

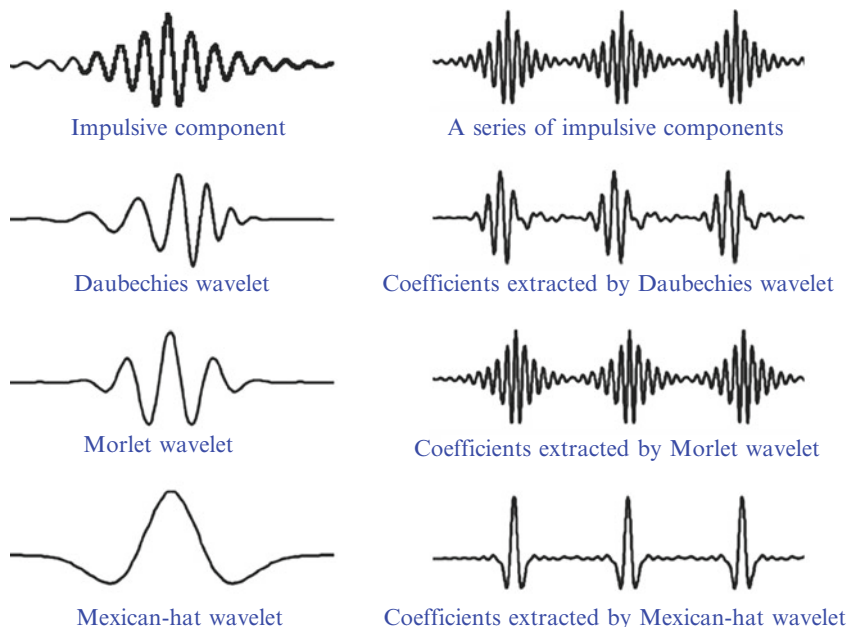
## Chapter 10

# Selection of Base Wavelet

One of the advantages of wavelet transform for signal analysis is the abundance of the base wavelets developed over the past decades – there are a total of 13 wavelet families documented in the MATLAB library. From such abundance arises a natural question of how to choose a base wavelet that is best suited for analyzing a specific signal. The question is valid, since the choice in the first place may affect the result of wavelet transform at the end. As an example, Fig. 11.1 (top row, left) illustrates an impulsive signal and how it may appear as a time series (top row, right) in real-world applications. The three rows below illustrate three representative base wavelets and the results of using them to analyze the impulsive signal: (1) Daubechies wavelet (Daubechies 1992), (2) Morlet wavelet, and (3) Mexican hat wavelet. These base wavelets have been used for machine condition monitoring and health diagnosis studies, as reported in Shao and Nezu (2004), Li et al. (2000), and Abu-Mahfouz (2005). Comparing the wavelet transform results using these wavelets (shown in the right column, Fig. 10.1), it is apparent that only the Morlet wavelet is effective in extracting the impulsive component from the signal, as illustrated by the similarity in the waveform between the corresponding wavelet coefficient and the impulsive component. The Daubechies and Mexican-hat wavelets, in comparison, did not fully reveal the characteristics of impulsive component. Such an example motivates the study of base wavelet selection to achieve optimal result in feature extraction from a signal. In this chapter, we first present a general strategy for base wavelet selection, from both a qualitative and a quantitative aspect. Subsequently, we introduce several quantitative measures that can be used as guidelines for wavelet selection, to guarantee effective extraction of signal features.

### 10.1 Overview of Base Wavelet Selection

The topic of base wavelet selection has been addressed by researchers from different aspects. These prior approaches can be categorized as either qualitative or quantitative, and are reviewed in the following two sections.



**Fig. 10.1** Impulsive feature extraction using different base wavelets

### 10.1.1 Qualitative Measure

Base wavelets are characterized by a number of properties, such as *orthogonality*, *symmetry*, and *compact support*. Understanding these properties will be helpful for choosing a candidate base wavelet from the wavelet families for analyzing a specific signal. For example, the *orthogonality* property indicates that the inner product of the base wavelet is unity with itself, and zero with other scaled and shifted wavelets. As a result, using an orthogonal wavelet will result in efficient signal decomposition into nonoverlapping subfrequency bands. High computational efficiency can be achieved when orthogonal wavelets are used for implementing the discrete wavelet transform (DWT, see Chap. 4 for details) and wavelet packet transform (WPT, refer to Chap. 5). The symmetric property ensures that a base wavelet can serve as a linear phase filter. This is an important property in filtering operations, as the absence of it can lead to phase distortion. A *compact support* wavelet is one whose basis function is nonzero only within a finite interval. This allows the wavelet transform to efficiently represent signals that have localized features. The efficiency of such representation is important for data compression.

In recent years, the basic properties of wavelets have been extensively investigated to determine the suitability of a wavelet for specific applications. For example, based on the experiments conducted on a total of 23 Brodatz textures (Mojsilović

et al. 2000), it was concluded that the Biorthogonal wavelets with symmetry property enabled higher texture classification rate than the Daubechies wavelets, which is asymmetrical (e.g., 64.34% for Db3 vs. 82.17% for Bior3.3r). Similarly, the symmetric property of five wavelets (i.e., Haar, Db6, Coif4, Bior5.5, and Bior6.8) were reviewed (Fu et al. 2003), from which the Bior6.8 wavelet was chosen as the best-suited wavelet to separate the roughness, waviness, and geometrical form of an engineering surface into different frequency bands for both functional correlation and process diagnosis in manufacturing. In the area of biomedical engineering, the *regularity* and *symmetry* of base wavelets were considered as essential features for auditory-evoked potentials (AEP) signal analysis (Bradley and Wilson 2004). The morphology and latency of peaks, which characterize the AEP signal, were preserved when using a symmetric base wavelet, and the smooth peaks contained in the AEP signal were well matched when regularity of a base wavelet is greater than two. By taking into account the properties of *compact support*, *vanishing moment*, and *orthogonality*, the Coiflet 4 wavelet was selected to effectively separate burst and tonic components in the compound surface electromyogram (EMG) signals recorded from patients with dystonia (Wang et al. 2004). In addition to *orthogonality*, the property of complex or real basis was used to guide the choice of the base wavelet for electrocardiogram (ECG) signal analysis (Bhatia et al. 2006). The Morlet wavelet, Gaussian wavelet, Paul wavelet at order 4, and quadratic B-Spline wavelet were preselected as the candidates for ECG events detection and segmentation. In the area of image processing, the properties of *regularity*, *compact support*, *symmetry*, *orthogonality*, and *explicit expression* were used for recommending base wavelet for image sequence superresolution (Ahuja et al. 2005). It was concluded that the B-Spline family represents the most suitable base wavelet among the four candidates (i.e., Daubechies, Symlet, Coiflet, and B-Spline wavelets) for image sequence superresolution, as it is orthogonal, symmetric, and has the highest regularity, smallest support size, and explicit expression. In analyzing power system transients (Safavian et al. 2005), the Db4, Coiflet, and B-Spline wavelet were shown to be equally well-performing for the transient detection in a power system, as they share the same basic properties: finite support size and low vanishing moment.

