Introduction

Wavelets are mathematical functions that cut up data into different frequency components, and then study each component with a resolution matched to its scale. They have advantages over traditional Fourier methods in analyzing physical situations where the signal contains discontinuities and sharp spikes [1]. This paper introduces wavelets to the interested technical person outside of the digital signal processing field. In this paper is also compared wavelet transforms with Fourier transforms, state properties and other special aspects of wavelets, and finish with some interesting applications such as high voltage processes.

The wavelet analysis procedure is to adopt a wavelet prototype function, called an analyzing wavelet or mother wavelet. Temporal analysis is performed with a contracted, high-frequency version of the prototype wavelet, while frequency analysis is performed with a dilated, low-frequency version of the same wavelet. Because the original signal or function can be represented in terms of a wavelet expansion (using coefficients in a linear combination of the wavelet functions), data operations can be performed using just the corresponding wavelet coefficients. And if you further choose the best wavelets adapted to your data, or truncate the coefficients below a threshold, your data is sparsely represented. This sparse coding makes wavelets an excellent tool in the field of data compression [1].

Other applied fields that are making use of wavelets include astronomy, acoustics, nuclear engineering, sub-band coding, signal and image processing, neurophysiology, music, magnetic resonance imaging, speech discrimination, optics, fractals, turbulence, earthquake-prediction, radar, human vision, and pure mathematics applications such as solving partial differential equations.

Fourier Analysis

Fourier's representation of functions as a superposition of sines and cosines has become common for both the analytic and numerical solution of differential equations and for the analysis and treatment of communication signals. Fourier and wavelet analysis have some very strong links.

Before 1930, the main branch of mathematics leading to wavelets began with Joseph Fourier (1807) with his theories of frequency analysis, now often referred to as Fourier synthesis. He asserted that any 2π-periodic function f(x) is the sum

\[ a_0 + \sum_{k=1}^\infty (a_k \cdot \cos(k \cdot x) + b_k \cdot \sin(k \cdot x)) \]

of its Fourier series. The coefficients \(a_0, a_k, b_k\) are calculated by

\[ a_0 = \frac{1}{2\pi} \int_{0}^{2\pi} f(x) \cdot \cos(x) \cdot dx \]

\[ a_k = \frac{1}{\pi} \int_{0}^{2\pi} f(x) \cdot \cos(k \cdot x) \cdot dx \]

\[ b_k = \frac{1}{\pi} \int_{0}^{2\pi} f(x) \cdot \sin(k \cdot x) \cdot dx \]

Fourier's assertion played an essential role in the evolution of the ideas mathematicians about the functions. He opened up the door to a new functional universe.

After 1807, by exploring the meaning of functions, Fourier series convergence, and orthogonal systems, mathematicians gradually were led from their previous notion of frequency analysis to the notion of scale analysis. That is, analyzing f(x) by creating mathematical structures that vary in scale. That is, constructing a function, shift it by some amount, and change its scale. Apply that structure in approximating a signal. Now repeat the procedure. Take that basic structure, shift it, and scale it again. Apply it to the same signal to get a new approximation. And so on. It turns out that this sort of scale analysis is less sensitive to noise because it measures the average fluctuations of the signal at different scales [1].

Fourier Transforms

The Fourier transform's utility lies in its ability to analyze a signal in the time domain for its frequency content. The transform works by first translating a function in the time domain into a function in the frequency domain. The signal can then be analyzed for its frequency content because the
Fourier coefficients of the transformed function represent the contribution of each sine and cosine function at each frequency. An inverse Fourier transform does just what you'd expect; transform data from the frequency domain into the time domain [1].

**Discrete Fourier Transforms**

The discrete Fourier transform (DFT) estimates the Fourier transform of a function from a finite number of its sampled points. The sampled points are supposed to be typical of what the signal looks like at all other times. The DFT has symmetry properties almost exactly the same as the continuous Fourier transform. In addition, the formula for the inverse discrete Fourier transform is easily calculated using the one for the discrete Fourier transform because the two formulas are almost identical.

