Introduction to Discrete-Time Wavelet Transform

Selin Aviyente
Department of Electrical and Computer Engineering
Michigan State University

February 9, 2010
Definition of a Wavelet

- A wave is usually defined as an oscillating function of time (such as a sinusoid).
- A wavelet is a 'small wave', which has its energy concentrated in time for analysis of transient, non-stationary or time-varying phenomena.
- We will use wavelet for expanding signals the same way Fourier series used the sinusoids.

\[ f(t) = \sum_l a_l \psi_l(t) \]  

(1)

where \(a_l\) are the expansion coefficients and \(\psi_l(t)\) are the expansion set.

- If \(\psi_l(t)\) form an orthonormal basis for the function space then \(a_l = \langle f(t), \psi_l(t) \rangle\).
Characteristics of a Wavelet System

- For the wavelet expansion, a two parameter system is constructed \( f(t) = \sum_k \sum_j a_{j,k} \psi_{j,k}(t) \).
- \( a_{j,k} \) are called the discrete wavelet transform (DWT), \( \psi_{j,k} \) are the wavelet expansion functions.
- The wavelet expansion set is not unique.
- A wavelet system is a set of building blocks to represent a signal.
- The wavelet expansion gives a time-frequency localization of the signal. Most of the energy of the signal is well represented by a few expansion coefficients.
- The calculation of the coefficients can be done efficiently, \( O(N) \).
All wavelet systems are generated from a single scaling function or wavelet (mother wavelet) by scaling and translation, \( \psi_{j,k}(t) = 2^{j/2}\psi(2^jt - k), j, k \in \mathbb{Z} \).

Multiresolution conditions: If a set of signals can be represented by a weighted sum of \( \psi(t - k) \) then a larger set (including the original) can be represented by a weighted sum \( \psi(2t - k) \).

The lower resolution coefficients can be calculated from the higher resolution coefficients through a filter bank. (efficient computation)

The magnitude of the expansion coefficients drop off rapidly, i.e. only a few coefficients are significant, good for applications like compression and denoising.
The multiresolution formulation requires two basic functions: scaling ($\phi(t)$) and wavelet functions ($\psi(t)$).

The simplest possible orthogonal wavelet system is generated from the Haar scaling function and wavelet.

$$f(t) = \sum_{k} c_k \phi(t - k) + \sum_{j=0}^{\infty} \sum_{k} d_{j,k} \psi(2^j t - k)$$

Haar wavelets form an orthonormal basis for $L_2(\mathbb{R})$: $\int |\psi_{j,k}(t)|^2 dt = 1$ and $< \psi_{j,k}(t), \psi_{j',k'}(t) > = \delta[j - j']\delta[k - k']$.

Well localized in time, as $j \to \infty$, close to delta functions, can detect abrupt changes and transient activity.
The scaling function, $\phi_k(t) = \phi(t - k), k \in \mathbb{Z}$. The subspace of $L_2(\mathbb{R})$ spanned by these functions is defined as $V_0 = \text{Span}_k \{\phi_k(t)\}$. If $f(t) \in V_0$ then $f(t) = \sum_k a_k \phi_k(t)$. 

$V_j = \text{Span}_k \{\phi(2^j t - k)\}$. 

For $j > 0$, the span can be larger since $\phi_{j,k}(t)$ gets narrower and is translated in smaller steps (finer details). For $j < 0$, the span is smaller (coarse information).
Properties of multiresolution systems

A sequence of \( \{ V_j \}_{j \in \mathbb{Z}} \) of closed subspaces of \( L_2(\mathbb{R}) \) is a multiresolution approximation if the following properties are satisfied:

1. \( f(t) \in V_j \) iff \( f(t - 2^{-j}k) \in V_j \).
2. Nested Subspaces: \( V_j \subset V_{j+1} \).
3. \( f(t) \in V_j \) iff \( f(2t) \in V_{j+1} \).
4. \( \lim_{j \to \infty} V_j = L_2(\mathbb{R}) \).
5. \( \lim_{j \to -\infty} V_j = \{0\} \).
6. There exists \( \phi \) such that \( \{ \phi(t - k) \}_{k \in \mathbb{Z}} \) is a basis of \( V_0 \).

