Chapter 4
Discrete Wavelet Transform

According to the definition of the continuous wavelet transform (CWT) given in (3.7), Chap. 3, the scale parameter s and translation parameter τ can be varied continuously. As a result, performing the CWT on a signal will lead to the generation of redundant information. Although the redundancy is useful in some applications, such as signal denoising and feature extraction where desired performance is achieved at the cost of increased computational time and memory size, other applications may need to emphasize reduced computational time and data size, for example, in image compression and numerical computation. Such requirements illustrate the need for reducing redundancy in the wavelet coefficients among different scales as much as possible, while at the same time, avoiding sacrificing the information contained in the original signal. This can be achieved by parameter discretization, as described in the following section.

4.1 Discretization of Scale and Translation Parameters

The approach to reducing redundancy is to use discrete values of scale and translation parameters. A natural way to implement this is to use a logarithmic discretization of the scale s and then link it to step size taken between the values of translation parameter τ. This type of discretization is expressed as

$$\begin{cases} s = s_0^j \\ \tau = k\tau_0 s_0^j \end{cases} \quad s_0 < 1,\ \tau_0 \neq 0,\ j \in Z,\ k \in Z \tag{4.1}$$

where the symbol Z denotes an integer. The corresponding family of the base wavelet is then expressed as

$$\psi_{j,k}(t) = \frac{1}{\sqrt{s_0^j}} \psi\left(\frac{t - k\tau_0 s_0^j}{s_0^j}\right) \tag{4.2}$$

R.X. Gao and R. Yan, *Wavelets: Theory and Applications for Manufacturing*,
DOI 10.1007/978-1-4419-1545-0_4, © Springer Science+Business Media, LLC 2011

Generally, the values of $s_0 = 2$ and $\tau_0 = 1$ are adopted (Addison 2002). Consequently, (4.2) is expressed as

$$\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t - k2^j}{2^j}\right) \tag{4.3}$$

As a result, the wavelet transform of a given signal $x(t)$ is obtained as

$$wt(j,k) = \langle x(t), \psi_{j,k}(t)\rangle = \frac{1}{\sqrt{2^j}}\int_{-\infty}^{\infty} x(t)\psi^*\left(\frac{t - k2^j}{2^j}\right)\, dt \tag{4.4}$$

where the symbol $\langle\cdot\rangle$ denotes inner product operation. Equation (4.4) poses the following two questions:

1. Can the results of the discretized wavelet transform represent the entire information content of the signal $x(t)$? In other words, can the wavelet coefficients obtained as a result of the wavelet transform be used to *perfectly* reconstruct the original signal $x(t)$?
2. Can any signal $x(t)$ be expressed as the summation of $\psi_{j,k}(t)$, in the form of the following equation?

$$x(t) = \sum_{j,k} C_{j,k}\psi_{j,k}(t) \tag{4.5}$$

In (4.5), $C_{j,k}$ represents the coefficient of the discrete wavelet transform (DWT), which corresponds to $wt(j, k)$ in (4.4). Finally, if the answer to question (2) is "yes," then how can we calculate the coefficient $C_{j,k}$?

Assume that the answer to question (1) is "yes," and we can select $\psi_{s,\tau}(t)$ and discretize s and τ properly. Then, there must exist a function $\tilde{\psi}_{j,k}(t)$, defined as the dual function of $\psi_{j,k}(t)$, that can be used for reconstructing the signal $x(t)$ described in question (1), as follows:

$$x(t) = \sum_{j,k} \langle x(t), \psi_{j,k}(t)\rangle\tilde{\psi}_{j,k}(t) \tag{4.6}$$

where the term $\tilde{\psi}_{j,k}(t)$ can be obtained by performing the scaling and translation operations on $\tilde{\psi}(t)$ as

$$\tilde{\psi}_{j,k}(t) = \frac{1}{\sqrt{2^j}}\tilde{\psi}\left(\frac{t - k2^j}{2^j}\right) \tag{4.7}$$

On the basis of the above assumption, if there exists another signal $y(t)$, we can obtain the inner product of the signals $x(t)$ and $y(t)$ as shown in (4.8). Note that the symbol * indicates the complex conjugate operator:

$$\begin{aligned}\langle y(t), x(t)\rangle =& \langle x(t), y(t)\rangle^* = \left\langle \sum_{j,k} \langle x(t), \psi_{j,k}(t)\rangle \tilde{\psi}_{j,k}(t), y(t) \right\rangle^* \\ =& \left(\sum_{j,k} \langle x(t), \psi_{j,k}(t)\rangle \langle \tilde{\psi}_{j,k}(t), y(t)\rangle \right)^* \\ =& \sum_{j,k} \langle y(t), \tilde{\psi}_{j,k}(t)\rangle \langle \psi_{j,k}(t), x(t)\rangle \\ =& \left\langle \sum_{j,k} \langle y(t), \tilde{\psi}_{j,k}(t)\rangle \psi_{j,k}(t), x(t) \right\rangle \end{aligned} \tag{4.8}$$

Equation (4.8) implies that

$$y(t) = \sum_{j,k} \langle y(t), \tilde{\psi}_{j,k}(t)\rangle \psi_{j,k}(t) \tag{4.9}$$

which means that the answer to question (2) is also positive. It further implies that the coefficient $C_{j,k}$ can be calculated as

$$C_{j,k} = \langle y(t), \tilde{\psi}_{j,k}(t)\rangle \tag{4.10}$$

Therefore, once question (1) is answered, the answer to question (2) can be readily derived from it. The answer to the question (1) can be presented in mathematical terms as follows.

If a set of wavelet coefficients $\langle x(t), \psi_{j,k}(t)\rangle$ exists that describes complete information of the signal $x(t)$, then the following statements must hold:

1. When $x_1(t) = x_2(t)$, the inner product of $x_1(t)$ and the scaled and translated wavelet $\psi_{j,k}(t)$ can be expressed as

$$\langle x_1(t), \psi_{j,k}(t)\rangle = \langle x_2(t), \psi_{j,k}(t)\rangle \tag{4.11}$$

2. For $x(t) = 0$, we have

$$\langle x(t), \psi_{j,k}(t)\rangle = 0 \tag{4.12}$$

3. When $x_1(t)$ is very close to $x_2(t)$, the corresponding wavelet coefficients $\langle x_1(t), \psi_{j,k}(t)\rangle$ must be close to $\langle x_2(t), \psi_{j,k}(t)\rangle$. In other words, if $\|x_1(t) - x_2(t)\|$ is very small, then $\sum_{j,k} \left|\langle x_1(t), \psi_{j,k}(t)\rangle - \langle x_2(t), \psi_{j,k}(t)\rangle\right|^2$ must be very small, too. Mathematically, this can be expressed as

$$\sum_{j,k} \left|\langle x_1(t), \psi_{j,k}(t)\rangle - \langle x_2(t), \psi_{j,k}(t)\rangle\right|^2 \le B\|x_1(t) - x_2(t)\|^2, \quad B \in R^+ \tag{4.13}$$

that is,

$$\sum_{j,k} \left|\langle x(t), \psi_{j,k}(t)\rangle\right|^2 \leq B\|x(t)\|^2 \tag{4.14}$$

In (4.13), the symbol R^+ denotes the set of positive real numbers, and B is a positive real number.