Shape matching has been studied as an alternative approach to wavelet selection. For example, to measure the timing of multiunit bursts in surface EMGs from single trials (Flanders 2002), wavelets of different shapes, such as square, triangular, Gaussian and Mexican Hat, were investigated. The Db2 wavelet was chosen for its similarity to the shape of motor unit potentials hidden in the EMG signal. Also, base wavelets of different shapes were compared with ECG signals to determine their appropriateness for extracting a reference base from corrupted ECG, for magnetic resonance imaging (MRI) sequence triggering (Fokapu et al. 2005; Abi-Abdallah et al. 2006). To analyze impulses in vibration signals, researchers looked at the geometric shape of wavelets to determine the optimal choice (Yang and Ren 2004). It was found that components in a signal may be extracted effectively when a base wavelet with similar shape as the component is employed.

### 10.1.2 Quantitative Measure

The various approaches described in Sect. 10.1.1 illustrate the importance of choosing an appropriate base wavelet for effective signal processing. 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. These deficiencies motivate the study of quantitative measures for base wavelet selection.

The measures of *inequality* (Goel and Vidakovic 1995), which includes the Schur concave functions such as *Shannon entropy* and *Fishlow's measure* (Marshall and Olkin 1979), *Emelen's modified entropy measure* (Emlen 1973), and Schur convex functions (*Gini's coefficient* and *Schutz's coefficient* by Marshall and Olkin 1979) were proposed for wavelet selection in data compression and data denoising. A time-series, which was constructed by adding white noise into the sampled Db3 wavelet function, was used to evaluate each of the *inequality* measures. All of them, except for the *Fishlow's measure*, recognized the Db3 wavelet as the best base wavelet among a set of wavelets (Db1–Db20, Db30, Coif8, Coif12, and Coif18).

The *Shannon entropy* was also utilized to identify optimal base wavelet for velocity and temperature time series analysis in atmospheric surface layer (ASL) (Katul and Vidakovic 1996). The large scale eddy motion and small scale fluctuations in the ASL were successfully separated with the chosen Daubechies wavelet. In another study (Bedekar et al. 2005), the *Shannon entropy* was used to choose the Daubechies wavelet of order 3 from 23 preselected wavelets as the optimal wavelet for radio-frequency intravascular ultrasound (IVUS) data decomposition. It accurately decomposed 29 out of 30 IVUS data at all levels. The rest of the wavelets only decomposed less than 21 IVUS data.

In the field of biomedical engineering, study on horse gait classification has discussed an *uncertainty* model for wavelet selection (Arafat et al. 2003). The model combines the *fuzzy uncertainty* with the *probabilistic uncertainty* to provide a better measure, when compared with using either fuzzy or probabilistic uncertainty alone, for choosing an appropriate base wavelet to improve correct classification of different horse gait signals.

Study on assessing hypnotic state of anesthetized patients undergoing surgery (Bibian et al. 2001) has used the *discrimination power*, which is defined as the difference between the statistical features, such as *probability density function*, in the awake as well as in the anesthetized states, to select the appropriate base wavelet. It was found that among the Daubechies, Coiflet, Symlet, biorthogonal, and reverse biorthogonal wavelets, the Daubechies wavelet at order 8 provided the highest discrimination power, thus effectively estimated the hypnotic state. In another study on diagnosing cardiovascular ailments in patients (Singh and Tiwari 2006), experimental results have revealed the suitability of the Daubechies base wavelet at order 8 for the ECG signal denoising, as it

has the maximum *cross correlation* coefficient between the ECG signal and the chosen base wavelets, (Daubechies, Symlet, and Coiflet wavelets). The *cross correlation* measure has also been used in evaluating a base wavelet for detecting and locating the partial discharge (PD) occurred in operational transformers (Ma et al. 2002a, b; Yang et al. 2004). It was found that an optimal base wavelet would maximize the correlation coefficient between the signal of interest and the base wavelet, resulting in PD pulses being successfully separated from electrical noise.

For image denoising, two criteria, the *signal information extraction criterion* and the *distribution error criterion*, were proposed to select an optimal wavelet for improving the denoising performance (Zhang et al. 2005). The first criterion was implemented by calculating the *mutual information* of wavelet coefficients of the image without noise contamination and those of the image with noise contamination. The second criterion was the difference between the Gaussian and the actual distribution of the wavelet coefficients of the image without noise contamination. It was reasoned that the smaller the difference is the better the denoising performance will be, as the denoising performance is optimal only if the underlying signal distribution is Gaussian. Using these two criteria, it was found that the Bior1.3 wavelet provided the best performance among the eight wavelets (Bior1.1, Bior1.3, Bior2.2, Bior2.4, Bior3.3, Db2, Db3, and Db4) investigated when testing four benchmark images.

For automatic ultrasound nondestructive foreign body (FB) detection and classification in nonflat surface containers (Tsui and Basir 2006), the *relative entropy* was employed as a wavelet coefficient similarity measure to select the best base wavelet. The results have shown that the best base wavelet for FB shape classification is Bior3.1, while Haar (or Sym1 or reverse Bior1.1) and reverse Bior3.9 are the best for spherical and rectangular FB material classifications, respectively.

For analysis of impulses in vibration signals (Schukin et al. 2004), the *minimum total error* and *time-frequency resolution* were devised to evaluate different base wavelets on impulsive parameter identification of a single-degree-of-freedom system model. A comprehensive comparison among ten base wavelets (complex B-Spline, Gaussian, Shannon, etc.) indicated that the impulse wavelet is the most appropriate base wavelet for the analysis of impulses.

## 10.2 Wavelet Selection Criteria

The importance of the base wavelet has been addressed by various researchers, as summarized earlier. This section introduces several quantitative measures in evaluating the performance of base wavelets for the specific application domain of condition monitoring and health diagnosis in manufacturing.

### 10.2.1 Energy and Shannon Entropy

The energy content of a signal is a measure that uniquely characterizes the signal, thus can be used for base wavelet selection. The amount of energy contained in a signal  $x(t)$  is expressed as:

$$E_{x(t)} = \int |x(t)|^2 dt \quad (10.1)$$

Similarly, when the signal is represented by discrete sample values  $x(i) (i = 1, 2, \dots, N)$ , the amount of energy is given by:

$$E_{x(i)} = \sum_{i=1}^N |x(i)|^2 \quad (10.2)$$

In (10.2),  $N$  is the length of the signal expressed by the number of data points, and  $x(i)$  is the amplitude of the signal.

The energy content of a signal can also be calculated from its wavelet coefficients, and is expressed as:

$$E_{energy} = \iint |wt(s, \tau)|^2 ds d\tau \quad (10.3)$$

The corresponding sampled version is given by:

$$E_{energy} = \sum_s \sum_i |wt(s, i)|^2 \quad (10.4)$$

Equations (10.3) and (10.4) indicate that the energy associated with each particular scaling parameter  $s$  is expressed as:

$$E_{energy}(s) = \int |wt(s, \tau)|^2 d\tau \quad (10.5)$$

And the energy of the corresponding sampled version is described as:

$$E_{energy}(s) = \sum_{i=1}^N |wt(s, i)|^2 \quad (10.6)$$

where  $N$  is the number of wavelet coefficients and  $wt(s, i)$  represents the wavelet coefficients.

