WAVELETS: GETTING PERSPECTIVE

Lori Mann Bruce, Anil Cheriyadat and Mathew Burns

When a person becomes so involved in details that they begin to lose sight of the big picture, it is said "They can't see the forest for the trees!" This saying illustrates perspective. Far from the forest, it is difficult to discern the individual trees. Close up or actually in the forest, it is easy to see the individual trees, but now the total forest cannot be seen. The same can be true in digital signal processing. When we try to analyze the small details (high frequency components) of the signal, we can lose sight of the big picture (low frequency components). However, if we try to stand back from the signal such that we can see the big picture, the details often become too blurred to be useful. The wavelet transform is a relatively new signal processing tool that allows us to efficiently analyze the small details and the big picture.

What are wavelets?

When someone refers to wavelets, they could be talking about functions, filters, or transforms. A wavelet transform is similar to a Fourier transform. With a Fourier transform, a function (or signal) is decomposed into a weighted sum of sinusoids. With a wavelet transform, a function (or signal) is decomposed into a weighted sum of wavelet functions. Figure 1 shows some of the commonly used wavelet functions. The decomposition involves many convolutions (or inner products), so it can be very expensive computationally. However, algorithms for fast wavelet transforms have been developed, like those for fast Fourier transforms (FFTs). For many of these algorithms, filters associated with particular wavelet functions (referred as wavelet filters) are used. Thus, wavelet functions can be used to compute wavelet transforms, or wavelet filters can be used to compute wavelet transforms.

The continuous wavelet transform (CWT) is simply the correlation of a set of wavelet functions with an input function. The set of wavelet functions is generated from a single wavelet called the *mother wavelet*. Figure 1 shows some of the commonly used mother wavelets. Given a real mother wavelet, $\psi(t)$, the set of wavelet functions is generated by scaling (dilating or compressing) the mother wavelet (eq. 1),

$$\psi_a(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t}{a}\right), \qquad 1)$$

where a is referred to as scale. This set of wavelet functions are referred as *wavelet basis*. The CWT of an input function, $x(t)$, is defined as a set of correlations for varying a's (eq. 2),

$$W_x(a,t) = x(t) \otimes \psi_a(t) = x(t) * \psi_a(-t) \qquad 2),$$

where $\otimes$ and $*$ represent correlation and convolution, respectively. Think of convolution in terms of flip-shift-multiply-and-add. Then, correlation can be thought of as convolution without the flip or convolution with one of the functions flipped beforehand. It is also useful to think of correlation in terms of inner products. For each shift, the output of the correlation is the inner product of the wavelet and the input function. The results of the inner products are referred to as *wavelet coefficients*. If we are shifting the wavelet function, then for each shift, b, of the correlation the output would be (eq. 3)

$$W_x(a,b) = \langle x, \psi_{a,b} \rangle, \qquad \text{where } \psi_{a,b} = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right) \qquad 3)$$

Also note that not just any function can be called a mother wavelet. It must satisfy an admissibility condition.

When dealing with Fourier transforms, we often use the discrete Fourier transform (DFT), such that discrete-time signals may be analyzed and the computational expense is decreased. Similarly, discrete wavelet transforms (DWTs) have been developed. One type of DWT uses *dyadic* wavelets (with a dyad referring to a pair, objects that come in two's). In this case, the set of wavelet functions is again formed from scaling the mother wavelet. However, we only use scales that are powers of two, 2^{-j}. Also, when computing the correlations, not all shifts are used. Only dyadic shifts are used. A dyadic shift is a translation of the wavelet by the amount $k/2^j$, which is an integer multiple of the dyadic scale. Figure 2 shows an example of a wavelet at various dyadic scales and shifts. For the example shown, at scale $a=2^0$ there exists only one shift and thus one inner product. Similarly at scale $a=2^{-3}$ there are 8 shifts, and thus 8 inner products are required to compute the wavelet transform for that scale. So for our example, 8 wavelet coefficients are produced at scale 2^{-3}.

