Entropy **2015**, *17*, 6093-6109; doi:10.3390/e17096093

OPEN ACCESS

entropy

ISSN 1099-4300

www.mdpi.com/journal/entropy

Article

Optimal Base Wavelet Selection for ECG Noise Reduction Using a Comprehensive Entropy Criterion

Hong He, Yonghong Tan * and Yuexia Wang

Department of Electrical Engineering, College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai 200234, China; E-Mail: heh@shnu.edu.cn

* Author to whom correspondence should be addressed; E-Mail: tany@shnu.edu.cn; Fax: +86-21-3702-2501.

Academic Editor: Raúl Alcaraz Martínez

Received: 29 May 2015 / Accepted: 17 August 2015 / Published: 1 September 2015

Abstract: The selection of an appropriate wavelet is an essential issue that should be addressed in the wavelet-based filtering of electrocardiogram (ECG) signals. Since entropy can measure the features of uncertainty associated with the ECG signal, a novel comprehensive entropy criterion E_{com} based on multiple criteria related to entropy and energy is proposed in this paper to search for an optimal base wavelet for a specific ECG signal. Taking account of the decomposition capability of wavelets and the similarity in information between the decomposed coefficients and the analyzed signal, the proposed E_{com} criterion integrates eight criteria, *i.e.*, energy, entropy, energy-to-entropy ratio, joint entropy, conditional entropy, mutual information, relative entropy, as well as comparison information entropy for optimal wavelet selection. The experimental validation is conducted on the basis of ECG signals of sixteen subjects selected from the MIT-BIH Arrhythmia Database. The E_{com} is compared with each of these eight criteria through four filtering performance indexes, *i.e.*, output signal to noise ratio (SNR_o), root mean square error (RMSE), percent root mean-square difference (PRD) and correlation coefficients. The filtering results of ninety-six ECG signals contaminated by noise have verified that E_{com} has outperformed the other eight criteria in the selection of best base wavelets for ECG signal filtering. The wavelet identified by the E_{com} has achieved the best filtering performance than the other comparative criteria. A hypothesis test also validates that SNR_o, RMSE, PRD and correlation coefficients of E_{com} are significantly different from those of the shape-matched approach ($\alpha = 0.05$, two-sided *t*- test).

Keywords: base wavelet; thresholding filtering; entropy; optimal selection

1. Introduction

Cardiovascular disease is one of the most causes of death in the world. With the aging trend of the population, people are paying more and more attention to research into telemedicine systems for the immediate and accurate detection of cardiac diseases [1]. As a noninvasive test for recording the electric activity of the heart, electrocardiogram (ECG) plays a vital role in cardiac telemedicine systems. The assessment of alterations in the features of ECG signals provides useful information for the detection, diagnosis and treatment of cardiac diseases. However, during the ECG signal acquisition and transmission procedures, the sampled ECG signal is inevitably corrupted by various noises, such as baseline wander, electrode motion, power line interference, motion artifact and so on [2]. Usually, some specific measures such as median filter and band-stop filter can be implemented to suppress the influence of baseline wander and power line interference existing in ECG signals, respectively. However, electromagnetic disturbances such as thermal noise existing in measurement circuits have a significant influence on ECG signals. Thus, the noise reduction of ECG signals is a key requirement prior to pathological feature analysis [3].

Among a variety of filtering techniques, the wavelet transform has been proven as a useful tool for ECG signal denoising due to its powerful analysis ability in both the time and frequency domains. Moreover, the abundance of the base wavelets developed over the past decades is also another prominent advantage for the enhancement of various ECG signals. Nevertheless, since ECG signals have diverse wave shapes from individual to individual, how to choose an optimal base wavelet for analyzing a specific signal has become a crucial problem in the wavelet transform application domain of [4].

In recent years, the topic of base wavelet selection has been addressed by researchers from different points of view. Theoretically, a good base wavelet for noise reduction should satisfy the properties of orthogonality, symmetry, compact support and regularity so as to ensure the smoothness of the denoised signal without distortion. The orthogonal wavelets are not redundant and are suitable for signal or image denoising and compression [5], whilst the regularity is essential for getting nice features, e.g., the smoothness of the reconstructed signal or image, and function estimation in nonlinear regression analysis [6]. In addition, shape matching is also an important issue investigated in the stage of base wavelet selection. Bhatia *et al.* [7] used both orthogonality and the property of complexity to guide the base wavelet choosing procedure for ECG signal analysis. Singh *et al.* [6] introduced a cross correlation coefficient between the ECG signal and a selected wavelet filter for the selection of an appropriate base wavelet. In their method, in terms of root mean square error (RMSE), root means square bias (RMSB), and L_1 norm, the Daubechies mother wavelet of order 8 was validated as the most appropriate wavelet basis function for denoising applications. However, the basis properties of a wavelet only qualitatively determine its suitability for a particular application. As far as shape matching is concerned, it is generally difficult to accurately match the shape of a signal to that of a base wavelet through a visual comparison [4]. As a result, quite a few of ECG denoising schemes implement the choice of the best base wavelet through the noise suppression evaluation criteria, such as percentage root mean square difference (PRD) [8], signal to noise ratio (SNR) [9,10] and mean square error (MSE) [11]. Furthermore, in order to maximize the filtering performance, a genetic algorithm [12] based on SNR and PRD is also applied to the wavelet transform for base selection.

However, these filtering performance measures can only be obtained by comparing the noisy ECG signals and denoised ones. A great deal of filtering tests has to be done before removing the noise from all the signals, which inevitably increases the computational complexity of noise reduction. It is well known that entropy is a measure that can be used to evaluate the uncertainty associated with a random variable, *i.e.*, the expected value of information in a message. Besides, Shannon entropy is one of the most important metrics in information theory. Regarding a sampled ECG signal as a message sequence, the characteristics of the ECG signal can be defined by its entropy. Therefore, in order to automatically find the best base wavelet in a simple and easy way, a comprehensive entropy measure (E_{com}) based on the integration of multiple entropy criteria including the energy to entropy ratio, relative entropy, joint entropy, conditional entropy, mutual information and comparison information entropy is proposed in this paper. In terms of the proposed index E_{com}, the best base wavelet is chosen from 22 widely used wavelets for wavelet thresholding filtering (WTF) of ECG signals. Then the comprehensive analysis of denoising results will also be conducted for the evaluation of the proposed approach through the ECG data derived from the MIT database. Meanwhile, four filtering measures, *i.e.*, signal to noise ratio (SNR_o), Root Mean square Error (RMSE), Percent Root Difference (PRD), as well as Correlation coefficient (r) are utilized for comparison so as to evaluate the performance of the proposed approach.

This paper is organized as follows: the principle of wavelet thresholding filtering is briefly described in Section 2. Then, the proposed comprehensive entropy measure is illustrated in Section 3. Subsequently, the analysis and comparison of ECG signal filtering based on multiple criteria is presented in Section 4. Finally, concluding remarks are provided in Section 5.