If a major frequency component corresponding to a particular scale  $s$  exists in the signal, then the wavelet coefficients at that scale will have relatively high

magnitudes at the time when this major frequency component occurs. As a result, the energy related to such frequency component will be extracted from the signal when applying the wavelet transform to the signal. For purpose of condition monitoring and health diagnosis, the higher the energy content extracted from the defect-induced transient vibrations is the more effective the wavelet transform of the signal will be. Therefore, the energy content can serve as a criterion for selecting the base wavelet. This is formulated in the following criterion.

1. *Maximum energy criterion:* The base wavelet that extracts the largest amount of energy from the signal being analyzed represents the most appropriate wavelet for extracting features from defect-induced transient vibrations.

Given that for the same amount of energy within a subfrequency band, the specific condition of the signal may be significantly different (e.g., only several frequency components with high magnitude and others with negligible magnitude vs. a widespread spectrum), the spectral distribution (or concentration) of the energy needs to be considered also to ensure effective feature extraction. The energy distribution of the wavelet coefficients is quantitatively described by the Shannon entropy (Cover and Thomas 1991):

$$E_{entropy}(s) = - \sum_{i=1}^N p_i \cdot \log_2 p_i \quad (10.7)$$

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

$$p_i = \frac{|wt(s, i)|^2}{E_{energy}(s)} \quad (10.8)$$

with  $\sum_{i=1}^N p_i = 1$ , and  $p_i \cdot \log_2 p_i = 0$  if  $p_i = 0$ .

Equations (10.7) and (10.8) indicate that the entropy of the wavelet coefficients is bounded by:

$$0 \leq E_{entropy}(s) \leq \log_2 N \quad (10.9)$$

in which  $E_{entropy}(s)$  will be equal to (1) zero, if all other wavelet coefficients are equal to zero except for one wavelet coefficient, and (2)  $\log_2 N$ , if the probability of energy distribution for all the wavelet coefficients are the same (i.e.,  $1/N$ ). This leads to the conclusion that the lower the entropy value is, the higher the energy concentration will be. Therefore, an appropriate base wavelet should yield large magnitude at a few wavelet coefficients and negligible magnitude at others when the signal is decomposed into various scales, leading to the minimum Shannon entropy. The corresponding Shannon entropy-based wavelet selection criterion can thus be designed as:

2. *Minimum Shannon entropy criterion:* The base wavelet that minimizes the entropy of the wavelet coefficients represents the most appropriate wavelet for defect-induced transient vibration analysis.

Combining the strengths of the two criteria described earlier, we note that 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. This lead to the *energy-to-Shannon entropy* ratio, which is defined as:

$$R(s) = \frac{E_{\text{energy}}(s)}{E_{\text{entropy}}(s)} \quad (10.10)$$

In (10.10), the energy  $E_{\text{energy}}(s)$  and the entropy  $E_{\text{entropy}}(s)$  are calculated from (10.6) and (10.7), respectively. By maximizing the *energy-to-Shannon entropy* ratio  $R(s)$ , an appropriate base wavelet can be selected from a set of candidate base wavelets. This leads to the following criterion for wavelet selection:

3. *Energy-to-Shannon entropy ratio measure*: The base wavelet that has produced the maximum energy-to-Shannon entropy ratio should be chosen as the most appropriate wavelet for defect-induced transient vibration signal analysis.

## 10.2.2 Information Theoretic Measure

The *energy* and *Shannon entropy*-related criteria are solely based on the content of the wavelet coefficients themselves. Since the coefficients of a signal's wavelet transformation are inherently related to the signal, information theoretic measures, which describes the relationship between a pair of data sequence, can be explored for best suited base wavelet selection. These are introduced in the following sections.

### 10.2.2.1 Joint Entropy

The joint entropy  $H(X, Y)$  between two data sequences  $X$  and  $Y$  is defined to measure information associated with them as a whole (Cover and Thomas 1991). This is expressed as

$$H(X, Y) = - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y) \quad (10.11)$$

where  $p(x, y)$  is the joint probability distribution of the two data sequences.

### 10.2.2.2 Conditional Entropy

With probability distribution of the data sequence  $X$  known, the amount of information contained in the other data sequence  $Y$  can be measured by the condition entropy  $H(Y|X)$  as (Cover and Thomas 1991):



$$\begin{aligned}
H(Y|X) &= - \sum_{x \in X} p(x) H(Y|X = x) \\
&= - \sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \log p(y|x)
\end{aligned} \tag{10.12}$$

In (10.12),  $p(x)$  is the probability distribution of the data sequence  $X$ , and  $p(y|x)$  denotes the conditional probability distribution of the data sequence  $Y$  when the data sequence  $X$  is known. The conditional probability distribution  $p(y|x)$  is expressed as (Mendenhall and Sincich 1995):

$$p(y|x) = \frac{p(x, y)}{p(x)} \tag{10.13}$$

with  $p(x, y)$  being the joint probability distribution of the two data sequence  $X$  and  $Y$ . As a result, (10.12) can be further expressed as:

$$\begin{aligned}
H(Y|X) &= - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x)} \\
&= - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y) + \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x) \\
&= H(X, Y) + \sum_{x \in X} p(x) \log p(x) = H(X, Y) - H(X)
\end{aligned} \tag{10.14}$$

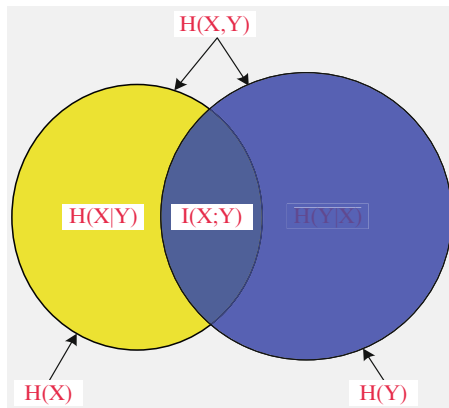
Equation (10.14) indicates that, given the data sequence  $X$ , the condition entropy of data sequence  $Y$  can be calculated by the joint entropy between the two data sequences, minus the entropy of the data sequence  $X$ .

### 10.2.2.3 Mutual Information

The mutual information  $I(X; Y)$  measures the amount of information that data sequence  $X$  contains about data sequence  $Y$ , which is defined as (Cover and Thomas 1991):

$$\begin{aligned}
I(X; Y) &= \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \\
&= \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y) - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log [p(x)p(y)] \\
&= -H(X, Y) - \sum_{x \in X} p(x) \log p(x) - \sum_{y \in Y} p(y) \log p(y) \\
&= -H(X, Y) + H(X) + H(Y)
\end{aligned} \tag{10.15}$$

**Fig. 10.2** Relationships among entropies and mutual information



Equation (10.15) indicates that the mutual information is the sum of the entropies  $H(X)$  and  $H(Y)$ , minus the joint entropy  $H(X, Y)$ .