Furthermore, if we want to reconstruct $x(t)$ from the wavelet coefficient $\langle x(t), \psi_{j,k}(t)\rangle$, the following condition must hold:

When $\langle x_1(t), \psi_{j,k}(t)\rangle$ is very close to $\langle x_2(t), \psi_{j,k}(t)\rangle$, $x_1(t)$ must be very close to $x_2(t)$, too, which leads to

$$A\|x(t)\|^2 \leq \sum_{j,k} \left|\langle x(t), \psi_{j,k}(t)\rangle\right|^2, \quad A \in R^+ \tag{4.15}$$

where A is a positive real number.

Combining (4.15) with (4.14), we obtain the following equation:

$$A\|x(t)\|^2 \leq \sum_{j,k} \left|\langle x(t), \psi_{j,k}(t)\rangle\right|^2 \leq B\|x(t)\|^2, \quad A, B \in R^+ \tag{4.16}$$

This ensures that the DWT of a signal $x(t)$ can be obtained. Equation (4.16) is called a *wavelet frame* (Addison 2002). The values of the wavelet frame bounds, A and B, depending on both the scale parameter s and the translation parameter τ that are chosen for analysis and the base wavelet function used (Daubechies 1992). Particularly, if $A = B$, the wavelet frame is known as a *tight* frame. In such a case, the signal $x(t)$ can be reconstructed through the inverse discretized wavelet transform as

$$x(t) = \frac{1}{A} \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} wt(j,k)\psi_{j,k}(t) \tag{4.17}$$

If $A \neq B$, but the difference between A and B is not too large (Addison 2002), the signal $x(t)$ can still be reconstructed as

$$x'(t) = \frac{2}{A+B} \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} wt(j,k)\psi_{j,k}(t) \tag{4.18}$$

The difference between $x(t)$ and $x'(t)$ is determined by the values of A and B, and becomes small in practice when the ratio of B/A is approaching the value of one.

4.2 Multiresolution Analysis and Orthogonal Wavelet Transform

Of the various forms of wavelet discretization, the *dyadic* discretization with $s_0 = 2$ and $\tau_0 = 1$ has been widely used, as shown in (4.3). This is because it allows the selection of the base wavelet to be made in such a way that its corresponding family set $\psi_{j,k}(t)$ constitutes an *orthogonal* basis within the tight wavelet frame, characterized by $A = 1$. To construct a base wavelet having the characteristics of orthogonality, the multiresolution analysis (MRA) is presented here as the theoretical foundation.

4.2.1 Multiresolution Analysis

The concept of MRA was formed when Mallat was working on image processing in the 1980s (Mallat 1989a, b). At that time, the idea of studying images simultaneously at different scales had been popular for years already (Witkin 1983; Burt and Adelson 1983). This provided the background for using orthogonal wavelet bases as a tool to describe the information contained in the image, from coarse approximation to high-resolution approximation, and led to the formulation of MRA (Mallat 1989a, b). Theoretically, a MRA of the space $L^2(R)$ consists of a sequence of successive approximation subspaces $\{V_j, j \in Z\}$ that satisfies the following properties:

1. Monotonicity, that is, $\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots$ (where the symbol $\subset$ denotes a subset operator). This means that the subspace $\{V_j, j \in Z\}$ holds the successive inclusion relationship.
2. Completeness, that is, $\bigcap_{j \in Z} V_j = \{0\}$; $\bigcup_{j \in Z} V_j = L^2(R)$, where $\cap$ denotes the intersect operator, and $\cup$ denotes the union operator. This property indicates that all the subspaces together form a complete $L^2(R)$.
3. Dilation regularity, that is, $x(t) \in V_j \Leftrightarrow x(2^j t) \in V_0$, where $\Leftrightarrow$ denotes "if and only if," and $\in$ denotes "is an element of." The term $j \in Z$ indicates the multiresolution aspect of the subspaces $\{V_j, j \in Z\}$.
4. Translation invariance, that is, $x(t) \in V_0 \Rightarrow x(t-n) \in V_0$, for all $n \in Z$ (with $\Rightarrow$ denotes "imply").
5. Existence of orthogonal basis: there exists a function $\phi(t) \in V_0$, whose corresponding closed subspaces $\{\phi(t-n)\}_{n \in Z}$ form an orthogonal basis of the zero-scale space V_0; that is, $\int_R \phi(t-n)\phi(t-m)\,dt = \delta_{m,n}$.

The function $\phi(t)$ is the *scale* function, whose translated version $\phi_k(t) = \phi(t-k)$ satisfies the condition of $\langle \phi_k(t), \phi_{k'}(t) \rangle = \delta_{k,k'}(k, k' \in Z)$. The zero-scale space V_0 is composed of a set of closed subspaces, formed by $\phi_k(t)$ and is denoted as $V_0 = \overline{\operatorname{span}_k \{\phi(t-k)\}}$.

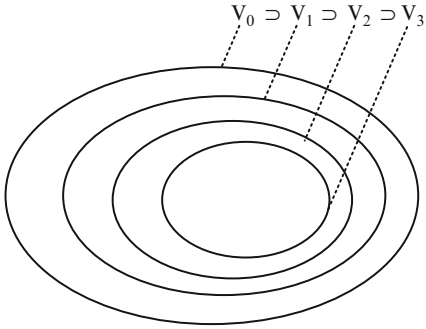


Fig. 4.1 Inclusion relationship among closed subspaces $\{V_j, j \in Z\}$

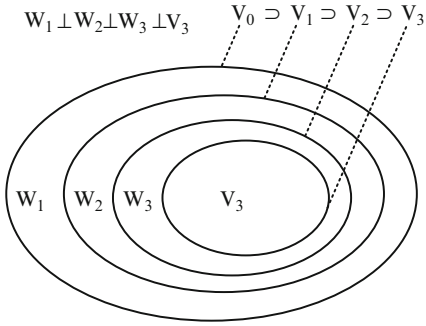


Fig. 4.2 Illustration of wavelet subspaces

From the above description, we know that all the closed subspaces $\{V_j, j \in Z\}$ are formed from the same scale function $\phi(t)$ with different translation values, and the relationship among all the subspaces is illustrated in Fig. 4.1. It can be seen that the closed subspaces $\{V_j, j \in Z\}$ hold the inclusion relationship, and they are not orthogonal. As a result, the scale function family $\phi_{j,k}(t) = 2^{-j/2}\phi(2^{-j}t - k)$ does not hold the orthogonal property; that is, $\{\phi_{j,k}(t)\}_{j \in Z, k \in Z}$ cannot be used as an orthogonal basis in $L^2(R)$ space.