2. Discrete Wavelet Thresholding Denoising of ECG Signals

Suppose $\tilde{S}(n)=[\tilde{s}_0,\tilde{s}_1,...,\tilde{s}_N]$ is the observed data vector of a noisy ECG signal, and $\tilde{s}_i$ is given by: $\tilde{s}_i=s_i+n_i$, i = 0, 1, 2, ... , N−1, where s_i is a signal function S to be recovered and n is the Gaussian white noise with independent and identical distribution $N(0,\sigma)$, which mimics the electromagnetic noise in measurement environment $N(0,\sigma)$. The objective of wavelet filtering is to achieve enhanced signal by remaining useful components and removing the noise. Through discrete wavelet transform (DWT), a discrete signal $\tilde{S}(n)$ can be expanded as follows:

$$\tilde{S}(n)=\sum_{k\in Z}a_{2^j}(k)\phi_{2^j}(n-2^jk)+\sum_{k\in Z}d_{2^j}(k)\psi_{2^j}(n-2^jk) \tag{1}$$

where n, k, j represent the sample number, the number of wavelet coefficients and the decomposition scale, respectively. The frequency spectrum of $\tilde{S}(n)$ is divided into a high frequency sub-band and low frequency sub-band as the decomposition scale increases. The approximated component of $\tilde{S}(n)$ is calculated by multiplying the scale function $\phi_{2^j}(n-2^jk)$ by scale factors $a_{2^j}(k)$. Similarly, the product of wavelet function $\psi_{2^j}(n-2^jk)$ and translation factors $d_{2^j}(k)$ represents the detailed component of $\tilde{S}(n)$.

After wavelet decomposition, the WTF is also involved in deleting the wavelet coefficients of the noisy signal whose modulus is below a threshold and reconstructing the denoised signal from the remaining coefficients. To summarize, the conventional wavelet thresholding filtering method includes the following three steps [3,13]:

(1) Choose an optimal base wavelet according to the noisy signal. Then, decompose the noisy signal into sub-bands by wavelet transform to yield coarse level scaling coefficients and wavelet coefficients.
(2) Estimate the noise threshold T and design the threshold shrinkage function according to the requirements of noise reduction. Then, filter the noise by applying the shrinkage function to every wavelet coefficient.
(3) Reconstruct the signal with new coefficients to obtain the estimated signal $\hat{S}$.

At present, soft thresholding is a commonly-used wavelet shrinkage function [14–16] which has better signal filtering performance than hard thresholding. It can be described as follows:

$$y_S(c) = \begin{cases} 0 & |c| \le T \\ \operatorname{sgn}(c) \cdot (|c| - T) & |c| > T \end{cases} \tag{2}$$

where c is the wavelet coefficient of $\hat{S}$ after wavelet decomposition. T is the *Universal threshold* proposed by Donoho and Johnstone [17]. It can be expressed as:

$$T = \hat{\sigma}\sqrt{2\log(N)} \tag{3}$$

where N is the number of samples in the signal vector. $\hat{\sigma}$ is the estimate of standard deviation of noise and $\hat{\sigma} = MAD / 0.6745$, where the MAD denotes the median absolute deviation of wavelet coefficients. It should be noted that no matter what kind of shrinkage function is applied, the selection of an optimal wavelet function for a given ECG will directly affect the results of wavelet thresholding in the end. On account of the diversity of wavelets, how to choose a base wavelet that is best suited for analyzing a special signal is an important issue in the filtering procedure. Therefore, the base wavelet selection based on a comprehensive entropy criterion is proposed in this paper for removing noise from ECG signals.

3. Comprehensive Entropy Criterion for Optimal Wavelet Selection

From the viewpoint of signal filtering, the objective of noise cancellation is to estimate a function S with minimum mean square error (MSE), *i.e.*, to minimize L_2 risk for a given noisy function:

$$R(\hat{S}, S) = \tfrac{1}{N}\left\|\hat{S} - S\right\|^2 = \tfrac{1}{N}\sum_{i=0}^{N-1}(\hat{s}_i - s_i)^2 \tag{4}$$

By considering representations of such functions in orthonormal bases such as wavelets, the Parseval relation establishes the equivalence between the L_2 risk in the function space and the l_2 risk of the wavelet coefficients of the functions, and consequently the related theorems can be proved in the sequence space of wavelet coefficients rather than the function space itself [18]. Moreover, when a signal to be analyzed is decomposed by the discrete wavelet transform, its characteristics are mainly depicted by its wavelet coefficients. Since an ECG signal is a periodic cardiac signal, the corresponding ECG wavelet coefficients generated by the best base wavelet should reveal a certain regularity in its distribution. Simultaneously, the wavelet coefficients of the decomposed ECG signal should be closely related to the sample sequence of the original signal so as to recover the main variation of an ECG signal. Because of the diverse characteristics and applicability of base wavelets, not all wavelets are suitable for denoising a specific ECG signal. Hence, an optimal base wavelet for a specific ECG signal can be determined by the evaluation of decomposing ability of wavelets and information comparison between the wavelet decomposed coefficients and the sample sequence of the original ECG signal.

3.1. Decomposing Capability Measure

The energy of a signal is one of the measures that characterize the feature of a signal. For a noisy ECG signal $\tilde{S}(n) = [\tilde{s}_0, \tilde{s}_1, ..., \tilde{s}_N]$, suppose $C(n) = [c_0, c_1, ..., c_N]$ is its wavelet coefficient sequence at decomposing scale j after the DWT. The energy of $C(n)$ can be defined by its wavelet coefficients as follows:

$$E_{energy}(c) = \sum_{i=1}^{N} \left| c(j,i) \right|^2 \tag{5}$$

where $c(i, j)$ is the ith coefficient of the jth level of decomposed $\tilde{S}(n)$. For the same amount of energy within a frequency sub-band, the specific features of the signal may be significantly different. The spectral distribution of the energy needs to be considered to ensure the effective feature extraction. In information theory, Shannon entropy is a measure of uncertainty associated with random variables. Thus, the energy distribution of wavelet coefficients can be described by Shannon entropy as follows:

$$E_{entropy}(c) = -\sum_{i=1}^{N} p_i \cdot \log_2 p_i \tag{6}$$

where p_i is the energy probability distribution of wavelet coefficients, expressed as:

$$P_i = \frac{\left| c(i) \right|^2}{E_{energy}(c)} \tag{7}$$

and $\sum_{i=1}^{n} p_i = 1$. If $p_i = 0$, then $p_i \cdot \log_2 p_i = 0$. $E_{entropy}(c)$ is bounded by $0 \le E_{entropy}(c) \le \log_2 N$.

Note that the ECG signal has prominent characteristics of small amplitude, usually 10 μV ~ 5 mV, and with frequencies of interest in the range 0.05–100 Hz. Hence, when the ECG signal is decomposed into scales, an appropriate base wavelet should yield large magnitude coefficients at low frequency and negligible magnitude coefficients at the others. Moreover, the higher the energy extracted from the corrupted ECG signal, the more effective the wavelet transform of the signal will be. The lower the entropy is, the higher the energy concentration will be. Therefore, a self-evaluation criterion for the wavelet decomposing coefficients can be defined by the energy-to-Shannon entropy ratio:

$$E_r = \frac{E_{energy}(c)}{E_{entropy}(c)} \tag{8}$$

By combining the energy measure with the Shannon entropy, an appropriate base wavelet should extract the maximum amount of energy from the signal being analyzed, while minimizing the Shannon entropy of the corresponding wavelet coefficients. The bigger E_r is, the more suitable a base wavelet will be for filtering.