By the nesting subspaces property of MRA, if \( \phi(t) \in V_0 \) it is also in \( V_1 \). This means that \( \phi(t) \) can be expressed in terms of a weighted sum of shifted \( \phi(2t) \) as \( \phi(t) = \sum_n h(n) \sqrt{2} \phi(2t - n) \) (dilation or scaling equation).
Scaling Equation

- \( \phi(t) = \sum_n h(n) \sqrt{2} \phi(2t - n) \), the coefficients \( h(n) \) are a sequence of real or complex numbers called the scaling function coefficients (or the scaling filter).

- This recursive equation is fundamental to the theory of the scaling function. Designing a scaling function reduces down to designing a FIR filter.

- Example (Haar): \( \phi(t) = \phi(2t) + \phi(2t - 1) \), which means that \( h(0) = \frac{1}{\sqrt{2}}, h(1) = \frac{1}{\sqrt{2}} \).
Introduction to Discrete-Time Wavelet Transform

Wavelet Function

- Sometimes the important features of a signal can be better described not by using $\phi_{j,k}(t)$ but by defining a different set of functions that span the differences between $V_j$ and $V_{j+1}$. These functions are the wavelets $\psi_{j,k}(t)$.

- $W_j$ is defined as the orthogonal complement of $V_j$ in $V_{j+1}$. This means that all members of $V_j$ are orthogonal to all members of $W_j$.

- $V_{j+1} = V_j \oplus W_j$.

- If $V_j$ is a Hilbert space and subspace of $V_{j+1}$, each vector $v \in V_{j+1}$ can be written uniquely as $v = w + z$, $w \in V_j, z \in V_j^\perp$, that is $V_{j+1} = V_j \oplus V_j^\perp$.

- $\psi_{j,k}(t)$ span $W_j$. 
Properties of the Wavelet Function

1. \( V_1 = V_0 \oplus W_0, L_2 = V_0 \oplus W_0 \oplus W_1 \oplus \ldots \)

2. The scaling function and the wavelet function at the same scale are orthogonal to each other: \( \phi_{j,k}(t), \psi_{j,l}(t) \geq 0 \).

3. \( V_0 = W_{-\infty} \oplus \ldots \oplus W_{-1} \). Similarly, \( L_2 \) can be written as the direct sum of only wavelet subspaces.

4. Since \( \psi(t) \in V_1, \psi(t) = \sum h_1(n)\sqrt{2}\phi(2t - n). \) \( h_1(n) \) is the wavelet filter. For example, for the Haar wavelet \( \psi(t) = \phi(2t) - \phi(2t - 1), h_1(0) = \frac{1}{\sqrt{2}}, h_1(1) = -\frac{1}{\sqrt{2}}. \)
There are infinitely many ways to decompose \( L_2(\mathbb{R}) \) and thus different ways to expand any function.

For example:
\[
L_2(\mathbb{R}) = V_0 \oplus W_0 \oplus W_1 \oplus \ldots
\]

\[
f(t) = \sum_k c_k \phi_k(t) + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} d_{j,k} \psi_{j,k}(t)
\]  \hspace{1cm} (3)

For the case that \( \psi_{j,k}(t) \) are orthonormal and \( \phi_k(t) \) is orthogonal to \( \psi_{j,k}(t) \), the coefficients can be found:

\[
c_k = < f(t), \phi_k(t) > = \int f(t) \phi(t - k) dt
\]

\[
d_{j,k} = < f(t), \psi_{j,k}(t) > = \int f(t) 2^{j/2} \psi(2^j t - k) dt
\]  \hspace{1cm} (4)
In general, \( L_2(\mathbb{R}) = V_{j_0} \oplus W_{j_0} \oplus W_{j_0+1} \oplus \ldots \):

\[
f(t) = \sum \sum_{k} c_{j_0}(k)2^{j_0/2} \phi(2^{j_0}t - k) + \sum_{k} \sum_{j=j_0}^{\infty} d_j(k)2^{j/2} \psi(2^{j/2}t - k)
\]

The choice of \( j_0 \) depends on the signal and sets the coarsest scale whose space is spanned by \( \phi_{j_0,k}(t) \).

The coefficients are called the discrete wavelet transform of \( f(t) \). If certain conditions are satisfied, these coefficients completely describe the original signal.

