Discrete Wavelet Transform Decomposition Level Determination Exploiting Sparseness Measurement

Lei Lei,   Chao Wang,   and Xin Liu

Abstract—Discrete wavelet transform (DWT) has been widely adopted in biomedical signal processing for denoising, compression and so on. Choosing a suitable decomposition level (DL) in DWT is of paramount importance to its performance. In this paper, we propose to exploit sparseness of the transformed signals to determine the appropriate DL. Simulation results have shown that the sparseness of transformed signals after DWT increases with the increasing DLs. Additional Monte-Carlo simulation results have verified the effectiveness of sparseness measure in determining the DL.

Keywords—Sparseness, DWT, decomposition level, ECG.

I. INTRODUCTION

Time series analysis has generated significant interest in audio, speech, image as well as biomedical signal processing for the purpose of identifying the various characteristics of the variable studied. Typical approaches include filtering, Fourier transform (FT), wavelet transform (WT) and so on [1]. These transformations aim at revealing hidden information that is not readily available in the raw data. The FT is probably one of the most popular transformations for obtaining the frequency component of the target signal. However, the main drawback of FT is that it can only provides the frequency information, which means that the time-frequency information cannot been seen at the same time. As most of the signals dealt in real applications are not stationary, we need to know at what time the frequency components occur [2]. One possible solution is to adopt short-time Fourier transform (STFT) which analyzes only a small section of the signal at a time. Effectively, the windowing technique is explored by STFT and each windowed signal is assumed to be stationary. However, STFT encounters the problem of dilemma of resolution, which means that a narrow window will produce poor frequency resolution while a wide window will result in poor time resolution.

An alternative approach to STFT is WT which analyzes the signal at different frequencies with different resolutions. The advantage of using WT is that it can provide good time resolution and relatively poor frequency resolution at high frequencies while good frequency resolution and poor time resolution at low frequencies. This important characteristic is useful as most natural signals, such as electrocardiography (ECG) and electroencephalography (EEG) signals, have low
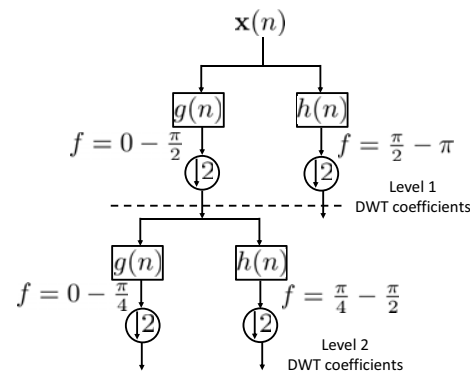
Fig. 1.   Illustration of a two-level DWT process.

frequency content spread over long duration and high frequency content for short durations [2]. For these reasons, discrete WT (DWT) has drawn intensive attention in biomedical signal processing, for example, in ECG [3]. In general, ECG signals have unique P-QRS-T complex waveforms and it is much more significant than other biological signals [4]. It is possible to diagnose many cardiac diseases by visualizing the variations of its morphological characteristics [4]. However, the presence of noise, such as muscle noise, baseline artifacts, the 50 Hz power-line interface, will degrade the accuracy of such diagnosation [5][6].

Not only DWT can be applied to extract the time-frequency information, it also can be exploited for noise suppressing. The main idea of noise reduction via DWT is to compare the DWT coefficients with a pre-determined threshold to determine if it can be seen as a desirable part of the original signal [7]. This approach is well known as *wavelet thresholding*, which can be further categorized into hard and soft thresholding. In hard thresholding, a DWT coefficient is kept unchanged if its absolute value is greater or equal than the threshold; otherwise, it is set to zero [7]. The soft thresholding technique not only zeros the small DWT coefficients, it also shrinks the coefficients with large amplitudes towards zero [8][9].

The performance of wavelet thresholding depends largely on the following four factors [10], the choice of wavelet type, decomposition level (DL), threshold estimation and thresholding rules. It is also important to note that the threshold value is dependent on the DL in DWT. Therefore, it is often regarded as *level-dependent threshold* [7]. As a result, to obtain a suitable DL in DWT becomes a crucial issue. In [10], the authors proposed to measure the entropy of transformed signal and let the decomposing process stop when the resultant entropy

becomes significantly different from that of an artificially generated noise series. The main drawback of this method is that it requires prior information of the distribution model of the noise which is generally unknown in practice.