3.2. Information Comparison Measure

The energy-to-Shannon entropy ratio only evaluates the contents of wavelet coefficients themselves. It is necessary to compare $C(n)$ with $\tilde{S}(n)$ in order to ensure the coefficients of the decomposed signal are inherently related to the signal. Thus, several information theoretic criteria are used to further evaluate the base wavelet [19], *i.e.*:

(1) Joint entropy

$$E_{joint} = H(\tilde{S}, C) = -\sum_{\tilde{s}\in\tilde{S}}\sum_{c\in C} p(\tilde{s}, c)\log p(\tilde{s}, c) \tag{9}$$

where $p(\tilde{s}, c)$ is the joint probability distribution of two data sequences. The joint entropy measures information associate with $\tilde{S}$ and C as a whole.

(2) Conditional Entropy

$$\begin{aligned} E_{con} = H(C\mid\tilde{S}) &= -\sum_{\tilde{s}\in\tilde{S}} p(\tilde{s}) H(C\mid\tilde{S}=\tilde{s}) = -\sum_{\tilde{s}\in\tilde{S}} p(\tilde{s})\sum_{c\in C} p(c\mid\tilde{s})\log p(c\mid\tilde{s}) \\ &= -\sum_{\tilde{s}\in\tilde{S}}\sum_{c\in C} p(\tilde{s}, c)\log\frac{p(\tilde{s}, c)}{p(\tilde{s})} = H(\tilde{S}, C) - H(\tilde{S}) \end{aligned} \tag{10}$$

where $p(\tilde{s})$ is the probability distribution of the noisy ECG sequence $\tilde{S}$, $p(c\mid\tilde{s})$ denotes the conditional probability distribution of the wavelet coefficient sequence C when $\tilde{S}$ is known. The conditional entropy indicates the information that is particular to each corresponding data sequence itself.

(3) Mutual Information

$$\begin{aligned} I_{mu} = I(\tilde{S}; C) &= \sum_{\tilde{s}\in\tilde{S}}\sum_{c\in C} p(\tilde{s}, c)\log\frac{p(\tilde{s}, c)}{p(\tilde{s})p(c)} \\ &= \sum_{\tilde{s}\in\tilde{S}}\sum_{c\in C} p(\tilde{s}, c)\log p(\tilde{s}, c) - \sum_{\tilde{s}\in\tilde{S}}\sum_{c\in C} p(\tilde{s}, c)\log[p(\tilde{s})p(c)] \\ &= -H(\tilde{S}, C) + H(\tilde{S}) + H(C) \end{aligned} \tag{11}$$

Mutual information is a measure of the information contained in one process related to another process. The average mutual information between the two processes can be expressed as the sum of two self entropies minus the entropy of the pair [19]. In the wavelet thresholding filtering, the mutual information represents the amount of shared information contained in both $\tilde{S}$ and C. Usually, the larger value of the mutual information of $\tilde{S}$ and C indicates that the decomposed wavelet coefficients contain more information of the signal to be analyzed.

(4) Relative Entropy

$$E_{re} = D(\tilde{S}\,\|\,C) = \sum_{\tilde{s}\in\tilde{S}} p(\tilde{s})\log\frac{p(\tilde{s})}{p(c)} \tag{12}$$

with $p(\tilde{s})\log\frac{p(\tilde{s})}{p(c)} = 0$ when $p(\tilde{s}) = 0$, and $p(\tilde{s})\log\frac{p(\tilde{s})}{p(c)} = \infty$ when $p(c) = 0$. The relative entropy measures the distance between probability distributions of data sequences $\tilde{S}$ and C. It is zero if and only if both probability distributions are equivalent to each other.

(5) Comparison information entropy

To recover the main characteristics of the ECG signal, the decomposed coefficients obtained by the best base wavelet should have shared information with the ECG signal as much as possible and exclusive information as little as possible. Therefore, an appropriate wavelet should minimize the joint entropy, conditional entropy, and relative entropy of $\tilde{S}$ and C while maximizing their mutual information. Taking four measures E_{joint}, E_{con}, I_{mu} as well as E_{re} into account, a comparison information criterion is obtained as follows:

$$I_c = \frac{I(\tilde{S};C)}{H(\tilde{S},C)\times H(C\,|\,\tilde{S})\times D(\tilde{S}\,||\,C)} = \frac{I_{mu}}{E_{joint}E_{con}E_{re}} \tag{13}$$

It can be seen that the bigger I_c is, the more suitable a base wavelet will be for filtering.

3.3. Comprehensive Entropy Criterion

According to the decomposition capability measures and information comparison measures, by integrating the advantages of the aforementioned entropy criteria, a novel comprehensive entropy criterion for the optimal base wavelet selection is proposed, which is defined as:

$$E_{com} = \left(\frac{E_{energy}}{E_{entropy}}\right)\left(\frac{I_{mu}}{E_{joint}E_{con}E_{re}}\right) = E_r I_c \tag{14}$$

It is noticeable that the comprehensive entropy criterion not only measures the decomposition capability of a base wavelet, but also measures the similarity between a noisy signal and its wavelet decomposition coefficients. It comprehensively depicts the suitability of a base wavelet for a specific noisy signal. Generally, the base wavelet that produces the maximum of E_{com} should be chosen as the most appropriate wavelet for the noise reduction of an ECG signal.

4. Experiments and Comparison

4.1. ECG Signals

In order to validate the proposed criterion, ECG signals in the MIT-BIH Arrhythmia Database are used for the analysis. The recordings in the MIT-BIH Arrhythmia Database are sampled at 360 Hz per channel with an 11-bit resolution over the 10 mV range with lead II configuration [20,21]. In the noise filtering procedure, the ECG signals in the MIT-BIH Arrhythmia database are regarded as the original clean ECG signals, but the ECG signals of some subjects in the MIT-BIH Arrhythmia Database are still noisy. They cannot be utilized as clean ECG signals and are not suitable for filtering comparison. Therefore, after careful investigation, the ECG signals of sixteen subjects (Subject No. 100, 103, 105, 106, 107, 113, 115, 116, 117, 123, 124, 217, 219, 220, 223, and 234) from the MIT-BIH Arrhythmia Database are selected for evaluating the performance of wavelet thresholding filtering. Moreover, for the same reason, only 10,000 consecutive samples of an analyzed ECG signal with less noise were chosen as the original clean ECG signal. Since electromagnetic noise such as thermal noise is one kind of noise difficult to filter from ECG signals, Gaussian white noise which can mimic the electromagnetic noise usually happening in measurements was added to contaminate the clean ECG signals with noise level SNR of 15, 20, 25, 30, 35 and 40 dB, respectively. Thus, ninety-six noisy ECG signals are generated for the comparison of various criteria for the best base wavelet selection.