The relationships among the joint entropy, condition entropy, and mutual information are illustrated in a Venn diagram (Cover and Thomas 1991), as shown in Fig. 10.2. It is noted that the mutual information  $I(X; Y)$  is represented by the intersection of the two data sequences. The greater the mutual information is, the more similar the two data sequences will be. The condition entropy  $H(X|Y)$  or  $H(Y|X)$  expresses the information that is particular to each corresponding data sequence itself, while the joint entropy  $H(X, Y)$  includes all information of the two data sequences.

The above-described relationship can be applied to base wavelet selection by taking the signal to be analyzed and its corresponding wavelet coefficients as two data sequences  $X$  and  $Y$ , respectively. Since defect-induced transient features are represented by the wavelet coefficients, a high value of mutual information between the vibration signal and wavelet coefficients can be expected when an appropriate wavelet is chosen. It should be noted that, when a vibration signal is obtained, the information  $H(X)$  is fixed. Similarly, the information  $H(Y)$  of the wavelet coefficients is fixed once a base wavelet is chosen. On the basis of the relationship described earlier, low values of both the joint entropy and condition entropy are desired for choosing an appropriate wavelet for characterizing defect-induced transient vibrations.

Following are several criteria for base wavelet selection, based on the information theoretic measures.

4. *Minimum joint entropy criterion:* The base wavelet that minimizes the joint entropy between the signal and the wavelet coefficients represents the most appropriate wavelet for defect-induced transient feature extraction.
5. *Minimum condition entropy criterion:* The base wavelet that minimizes the condition entropy between the signal and the wavelet coefficients represents the most appropriate wavelet for defect-induced transient feature extraction.

6. *Maximum mutual information criterion*: The base wavelet that maximizes the mutual information between the signal and the wavelet coefficients represents the most appropriate wavelet for defect-induced transient feature extraction.

#### 10.2.2.4 Relative Entropy

In contrast to the mutual information, which measures shared information between two data sequences, the relative entropy (also known as *Kullback Leibler distance* or the *divergence*) is a measure of the distance between probability distributions of data sequences  $X$  and  $Y$  (Cover and Thomas 1991). The relative entropy is defined as

$$D(X||Y) = \sum_{x \in X} p(x) \log \frac{p(x)}{p(y)} \quad (10.16)$$

with  $p(x) \log \frac{p(x)}{p(y)} = 0$  if  $p(x) = 0$ , and  $p(x) \log \frac{p(x)}{p(y)} = \infty$  if  $p(y) = 0$ .

Equation (10.16) states that the relative entropy value is always nonnegative, and it is zero if and only if both probability distributions are equivalent [i.e.,  $p(x) = p(y)$ ]. The smaller the relative entropy is, the more similar the distributions of the two data sequences will be. For applications in machine condition monitoring and health diagnosis, it is expected that an appropriately chosen base wavelet will be able to extract features related to defect-induced transient vibrations completely. Consequently, a small relative entropy value between the signal (i.e., data sequence  $X$ ) and its corresponding wavelet coefficients (i.e., data sequence  $Y$ ) is desired. The following criterion reflects this consideration:

7. *Minimum relative entropy criterion*: The base wavelet that minimizes the relative entropy between the signal and the wavelet coefficients represents the most appropriate wavelet for defect-induced transient feature extraction.

Synthesizing the above criteria, an appropriate wavelet should minimize the joint entropy, condition entropy, and relative entropy while maximizing the mutual information. Such consideration is captured in the following information measure:

$$Info(s) = \frac{I(X; Y)}{H(X, Y) \times H(Y|X) \times D(X||Y)} \quad (10.17)$$

In (10.17), the joint entropy  $H(X, Y)$ , condition entropy  $H(Y|X)$ , relative entropy  $D(X||Y)$ , and mutual information  $I(X; Y)$  are calculated using (10.11), (10.14), (10.16), and (10.15), respectively. Maximizing the information measure  $Info(s)$  leads to the following comprehensive criterion:

8. *Maximum information criterion*: The base wavelet that has produced the maximum information value should be chosen to be the most appropriate wavelet for defect-induced transient feature extraction.

### 10.3 Numerical Study on Base Wavelet Selection

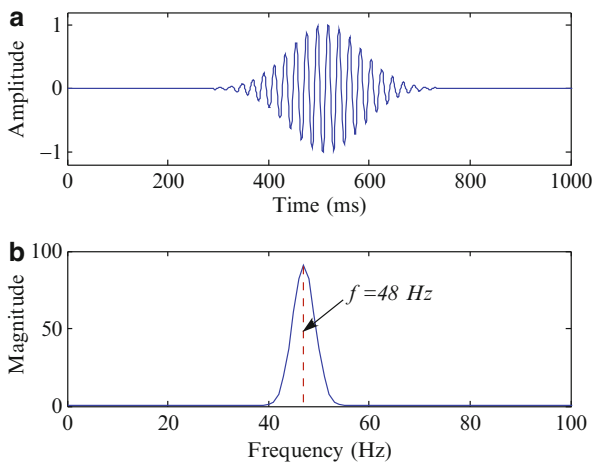
To quantitatively evaluate the base wavelet selection criteria described earlier, a Gaussian-modulated sinusoidal test signal is numerically simulated. Mathematically such a signal can be expressed as

$$x(t) = e^{-\beta(t-t_0)^2} \sin[2\pi f(t - t_0)] \quad (10.18)$$

The symbol  $\beta$  denotes the attenuation factor, and  $t_0$  is the time delay of the signal. This type of signal has been widely used for simulating transient vibrations involved in mechanical systems (Ho and Randall 2000; Schukin et al. 2004; Yang and Ren 2004). Figure 10.3 illustrates the test signal, in which the center frequency is 48 Hz, and the sampling frequency is 1,024 Hz. In the following, the criteria presented in the above sections are evaluated for choosing best suited base wavelet from both real-valued and complex-valued wavelets.

#### 10.3.1 Evaluation Using Real-Valued Wavelets

The performance of real-valued wavelets on processing the test signal is evaluated first, for which the DWT is performed to decompose the test signal. The decomposition level  $L$  of the wavelet transform is determined by the sampling frequency  $f_q$  and



**Fig. 10.3** Test signal: Gaussian-modulated sinusoidal signal

frequency component to be identified in the signal, as expressed in the following equation:

$$\frac{f_q}{2^{L+1}} \leq f_{char} \leq \frac{f_q}{2^L} \quad (10.19)$$

In (10.19),  $f_q$  is the sampling frequency, and  $f_{char}$  is related to the characteristic frequency component of the signal (e.g.,  $f_{char} = 48$  Hz for the test signal). In Table 10.1, the respective frequency ranges covered by each of the decomposition levels under the sampling rate of 1,024 Hz are shown. Since the center frequency (48 Hz) of the test signal falls within the frequency range of 32–64 Hz, which is covered by the decomposition level 4 (corresponding to scale  $s = 2^4 = 16$ ), this level is chosen for the evaluation of each wavelet.