In this work, we propose a new method to obtain a suitable DL for DWT with application to ECG signal processing. As opposed to the method presented in [10], our method does not require any prior information of the signal distribution model. More specifically, our method utilizes the sparseness measure of the transformed signals at each DL of DWT. Monte-Carlo simulation results have verified the effectiveness of our method.

II. BACKGROUND AND PROBLEM FORMULATION

A. Review of DWT

A wavelet, in the sense of DWT, is an orthogonal function which can be applied to an infinite group of data. A typical DWT decomposition equation can be formulated as [11]

$$\text{DWT}(m,k) = \frac{1}{a} \sum_{n=0}^{N-1} s(n)g\left(\frac{k-b}{a}\right) \qquad (1)$$

where $s(n)$ is the original signal, $a = a_0^m$, $b = na_0^m$, N is the number of samples in the windowed signal, function $g(\cdot)$ is called the mother wavelet, m is the DL index while a and b are called the scaling and translation parameter, respectively. In the case of $a_0 = 2$, DWT can be interpreted as a multi-stage filter banks with high-pass (HP) and low-pass (LP) filters performing a series of dilations. Coefficients obtained after the HP filters are called detail coefficients while those after the LP filter are called the approximate coefficients. Throughout this paper, $a_0 = 2$ is adopted. Figure 1 illustrates the process of a 2-level DWT decomposition. At each level, the approximate/detail coefficients represent a filtered signal spanning only half of the frequency band. This improves the frequency resolution as the frequency uncertainty is reduced by half. Following the Nyquist's theorem, the output of each LP and HP filter can be decimated by a factor of two. This also explains why DWT provides arbitrary good time resolution at high frequencies and arbitrary good frequency resolution at low frequencies.

B. Review of entropy based DL determination in DWT [10]

We first denote the additive noise as $v(n)$. The noisy signal $x(n)$ can be written as $x(n) = s(n) + v(n)$. A frame of the clean signal with length N can be expressed as $\mathbf{s}(n) = [s(n)\ s(n-1)\cdots s(n-N+1)]$ while

$$\mathbf{x}(n) = \mathbf{s}(n) + \mathbf{v}(n) \qquad (2)$$

is a frame of the noisy signal such that $\mathbf{v}(n) = [v(n)\ v(n-1)\cdots v(n-N+1)]$.

The motivation of using entropy to determine DL in DWT is that the energy of the clean signal $\mathbf{s}(n)$ are concentrated on several DLs while the energy of noise scatters in the whole temporal scales which decays rapidly with DL [10]. The authors of [10] proposed to use a certain small DL initially and apply dyadic DWT to denoise $\mathbf{x}(n)$. It has been pointed

out in [10] that as the DL increases, the wavelet entropy energy (WEE) of denoised $\mathbf{x}(n)$ should be significantly different from that of a pure noise frame $\mathbf{v}(n)$ at a certain level m^*. It has also been shown in [10] that once the number of DL exceeds m^*, some clean signal $s(n)$ would be removed in the denoising process and therefore, the suitable DL should be m^* minus 1. The analytic steps proposed in [10] are summarized as follows:

1) Determine the theoretical maximum DL, i.e., $\text{DL}_{\max} = \log_2 N$.
2) Normalize the noisy signal by $x'(n) = \frac{x(n)-\mu}{\sigma}$, where μ and σ are the mean and standard deviation of $\mathbf{x}(n)$, respectively.
3) Apply dyadic DWT to $\mathbf{x}(n)$ from DL 1 to $\log_2 N$ and compute the WEE at each DL. Subsequently, plot the WEE curve of $\mathbf{x}(n)$.
4) According to practical situations and experiences, an appropriate probability distribution model is chosen to generate a "normalized" noise series with the same length of $\mathbf{x}(n)$. The WEE curve of this noise series is determined by Monte-Carlo simulation.
5) Finally, compare the WEE of $\mathbf{x}(n)$ with that of the pure noise series from DL of 1 to $\log_2 N$. Once at a certain level m^*, the WEE of $\mathbf{x}(n)$ is significantly different from that of the pure noise series, the best DL can be chosen as $(m^* - 1)$.

One of the main drawbacks of this approach is the unknown distribution model of the noise. It is important to note that we do not have any prior information of the noise in a real-time application. As a result, the implementation of this method is difficult and it leads to poor performance when an inaccurate distribution model is chosen. Another weakness of this method is the computational load as it requires to regenerate a noisy series and compute its WEE at each DL.