To find the orthogonal bases in the $L^2(R)$ space, we can define W_j $(j \in Z)$ as the orthogonal complement of V_j in V_{j-1}, as illustrated in Fig. 4.2.

We can then write as follows:

$$V_{j-1} = V_j \oplus W_j \tag{4.19}$$

and

$$W_j \perp W_{j'}, \quad \text{for } j \neq j' \tag{4.20}$$

where the symbol $\oplus$ denotes direct summation operator, and $\perp$ denotes the orthogonal operator.

It follows that, for $j < J$, we can have the following relationship:

$$V_j = V_J \oplus \overset{J-j-1}{\underset{k=0}{\oplus}} W_{J-k} \tag{4.21}$$

where all the subspaces W_j ($j \in Z$) are orthogonal, and they form the $L^2(R)$ space as

$$L^2(R) = \underset{j \in Z}{\oplus} W_j \tag{4.22}$$

Furthermore, the W_j spaces inherit the scaling property from the V_j (Daubechies 1992); that is,

$$x(t) \in W_0 \Leftrightarrow x(2^{-j}t) \in W_j \tag{4.23}$$

Therefore, if $\{\psi_{0,k}\, k \in Z\}$ is a set of orthogonal bases in W_0 space, then according to (4.23), for the scale $j \in Z$, $\{\psi_{j,k} = 2^{-i/2}\psi(2^{-j}t - k); \quad k \in Z\}$ is a collection of orthogonal bases in the W_j space. Accordingly, the entire collection of $\{\psi_{j,k}; \quad j \in Z,\ k \in Z\}$ forms the sets of orthogonal bases in $L^2(R)$ space, and we call the function $\psi(t)$ the wavelet function, and the W_j space in (4.23) denotes the wavelet space in scale j.

4.2.2 *Orthogonal Wavelet Transform*

From the definition of the MRA, we know that

$$V_0 = V_1 \oplus W_1 = V_2 \oplus W_2 + W_1 = V_3 \oplus W_3 \oplus W_2 + W_1 = \ldots \tag{4.24}$$

Therefore, for a given signal $x(t) \in V_0$, where V_0 is defined as zero-scale space, we can decompose it into two parts (the detailed information in W_1 and the approximate information in V_1). The approximate information in V_1 can then be further decomposed to get the next level of detailed information in W_2 and approximate information in V_2, respectively. Such a decomposition process can be repeated until the designed scale j is reached. This, in a nutshell, is how a DWT is implemented.

Mathematically, we can define $x_a^j(t)$ as the approximate information at scale j after the signal $x(t)$ is projected onto the V_j space:

$$x_a^j(t) = \sum_k a_{j,k}\phi_k(2^{-j}t) = \sum_k a_{j,k}\phi_{j,k}(t), \quad k \in Z \tag{4.25}$$

where

$$a_{j,k} = \langle x(t), \phi_{j,k}(t) \rangle \tag{4.26}$$

are called the *approximate* coefficients.

Similarly, when the signal $x(t)$ is projected onto the W_j space, the detailed information at scale j is obtained as

$$x_{\mathrm{d}}^{j}(t) = \sum_{k} d_{j,k}\psi_k(2^{-j}t) = \sum_{k} d_{j,k}\psi_{j,k}(t), \quad k \in Z \tag{4.27}$$

where

$$d_{j,k} = \langle x(t), \psi_{j,k}(t) \rangle \tag{4.28}$$

are called *detailed* coefficients.

Consequently, when a given signal $x(t) \in L^2(R)$ is decomposed into the set of subspaces,

$$L^2(R) = \sum_{j=-\infty}^{J} W_j \oplus V_J \tag{4.29}$$

with J being any predetermined scale, we will have

$$x(t) = \sum_{j=-\infty}^{J} \sum_{k=-\infty}^{\infty} d_{j,k}\psi_{j,k}(t) + \sum_{k=-\infty}^{\infty} a_{J,k}\phi_{j,k}(t) \tag{4.30}$$

If $J \rightarrow \infty$, (4.30) can be simplified as

$$x(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} d_{j,k}\psi_{j,k}(t) \tag{4.31}$$

Equation (4.31) is equivalent to (4.17) when $A = B = 1$. We know that in such cases the wavelet bases are orthogonal (Daubechies 1992). As a result, (4.30) and (4.31) express the inverse orthogonal wavelet transform, and (4. 26) and (4.28) express the orthogonal wavelet transform. From the above description, we see that the idea of orthogonal wavelet transform and MRA has followed the same path; thus, MRA provides the theoretical basis for orthogonal wavelet transform.

4.3 Dual-Scale Equation and Multiresolution Filters

The inherent relationship between the scale function $\phi(t)$ and wavelet function $\psi(t)$ can be expressed in a dual-scale equation as

$$\phi(t) = \sum_{n} h(n)\phi_{-1,n}(t) = \sqrt{2}\sum_{n} h(n)\phi(2t - n) \tag{4.32}$$

$$\psi(t) = \sum_n g(n)\phi_{-1,n}(t) = \sqrt{2}\sum_n g(n)\phi(2t-n) \tag{4.33}$$

where

$$\begin{cases} h(n) = \langle \phi, \phi_{-1,n} \rangle \\ g(n) = \langle \psi, \phi_{-1,n} \rangle \end{cases} \tag{4.34}$$