**Windowed Fourier Transforms**

If \( f(t) \) is a non-periodic signal, the summation of the periodic functions, sine and cosine, does not accurately represent the signal. You could artificially extend the signal to make it periodic but it would require additional continuity at the endpoints. The windowed Fourier transform (WFT) is one solution to the problem of better representing the non-periodic signal. The WFT can be used to give information about signals simultaneously in the time domain and in the frequency domain.

With the WFT, the input signal \( f(t) \) is chopped up into sections, and each section is analyzed for its frequency content separately. If the signal has sharp transitions, the input data are windowed so that the sections converge to zero at the endpoints [2]. This windowing is accomplished via a weight function that places less emphasis near the interval's endpoints than in the middle. The effect of the window is to localize the signal in time.

**Fast Fourier Transforms**

To approximate a function by samples, and to approximate the Fourier integral by the discrete Fourier transform, requires applying a matrix whose order is the number sample points \( n \). Since multiplying a matrix by a vector costs on the order of arithmetic operations, the problem gets quickly worse as the number of sample points increases. However, if the samples are uniformly spaced, then the Fourier matrix can be factored into a product of just a few sparse matrices, and the resulting factors can be applied to a vector in a total of order arithmetic operations. This is the so-called fast Fourier transform or FFT [3].

**Wavelet vs. Fourier Transforms**

**Similarities between Fourier and Wavelet Transforms**

The fast Fourier transform (FFT) and the discrete wavelet transform (DWT) are both linear operations that generate a data structure that contains segments of various lengths, usually filling and transforming it into a different data vector of length.

The mathematical properties of the matrices involved in the transforms are similar as well. The inverse transform matrix for both the FFT and the DWT is the transpose of the original. As a result, both transforms can be viewed as a rotation in function space to a different domain. For the FFT, this new domain contains basis functions that are sines and cosines. For the wavelet transform, this new domain contains more complicated basis functions called wavelets, mother wavelets, or analyzing wavelets.

Both transforms have another similarity. The basis functions are localized in frequency, making mathematical tools such as power spectra (how much power is contained in a frequency interval) and scalograms, useful at picking out frequencies and calculating power distributions [1].

Dissimilarities between Fourier and Wavelet Transforms

The most interesting dissimilarity between these two kinds of transforms is that individual wavelet functions are localized in space. Fourier sine and cosine functions are not. This localization feature, along with wavelets' localization of frequency, makes many functions and operators using wavelets "sparse" when transformed into the wavelet domain. This sparseness, in turn, results in a number of useful applications such as data compression, detecting features in images, and removing noise from time series [1].

One way to see the time-frequency resolution differences between the Fourier transform and the wavelet transform is to look at the basis function coverage of the time-frequency plane [4]. Figure 1 shows a windowed Fourier transform, where the window is simply a square wave. The square wave window truncates the sine or cosine function to fit a window of a particular width. Because a single window is used for all frequencies in the WFT, the resolution of the analysis is the same at all locations in the time-frequency plane.

An advantage of wavelet transforms is that the windows vary. In order to isolate signal discontinuities, one would like to have some very short basis functions. At the same time, in order to obtain detailed frequency analysis, one would like to have some very long basis functions. A way to achieve this is to have short high-frequency basis functions and long low-frequency ones. This happy medium is exactly what you get with wavelet transforms. Figure 2 shows the coverage in the time-frequency plane with one wavelet function, the Daubechies wavelet.

One thing to remember is that wavelet transforms do not have a single set of basis functions like the Fourier transform, which utilizes just the sine and cosine functions. Instead, wavelet transforms have an infinite set of possible basis functions. Thus wavelet analysis provides immediate access to information that can be obscured by other time-frequency methods such as Fourier analysis.

**Shape of Some Wavelets**

Wavelet transforms comprise an infinite set. The different wavelet families make different trade-offs between how compactly the basis functions are localized in space and how smooth they are.