Thirty candidate base wavelets were preselected from seven wavelet families. The energy extracted from the Gaussian-modulated sinusoidal signal by these wavelets is listed in Table 10.2. It is shown that the *Meyer* wavelet has extracted the highest amount of energy, thus is considered the most appropriate base wavelet for analyzing the Gaussian-modulated sinusoidal signal with the given parameters. It is also found that the amount of energy that is extracted from the signal increases with increasing order of the base wavelet, for each wavelet family. This is because base wavelets of higher order within a wavelet family possess higher degree of regularity. As a result, they are better suited for extracting energy from the Gaussian-modulated sinusoidal test signal than their lower-ordered counterparts in the same wavelet family.

The Shannon entropy of the extracted Gaussian-modulated sinusoidal signal is then calculated, as listed in Table 10.3. On the basis of the minimum Shannon entropy, the *Symlet 3* wavelet is considered as the most appropriate base wavelet.

**Table 10.1** Frequency range for each decomposition level under a 1,024 Hz sampling rate

Decomposition level ( $L$ )	Frequency range (Hz)	Decomposition level ( $L$ )	Frequency range (Hz)
1	256–512	4	32–64
2	128–256	5	16–32
3	64–128	6	8–16

**Table 10.2** Energy extracted from the test signal: real-valued wavelets

Base wavelet	Energy (J)	Base wavelet	Energy (J)	Base wavelet	Energy (J)
Haar	33.855	Coif4	60.662	Bior2.6	53.645
Db2	45.546	Coif5	61.856	Bior4.4	52.310
Db4	54.433	Sym2	45.546	Bior5.5	54.614
Db6	58.167	Sym3	51.143	Bior6.8	58.415
Db8	60.207	Sym4	54.433	rBio1.3	45.326
Db10	61.471	Sym6	58.167	rBio2.4	55.138
DB20	63.687	Sym8	60.217	rBio2.6	55.546
Coif1	46.065	Meyr	64.146	rBio4.4	59.235
Coif2	55.038	Bior1.3	53.481	rBio5.5	61.123
Coif3	58.692	Bior2.4	49.198	rBio6.8	60.464

This conclusion is not consistent with the Meyer wavelet selected by the maximum energy criterion. To resolve such conflict, the energy-to-Shannon entropy ratio is calculated and the results are listed in Table 10.4. From the *maximum energy-to-Shannon entropy ratio* criterion, the *Meyer* wavelet possesses the highest values, thus is considered the most appropriate wavelet to analyze the Gaussian-modulated sinusoidal signal.

Various criteria based on information theoretic measures have also been studied to evaluate the performance of each of the real-valued candidate wavelets, as listed in Table 10.5–10.8. It is noted that all of the four measures (i.e., joint entropy, condition entropy, mutual information, and relative entropy) point to the *Meyer* wavelet as the most suited wavelet when analyzing the Gaussian-modulated sinusoidal signal. This is because it maximizes the maximum mutual information, while minimizing the joint entropy, condition entropy, and relative entropy. The comprehensive criterion “*maximum information*” that integrates the effect of these four measures, as illustrated in (10.17), has also shown that the Meyer wavelet is the most appropriate wavelet. This is verified in Table 10.9, in which a maximum information value is obtained when the Meyer wavelet is chosen as the base wavelet.

**Table 10.3** Shannon entropy of the extracted signal: real-valued wavelets

Base wavelet	Shannon entropy	Base wavelet	Shannon entropy	Base wavelet	Shannon entropy
Haar	3.667	Coif4	2.945	Bior2.6	5.959
Db2	3.137	Coif5	2.985	Bior4.4	5.673
Db4	3.475	Sym2	3.137	Bior5.5	6.042
Db6	3.171	Sym3	2.800	Bior6.8	4.069
Db8	3.491	Sym4	3.011	rBio1.3	4.579
Db10	3.121	Sym6	3.598	rBio2.4	4.665
DB20	3.653	Sym8	3.613	rBio2.6	4.949
Coif1	2.856	Meyr	2.959	rBio4.4	4.664
Coif2	3.609	Bior1.3	6.197	rBio5.5	5.034
Coif3	3.617	Bior2.4	6.214	rBio6.8	4.069

**Table 10.4** Energy-to-Shannon entropy ratio of the extracted signal: real-valued wavelets

Base wavelet	Energy-to-Shannon entropy ratio	Base wavelet	Energy-to-Shannon entropy ratio	Base wavelet	Energy-to-Shannon entropy ratio
Haar	9.229	Coif4	20.594	Bior2.6	9.002
Db2	14.512	Coif5	20.719	Bior4.4	9.220
Db4	15.662	Sym2	14.512	Bior5.5	9.047
Db6	18.341	Sym3	18.265	Bior6.8	14.356
Db8	17.246	Sym4	18.079	rBio1.3	9.896
Db10	19.693	Sym6	16.162	rBio2.4	11.817
DB20	17.428	Sym8	16.662	rBio2.6	11.221
Coif1	16.124	Meyr	21.678	rBio4.4	12.699
Coif2	15.244	Bior1.3	8.629	rBio5.5	12.141
Coif3	16.224	Bior2.4	7.915	rBio6.8	14.858

**Table 10.5** Joint entropy of the extracted signal: real-valued wavelets

Base wavelet	Joint entropy	Base wavelet	Joint entropy	Base wavelet	Joint entropy
Haar	4.086	Coif4	3.261	Bior2.6	3.414
Db2	3.409	Coif5	3.246	Bior4.4	3.338
Db4	3.394	Sym2	3.398	Bior5.5	3.415
Db6	3.495	Sym3	3.291	Bior6.8	3.379
Db8	3.358	Sym4	3.250	rBio1.3	3.603
Db10	3.256	Sym6	3.441	rBio2.4	3.329
DB20	3.086	Sym8	3.336	rBio2.6	3.619
Coif1	3.082	Meyr	2.957	rBio4.4	3.341
Coif2	3.614	Bior1.3	3.682	rBio5.5	3.348
Coif3	3.366	Bior2.4	3.313	rBio6.8	3.453

**Table 10.6** Condition entropy of the extracted signal: real-valued wavelets

Base wavelet	Condition entropy	Base wavelet	Condition entropy	Base wavelet	Condition entropy
Haar	1.539	Coif4	0.715	Bior2.6	0.792
Db2	0.851	Coif5	0.699	Bior4.4	0.868
Db4	0.847	Sym2	0.851	Bior5.5	0.833
Db6	0.948	Sym3	0.745	Bior6.8	1.057
Db8	0.812	Sym4	0.704	rBio1.3	0.783
Db10	0.710	Sym6	0.895	rBio2.4	1.073
DB20	0.539	Sym8	0.790	rBio2.6	0.795
Coif1	0.536	Meyr	0.411	rBio4.4	0.802
Coif2	1.067	Bior1.3	1.136	rBio5.5	0.907
Coif3	0.819	Bior2.4	0.767	rBio6.8	0.792

