR(dB)</th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hard threshold</td>
<td>27.692</td>
<td>34.126</td>
<td>0.913</td>
<td>0.0738</td>
</tr>
<tr>
<td>Soft threshold</td>
<td>30.588</td>
<td>38.921</td>
<td>0.701</td>
<td>0.0379</td>
</tr>
<tr>
<td>Semi-soft Threshold</td>
<td>31.796</td>
<td>39.963</td>
<td>0.575</td>
<td>0.0206</td>
</tr>
</tbody>
</table>

**4. EXPERIMENT RESULTS**

A synthetic GPR data was used to check the feasibility and validity of EWT algorithm. The comparation of CEEMD, SWT and the proposed algorithms will prove the improvement on previous work. The field GPR data was used to check the practicability and reliability of EWT algorithm.

### 4.1 Synthetic example

To check the feasibility and validity of EWT algorithm, we use a synthetic GPR data with the dominant frequency of 200 Hz, and the wave velocity is about 1.11×10^5 km/s. The synthetic GPR data was generated by finite difference-time domain method. The simulative model is as shown in Figure 7a. The sampling frequency is about 10 GHz, and the SNR is -50dB (Figure 7b). The EWT algorithm is compared with CEEMD and SWT algorithm.

![Figure 6. Mixed signal denoising via SWT. (a) Mixed signal. (b) Denoised signal. (c) Residual signal](image)
For A-Scan GPR data (trace number is 180), the results based on three types of methods are shown in Figure 8.

The EWT algorithm is compared with CEEMD and SWT algorithm. For GPR A-Scan signal, the results based on three types of methods are shown in Figure 9. In the result of CEEMD, there is a lot of noise residue near useful signal. In the enlarged part, CEEMD and SWT are all leak signal. In contrast, the proposed algorithm can preserve valid signals better, including peaks and troughs. Meanwhile, a larger amount of random noise is eliminated by using proposed algorithm in smooth regions.

As far as spectrum is concerned, the spectrum of CEEMD almost concentrates in low frequency constituents, and the spectrum of SWT almost concentrates in middle frequency constituents. The spectrum of EWT is closer to the high frequency and low frequency constituents of pure signal.

<table>
<thead>
<tr>
<th>Method</th>
<th>SNR(dB)</th>
<th>PSNR(dB)</th>
<th>MSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed signal</td>
<td>-50.000</td>
<td>-0.858</td>
<td>3068.4</td>
<td>216.54</td>
</tr>
<tr>
<td>CEEMD</td>
<td>-47.966</td>
<td>3.361</td>
<td>2395.0</td>
<td>169.94</td>
</tr>
<tr>
<td>SWT</td>
<td>-43.315</td>
<td>5.225</td>
<td>1587.7</td>
<td>113.35</td>
</tr>
<tr>
<td>EWT</td>
<td>-42.027</td>
<td>6.018</td>
<td>1366.7</td>
<td>93.43</td>
</tr>
</tbody>
</table>

As far as field results are concerned, the results of CEEMD, SWT and EWT are shown in Figure 11. Although SWT and CEEMD remove a large amount of random noise, the waveform information is not restored well. In comparison, EWT can better retain waveform features and improve waveform clarity.

Table 3. Synthetic GPR data denoising results

Figure 12. SNR results of different denoising method
The SNR, PSNR, MSE and MAE are four objective indexes in GPR signal processing. Table 3 shows the denoising results based on the synthetic GPR data. Comparing with the CEEMD and SWT, the EWT gets a higher SNR and smaller MSE in a variety of situations. We can see that the denoising result weakened when the SNR is below −60dB from Figure 13. Therefore, the SNR of the recorded data is required no less than −60dB for efficient performance of the presented EWT method. When the pure data are unknown for field data denoising tests, the SNR-based evaluation becomes unavailable.

4.2 GPR field data

![Figure 14. 300MHz GPR was used to detect the urban roads](image)

In order to prove the practicability and reliability of proposed method, it is necessary to apply this method to actual data. We use the field signals from a GPR with the dominant frequency of 300 MHz (Figure 14). After being enhanced, the GPR weak signals always include a lot of noise (Figure 15a).

In Figure 15, three algorithms are applied to B-Scan GPR signals. It clearly shows that EWT method (Figure 15d) obtains the most satisfactory output, and the events become clearer. At the reflection part, our algorithm keeps the important details of the input signal. The waveform of SWT results (Figure 15b) is distorted. The noise and signal are all reduced in the results of CEEMD (Figure 15c). There are effective signals in the error images of SWT (Figure 15b) and CEEMD (Figure 15c).

![Figure 15. Field results of different denoising method. (a) GPR data. (b) Denoising results of SWT. (c) Denoising results of CEEMD. (d) Denoising results of EWT](image)

![Figure 16. The error images of different denoising method. (a) Two regions of GPR data. (b) The error image of ECEEMD. (c) The error image of SWT. (d) The error image of EWT](image)

To display merits of EWT algorithm prominently, three methods are compared in regions A and B. Region A contains a large number of valid signals, and region B contains a lot of noise. Figure 15b, Figure 15c, and Figure 15d are the denoising results of SWT, CEEMD, and EWT methods, respectively. Region B shows that CEEMD denoised signal distortion is greater than SWT. Whereas EWT denoised signal distortion is minimal. Region A shows that there is a great deal of noise residual. The noise contained in the SWT denoising results is the most, followed by CEEMD, but the noise in the EWT is the least. From the box of residual graph, we can see that CEEMD and SWT algorithms lose the useful signal parts (Figure 16b, Figure 16c). The signal leakage through EWT is less than other two methods (Figure 16d). In suppression noise, the result obtained by proposed method outperforms the results from the CEEMD and SWT. By contrast, we can find that the EWT algorithm can improve the SNR and the continuity of events is better. Therefore, the EWT denoising performance is better than other two methods.

5. CONCLUSION

This paper presents a new algorithm for GPR weak signal denoising based on EWT. It builds a wavelet filter bank based
on Fourier supports detected from the information contained in the processed signal spectrum. In order to increase denoising ability of EWT, it is combined with improved threshold functions to retain effective signals. The experimental results on synthetic and field data show that the CEEMD and SWT algorithm leads to distortion of the reconstructed signal. However, EWT can retain the effective signal while suppressing the noise. Furthermore, we analyze quantitatively the denoising results of the three methods with the help of SNR, PSNR, MSE and MAE. Experiments show that the proposed algorithm have better performance than CEEMD and SWT methods in keeping detail information and improving SNR of reconstruction signal.

ACKNOWLEDGMENT


REFERENCES