Some of the wavelet bases have fractal structure. The Daubechies wavelet family is one example (see Figure 3).
Fig. 3 The fractal self-similarity of the Daubechies mother wavelet. This figure was generated using the WaveLab [1].

Within each family of wavelets (such as the Daubechies family) are wavelet subclasses distinguished by the number of coefficients and by the level of iteration. Wavelets are classified within a family most often by the number of vanishing moments. This is an extra set of mathematical relationships for the coefficients that must be satisfied, and is directly related to the number of coefficients [5]. For example, within the Coiflet wavelet family are Coiflets with two vanishing moments, and Coiflets with three vanishing moments. In Figure 4, I illustrate several different wavelet families.

Fig. 4. Several different families of wavelets. The number next to the wavelet name represents the number of vanishing moments (A stringent mathematical definition related to the number of wavelet coefficients) for the subclass of wavelet. These figures were created using WaveLab [1].

Wavelet Analysis
The Discrete Wavelet Transform

Dilations and translations of the "Mother function," or "analyzing wavelet" define an orthogonal basis, our wavelet basis:

\[
\Phi_{(s,l)}(x) = 2^{\frac{s}{2}} \cdot \Phi(2^{-s} \cdot x - l)
\]

The variables \( s \) and \( l \) are integers that scale and dilate the mother function \( \Phi(x) \) to generate wavelets, such as a Daubechies wavelet family. The scale index \( s \) indicates the wavelet's width, and the location index \( l \) gives its position. Notice that the mother functions are rescaled, or "dilated" by powers of two, and translated by integers. What makes wavelet bases especially interesting is the self-similarity caused by the scales and dilations. Once we know about the mother functions, we know everything about the basis.

To span our data domain at different resolutions, the analyzing wavelet is used in a scaling equation:

\[
W(x) = \sum_{k=0}^{N-1} (-1)^k \cdot c_{k+1} \cdot \Phi(2 \cdot x + k)
\]

where \( W(x) \) is the scaling function for the mother function \( \Phi(x) \), and \( c_k \) are the wavelet coefficients. The wavelet coefficients must satisfy linear and quadratic constraints of the form

\[
\sum_{k=0}^{N-1} c_k = 2 \quad \text{and} \quad \sum_{k=0}^{N-1} c_k \cdot c_{k+2l} = 2 \cdot \delta_{l,0}
\]

where \( \delta \) is the delta function and \( l \) is the location index.

One of the most useful features of wavelets is the ease with which a scientist can choose the defining coefficients for a given wavelet system to be adapted for a given problem. In Daubechies' original paper [6], she developed specific families of wavelet systems that were very good for representing polynomial behavior. The Haar wavelet is even simpler, and it is often used for educational purposes.

It is helpful to think of the coefficients as a filter. The filter or coefficients are placed in a transformation matrix, which is applied to a raw data vector. The coefficients are ordered using two dominant patterns, one that works as a smoothing filter (like a moving average), and one pattern that works to bring out the data's "detail" information. These two orderings of the coefficients are called a quadrature mirror filter pair in signal processing parlance. A more detailed description of the transformation matrix can be found elsewhere [3].

To complete our discussion of the DWT, let's look at how the wavelet coefficient matrix is applied to the data vector. The matrix is applied in a hierarchical algorithm, sometimes called a pyramidal algorithm. The wavelet coefficients are arranged so that odd rows contain an ordering of wavelet coefficients that act as the smoothing filter, and the even rows contain an ordering of wavelet coefficient with different signs that act to bring out the data's detail. The matrix is first applied to the original, full-length vector. Then the vector is smoothed and decimated by half and the matrix is applied again. Then the smoothed, halved vector is smoothed and halved again, and the matrix applied once more. This process continues until a trivial number of "smooth-smooth-smooth..." data remain. That is, each matrix application brings out a higher resolution of the data while at the same time smoothing the remaining data. The output of the DWT consists of the remaining "smooth (etc.)" components, and all of the accumulated "detail" components.