It should be noted that the dual-scale relationship only exists between two successive scales j and $j-1$; that is,

$$\phi_{j,0}(t) = \sum_n h(n)\phi_{j-1,n}(t) \tag{4.35}$$

$$\psi_{j,0}(t) = \sum_n g(n)\phi_{j-1,n}(t) \tag{4.36}$$

Furthermore, the coefficients $h(n)$ and $g(n)$ will not change with the scale j. This can be proved as follows:
Proof:

$$\begin{aligned} \langle \phi_{j,0}(t), \phi_{j-1,n}(t) \rangle &= \int_R [2^{-j/2}\phi(2^{-j}t)][2^{-j-1/2}\phi^*(2^{-j+1}t-n)]\,dt \\ &= \sqrt{2}\int \phi(t')\phi^*(2t'-n)\,dt' (\text{let } t' = 2^{-j}t) \\ &= \langle \phi(t), \phi_{-1,n}(t) \rangle = h(n) \end{aligned} \tag{4.37}$$

Similarly, we can prove that $\langle \psi_{j,0}(t), \phi_{j-1,n}(t) \rangle = g(n)$. This means that the coefficients $h(n)$ and $g(n)$ are determined by the scaling function $\phi(t)$ and wavelet function $\psi(t)$, respectively, and are not related to how we choose the scale j. Furthermore, if we perform an integral operation on both sides of (4.35), we obtain the following:

$$\int_R \phi_{j,0}(t)\,dt = \sum_n h(n)\int_R \phi_{j-1,n}(t)\,dt \tag{4.38}$$

As

$$\begin{aligned} \int_R \phi_{j-1,n}(t)dt &= 2^{-\frac{j-1}{2}}\int_R \phi(2^{-j+1}t-n)dt \\ &\overset{t'=2t}{=} \sqrt{2}\int_R 2^{-\frac{j}{2}}\phi\left(2^{-j}t'-n\right)\frac{1}{2}\,dt' \\ &= \frac{1}{\sqrt{2}}\int_R \phi_{j,n}(t)\,dt \\ &= \frac{1}{\sqrt{2}}\int_R \phi_{j,0}(t)\,dt \end{aligned} \tag{4.39}$$

substituting (4.39) in (4.38) leads to

$$\sum_n h(n) = \sqrt{2} \tag{4.40}$$

Similarly, we can perform an integral operation on both sides of (4.36) as

$$\int_R \psi_{j,0}(t)\,dt = \sum_n g(n) \int_R \phi_{j-1,n}(t)\,dt \tag{4.41}$$

Given that $\int_R \psi(t)\,\mathrm{d}t = 0$, (4.41) can be simplified as

$$\sum_n g(n) = 0 \tag{4.42}$$

The coefficients $h(n)$ and $g(n)$ are called a pair of *low-pass* and *high-pass* wavelet filters, which are used to realize the DWT, on the basis of the Mallat algorithm, as described below.

4.4 The Mallat Algorithm

The dual-scale (4.32) can be rewritten as

$$\phi(t) = \sum_n h(n)\sqrt{2}\phi(2t - n) \tag{4.43}$$

Accordingly, the scaled and translation version of $\phi(t)$ can then be expressed as

$$\begin{aligned} \phi(2^{-j}t - k) &= \sum_n h(n)\sqrt{2}\phi(2(2^{-j}t - k) - n) \\ &= \sum_n h(n)\sqrt{2}\phi(2^{-j+1}t - 2k - n) \end{aligned} \tag{4.44}$$

Let $m = 2k + m$; then (4.44) can be rewritten as

$$\phi(2^{-j}t - k) = \sum_n h(m - 2k)\sqrt{2}\phi(2^{-j+1}t - m) \tag{4.45}$$

On the basis of the theory of MRA, we can define the following:

$$V_{j-1} = \overline{\operatorname*{span}_k \{2^{(-j+1)/2}\phi(2^{-j+1}t - k)\}} \tag{4.46}$$

As a result, a given signal $x(t)$ in the V_{j-1} space can be expressed as

$$x(t) = \sum_k a_{j-1,k} 2^{(-j+1)/2} \phi(2^{-j+1}t - k) \tag{4.47}$$

If such a signal is projected (i.e., decomposed) onto the V_j and W_j spaces, the result can be expressed as

$$x(t) = \sum_k a_{j,k} 2^{-j/2} \phi(2^{-j}t - k) + \sum_k d_{j,k} 2^{-j/2} \psi(2^{-j}t - k) \tag{4.48}$$

where $a_{j,k}$ and $d_{j,k}$ are calculated as

$$a_{j,k} = \langle x(t), \phi_{j,k}(t) \rangle = \int_R x(t) 2^{-j/2} \phi^*(2^{-j}t - k)\, dt \tag{4.49}$$

$$d_{j,k} = \langle x(t), \psi_{j,k}(t) \rangle = \int_R x(t) 2^{-j/2} \psi^*(2^{-j}t - k)\, dt \tag{4.50}$$

Substituting (4.45) in(4.49) results in

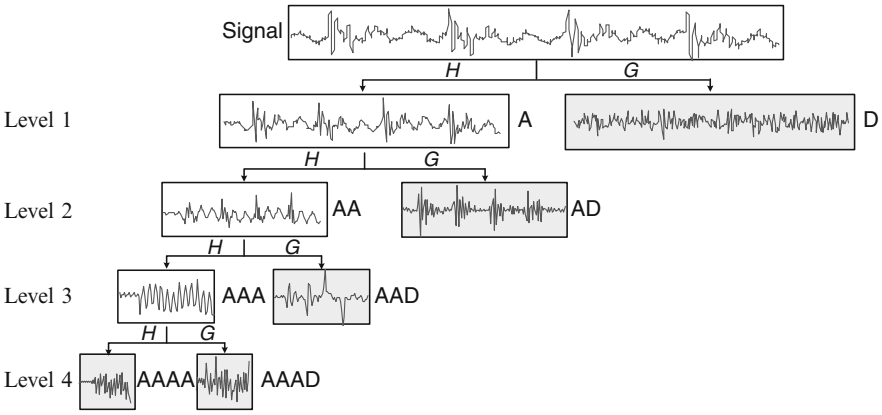
$$\begin{aligned} a_{j,k} &= \sum_m h(m - 2k) \int_R x(t) 2^{(-j+1)/2} \phi^*(2^{-j+1}t - m)\, dt \\ &= \sum_m h(m - 2k) \langle x(t), \phi_{j-1,m} \rangle \\ &= \sum_m h(m - 2k) a_{j-1,m} \end{aligned} \tag{4.51}$$

Similarly, (4.50) can be further rewritten as

$$\begin{aligned} d_{j,k} &= \sum_m g(m - 2k) \int_R x(t) 2^{(-j+1)/2} \psi^*(2^{-j+1}t - m)\, dt \\ &= \sum_m h(m - 2k) \langle x(t), \psi_{j-1,m} \rangle \\ &= \sum_m h(m - 2k) d_{j-1,m} \end{aligned} \tag{4.52}$$