III. PROPOSED SPARSENESS BASED DL DETERMINATION

To develop an algorithm which does not require prior information of the distribution model of the background noise, we propose to exploit the sparsity of the transformed data after DWT. It is well known that the coefficients of DWT are generally sparse [7]. One way to quantify its degree of sparseness is to evaluate the percentage of the number of zero/near-zero coefficients among the entire transformed coefficients [12], i.e.,

$$\text{sp} = \frac{N_o}{N} \qquad (3)$$

where N_o is the number of zero/near-zero coefficients and N is the length of the original signal. Based on (3), we propose to modify this definition slightly as

$$\text{sp}' = \frac{N_o}{N-1}. \qquad (4)$$

This is to constrain the value of sparseness between $[0,\ 1]$, i.e., $\text{sp}' = 1$ in a perfectly sparse case while $\text{sp}' = 0$ if only one coefficient is non-zero.

In order to study the relationship between sparseness and the number of DLs, a clean ECG signal from MIT-BIH Arrhythmia Database [13] is adopted as an illustrative example.
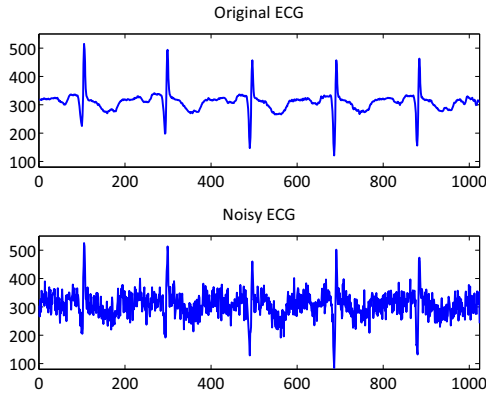
Fig. 2.   (a) A frame of clean ECG signal from [13] with $N = 1024$;
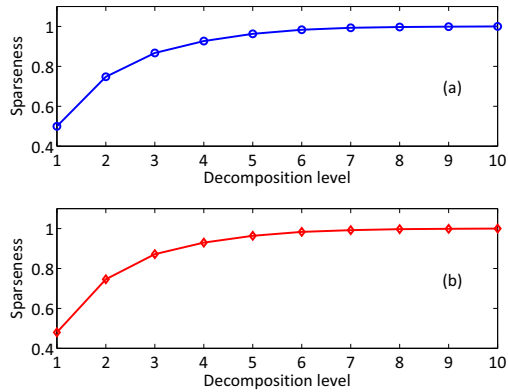(b) Noisy ECG where $\mathrm{SNR} = 20$ dB.

Fig. 3.   Sparseness of the decomposed signal based on (a) the clean
ECG; (b) the noisy ECG with $\mathrm{SNR} = 20$ dB.

Fig. 4.   DWT decomposed signal using Haar wavelet at (a) $m = 1$;
(b) $m = 2$; (c) $m = 8$; (d) $m = 10$.

after the 5th decomposition level. Figure 4 shows the signals
after DWT using Haar wavelet at decomposition level 1, 2, 8
and 10, respectively. It can be seen that the energy are clustered
towards the low frequency as the DL increases and also, the
waveform shapes at level 8 and 10 are quite similar. Therefore,
it is motivated that choosing an optimal DL is correlated to
evaluating the sparseness of the transformed signal. At this
junction, we propose to consider the sparseness defined in
(3) as a criterion to determine the 'best' DL. By choosing a
suitable threshold of sparseness, the 'best' DL can be defined
as at which level the sparseness of the transformed signal
exceeds such threshold.

IV. SIMULATION RESULTS

In order to investigate the effectiveness of our proposed
method, this section presents the results of mass simulations
under different scenarios. In these simulations, 219 sets of
clean ECG signals are adopted from [13], each of which is of
frame length $N = 1024$. The Haar wavelet is used through
out all DWT decompositions. In order to obtain a suitable
threshold value of sparseness, Monte-Carlo simulations are
performed under three different noise levels $\mathrm{SNR} = 15$, 20
and 25 dB. For each SNR scenario, a group of 219 simulations
are conducted.