4.2. Wavelet Set

Twenty-two widely-used wavelets are chosen as the base wavelet set for signal filtering, including Daubechies (db) series (db1, db2, db3,db4, db5, db6, db7, db8, db9, db10), sym series (sym2, sym3, sym4, sym5, sym6, sym7, sym8) and coif series (coif1,coif2, coif3, coif4,coif5). To compare the criteria described in Section 3, each noisy ECG signal was firstly decomposed up to eight levels by every wavelet in the set. Then, the best wavelet for a noisy ECG signal is chosen according to nine criteria: energy, entropy, energy-to-entropy ratio, joint entropy, conditional entropy, mutual information, relative entropy, comparison information entropy as well as the comprehensive entropy criterion, respectively. In wavelet thresholding filtering, the universal threshold expressed as Equation (3) is employed to estimate the noise in the noisy ECG signal. The soft thresholding function shown as Equation (2) is used to realize the transformation of wavelet decomposing coefficients. The filtering process is conducted according to three steps described in Section 2. All the analysis of ECG signals in this work are carried out through software Matlab7.10.0 (R2010a) (The MathWorks, Inc., Natick, MA, USA).

4.3. Denoising Performance Indexes

Suppose $S(n)=[s_0,s_1,...,s_N]$ is the original clean ECG signal from the MIT database, while $\hat{S}(n)=[\hat{s}_0,\hat{s}_1,...,\hat{s}_N]$ denotes the denoised ECG signal by using the selected optimal base wavelet and soft thresholding filtering. The filtering performance of different best wavelets are evaluated through comparing S and $\hat{S}$ based on four filtering indexes defined in the following:

The output signal to noise ratio SNR_o:

$$SNR_o = 10\log_{10}\frac{\sum_{i=1}^{N} S^2(i)}{\sum_{i=1}^{N} | S(i)-\hat{S}(i) |^2} \tag{15}$$

root mean square error RMSE:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[S(i)-\hat{S}(i)]^2} \tag{16}$$

percent root mean-square difference PRD:

$$PRD = 100\sqrt{\left\{\sum_{n=1}^{N}[S(n)-\hat{S}(n)]^2 / \sum_{n=1}^{N}[S(n)]^2\right\}} \tag{17}$$

and correlation coefficient r [22]:

$$r = \frac{\sum_{i=1}^{N}(s_i-\bar{s})(\hat{s}_i-\bar{\hat{s}})}{\sqrt{\sum_{i=1}^{N}(s_i-\bar{s})^2\sum_{i=1}^{N}(\hat{s}_i-\bar{\hat{s}})^2}} \tag{18}$$

It is noticed that the SNR_o is a measure of signal strength related to the background noise after wavelet thresholding denoising. Theoretically, the SNR_o of a filtered ECG signal should be large in amplitude in order to recover the useful signal. On the other hand, the RMSE reflects the distortion of the filtering result. The smaller the RMSE is, the closer the denoised signal is to the original signal and less distortion of the denoised signal after filtering. Moreover, the PRD gives the information about the percentage of distortion of the filtered signal. A small PRD value indicates the efficiency of the denoising procedure. In addition, the correlation coefficient r is a statistical concept that measures how well the denoised signals follow the actual signal. The larger value of r is, the closer the denoised signal is to the original signal.

4.4. Results and Discussion

Through the discrete wavelet decomposition, each of ninety-six noisy ECG signals is transformed into a wavelet coefficient sequence. There are twenty-two test wavelets in total. Hence, every noisy ECG signal is converted into 22 coefficient sequences in terms of 22 base wavelets. Nine criteria for the best base selection are calculated for every coefficient sequence of the same noisy signal. The best base wavelet based on the same criterion is finally selected and recorded through the comparison of criterion values. Because of the diversity of ECG signals of individuals, the selected best wavelet of all subjects reveals a great variety even through the same criterion. Moreover, for the same noisy signal, the selection results of the best base wavelets may also be different depending on the different criteria. Figure 1 shows a typical example of the best base wavelet selection for the ECG signals of all 16 subjects with same noise level SNR 20 dB. Each line represents the best wavelets of one noisy ECG signal obtained respectively by nine criteria. In total, there are 16 wavelet lines in Figure 1. It can be seen that each line of best wavelets fluctuates dramatically. Besides, the best wavelets selected by nine criteria are not identical to each other for all the subjects. Therefore, it is a significant issue to find an appropriate criterion for the best wavelet selection in ECG signal filtering.

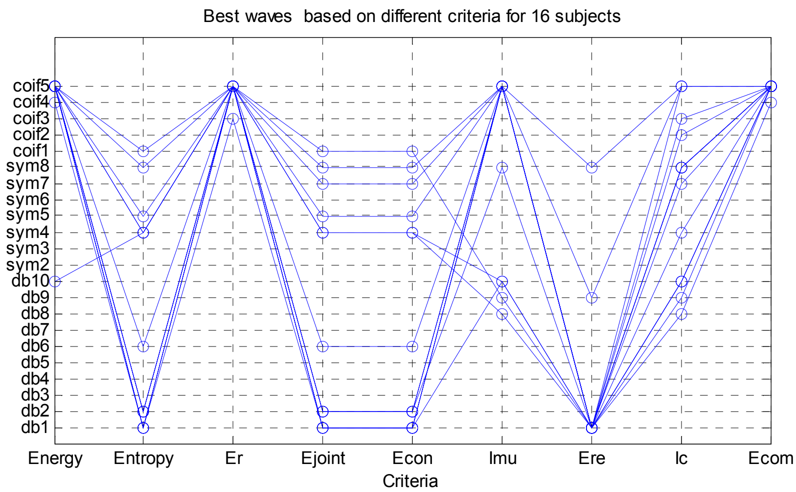

Figure 1. Best wavelets for ECG signals with input SNR = 20 dB of all subjects.

It can also be seen from Figure 1 that two criteria E_{joint} and E_{con} always choose the same wavelets for all noisy signals of 16 subjects. The optimal base wavelets for all subjects obtained by $E_{entropy}$, E_{joint}, E_{con}, I_{mu} and I_c are scattered in the wavelet family; while those selected by E_{energy}, E_r, E_{re} and E_{com} are relatively concentrated. In fact, for the same noisy ECG signal, obvious differences exist in filtering performance based on these nine criteria. Take the ECG signal with noise level of SNR = 20 dB of subject No. 116 as an example. In terms of the nine criteria, the nine best base wavelets and the corresponding criterion values are recorded in Table 1. Then, the corresponding nine denoised ECG signals are shown in Figure 2 for comparison with the original ECG signal and the corrupted ECG signal. Since the soft thresholding inherently shrinks large wavelet coefficients with larger bias [23], the amplitudes of all the denoised signals are lower than those of the original ECG signal. The difference of denoised signals based on nine criteria of the best base wavelet selection can be seen in Figure 3, which is an enlarged version of the denoised one-cycle ECG signals. It is noticeable that the wave shapes of the filtered signals in Figures 2 and 3 based on $E_{entropy}$, E_{joint}, E_{con} and E_{re} are more or less distorted.

Table 1. Best base wavelets and corresponding criterion values of ECG signal with input SNR 20 dB of subject 116.