If the scaling functions and the wavelets form an orthonormal basis, then Parseval’s theorem can be applied:

\[
\int |f(t)|^2 dt = \sum_{k} |c(k)|^2 + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} |d_j(k)|^2
\]
Example: Haar Decomposition

- Let \( f(t) \in V_j \), then \( f(t) = \sum c_k \phi(2^j t - k) \).
- Divide \( f(t) \) into even and odd terms.

\[
f(t) = \sum c_{2k} \phi(2^j t - 2k) + \sum c_{2k+1} \phi(2^j t - 2k - 1) \quad (7)
\]

- We can write:

\[
\phi(2^j t - 2k) = \frac{\psi(2^{j-1} t - k) + \phi(2^{j-1} t - k)}{2}
\]

\[
\phi(2^j t - 2k - 1) = \frac{\phi(2^{j-1} t - k) - \psi(2^{j-1} t - k)}{2} 
\quad (8)
\]
Therefore,

\[ f(t) = \sum c_{2k} \left( \frac{\psi(2^{j-1}t - k) + \phi(2^{j-1}t - k)}{2} \right) \]

\[ + \sum c_{2k+1} \left( \frac{\phi(2^{j-1}t - k) - \psi(2^{j-1}t - k)}{2} \right) \]

\[ = \sum_k \left( \frac{c_{2k} - c_{2k+1}}{2} \right) \psi(2^{j-1}t - k) \]

\[ + \left( \frac{c_{2k} + c_{2k+1}}{2} \right) \phi(2^{j-1}t - k) \]  \hspace{1cm} (9)

\[ f_j(t) = f_{j-1}(t) + w_{j-1}(t), \text{ where} \]

\[ w_{j-1}(t) = \sum_k \left( \frac{c_{2k} - c_{2k+1}}{2} \right) \psi(2^{j-1}t - k) \]

\[ f_{j-1}(t) = \sum_k \left( \frac{c_{2k} + c_{2k+1}}{2} \right) \phi(2^{j-1}t - k) \]  \hspace{1cm} (10)

Coefficients at the lower scale can be found from the coefficients at the higher scale and this can be repeated iteratively.
Steps for Haar Decomposition

- Discretize the signal such that $f_j \in V_j$, $f_j$ is an approximation to $f$.
- $c_k^j = f(k/2^j)$, samples of the signal become the scaling coefficients.
- Decompose $f_j = w_{j-1} + w_{j-2} + \ldots + w_0 + f_0$. 
Steps for Haar Reconstruction

- After decomposing a signal $f$ into its components in terms of $V_0$ and $W_j$, what do we do?
- The answer depends on the goal. If the goal is to filter out noise, then the wavelet coefficients at scales which correspond to the noise can be thrown out. If the goal is compression, the coefficients that are small can be thrown out.
- Given $f(t) = f_0(t) + w_0(t) + w_1(t) + \ldots + w_{j-1}(t)$ where $f_0(t) = \sum_k c_k \phi(t - k)$ and $w_l = \sum_k d_{l,k} \psi(2^l t - k)$, the goal is to write $f(t) = \sum_k c^j_k \phi(2^j t - k)$. 


We can write:

\[
\phi(2^{j-1} t) = \phi(2^j t) + \phi(2^j t - 1) \\
\psi(2^{j-1} t) = \phi(2^j t) - \phi(2^j t - 1)
\]  

(11)

Since

\[
f_0 = \sum_k c_k \phi(t - k) \\
= \sum_k c_k \phi(2t - 2k) + \sum_k c_k \phi(2t - 2k - 1) \\
= \sum_l \hat{c}_l^1 \phi(2t - l)
\]  

(12)

where \( \hat{c}_l^1 = c_k \).

Similarly, \( w_0(t) = \sum_l d_l^1 \phi(2t - l) \) where

\[
d_l^1 = \begin{cases} 
    d_k & l = 2k \\
    -d_k & l = 2k + 1
\end{cases}
\]
Combining the terms $f_0(t) + w_0(t) = \sum_l c_l^1 \phi(2t - l)$, where

$$c_l^1 = \begin{cases} c_k + d_k & l = 2k \\ c_k - d_k & l = 2k + 1 \end{cases}$$

This result can be generalized to obtain the scaling coefficients at scale $j'$ as:

$$c_l^{j'} = \begin{cases} c_k^{j'-1} + d_k^{j'-1} & l = 2k \\ c_k^{j'-1} - d_k^{j'-1} & l = 2k + 1 \end{cases}$$  (13)