For clarity, we only plot the results of $\mathrm{SNR} = 15$ dB in
Fig. 5. It can be observed from Fig. 5 (a) that the SP values
of these 219 sets of ECG signals overlap with each other
at each decomposition level ($1 \leq m \leq 10$). This implies
that the sparseness of these ECG signals are approximately
the same. Similar trend can be observed from Fig. 5 (b) for
the noisy signals. Table I summarizes the sparseness measure
of the signals after DWT at each decomposition level under
different noise levels in a statistical manner. The 'mean' is
the averaged sparseness value across the 219 ECG signals

An white Gaussian noise (WGN) is added to the original ECG
signal to achieve a signal-noise-ratio (SNR) of 20 dB such that
$\mathrm{SNR} = 10 \log_{10} \frac{\|\mathbf{s}(n)\|_2^2}{\|\mathbf{v}(n)\|_2^2}$, where $\|\cdot\|_2^2$ is the squared l_2-norm.
Figure 2 (a) shows a clean ECG signal with frame length
$N = 1024$ and Fig. 2 (b) illustrates the noise contaminated
ECG signal with $\mathrm{SNR} = 20$ dB. In each transformation, the
Haar wavelet is adopted. The DWT coefficients whose ampli-
tude are smaller or equal than $1/K$ of the largest coefficient,
i.e., $\mathrm{DWT}(m,k) \leq \max\left(\mathrm{DWT}(m,k)\right)/K$, are regarded
as zero/near-zero coefficients. Empirically, $5 \leq K \leq 10$.
The corresponding sparseness of the transformed signal at
each level, $1 \leq m \leq \log_2 1024$, are plotted in Fig. 3. It
can be observed from Fig. 3 that the value of SP increases
with the increasing DL, which implies that the transformed
signal becomes more and more sparse. It is important to
note from Fig. 3 that the increasing rate of SP diminishes
with the increasing DL and it approaches the theoretical
maximum value 1 asymptotically. It can also be observed from
Fig. 3 that the sparseness of the transformed signal becomes
approximately saturated once $m \geq 5$. This justifies that the
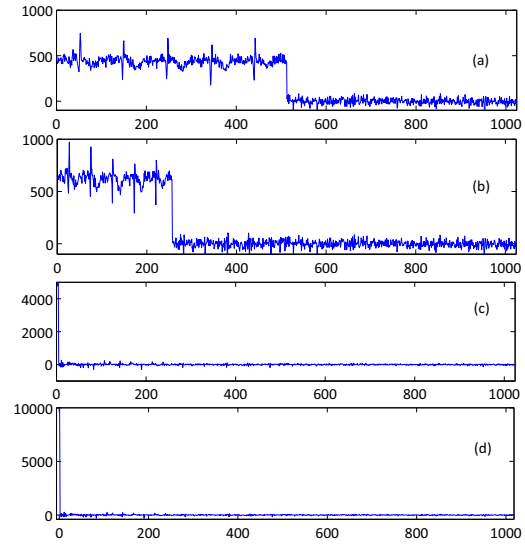statistical property of transformed signal does not vary much

TABLE I
MEAN AND VARIANCE OF THE SPARSENESS MEASURE IN THE MONTE-CARLO SIMULATIONS.

| Decomposition level		1	2	3	4	5	6	7	8	9	10
SNR=25	mean	0.4986	0.7416	0.8613	0.9227	0.9551	0.9768	0.9891	0.9963	0.9990	1.0000
	variance	1.0e-004 *									
		0.0301	0.0535	0.0663	0.0607	0.1063	0.0601	0.0246	0.0084	0.0000	0.0000
SNR=20	mean	0.4857	0.7237	0.8430	0.9048	0.9386	0.9586	0.9746	0.9878	0.9946	0.9991
	variance	1.0e-004 *									
		0.0513	0.1163	0.1771	0.1542	0.1636	0.1481	0.1551	0.0782	0.0264	0.0093
SNR=15	mean	0.4721	0.7083	0.8283	0.8862	0.9150	0.9375	0.9568	0.9720	0.9853	0.9933
	variance	1.0e-004 *									
		0.0822	0.1327	0.1724	0.2499	0.2521	0.2924	0.2207	0.1627	0.1244	0.0430
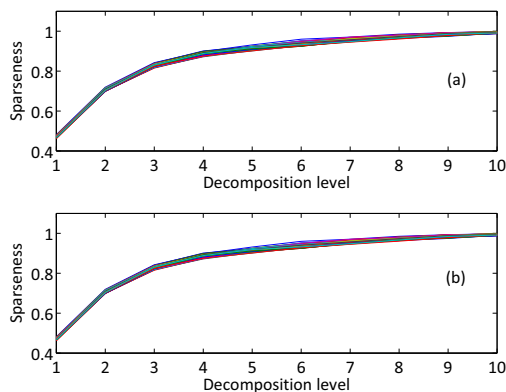
Fig. 5. Sparseness of the decomposed signal based on (a) the clean ECG; (b) the noisy ECG with $\mathrm{SNR} = 15$ dB using Monte-Carlo simulation with 219 iterations.