Criterion	E_{energy}	$E_{entropy}$	E_r	E_{joint}	E_{con}	I_{mu}	E_{re}	I_c	E_{com}
Wavelet	coif5	coif1	coif5	coif1	coif1	db9	db1	db8	coif4
Criterion value	8.3943×10^3	4.2224	1.5282×10^3	14.5307	4.2195	1.4138×10^{-2}	5.5464	3.4654×10^{-5}	2.6248×10^{-2}

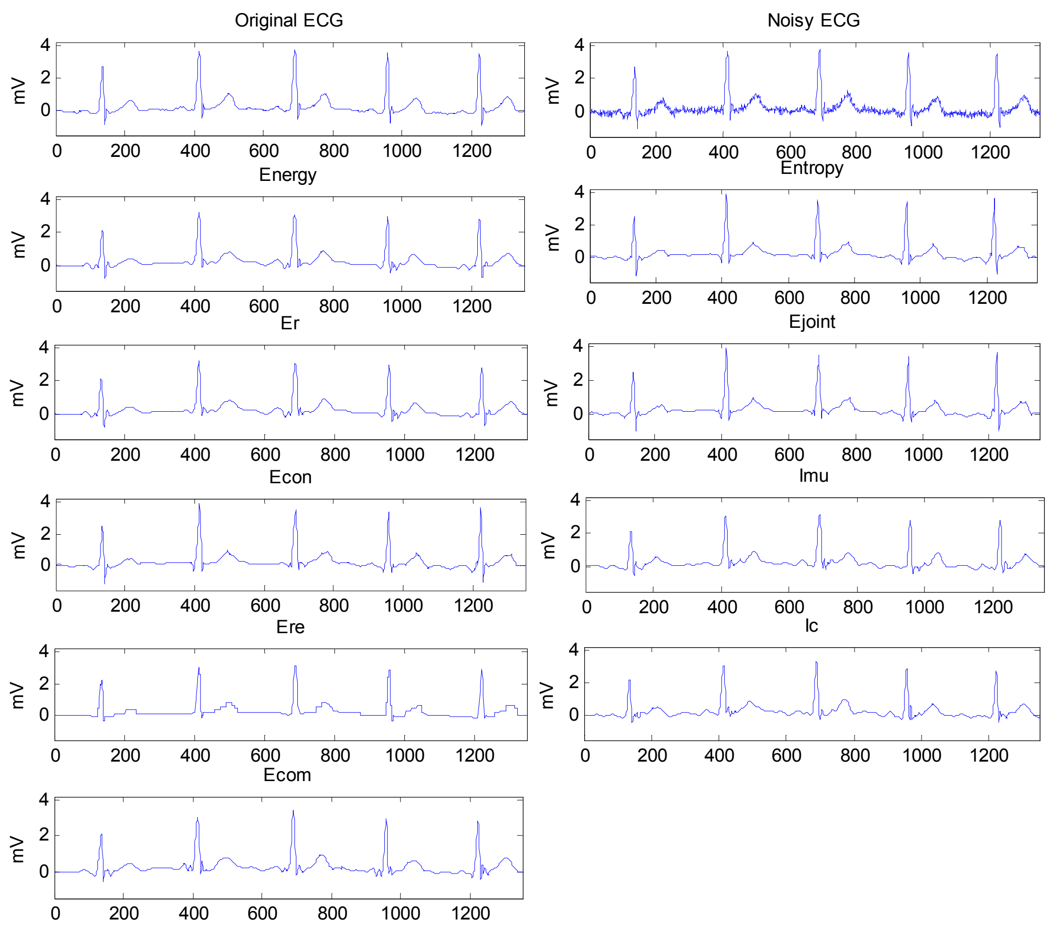

Figure 2. Filtering comparison of ECG signals with noise level SNR = 20 dB of subject No. 116.

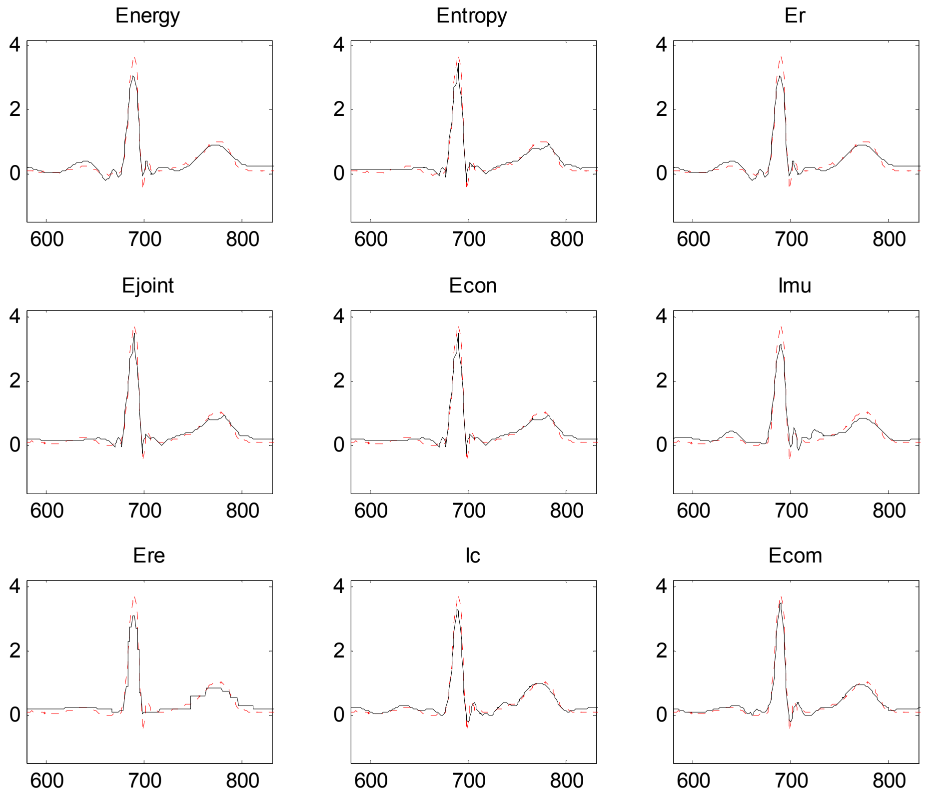

Figure 3. Filtering comparison of one-cycle ECG signal with noise SNR = 20 dB of subject No. 116. Notes: The dotted line represents the original clean ECG signal, while the solid line indicates the denoised ECG signal.