If the mother wavelet satisfies certain criteria related to multiresolution, then the DWT can be computed even more efficiently using a method known as the fast wavelet transform (FWT). Associated with the mother wavelet, there will be a lowpass and a highpass filter. By passing the input function (or signal) through the filters in a repetitive fashion, the wavelet coefficients can be obtained.

Figure 3 illustrates the FWT method; the figure is generally referred to as a dyadic filter tree. In Fig. 3, HP refers to highpass filter; LP refers to lowpass filter; and 2↓ refers to downsampling by a factor of 2 (removing every other sample). The input signal $x(t)$, is fed to the initial set of high pass and low pass filters. The output from the high pass filter represents the wavelet coefficients (high frequency information) at the highest level. The output from the low pass filter is down sampled by a factor of 2 and again fed to a high pass filter. The output of the high pass filter represents detailed information in the next level. Each level of the filter tree corresponds to a scale. This process is repeated. At the highest level, the output corresponds to the wavelet coefficients at the highest scale and represents high frequency information. The filter tree can be extended until the last level produces one coefficient. The wavelet coefficient at the last level (the output of the low pass filter)

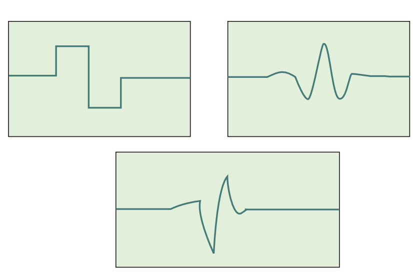

Fig. 1 Common wavelet functions

represents the least approximation of the original signal. This output corresponds to the wavelet coefficient at the lowest scale, $a=2^0$.

Why wavelets?

One might ask why decompose a signal into a weighted sum of wavelets rather than a weighted sum of sinusoids? To understand why, we will compare the wavelet transform to the Fourier transform. The well-known definition of the Fourier transform is (eq. 4)

$$X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt \,. \qquad 4)$$

From this equation, we can see the following: to analyze any particular frequency, *x(t)* is evaluated over all time $(-\infty,\infty)$. There are two immediate problems with this scenario: 1) evaluating a signal over infinite time is typically not practical, and 2) the signal could be changing with time (be *non-stationary*). For example, observe the original signal in Fig. 2. The first half of the signal is distinctly different from the second half of the signal; the first half has higher frequencies than the second half. If we observed the magnitude of the Fourier transform of this signal, we would not be able to resolve where in time the high frequencies occurred. These problems can be addressed by using a short-time Fourier transform (STFT).

Using the STFT, we analyze small windows (or segments) of the signal. We take the Fourier transform of a segment; then slide along to the next segment and do the same. Various types of windows can be used: rectangular, Bartlett (triangular), Hamming, or Hanning, for example. This begs the question how long or short (in time) are the individual windows to be. This depends on the nature of the signal being analyzed, whether it contains high frequencies, low frequencies, or a combination of frequencies. High frequency components take very little time to go through a complete oscillation ($T=1/f$), and low frequency components take a relatively long time to go through a complete oscillation.

Let's assume that the input signal has a combination of frequencies. If a short window is used, high frequency components can be located (or resolved) very well in time; however, short duration windows are insufficient for analyzing low frequency components. See Fig. 4. Thus, one might conclude that longer windows should be used. If a long window is used, low frequency components can be analyzed; that is, the signal can be resolved very well in frequency. Now the high frequency components can no longer be located very well in time. We sacrificed time resolution for frequency resolution.