while the 'variance' is the squared standard deviation of the results obtained from all these 219 signals. It can be observed from Table I that the sparseness value of the transformed signals increases gradually with the DL for each SNR scenario. In general, a higher noise level results in a more dispersive transformed signal which is in line with the dispersive nature of WGN. It is important to note from Table I that among all the simulations, the sparseness is nearly saturated when it reaches 0.9. Therefore, we can set the sparseness threshold as 0.9 empirically in order to determine the suitable DL. Following this, the 'best' DL can be determined as 4 for the cases of $\mathrm{SNR} = 25$ and 20 dB while for $\mathrm{SNR} = 15$ dB, it can be determined as 5.

V. CONCLUSION

In this paper, we proposed a computational efficient method to determine the suitable DL for DWT with application to ECG signal processing. The proposed method exploited the sparseness of the transformed signal at each DL. Monte-Carlo simulation results have demonstrated that sparseness is a significant characteristic of ECG signals after DWT. Specifically, the sparseness of transformed ECG signals increases gradually to the theoretical maximum value. More importantly, the increasing rate of sparseness measure reduces asymptotically when the DL increases. Unlike the method proposed in [10],

our proposed method does not require any prior information of the input which is suitable for real-time processing.

REFERENCES

[1] S. K. Mitra, *Digital Signal Processing: A Computer Based Approach*. McGraw-Hill, 2005.
[2] S. S. Patil and M. K. Pawar, "Quality advancement of EEG by wavelet denoising for biomedical analysis," in *Proc. Int. Conf. on Communication, Information and Computing Technology*, Oct. 2012.
[3] X. Liu, Y. J. Zheng, M. W. Phyu, B. Zhao, and X. J. Yuan, "Power & area efficient wavelet-based on-chip ECG processor for WBAN," in *Proc. IEEE Int. Conf. on Body Sensor Networks*, 2010, pp. 124–130.
[4] P. Karthikeyan, M. Murugappan, and S.Yaacob, "ECG signal denoising using wavelet thresholding techniques in human stress assessment," *International Journal on Electrical Engineering and Informatics*, vol. 4, no. 2, Jul. 2012.
[5] Z. Zhao and P. Min, "ECG denoising by sparse wavelet shrinkage," in *Proc. IEEE Conf. Bioinformatics and Biomedical Engineering*, 2007, pp. 786–789.
[6] X. Liu, Y. J. Zheng, M. W. Phyu, B. Zhao, M.-K. Y. Je, and X. J. Yuan, "Multiple functional ECG signal is processing for wearable applications for long-term cardiac monitoring," *IEEE Trans. Biomedical Engineering*, vol. 58, pp. 380–389, 2011.
[7] R. Cohen, "Signal denoising using wavelets," Technion, Israel Institute of Technology, Tech. Rep., 2011.
[8] D. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Information Theory*, vol. 41, pp. 613–627, 1995.
[9] Y. H. Peng, "De-noising by modified soft-thresholding," in *Proc. IEEE Asia-Pacific Conf. Circuits and Systems*, 2000, pp. 760–762.
[10] D. W. Yan-Fang Sang and J.-C. Wu, "Entropy-based method of choosing the decomposition level in wavelet threshold de-noising," *Journal of Entropy and Information Studies*, Jun. 2010.
[11] J. S. M. D. C. Robertson, O. I. Camps and W. B. Gish, "Wavelets and electromagnetic power system transient," *IEEE Trans. Power Delivery*, vol. 11-2, pp. 1050–1058, Apr. 1996.
[12] A. M. R. Dixon, G. Allstot, D. Gangopadhyay, and D. J. Allstot, "Compressed sensing system considerations for ECG and EMG wireless biosensors," *IEEE Trans. Biomedical Circuits and Systems*, vol. 6, no. 2, pp. 155–166, Apr. 2012.
[13] [Online]. Available: http://gerstner.felk.cvut.cz/biolab/newbiolab/teach/mitdat.htm