This means that, through such a pair of filters, the signal $x(t)$ is decomposed into low- and high-frequency components, respectively, as (Mallat 1998)

$$\begin{cases} a_{j,k} = \sum_m h(m - 2k) a_{j-1,m} \\ d_{j,k} = \sum_m g(m - 2k) a_{j-1,m} \end{cases} \tag{4.53}$$



Note: H - Low pass filter; G - High pass filter; A - Approximate information; D - Detailed information

Fig. 4.3 Procedure of a four-level signal decomposition using discrete wavelet transform. Note: H low-pass filter, G high-pass filter, A approximate information, D detailed information

In (4.53), $a_{j,k}$ is the *approximate* coefficient, which represents the low-frequency component of the signal, and $d_{j,k}$ is the *detailed* coefficient, which corresponds to the high-frequency component. The approximate coefficients at wavelet decomposition level j are obtained by convolving the approximate coefficients at the previous decomposition level $(j-1)$ with the low-pass filter coefficients. Similarly, the detailed coefficients at wavelet decomposition level j are obtained by convolving the approximate coefficients at the previous decomposition level $(j-1)$ with the high-pass filter coefficients. Such a process represents the idea of Mallat's algorithm to implement the DWT, and is schematically shown in Fig. 4.3.

From Fig. 4.3, we see that a signal is decomposed by a four-level DWT. After passing through the high-pass and low-pass filters on the first level (level 1), the output of the low-pass filter, denoted as the *approximate* coefficients of the level 1, is filtered again by the second-level filter banks. The process repeats itself, and at the end of the fourth level decomposition, the signal is decomposed into five feature groups: one group containing the lowest frequency components, denoted as the *approximate* information and labeled as *AAAA*, and four groups containing progressively higher frequency components, called the *detailed* information and labeled as *AAAD*, *AAD*, *AD*, and *D*. The levels 1–4 correspond to the wavelet scales $2^1 = 2$, $2^2 = 4$, $2^3 = 8$, and $2^4 = 16$, respectively.

4.5 Commonly Used Base Wavelets

This section introduces several commonly used orthogonal wavelets, which can be used as the basis for performing the DWT.

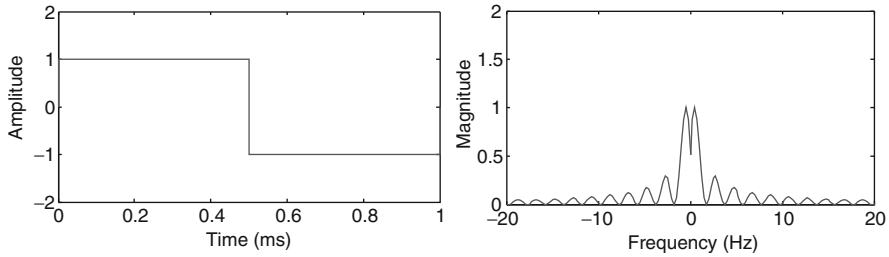


Fig. 4.4 Haar wavelet (*left*) and its magnitude spectrum (*right*)

4.5.1 Haar Wavelet

The Haar wavelet is mathematically defined as (Haar 1910)

$$\psi_{\text{Haar}}(t) = \begin{cases} 1 & 0 \le t < \frac{1}{2} \\ -1 & \frac{1}{2} \le t < 1 \\ 0 & \text{otherwise} \end{cases} \tag{4.54}$$

Its function and magnitude spectrum are illustrated in Fig. 4.4.

The Haar wavelet is orthogonal and symmetric in nature. The property of symmetry ensures that the Haar wavelet has linear phase characteristics, meaning that when a wavelet filtering operation is performed on a signal with this base wavelet, there will be no phase distortion in the filtered signal. Furthermore, it is the simplest base wavelet with the highest time resolution given by a compact support of one as shown in (4.54) (Daubechies 1992). However, the rectangular shape of the Haar wavelet determines its corresponding spectrum with slow decay characteristics, leading to a low frequency resolution. Examples of using the Haar wavelet for manufacturing related work include the stamping process monitoring (Zhou et al. 2006) and fault detection in dry etching process (Kim et al. 2010).

4.5.2 Daubechies Wavelet

The family of the Daubechies wavelets is orthogonal, however, asymmetric, which introduces a large phase distortion. This means that it cannot be used in applications where a signal's phase information needs to be kept. It is also a compactly supported base wavelet with a given support width of $2N$ - 1, in which N is the order of the base wavelet (Daubechies 1992). In theory, N can be up to infinity. In real-world applications, the Daubechies wavelets with order up to 20 have been used. The Daubechies wavelets do not have explicit expression except for the one with $N = 1$, which is actually the Haar wavelet as discussed above. With an increase of the support width (i.e., an increase of the base wavelet order), the Daubechies

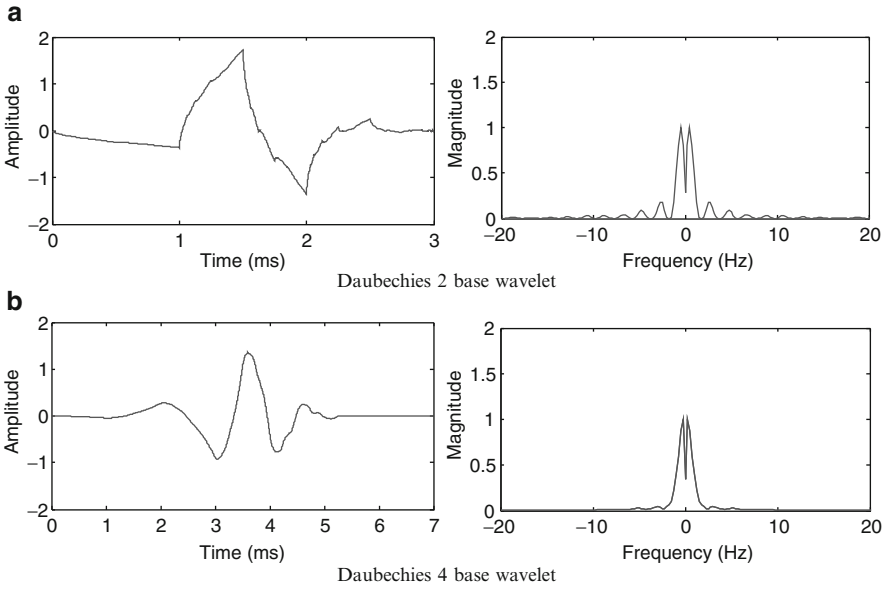


Fig. 4.5 Daubechies wavelet (*left*) and its magnitude spectrum (*right*). (**a**) Daubechies 2 base wavelet and (**b**) Daubechies 4 base wavelet

wavelet becomes increasingly smoother, leading to better frequency localization. Accordingly, the magnitude spectra for each of the Daubechies wavelets decay quickly, as illustrated in Fig. 4.5, where the Daubechies 2 base wavelet and Daubechies 4 base wavelet are used as examples.