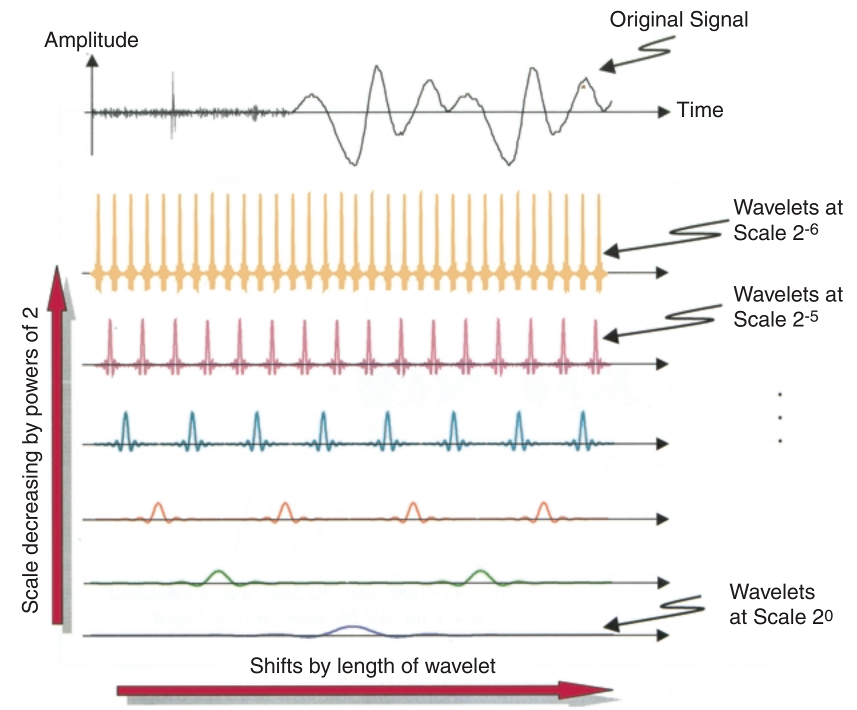

Fig. 2 Wavelets at dyadic scales and shifts

This trade-off between localization in time and frequency is referred to as the Heisenberg Uncertainty Principle. Simply put, just as one cannot know the exact momentum and location of an electron simultaneously, one cannot know the exact frequency and location of a signal component simultaneously. However, one can know the time intervals in which certain bands of frequencies exist.

For lower frequencies, we can choose longer time intervals (or windows). We gain knowledge about the frequency of the signal component, but we lose knowledge about the time location of the signal component. For higher frequencies, we can choose shorter time intervals. We gain knowledge about the time location of the signal component, but we lose knowledge about the frequency of the signal component. This varying of the time interval, or window length, is exactly what the wavelet transform accomplishes.

Referring back to Fig. 2, we can see that when the mother wavelet is dilated or stretched, it appears to contain low frequencies, and its duration in time is relatively long. However, when the mother wavelet is compressed, it appears to contain high frequencies, and its duration in time is relatively short.

Recall that the CWT is computed by correlating the scaled wavelets with the input signal. Also recall that when two signals are correlated with each other, we obtain a measure of the similarity between the two signals. Thus, when the wavelet transform is computed at a