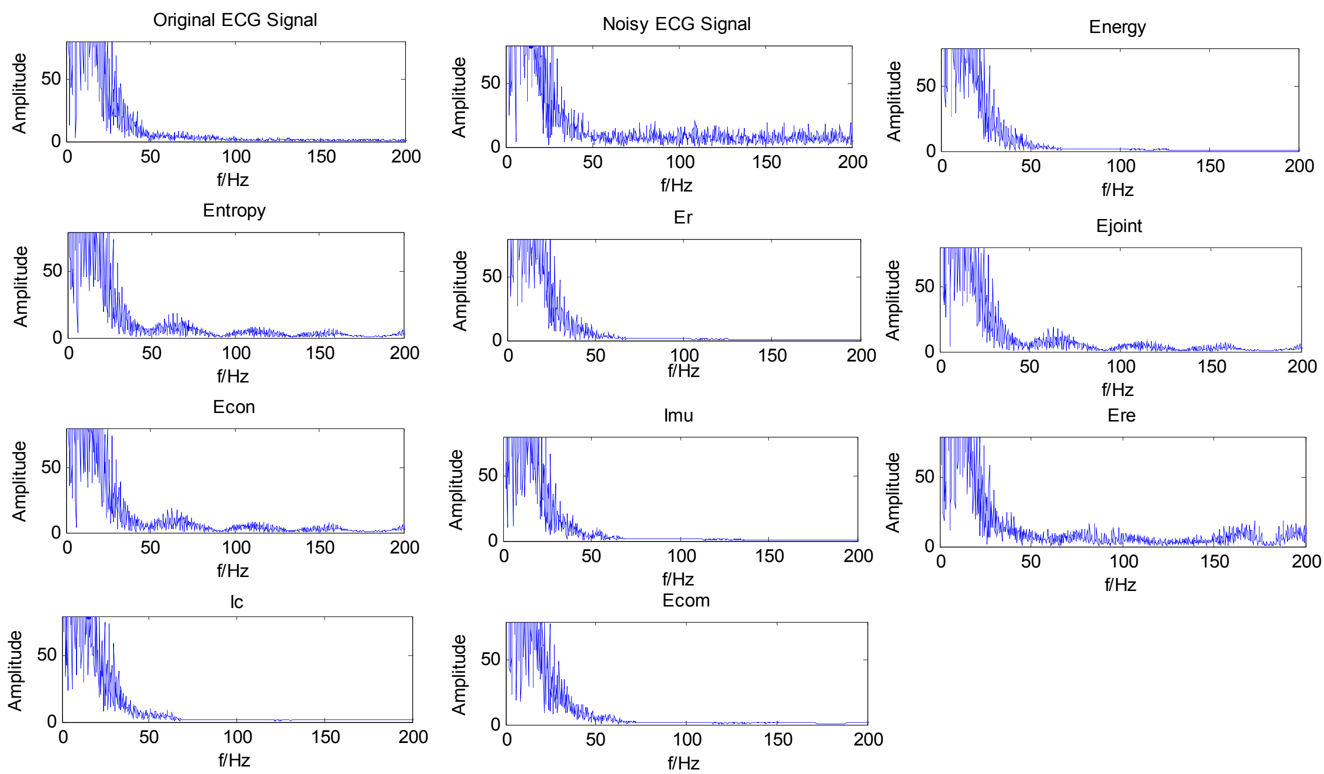

Figure 4. Frequency spectrum comparison of ECG signals with input SNR 20 dB of subject No. 116.

The corresponding frequency spectra shown in Figure 4 also indicate that some undesirable noise still remains in the filtered signals based on these four criteria. Furthermore, it is a common phenomenon for the most of noisy ECG signals with different noise levels. Hence, though $E_{entropy}$, E_{joint}, E_{con} and E_{re} reflect some features of an ECG signal, they cannot be used independently for the selection of best base wavelet. With the optimal wavelets, the filtering results of subject No. 116 are recorded in Table 2. The comprehensive entropy criterion chooses the wavelet Coif4 as the optimal base wavelet for the noisy ECG signal with noise level of SNR 20 dB of subject No. 116. By excluding undesirable criteria $E_{entropy}$, E_{joint}, E_{con} and E_{re}, criterion E_{com} obtains the best filtering indexes, *i.e.*, the highest values of SNR_o and r, as well as the lowest values of RMSE and PRD. The same conclusion can also be made according to the filtering results shown in Table 3, which indicates the mean values of four denoising indexes (*i.e.*, $\overline{SNR_o}, \overline{RMSE}, \overline{PRD}$ and $\overline{r}$) for the ECG signals with noise level SNR 20 dB of all subjects.

Furthermore, on the basis of nine best wavelet selection criteria, the mean filtering index values for all 96 ECG signals with six noise levels, *i.e.*, $\overline{\overline{SNR_o}}, \overline{\overline{RMSE}}, \overline{\overline{PRD}}$ and $\overline{\overline{r}}$, are shown in Table 4. For the comprehensive entropy criterion, the mean values of SNR_o, RMSE, PRD, and r of 96 ECG signals are 19.4922, 0.0763, 13.9420 and 0.9758, respectively. It is noted that, except $\overline{\overline{PRD}}$, indexes $\overline{\overline{SNR_o}}, \overline{\overline{RMSE}}$ and $\overline{\overline{r}}$ of E_{com} are better than those of the other eight criteria. Four filtering indexes of E_{re} are the worst. Besides, though the PRD of $E_{entropy}$ is 13.9304 which is the lowest value and less than that of E_{com}, too much noise remaining in the most of denoised signals greatly attenuates the efficacy of $E_{entropy}$.

In order to further validate the proposed approach, the shape-match (SM) approach based on cross correlation coefficient [6] is implemented since it is commonly used in the selection of optimal base wavelets. The procedure is described as follows: firstly, all the ECG signals are segmented into ECG cycle data. The cross correlation coefficients of cycle ECG segments and base wavelets are calculated. Then, the wavelet with highest correlation coefficient is chosen as the best base wavelet of an ECG signal. With the wavelet obtained by the SM, all the ECG signals are filtered by the same soft-thresholding method used in the former experiment. Moreover, the mean values of four filtering indexes of six denoised ECG signals of every subjects, *i.e.*, SNR_{om}, $RMSE_m$, PRD_m and r_m, are respectively computed for the E_{com} and the SM. The two-sided *t*-hypothesis test is conducted in the mean values of SNR_{om}, $RMSE_m$, PRD_m and r_m of two methods. The values of statistic variables of these four indexes for two methods are recorded in Table 5 and the corresponding box figures are illustrated in Figure 5. It can be clearly seen that the values of statistic variables of four indexes of E_{com} are better than those of the SM. Besides, the *t*-test results show that all the statistic values of the test are larger than the value of two-sided *t*-test standard deviation $u_{\frac{\alpha}{2}}$ = 1.6973 as the significance level $\alpha = 0.05$. Therefore, it can be concluded that the SNR_{om}, $RMSE_m$, PRD_m and r_m of E_{com} is significantly different from those of the shape-matched approach (n_1 = n_2 = 16, significance level $\alpha = 0.05$, two-sided *t*- test).

Finally, the mean value curves of SNR_o, RMSE, PRD and r for all the noisy signals based on the comprehensive entropy criterion are depicted in Figure 6. It is noted that both SNR_o and r varies proportionally to the increase of input SNR, while RMSE and PRD changes inversely as input SNR rises. Moreover, the relation between the output SNR_o and the input SNR is close to linear. It indicates that the proposed criterion E_{com} has the stable performance in the optimal base selection. The variation of RMSE, PRD and r tends to be slowly as the input SNR rises merely because there is less noise in the contaminated ECG signals to be reduced.

Table 2. Filtering performance comparison for ECG signal with input SNR 20 dB of subject No.116.