The Daubechies wavelets have been widely investigated for fault diagnosis of bearings (Nikolaou and Antoniadis 2002; Lou and Loparo 2004) and automatic gears (Rafiee et al. 2010)

4.5.3 Coiflet Wavelet

The family of the Coiflet wavelets is orthogonal (Daubechies 1992), and near symmetric. This property of near symmetry leads to the near linear phase characteristics of the Coiflet wavelet. They are designed to yield the highest number of vanishing moments (2*N*) for both the base wavelet of the order *N* and the scaling function, for a given support width of *6N* - 1. Figure 4.6 illustrates the sample waveforms of the Coiflet wavelets, with their corresponding magnitude spectra at orders 2 and 4, respectively. The Coiflet wavelet has been used for fault diagnosis of rolling bearings (Sugumaran and Ramachandran 2009).

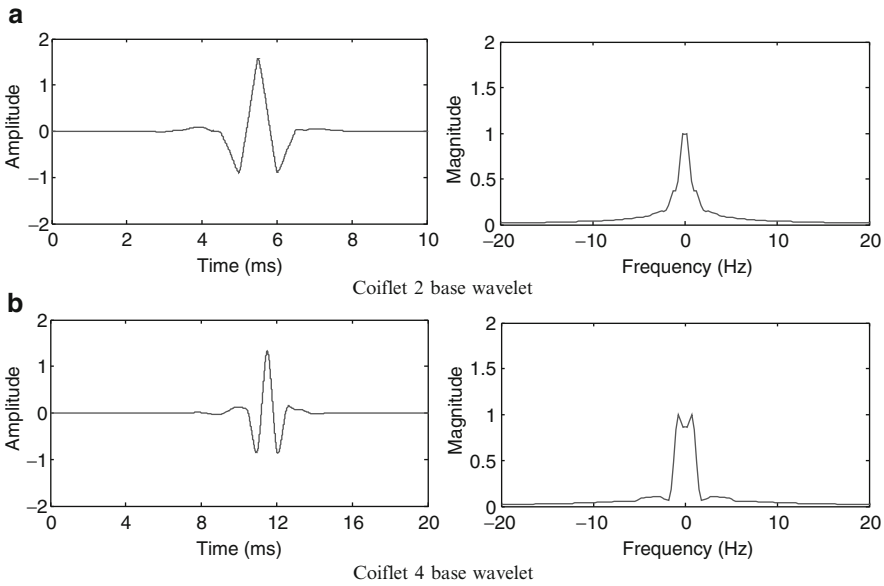


Fig. 4.6 Coiflet wavelet (*left*) and its magnitude spectrum (*right*). (**a**) Coiflet 2 base wavelet and (**b**) Coiflet 4 base wavelet

4.5.4 Symlet Wavelet

Symlet wavelets(Daubechies 1992) are orthogonal and near symmetric. This property ensures minimal phase distortion. A Symlet wavelet of order N has the number of vanishing moments N for a given support width of $2N$ - 1. They are similar to the Daubechies wavelet, except for better symmetry. Waveforms with their corresponding magnitude spectra for the Symlet wavelet at orders 2 and 4 are illustrated in Fig. 4.7a, b, respectively. Examples of using the Symlet wavelet for signal decomposition in manufacturing-related problems include characterization of fabric texture (Shakher et al. 2004) and health monitoring of rolling bearings (Gao and Yan 2006).

4.5.5 Biorthogonal and Reverse Biorthogonal Wavelets

The family of biorthogonal and reverse biorthogonal wavelets (Daubechies 1992) is biorthogonal and symmetric. The property of symmetry ensures that they have linear phase characteristics. This type of base wavelet can be constructed by the spline method (Cohen et al. 1992). Figures 4.8 and 4.9 illustrate sample waveforms with their magnitude spectrum for several biorthogonal and reverse biorthogonal wavelets, respectively. In practice, this group of wavelets has been used for surface profile filtering in manufacturing process monitoring and diagnostics (Fu et al. 2003).

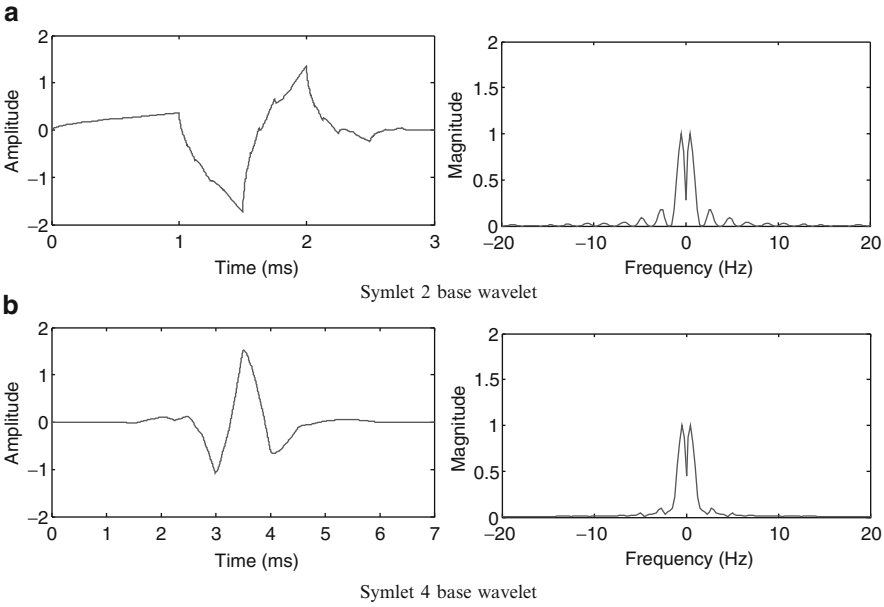


Fig. 4.7 Symlet wavelet (*left*) and its magnitude spectrum (*right*). (**a**) Symlet 2 base wavelet and (**b**) Symlet 4 base wavelet