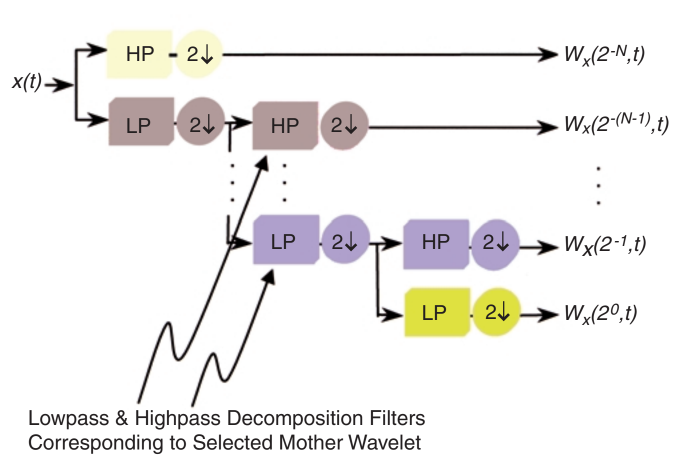

Fig. 3 Fast wavelet implementation, the dyadic filter tree

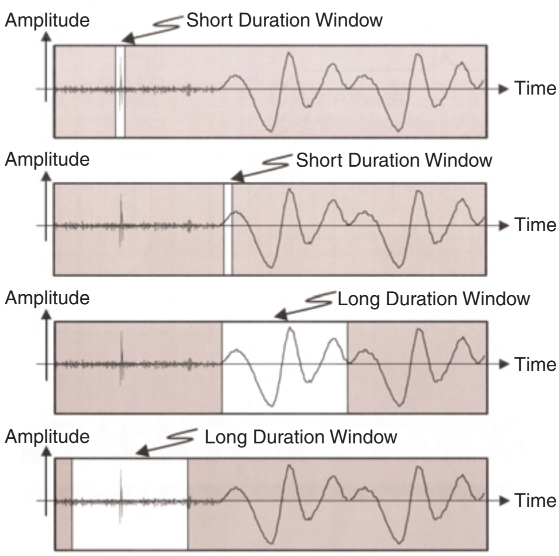

Fig. 4 Varying window sizes for analyzing the frequency content of signals

scale such that the wavelet is compressed, we obtain a measure of how similar the input signal is to the high frequency wavelet. Likewise, when the wavelet transform is computed at a scale such that the wavelet is dilated, we obtain a measure of how similar the input signal is to the low frequency wavelet. This kind of analysis is also referred as *multiresolution analysis*.

Let's suppose that the mother wavelet is chosen such that it is a symmetric finite impulse response (FIR) of an ideal bandpass filter. Then, its Fourier transform would be a perfect rectangle centered about some particular frequency. Further, the only difference between correlation and convolution is a flip of the impulse response. If the impulse response is symmetric, correlation is equivalent to convolution. When the wavelet is correlated with the input signal, it is equivalent to convolution. The result will be a filtered version of the input signal.

What happens when the wavelet is stretched to varying scales? The scaling property of the Fourier transform tells us that when a signal is stretched in time it is compressed in frequency. Thus, the varying scales of the wavelet transform would represent the input signal bandpass filtered at varying bands of frequency.

Now consider the DWT. It analyzes the input signal at dyadic scales using inner products at shifts equal to the length of the wavelet. See Fig. 2. For the scale 2^0, the wavelet is stretched such that only one inner product is computed; thus the DWT will produce just one coefficient at this scale. Since the wavelet is stretched in time, it is compressed in frequency. As a result, the one coefficient represents a very narrow band of frequencies.

Figure 5 is referred to as the dyadic tiling of the time-frequency plane. For scale 2^{-1}, the wavelet is compressed to half the original length, such that two inner products can be computed. The DWT will produce two coefficients for this scale. Since the wavelet is compressed to half the length in time, it is stretched to twice the length in frequency.

From Fig. 5, we see that this process is repeated, but how many times can it be repeated? The answer depends on the original sampling rate used for the input signal. Recall from Nyquist's Theorem that if a sampling rate of fs is used, the input signal will contain frequency information up to $f_s/2$. The repetition continues until $f_s/2$ is reached.

Another way of looking at this frequency division is by considering the DFT. Figure 6 illustrates the unit circle on the z-plane, which is where the coefficients of the DFT lie. The angle from the positive real axis around the z-plane denotes different frequency information. In the unit circle, the area near the 0° represents low frequency, the area close to 90° represents mid frequency and the area close to 180° represents high frequency.

Let's assume the original signal contained N samples, and we computed an N-point DFT. Then N equally spaced points along the unit circle represent the frequency content of the signal. The wavelet coefficient produced at scale $a=2^0$ represents the one point lying on the positive real-axis, the DC component (zero frequency component) of the input signal (the blue band on Fig. 6).

At scale $a=2^{-1}$, the wavelet transform produces two coefficients. The first coefficient represents the frequency content residing in the green bands on Fig. 6 for the first window (first half) of the input signal. The second coefficient represents the same frequency content for the second window (second half) of the input signal.

Likewise, at scale $a=2^{-2}$ the wavelet transform produces four coefficients. Each represents the frequency content within the peach-colored bands on Fig. 6, with each corresponding to a different time window of the input signal. This continues until N/2 points are produced at scale $a=2^{-j}$, where $j=\log_2(N)$. The wavelet coefficients contained in the yellow band on Fig. 6 represents the high frequency information of the input signal.