**Table 10.7** Mutual information of the extracted signal: real-valued wavelets

Base wavelet	Mutual information	Base wavelet	Mutual information	Base wavelet	Mutual information
Haar	1.243	Coif4	1.261	Bior2.6	1.045
Db2	0.69	Coif5	1.317	Bior4.4	1.101
Db4	1.074	Sym2	0.869	Bior5.5	1.165
Db6	1.151	Sym3	0.990	Bior6.8	1.559
Db8	1.174	Sym4	1.085	rBio1.3	0.969
Db10	1.291	Sym6	1.110	rBio2.4	0.873
DB20	1.502	Sym8	1.241	rBio2.6	1.011
Coif1	0.760	Meyr	1.721	rBio4.4	0.913
Coif2	1.078	Bior1.3	0.511	rBio5.5	0.979
Coif3	1.148	Bior2.4	0.997	rBio6.8	1.435

### 10.3.2 Evaluation Using Complex-Valued Wavelets

The criteria for choosing an appropriate complex-valued wavelet can be evaluated by applying the continuous wavelet transform to the test signal. The scale whose

**Table 10.8** Relative entropy of the extracted signal: real-valued wavelets

Base wavelet	Relative entropy	Base wavelet	Relative entropy	Base wavelet	Relative entropy
Haar	0.4851	Coif4	0.163	Bior2.6	1.155
Db2	1.09	Coif5	0.111	Bior4.4	0.564
Db4	0.506	Sym2	1.091	Bior5.5	0.423
Db6	0.290	Sym3	0.764	Bior6.8	0.371
Db8	0.194	Sym4	0.507	rBio1.3	0.240
Db10	0.125	Sym6	0.281	rBio2.4	0.182
DB20	0.022	Sym8	0.194	rBio2.6	1.146
Coif1	1.149	Meyr	0.002	rBio4.4	0.985
Coif2	0.435	Bior1.3	1.155	rBio5.5	0.515
Coif3	0.266	Bior2.4	0.564	rBio6.8	0.817

**Table 10.9** Information value of the extracted signal: real-valued wavelets

Base wavelet	Information value	Base wavelet	Information value	Base wavelet	Information value
Haar	0.296	Coif4	3.226	Bior2.6	0.773
Db2	0.232	Coif5	5.155	Bior4.4	1.054
Db4	0.679	Sym2	0.232	Bior5.5	1.550
Db6	1.139	Sym3	0.471	Bior6.8	2.915
Db8	2.146	Sym4	0.858	rBio1.3	0.187
Db10	4.348	Sym6	1.221	rBio2.4	0.292
DB20	40	Sym8	2.347	rBio2.6	0.461
Coif1	0.339	Meyr	1,000	rBio4.4	0.371
Coif2	0.594	Bior1.3	0.082	rBio5.5	0.537
Coif3	1.495	Bior2.4	0.633	rBio6.8	1.534

**Table 10.10** Energy extracted from the test signal: complex-valued wavelets

Base wavelet	Energy (J)
Morlet wavelet	96.243
Gaussian wavelet	58.942
B-Spline wavelet	57.257
Shannon wavelet	14.789
Harmonic wavelet	15.835

corresponding center frequency is equal to that of the frequency component of interest (e.g., 48 Hz in the test signal) is chosen for the wavelet transform. In general, the scale of the wavelet,  $s$ , and the corresponding center frequency of the scaled wavelet,  $f_{s-c}$ , are related by (Abry 1997):

$$s = \frac{f_q f_{b-c}}{f_{s-c}} \quad (10.20)$$



where  $f_q$  is the sampling rate,  $f_{b\_c}$  is the center frequency of the base wavelet, and  $f_{s\_c}$  is the center frequency of the scaled wavelet.

Table 10.10 lists the energy values extracted from the test signal, whereas Table 10.11 lists the corresponding Shannon entropy of the extracted signal. It is seen that the maximum energy criterion selected the *complex Morlet* wavelet as the most suited wavelet among the five complex-valued wavelets, while the *minimum Shannon entropy* criterion identified the *Complex Gaussian* wavelet. Such a conflict is resolved by the integrated *energy-to-Shannon entropy* ratio criterion. In Table 10.12, it is shown that the *Complex Morlet* wavelet led to the maximum energy-to-Shannon entropy ratio. As a result, the *Complex Morlet* wavelet is considered the most suited base wavelet.

The information theoretic measures can also be used to evaluate performance for each of the candidate wavelets, as listed in Table 10.13–10.16. It is noted that the *Complex Morlet* wavelet is again identified as the most suited wavelet by using the following three criteria: *minimum joint entropy*, the *minimum condition entropy*,

**Table 10.11** Shannon entropy of the extracted signal: complex-valued wavelets

Base wavelet	Shannon entropy
Morlet wavelet	7.322
Gaussian wavelet	7.290
B-Spline wavelet	7.365
Shannon wavelet	7.690
Harmonic wavelet	7.453

**Table 10.12** Energy-to-Shannon entropy ratio of the extracted signal: complex-valued wavelets

Base wavelet	Energy-to-Shannon entropy ratio
Morlet wavelet	13.143
Gaussian wavelet	8.085
B-Spline wavelet	7.772
Shannon wavelet	1.923
Harmonic wavelet	2.125

**Table 10.13** Joint entropy of the extracted signal: complex-valued wavelets

Base wavelet	Joint entropy
Morlet wavelet	3.121
Gaussian wavelet	3.280
B-Spline wavelet	3.248
Shannon wavelet	4.167
Harmonic wavelet	3.639

**Table 10.14** Condition entropy of the extracted signal: complex-valued wavelets

Base wavelet	Condition entropy
Morlet wavelet	0.575
Gaussian wavelet	0.734
B-Spline wavelet	0.702
Shannon wavelet	1.614
Harmonic wavelet	1.093

**Table 10.15** Mutual information of the extracted signal: complex-valued wavelets

Base wavelet	Mutual information
Morlet wavelet	1.665
Gaussian wavelet	1.808
B-Spline wavelet	1.705
Shannon wavelet	1.296
Harmonic wavelet	1.551

**Table 10.16** Relative entropy of the extracted signal: complex-valued wavelets

Base wavelet	Relative entropy
Morlet wavelet	0.009
Gaussian wavelet	0.060
B-Spline wavelet	0.011
Shannon wavelet	0.214
Harmonic wavelet	0.027

**Table 10.17** Information value of the extracted signal: complex-valued wavelets

Base wavelet	Information value
Morlet wavelet	111.111
Gaussian wavelet	12.346
B-Spline wavelet	66.667
Shannon wavelet	0.864
Harmonic wavelet	13.889

and the *minimum relative entropy*. However, when the *maximum mutual information* criterion is applied, the *Complex Gaussian* wavelet is identified as the winner. We apply once again the comprehensive criterion “maximum information,” and the conflict is successfully resolved. As shown in Table 10.17, a maximum information value is obtained when the complex Morlet wavelet is chosen as the base wavelet.