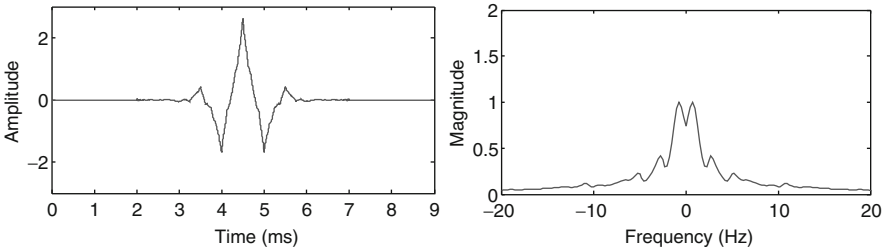


Fig. 4.8 Biorthogonal 2.4 wavelet (*left*) and its magnitude spectrum (*right*)

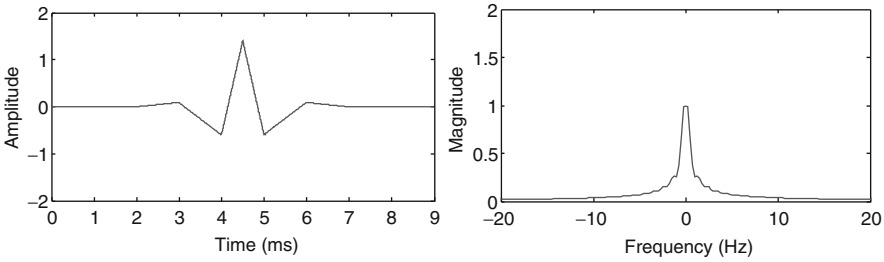


Fig. 4.9 Reverse biorthogonal 2.4 wavelet (*left*) and its magnitude spectrum (*right*)

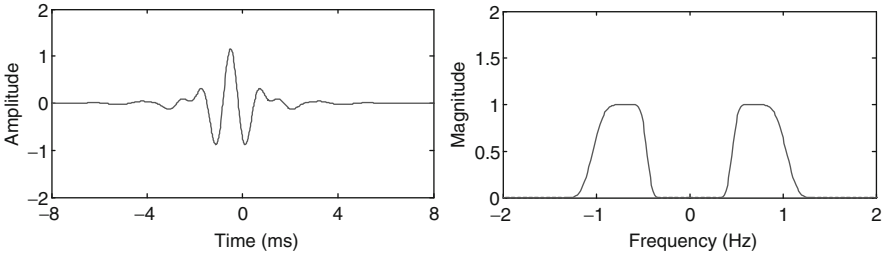


Fig. 4.10 Meyer wavelet (*left*) and its magnitude spectrum (*right*)

4.5.6 Meyer Wavelet

The Meyer wavelet is orthogonal and symmetric. However, it does not have a finite support. The Meyer wavelet has explicit expression and is defined in the frequency domain as follows:

$$\Psi_{Meyer}(f) = \begin{cases} \sqrt{2\pi}\, e^{i\pi f} \sin\left[\frac{\pi}{2} v(3|f|-1)\right] & \frac{1}{3} \le |f| \le \frac{2}{3} \\ \sqrt{2\pi}\, e^{i\pi f} \cos\left[\frac{\pi}{2} v\left(\frac{3}{2}|f|-1\right)\right] & \frac{2}{3} \le |f| \le \frac{4}{3} \\ 0 & |f| \notin \left\langle \frac{1}{3}, \frac{4}{3} \right\rangle \end{cases} \tag{4.55}$$

where $v(\cdot)$ is an auxiliary function, expressed as

$$v(\alpha) = \alpha^4(35 - 84\alpha + 70\alpha^2 - 20\alpha^3), \quad \alpha \in \langle 0, 1 \rangle \tag{4.56}$$

The Meyer wavelet with its magnitude spectrum is illustrated in Fig. 4.10.

Typical applications of Meyer wavelet in manufacturing-related problems include signal denoising and bearing fault diagnosis (Abbasion et al. 2007).

4.6 Application of Discrete Wavelet Transform

One of the most popular applications of the DWT is to remove noise contained in a signal. This is based on the observation that a signal's energy is often distributed over a few wavelet coefficients with high magnitude, while energy of the noise is distributed across most of the wavelet coefficients with low magnitude. A thresholding scheme can therefore be devised to remove the noise. Mathematically, assume a signal with noise contamination expressed as

$$y(t) = x(t) + \sigma e(t) \tag{4.57}$$

where $x(t)$ is the signal, $e(t)$ is a Gaussian white noise $N(0,1)$, and σ represents the noise level. The objective of denoising is to suppress the noise $e(t)$ and to recover the signal $x(t)$. Generally, the denoising procedure consists of three steps:

1. *Signal decomposition*: Choosing a base wavelet and a decomposition level J, and then performing DWT up to level J on the signal.
2. *Detailed coefficients thresholding*: For each decomposition level from 1 to J, selecting a threshold and applying it to the detailed coefficients.
3. *Signal reconstruction*: Performing wavelet reconstruction to obtain denoised signal, based on the original approximate coefficients of level J and the modified detailed coefficients of levels 1 to J.

It should be noted that two thresholding approaches (hard thresholding and soft thresholding) can be used in the denoising process (Donoho 1995; Donoho and Johnstone 1995). Hard thresholding can be described as the process of setting the value of the detailed coefficient $d_{j,k}$ to zero, if its absolute value is lower than the threshold (denoted as *thr*). This is mathematically expressed as

$$\hat{d}_{j,k} = \begin{cases} d_{j,k} & d_{j,k} \geq \text{thr} \\ 0 & d_{j,k} < \text{thr} \end{cases} \tag{4.58}$$

Soft thresholding can be considered as an extension of the hard thresholding, as shown in Fig. 4.11. It sets those detailed coefficients to zero if their absolute values are lower than the threshold, and then shrinks the nonzero coefficients toward zero. Mathematically, this can be expressed as