Figure 7 illustrates the frequency bands without consideration of the time duration needed for each band. We can see that the wavelets correspond to ideal bandpass filters. However, Gibb's phenomenon tells us that with FIR filters, ideal bandpass filters cannot be achieved. So practical filters must be

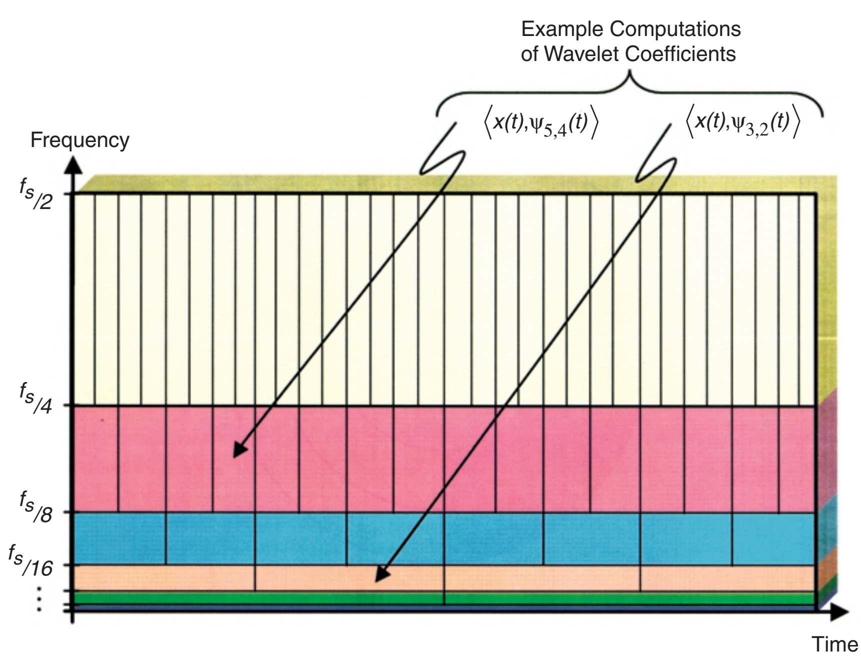

Fig. 5 Dyadic tiling of time-frequency plane generated by DWT

used, such as those illustrated in Fig. 7.

With the use of non-ideal filters, the wavelet coefficients no longer exactly represent the frequency content of the ideal bands. In fact, for some applications the goal is not to use wavelets that subdivide the frequency content into ideal bands. The goal may be to select a wavelet that represents a pattern for which someone is searching.

The wavelet transform correlates the scaled mother wavelet with the input signal. Thus, the wavelet transform can be used to search for similarities between a "pattern template" at varying scales and the input signal. If the pattern template does not closely represent a bandpass filter, the wavelet coefficients can be drastically different from the Fourier coefficients.

Applications

Wavelet analysis is a relatively new technique in signal processing. However, it has proven to be a powerful technique and has been extensively applied to diverse fields of engineering, such as medical imaging, aerial and satellite remote sensing, industrial robotics, seismology, and so on. The applications of wavelets in signal processing have included compression, denoising, and pattern recognition.

Wavelets have proven to be well suited for these applications for various reasons. One reason is its division of frequency into octaves. Natural signals, such as visual images and speech and audio signals, are often well suited for this type of analysis. In fact, experimental research has indicated that the human visual cortex uses a multifrequency channel decomposition when processing images. Some experiments have shown that the widths of the frequency channels vary on an octave scale. This may provide some insight into why wavelet processed images can appear more pleasing visually as compared to other processing schemes. Similarly, consider how audio signals are typically organized according to octaves. With the use of wavelets, frequency octaves, which contain more (or less) useful information, can be separated and processed accordingly.

Conclusion

This has been an extremely brief introduction to wavelets and their use in signal processing. Wavelets have been shown to be very useful for signal compression, denoising and pattern recognition. This should not be surprising since they provide an efficient method for analyzing non-stationary signals at varying frequencies. With wavelets, we can analyze low frequencies (the forest) at very fine frequency resolution, and we can analyze high frequencies (the trees) at a very fine time resolution. Indeed, wavelets do allow us to see the forest and the trees!