Criterion	SNR_o (dB)	RMSE (mV)	PRD (%)	R
E_{energy}	17.7889	0.1481	12.8990	0.9796
$E_{entropy}$	18.3061	0.1395	12.1533	0.9823
E_r	17.7889	0.1481	12.8990	0.9796
E_{joint}	18.3061	0.1395	12.1533	0.9823
E_{con}	18.3061	0.1395	12.1533	0.9823
I_{mu}	17.7078	0.1494	13.0199	0.9823
E_{re}	16.9774	0.1626	14.1622	0.9760
I_c	17.3693	0.1554	13.5374	0.9803
E_{com}	17.8896	0.1463	12.7502	0.9829

Table 3. Filtering performance comparison for ECG signal with input SNR 20 dB of all subject.

Criterion	$\overline{SNR_o}$ (dB)	$\overline{RMSE}$ (mV)	$\overline{PRD}$ (%)	$\overline{r}$
E_{energy}	14.6989	0.1093	19.9152	0.9662
$E_{entropy}$	14.6835	0.1091	19.7216	0.9625
E_r	14.6915	0.1094	19.9244	0.9658
E_{joint}	14.6596	0.1095	19.7784	0.9623
E_{con} E_{con}	14.6596	0.1095	19.7784	0.9623
I_{mu}	14.5886	0.1107	20.2360	0.9628
E_{re}	13.9541	0.1197	21.7119	0.9578
I_c	14.4673	0.1122	20.4741	0.9614
E_{com}	14.7314	0.1090	19.9128	0.9670

Table 4. Filtering performance comparison for all ECG signals of all subjects.

Criterion	$\overline{\overline{SNR}}$ (dB)	$\overline{\overline{RMSE}}$ (mV)	$\overline{\overline{PRD}}$ (%)	$\overline{\overline{r}}$
E_{energy}	19.4608	0.0766	13.9811	0.9757
$E_{entropy}$	19.1159	0.0773	13.9304	0.9738
E_r	19.4916	0.0763	13.9432	0.9757
E_{joint}	19.1047	0.0774	13.9744	0.9741
E_{con}	19.1047	0.0774	13.9744	0.9741
I_{mu}	19.4568	0.0767	14.0452	0.9745
E_{re}	17.7827 *	0.0882 *	15.7884 *	0.9712 *
I_c	19.4452	0.0770	14.0824	0.9743
E_{com}	19.4922	0.0763	13.9420	0.9758

Note: * denotes the worst filtering result of nine criteria.

Table 5. Statistic variables of SNR_{om}, $RMSE_m$, PRD_m and r_m of E_{com} and those of shape-matched approach.

Index	**SNR_{om}**		**$RMSE_m$**		**PRD_m**		**r_m**	
Statistic Variable	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**
E_{com}	19.4922	3.6429	0.0763	0.0107	13.9420	5.7904	0.9758	0.0163
Shape-Matched (SM)	17.3304	3.4844	0.0865	0.0131	17.5809	6.0617	0.9632	0.0137

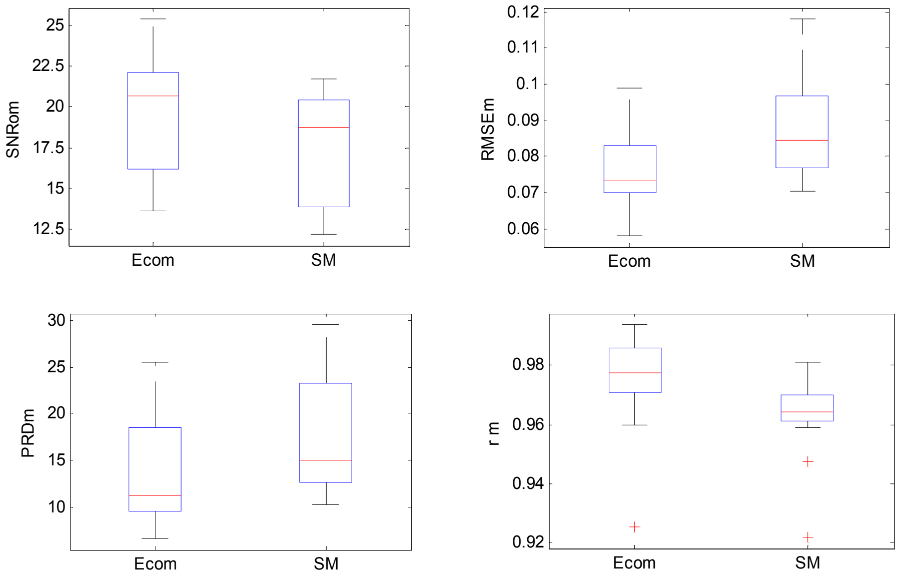

Figure 5. Box plots of SNR_{om}, $RMSE_m$, PRD_m, and r_m of E_{com} and shape-matched approach.

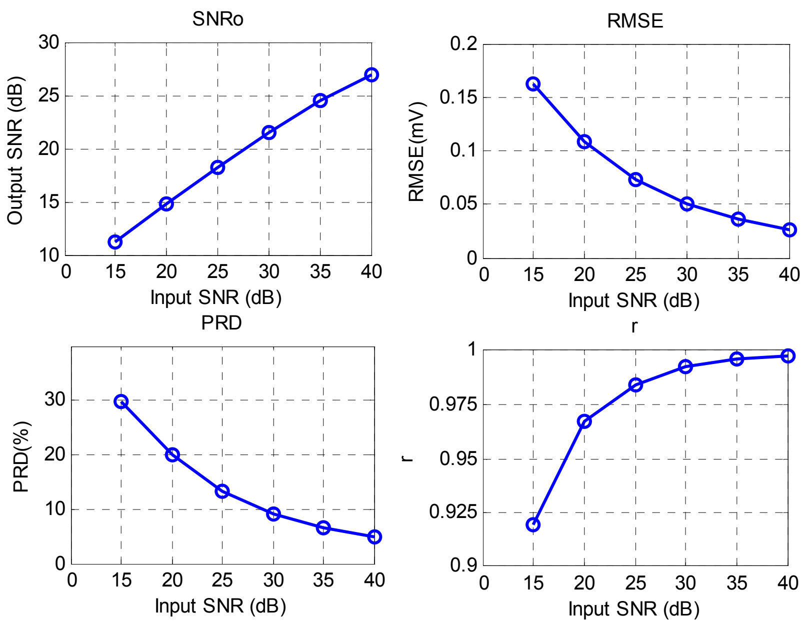

Figure 6. Variation of SNR_o, RMSE, PRD, and r of all ECG signals with different noise levels based on the comprehensive entropy criterion.