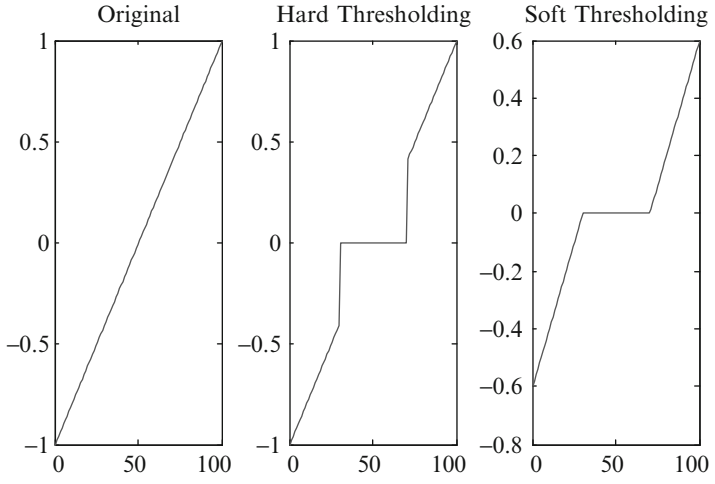


Fig. 4.11 Illustration of hard thresholding and soft thresholding

$$\hat{d}_{j,k} = \begin{cases} \text{sgn}(d_{j,k})(|d_{j,k}| - \text{thr}) & d_{j,k} \geq \text{thr} \\ 0 & d_{j,k} < \text{thr} \end{cases} \tag{4.59}$$

where

$$\text{sgn}(d_{j,k}) = \begin{cases} +1 & d_{j,k} \geq 0 \\ -1 & d_{j,k} < 0 \end{cases} \tag{4.60}$$

As an example, Fig. 4.12a shows a "blocks" test signal, and Fig. 4.12b shows that it is contaminated by a Gaussian white noise to make a signal-to-noise ratio of 4. The signal is decomposed up to level 3, with sym8 wavelet being the base wavelet. After performing soft thresholding to the detailed coefficients at each decomposition level, the signal is reconstructed as shown in Fig. 4.12c. As only a small number of large coefficients characterize the original "blocks" signal, this DWT-based denoising method performs well.

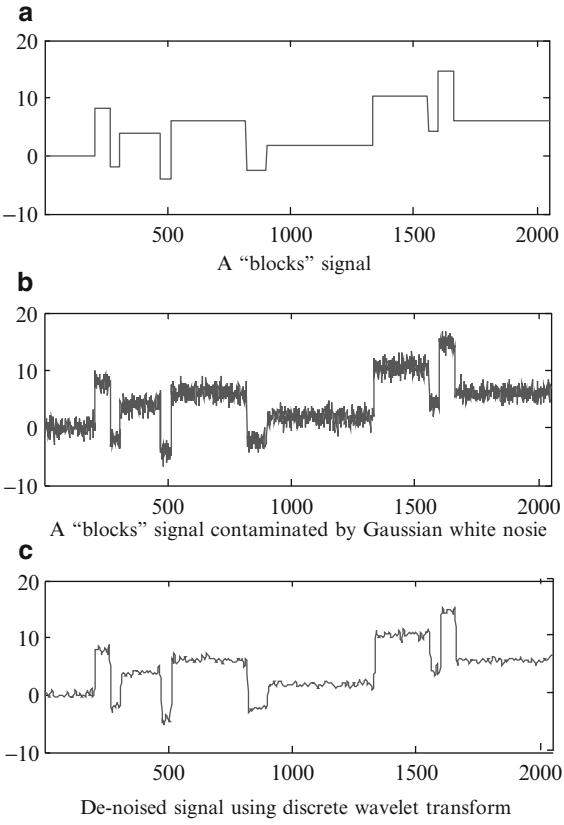


Fig. 4.12 Example of discrete wavelet transform for denoising. (**a**) A "blocks" signal, (**b**) a "blocks" signal contaminated by Gaussian white noise, and (**c**) denoised signal using discrete wavelet transform

4.7 Summary

This chapter begins with a description of the discretization of the scale and translation parameters. The MRA and orthogonal wavelet transform are then introduced in Sect. 4.2. After that, we describe in Sect. 4.3 the dual-scale equation and its associated wavelet filter pair. The Mallat algorithm for implementing the DWT is then discussed in Sect. 4.4, followed by the introduction of some commonly used wavelets in Sect. 4.5. Some typical applications of the DWT are shown in Sect. 4.6.

4.8 References

Abbasion S, Rafsanjani A, Farshidianfar A, Irani N (2007) Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine. Mech Syst Signal Process 21:2933–2945

Addison N (2002) The illustrated wavelet transform handbook. Taylor & Francis, New York

Burt P, Adelson E (1983) The Laplacian pyramid as a compact image code. IEEE Trans Commun 31:482–540

Cohen A, Daubechies I, Feauveau, JC (1992) Biorthogonal bases of compactly supported wavelets. Commun Pure Appl Math 45:485–560

Daubechies I (1992) Ten lectures on wavelets. SIAM, Philadelphia

Donoho DL (1995) De-noising by soft-thresholding. IEEE Trans Inform Theory, 41(3): 613–627

Donoho DL; Johnstone IM (1995) Adapting to unknown smoothness via wavelet shrinkage. J Am Stat Assoc 90(432):1200–1244

Fu S, Muralikrishnan, Raja J (2003) Engineering surface analysis with different wavelet bases. ASME J Manuf Sci Eng 125(6):844–852

Gao R, Yan R (2006) Non-stationary signal processing for bearing health monitoring. Int J Manuf Res 1(1):18–40

Haar A (1910) Zur theorie der orthgonalen funktionensysteme. Math Annalen 69:331–371

Kim JS, Lee JH, Kim JH, Baek J, Kim SS (2010) Fault detection of cycle-based signals using wavelet transform in FAB processes. Int J Precision Eng Manuf 11(2):237–246

Lou X, Loparo KA (2004) Bearing fault diagnosis based on wavelet transform and fuzzy inference. Mech Syst Signal Process 18:1077–1095

Mallat SG (1989a) A theory of multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Machine Intell 11(7) 674–693

Mallat SG (1989b) Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$. Trans Am Math Soc 315:69–87

Mallat SG (1998) A wavelet tour of signal processing. Academic, San Diego, CA

Nikolaou NG, Antoniadis IA (2002) Rolling element bearing fault diagnosis using wavelet packets. NDT&E Int 35:197–205

Rafiee J, Rafiee MA, Tse PW (2010) Application of mother wavelet functions for automatic gear and bearing fault diagnosis. Expert Syst Appl 37:4568–4579

Shakher C, Ishtiaque SM, Singh SK, Zaidi HN (2004) Application of wavelet transform in characterization of fabric texture. J Text Inst 95(1–6):107–120

Sugumaran V, Ramachandran KI (2009) Wavelet selection using decision tree for fault diagnosis of roller bearings. Int J Appl Eng Res 4(2):201–225

Witkin A (1983) Scale space filtering. In: Proceedings of international joint conference on artificial intelligence, Karlsruhe, Germany, pp 1019–1023

Zhou SY, Sun BC, Shi JJ (2006) An SPC monitoring system for cycle-based waveform signals using haar transform. IEEE Trans Automat Sci Eng 3(1):60–72