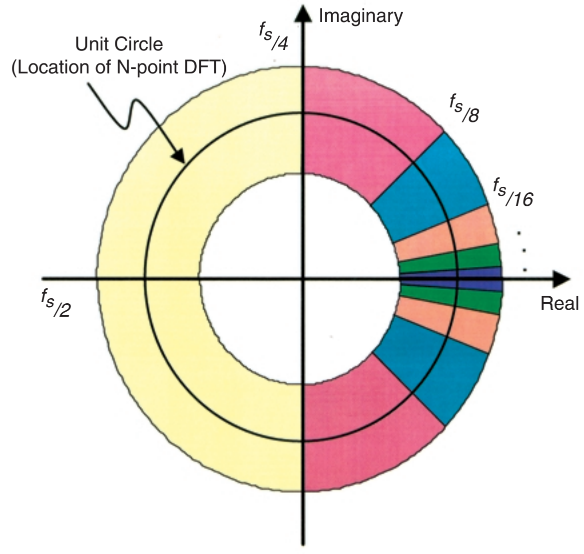

Fig. 6 Frequency decomposition Z-plane generated by DWT

Read more about it

• Kenneth R. Castleman, *Digital Image Processing*, Prentice-Hall, Englewood Cliffs, N.J., 1997. (Chap. 14)

• Stephane G. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, 1998.

• Oliver Rioul and Martin Vetterli, "Wavelets and Signal Processing," *IEEE Signal Processing Magazine*, vol. 8, no. 4, pp. 14-38, October 1991.

About the authors

Lori Mann Bruce received the B.S.E. from the University of Alabama in Huntsville in 1991, the M.S. from the Georgia Institute of Technology in 1992, and the Ph.D. from the University of Alabama in Huntsville in 1996, all in electrical and computer engineering. Dr. Bruce has served as a member of the technical staff at the U.S. Army Strategic Defense Command, and from 1996 to 2000, she was an assistant professor in the Electrical and Computer Engineering Department, Howard R. Hughes College of Engineering, University of Nevada Las Vegas. Currently, Dr. Bruce is an assistant professor in the Department of Electrical and Computer Engineering at Mississippi State University, where she is also affiliated with the Remote Sensing Technology Center and Geospatial Resources Institute. Dr. Bruce is a senior member of IEEE. Dr. Bruce is a member of the honor societies Eta Kappa Nu, Phi Kappa Phi, and Tau Beta Pi.

Anil Cheriyadat received his B.E. degree from Cochin University of Science and Technology, India in1997. He is currently pursuing his M.S. degree in the department of Electrical and Computer Engineering at Mississippi State University (Aug,2001-present). Anil works as research assistant in the Remote Sensing Technology Center under the supervision of Dr. Lori Mann Bruce. From 1997 to 2001 Anil has worked for several companies in India including Wipro Technologies, Xansa (formerly IIS Infotech), Pentafour Group. During this period he was responsible for development and maintenance of business applications running on high end IBM Mainframe servers. Anil has received the Barrier Fellowship scholarship for 2002-2003 in the Dept. of Electrical and Computer Engineering, Mississippi State University. He is a Student member of IEEE.

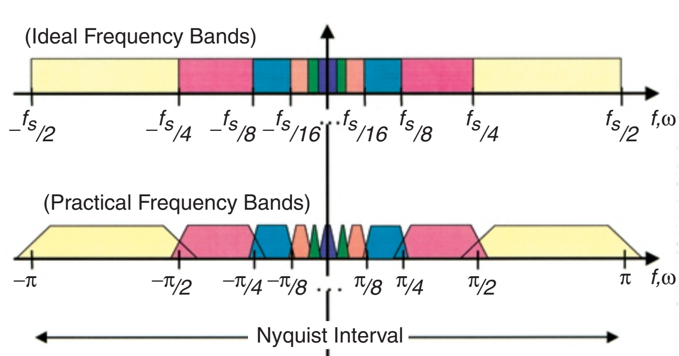

Fig. 7 Frequency decomposition generated by DWT