The reason why the *Complex Morlet* wavelet is the most suited base wavelet for analyzing the Gaussian-modulated sinusoidal signal can be explained from a physical point of view by comparing their corresponding analytical expressions, as shown in (10.18) and (10.21) below:

$$\psi_M(t) = \frac{1}{\sqrt{\pi f_b}} e^{j2\pi f_c t} e^{-\frac{t^2}{f_b}} \quad (10.21)$$

Tuning the bandwidth  $f_b$  and center frequency  $f_c$  of the Complex Morlet wavelet, the scaled Complex Morlet wavelet can be expressed as:

$$\psi(t) = \sqrt{\frac{120}{\pi}} e^{j2\pi 48t} e^{-120t^2} \quad (10.22)$$

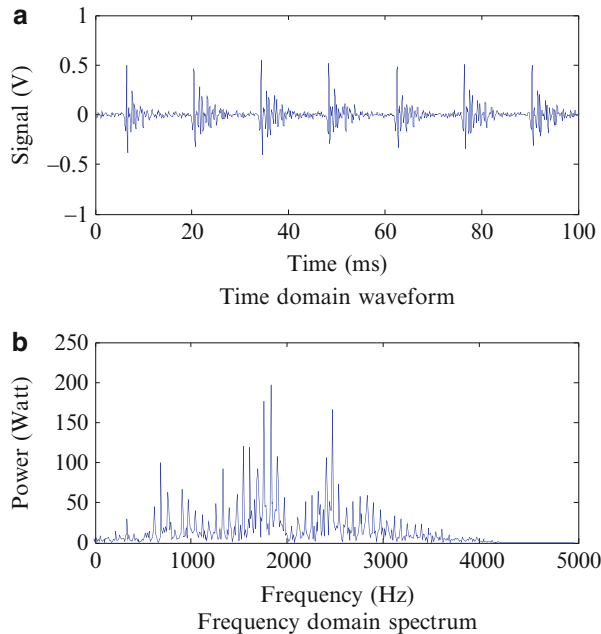
Equation (10.22) illustrates a perfect match of the scaled Complex Morlet wavelet to the Gaussian-modulated sinusoidal signal given in (10.18). As a result, its

wavelet coefficients best represent the test signal, which is why this wavelet has extracted the maximum amount of energy from the test signal.

In summary, using the Gaussian-modulated sinusoidal test signal, we have demonstrated how to systematically choose a base wavelet from a number of candidates by using the various quantitative measures. The two comprehensive criteria, i.e., *energy-to-Shannon entropy* measure and *maximum information* measure, have shown to be effective in choosing the most suited base wavelet for decomposing vibration signals for machine condition monitoring and health diagnosis.

### 10.4 Base Wavelet Selection for Bearing Vibration Signal

We now demonstrate how the two comprehensive wavelet selection criteria, *maximum energy-to-Shannon entropy* ratio and *maximum information* measure, have been applied to selecting base wavelet for bearing vibration signals. Figure 10.4a illustrates the waveform of a vibration signal measured from a ball bearing that contains a localized defect on its outer raceway. The sampling rate is 10,000 Hz. The spectrum in Fig. 10.4b indicates a major peak frequency component at 1,840 Hz. This component is used as the reference base for determining the decomposition



**Fig. 10.4** Bearing vibration signal and its corresponding spectrum. (a) Time domain waveform and (b) frequency domain spectrum

level (for DWT) as well as for the scale selection (when performing CWT). The two criteria have been applied to evaluating real-valued and complex-valued wavelets, respectively.

The real-valued wavelets are evaluated first. The decomposition level of the DWT is chosen to be 2, which corresponds to scale 4 ( $s = 2^2$ ). This scale covers the frequency range from 1,250 to 2,500 Hz, within which the major peak frequency component is located (at 1,840 Hz). After calculation of the energy and Shannon entropy values of the bearing vibration signal by each of the candidate wavelets, the *energy-to-Shannon entropy* ratio is calculated for each wavelet, and the results are listed in Table 10.18. On the basis of the *maximum energy-to-Shannon entropy* ratio criterion, the *reverse Biorthogonal* wavelet 5.5 (denoted as *rBio5.5*) was considered as the most appropriate wavelet for analyzing the bearing vibration signal.

Various similarity measures, including joint entropy, condition entropy, relative entropy, and mutual information, have also been calculated to evaluate the candidate base wavelets. By integrating these similarity measures into the *maximum*

**Table 10.18** Energy-to-Shannon entropy ratio of the extracted bearing vibration signal: real-valued wavelets

Base wavelet	Energy-to-Shannon entropy ratio	Base wavelet	Energy-to-Shannon entropy ratio	Base wavelet	Energy-to-Shannon entropy ratio
Haar	56.279	Coif4	75.980	Bior2.6	69.647
Db2	80.793	Coif5	76.473	Bior4.4	69.864
Db4	104.750	Sym2	80.120	Bior5.5	91.454
Db6	71.343	Sym3	73.969	Bior6.8	77.721
Db8	74.153	Sym4	59.229	rBio1.3	43.843
Db10	93.488	Sym6	77.946	rBio2.4	69.435
DB20	85.949	Sym8	68.515	rBio2.6	70.795
Coif1	66.550	Meyr	77.757	rBio4.4	76.204
Coif2	72.738	Bior1.3	39.720	rBio5.5	109.920
Coif3	75.050	Bior2.4	63.477	rBio6.8	78.777

**Table 10.19** Information value of the extracted bearing vibration signal: real-valued wavelets

Base wavelet	Information value	Base wavelet	Information value	Base wavelet	Information value
Haar	0.106	Coif4	0.143	Bior2.6	0.180
Db2	0.223	Coif5	0.142	Bior4.4	0.181
Db4	0.143	Sym2	0.223	Bior5.5	0.219
Db6	0.127	Sym3	0.180	Bior6.8	0.167
Db8	0.116	Sym4	0.108	rBio1.3	0.105
Db10	0.136	Sym6	0.171	rBio2.4	0.162
DB20	0.129	Sym8	0.122	rBio2.6	0.169
Coif1	0.173	Meyr	0.108	rBio4.4	0.140
Coif2	0.169	Bior1.3	0.101	rBio5.5	0.242
Coif3	0.124	Bior2.4	0.179	rBio6.8	0.158

**Table 10.20** Energy-to-Shannon entropy ratio of the extracted bearing vibration signal: complex-valued wavelets

Base wavelet	Energy-to-Shannon entropy ratio
Morlet wavelet	60.765
Gaussian wavelet	56.044
B-Spline wavelet	35.051
Shannon wavelet	12.476
Harmonic wavelet	14.504

**Table 10.21** Information value of the extracted bearing vibration signal: complex-valued wavelets

Base wavelet	Information value
Morlet wavelet	0.189
Gaussian wavelet	0.068
B-Spline wavelet	0.105
Shannon wavelet	0.017
Harmonic wavelet	0.091

*information* criterion, it is found that the *reverse Biorthogonal* wavelet 5.5 is the most suited wavelet for analyzing the bearing vibration signal. Details of the results are listed in Table 10.19.

Continuous wavelet transform is also applied to analyze the bearing signals, using the five commonly seen complex-valued wavelets. The energy and Shannon entropy values of the bearing vibration signal extracted by each wavelet are first calculated, and their corresponding *energy-to-Shannon* entropy ratios are then determined. As listed in Table 10.20, the *Complex Morlet* wavelet indicates the maximum *energy-to-Shannon entropy* ratio, thus is considered the most appropriate base wavelet for bearing signal analysis.