5. Conclusions

Wavelet transforms have become efficient analyzing tools for time-varying and non-stationary ECG signals due to their powerful decomposition ability of both high and low resolution signals. However, the selection of the best wavelet is a crucial issue in the wavelet filtering of ECG signals. Inspired by information entropy, a novel comprehensive entropy criterion based on the combination of multiple entropy measures is proposed in this paper to find an appropriate wavelet for a specific ECG signal. It is devised through evaluating both the decomposing ability of wavelet and the information similarity between decomposed coefficients and the signal to be analyzed. The ECG signals of sixteen subjects selected from the MIT-BIH Arrhythmia Database were used to validate the proposed criterion. Ninety-six noisy ECG signals were generated by artificially adding the Gaussian white noise with noise levels of SNR with 15, 20, 25, 30, 35 and 40 dB to the original ECG signal, respectively. A detailed comparison for the optimal wavelet selection was conducted between the comprehensive criterion and the other eight common criteria, *i.e.*, energy, entropy, energy-to-entropy ratio, joint entropy, conditional entropy, mutual information, relative entropy, and comparison information entropy. Four filtering indexes SNR_o, RMSE, PRD and correlation coefficient were used for performance comparison of the nine criteria. The filtering results have verified that entropy, joint entropy, conditional entropy, and relative entropy cannot be used independently as the criterion for best base wavelet selection due to the prominent distortion and residual noise in most of the filtered signals. Nevertheless, the comprehensive entropy criterion can find the best base wavelet for a specific ECG signal and obtain better filtering performance than those of the compared eight criteria. The statistical analysis results also show that the four filtering indexes of denoised ECG signals of E_{com} have significant differences from those of the shape-matched approach for base wavelet selection. By comparison to the shape-matched approach and other filtering-index-based methods, the comprehensive criterion chooses the optimal base wavelet only through decomposed wavelet coefficients without utilizing any other extra data or computations. Hence, the lesser complexity in computation of E_{com} will definitely shorten the computational time for optimal base wavelet selection. In terms of its efficiency, the comprehensive criterion can be further applied to the wavelet noise reduction of other biomedical or mechanical signals. However, it should be noted that the comprehensive criterion is based on the entropy which reflects the regularity of an ECG signal. Because of the diversity of cardiac diseases, the comprehensive criterion will be less effective for noise filtering of irregular burst ECG signals, such as heart attack signals. Hence, our research in the future will focus on the improvement of comprehensive entropy criterion for different accidental pathological ECG signals.

Acknowledgments

This work is partially supported by National Natural Science Foundation of China (Grant Nos. 61203108, 61171088, 61263016, and 61571302), The Natural Science Foundation of Shanghai (Grant No. 14ZR1430300), the Research Projects of Science and Technology Commission of Shanghai (Grant No. 14140711200) and the Innovation Program of Shanghai Municipal Education Commission (Grant No. 13YZ056).

Author Contributions

Hong He devised the new criterion for the best base wavelet selection and the validation filtering experiment, supervised data collection, processed signals, analyzed results and wrote the paper. Yonghong Tan revised and edited the paper. Yuexia Wang collected the data and devised noisy data.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Søensen, J.T.; Clemmensen, P.; Sejersten, M. Telecardiology: Past, Present and Future. *Rev. Esp. Cardiol*. **2013**, *66*, 212–218.
2. Luo, S.; Johnston, P. A Review of Electrocardiogram Filtering. *J. Electrocardiol.* **2010**, *43*, 486–496.
3. Gacek, A.; Pedrycz, W. *ECG Signal. Processing, Classification and Interpretation: A Comprehensive Framework of Computational Intelligence*; Springer: London, UK, 2012.
4. Gao, R.X.; Yan R. *Wavelets: Theory and Applications for Manufacturing*; Springer: Berlin, Germany, 2011.
5. Castillo, E.; Morales, D.P.; García, A.; Martínez-Martí, F.; Parrilla, L.; Palma, A.J. Noise Suppression in ECG Signals through Efficient One-Step Wavelet Processing Techniques. *J. Appl. Math.* **2013**, *2013*, doi: org/10.1155/2013/763903.
6. Singh, B.N.; Tiwari, A.K. Optimal Selection of Wavelet Basis Function Applied to ECG Signal Denoising. *Digit. Signal. Process.* **2006**, *16*, 275–287.
7. Bhatia, P.; Boudy, J.; Andreao, R.V. Wavelet Transformation and Pre-Selection of Mother Wavelet for ECG Signal Processing. In Proceedings of the 24th IASTED International Multiconference: Biomedical Engineering, Innsbruck, Austria, 15–17 Februray 2006; pp. 390–395.
8. Amit, K.; Mandeep, S. Optimal Selection of Wavelet Function and Decomposition Level for Removal of ECG Signal Artifacts. *J. Med. Imaging Health Inf.* **2015**, *5*, 138–146.
9. Tan, H.G.R.; Tan, A.C.; Khong, P.Y.; Mok, V.H. Best Wavelet Function Identification System for ECG Signal Denoise Applications. In Proceedings of the IEEE International Conference on Intelligent and Advanced Systems, ICIAS 2007, Kuala Lumpur, Malaysia, 25–28 November 2007; pp. 631–634.
10. Stantic, D.; Jo, J. Selection of Optimal Parameters for ECG Signal Smoothing and Baseline Drift Removal. *Comput. Inf. Sci.* **2014**, *7*, 99–110.
11. Ranjeet, K.; Farida, J. Retained Signal Energy Based Optimal Wavelet Selection for Denoising of ECG Signal Using Modifide Thresholding. In Proceedings of the IEEE 2011 International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT), Aligarh, India, 17–19 December 2011; pp. 196–199.
12. El-Sayed, A.; Dahshan, E. Genetic Algorithm and Wavelet Hybrid Scheme for ECG Signal Denoising. *Telecommun. Syst.* **2011**, *46*, 209–215.
13. Addison, P.S. Wavelet Transforms and the ECG: A Review. *Physiol. Meas.* **2005**, *26*, 155–199.

14. Mallat, S. *A Wavelet Tour of Signal Processing*, 3rd ed.; Academic Press: Waltham, MA, USA, 2009.
15. Mateo, J.; Torres, A.M.; Soria, C.; Santos, J.L. A Method for Removing Noise from Continuous Brain Signal Recordings. *Comput. Electr. Eng.* **2013**, *39*, 1561–1570.
16. Donoho, D.L.; Johnstone, I.M. Ideal Spatial Adaptation via Wavelet Shrinkage. *Biometrika* **1994**, *81*, 425–455.
17. Donoho, D.L. Denoising by Soft-thresholding. *IEEE Trans. Inform. Theory* **1995**, *41*, 613–627.
18. Von Sachs, R.; MacGibbon, B. Nonparametric Curve Estimation by Wavelet Thresholding with Locally Stationary Errors. *Scand. J. Stat.* **2000**, *27*, 475–499.
19. Gray, R.M. *Entropy and Information Theory*, 2nd ed.; Springer: New York, NY, USA, 2011.
20. Moody, G.B.; Mark, R.G. The Impact of the MIT-BIH Arrhythmia Database. *IEEE Eng. Med. Biol.* **2001**, *20*, 45–50.
21. Goldberger, A.L.; Amaral, L.A.N.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.-K.; Stanley, H.E. Physiobank, Physiotoolkit, and Physionet: Components of a New Research. *Circulation* **2000**, *101*, e215–e220.
22. Üstundag, M.; Gökbulut, M.; Sengür, A.; Ata, F. Denoising of Weak ECG Signals by Using Wavelet Analysis and Fuzzy Thresholding. *Netw. Model. Anal. Health Inform. Bioinfor.* **2012**, *1*, 1–6.
23. Bruce, A.G.; Gao, H.Y. WaveShrink with Firm Shrinkage. *Stat. Sin.* **1997**, *7*, 855–874.