In addition, the value of the information measure of the bearing vibration signal extracted by each candidate wavelet is calculated, and the result is shown in Table 10.21. Based on the *maximum information* criterion, the *Complex Morlet* wavelet is again identified as the most suited wavelet, since it demonstrates the highest information value compared to other four candidate wavelets.

## 10.5 Summary

Using a number of quantitative measures, we presented a systematic approach in selecting a base wavelet that is best suited for analyzing nonstationary signals, typically seen in manufacturing. These measures are examined from two difference aspects: (1) their corresponding wavelet coefficients (i.e., the energy and Shannon entropy measures) and (2) the relationship between the signal being analyzed and the

coefficients of the base wavelet used for the analysis (i.e., joint entropy, condition entropy, mutual information, and relative entropy). Based on these measures, two comprehensive base wavelet selection criteria (i.e., the *maximum energy-to-Shannon entropy ratio* and the *maximum information measure*) are identified as the quantitative measure for determining the best suited wavelet. Both numerical study and experimental data analysis have shown that these two criteria provide quantitative guidance to base wavelet selection for effective signal analysis.

## 10.6 References

- Abi-Abdallah D, Chauvet E, Bouchet-Fakri L, Bataillard A, Briguet A, Fokapu O (2006) Reference signal extraction from corrupted ECG using wavelet decomposition for MRI sequence triggering: application to small animals. *BioMed Eng Online* 5(11):1–12
- Abry P (1997) Wavelet and turbulence – multi-resolutions, algorithms of decomposition, invariance of scale and signals of pressure. Diderot Editeur, Paris
- Abu-Mahfouz I (2005) Drill flank wear estimation using supervised vector quantization neural networks. *Neural Comput Appl* 14(3):167–175
- Ahuja N, Lertrattanapanich S, Bose NK (2005) Properties determining choice of mother wavelet. *IEEE Proc Vis Image Signal Process* 152:659–664
- Arafat S, Skubic M, Keegan K (2003) Combined uncertainty model for best wavelet selection. In: *The IEEE international conference on fuzzy systems*, pp 1195–1199
- Bedekar D, Nair A, Vince DG (2005) Choosing the optimal mother wavelet for decomposition of radio-frequency intravascular ultrasound data for characterization of atherosclerotic plaque lesions. *Proc SPIE* 5750:490–502
- Bhatia P, Boudy J, Andreão RV (2006) Wavelet transformation and pre-selection of mother wavelet for ECG signal processing. In: *Proceedings of the 24th IASTED international multi-conference: biomedical engineering*, Innsbruck, Austria, 15–17 February, pp 390–395
- Bibian S, Zikov T, Dumont GA, Ries CR, Puil E, Ahmadi H, Huzmenzan M, Macleod BA (2001) Estimation of an anesthetic depth using wavelet analysis of electroencephalogram. In: *23rd International conference of the IEEE engineering in medicine and biology society*, Istanbul, Turkey, October
- Bradley AP, Wilson WJ (2004) On wavelet analysis of auditory evoked potentials. *Clin Neurophysiol* 115:1114–1128
- Cover TM, Thomas JA (1991) *Elements of information theory*. Wiley, New York
- Daubechies I (1992) *Ten lectures on wavelets*. SIAM, Philadelphia, PA.
- Emlen JM (1973) *Ecology: an evolutionary approach*. Addison-Wesley, Reading, MA
- Flanders M (2002) Choosing a wavelet for single-trial EMG. *J Neurosci Methods* 116:165–177
- Fokapu O, Abi-Abdallah D, Briguet A (2005) Extracting a reference signal for cardia MRI gating: experimental study for wavelet functions choice. *Proceedings of 12th international workshop on systems, signals & image processing*, Chalkida, Greece, pp 419–423
- Fu S, Muralikrishnan B, Raja J (2003) Engineering surface analysis with different wavelet bases. *ASME J Manuf Sci Eng* 125:844–852
- Goel P, Vidakovic B (1995) Wavelet transformations as diversity enhancers. *Proc SPIE Int Soc Opt Eng*, 2569:845–857.
- Ho D, Randall RB (2000) Optimization of bearing diagnostic techniques using simulated and actual bearing fault signals. *Mech Syst Signal Process* 14(5):763–788
- Katul G, Vidakovic B (1996) The partitioning of attached and detached eddy motion in the atmospheric surface layer using Lorentz wavelet filtering. *Bound Layer Meteorol* 77(2):153–172

- Li X, Tso SK, Wang J (2000) Real-time tool condition monitoring using wavelet transforms and fuzzy techniques. *IEEE Trans Syst Man Cybern C Appl Rev* 30(3):352–357
- Ma X, Zhou C, Kemp IJ (2002a) Automated wavelet selection and thresholding for PD detection. *IEEE Electr Insul Mag* 18(2):37–45
- Ma X, Zhou C, Kemp IJ (2002b) Interpretation of wavelet analysis and its application in partial discharge detection. *IEEE Trans Dielectr Electr Insul* 9(3):446–457
- Marshall AW, Olkin I (1979) *Inequalities: theory of majorization and its application*. Academic, New York
- Mendenhall W, Sincich TL (1995) *Statistics for engineering and the sciences*, 4th edn. Prentice Hall, Englewood Cliffs, NJ
- Mojsilović A, Popović MV, Rackov DM (2000) On the selection of an optimal wavelet basis for texture characterization. *IEEE Trans Image Process* 9(12):2043–2050
- Safavian LS, Kinsner W, Turanli H (2005) A quantitative comparison of different mother wavelets for characterizing transients in power systems. In: *Canadian Conference on Electrical and Computer Engineering*, Saskatoon, Canada, May, pp 1453–1456
- Schukin EL, Zamaraev RU, Schukin LI (2004) The optimization of wavelet transform for the impulsive analysis in vibration signals. *Mech Syst Signal Process* 18(6):1315–1333
- Shao Y, Nezu K (2004) Extracting symptoms of bearing faults in the wavelet domain. *Proc Inst Mech Eng I J Syst Control Eng* 218(1):39–51
- Singh BN, Tiwari AK (2006) Optimal selection of wavelet basis function applied to ECG signal denoising. *Digit Signal Process* 16:275–287
- Tsui PPC, Basir OA (2006) Wavelet basis selection and feature extraction for shift invariant ultrasound foreign body classification. *Ultrasonics* 45:1–14
- Wang S, Liu X, Yianni J, Aziz TZ, Stein JF (2004) Extracting burst and tonic components from surface electromyograms in dystonia using adaptive wavelet shrinkage. *J Neurosci Methods* 139:174–184
- Yang L, Judd MD, Bennoch CJ (2004) Denoising UHF signal for PD detection in transformers based on wavelet technique. *IEEE conference on electrical insulation and dielectric phenomena*, Boulder, CO, October 17–20, pp 166–169
- Yang WX, Ren XM (2004) Detecting impulses in mechanical signals by wavelets. *EURASIP J Appl Signal Process* 8:1156–1162
- Zhang L, Bao P, Wu X (2005) Multiscale LMMSE-based image denoising with optimal wavelet selection. *IEEE Trans Circuits Syst Video Technol* 15(4):469–481