The Fast Wavelet Transform

The DWT matrix is not sparse in general, so we face the same complexity issues that we had previously faced for the discrete Fourier transform [7]. We solve it as we did for the FFT, by factoring the DWT into a product of a few sparse matrices using self-similarity properties. The result is an algorithm that requires only order \( n \) operations to transform an \( n \)-sample vector. This is the "fast" DWT of Mallat and Daubechies.

Wavelet Packets

The wavelet transform is actually a subset of a far more versatile transform, the wavelet packet transform [8].

Wavelet packets are particular linear combinations of wavelets [7]. They form bases which retain many of the orthogonality, smoothness, and localization properties of their parent wavelets. The coefficients in the linear combinations are computed by a recursive algorithm making each newly computed wavelet packet coefficient sequence the root of its own analysis tree.

Adapted Waveforms

Because we have a choice among an infinite set of basis functions, we may wish to find the best basis function for a given representation of a signal [7]. A basis of adapted waveform is the best basis function for a given signal representation. The chosen basis carries substantial information about the signal, and if the basis description is efficient (that is, very few terms in the expansion are needed to represent the signal), then that signal information has been compressed.
According to Wickerhauser [7] some desirable properties for adapted wavelet bases are:
- speedy computation of inner products with the other basis functions;
- speedy superposition of the basis functions;
- good spatial localization, so researchers can identify the position of a signal that is contributing a large component;
- good frequency localization, so researchers can identify signal oscillations; and
- independence, so that not too many basis elements match the same portion of the signal.

For adapted waveform analysis, researchers seek a basis in which the coefficients, when rearranged in decreasing order, decrease as rapidly as possible. To measure rates of decrease, they use tools from classical harmonic analysis including calculation of information cost functions. This is defined as the expense of storing the chosen representation. Examples of such functions include the number above a threshold, concentration, entropy, logarithm of energy, Gauss-Markov calculations, and the theoretical dimension of a sequence.

Wavelet Applications

Applications of Wavelets in Signal Processing

From the derivation of the wavelet transform as an alternative to the short-time Fourier transform (STFT), it is clear that one of the main applications will be in non-stationary signal analysis.

Applications of wavelet decompositions in numerical analysis, e.g. for solving partial differential equations, seem very promising because of the “zooming” property which allows a very good representation of discontinuities, unlike the Fourier transform.

Perhaps the biggest potential of wavelets has been needed for signal compression. Since discrete wavelet transforms are essentially subband coding systems, and since subband coders have been successful in speech and image compression, it is clear that wavelets will find immediate application in compression problems. The only difference with traditional subband coders is the fact that filters are designed to be regular (that is, they have many zeroes at $z = 0$ or $z = \pi$). Note that although classical subband filters are not regular, they have been designed to have good stopbands and thus are close to being “regular”, at least for the first few octaves of subband decomposition.

Denoising Noisy Data

In diverse fields from planetary science to molecular spectroscopy, scientists are faced with the problem of recovering a true signal from incomplete, indirect or noisy data. Can wavelets help solve this problem? The answer is certainly “yes,” through a technique called wavelet shrinkage and thresholding methods that David Donoho has worked on for several years [9].

The technique works in the following way. When you decompose a data set using wavelets, you use filters that act as averaging filters and others that produce details [10]. Some of the resulting wavelet coefficients correspond to details in the data set. If the details are small, they might be omitted without substantially affecting the main features of the data set. The idea of thresholding, then, is to set to zero all coefficients that are less than a particular threshold. These coefficients are used in an inverse wavelet transformation to reconstruct the data set. Figure 6 is a set of “before” and “after” illustrations of a partial discharge signal in a high voltage cable. The signal is transformed, thresholded and inverse-transformed. The technique is a significant step forward in handling noisy data because the denoising is carried out without smoothing out the sharp structures. The result is cleaned-up signal that still shows important details.

To improve the accuracy, wavelet decomposition can be used to remove the high-frequency noise from the signal. Successive approximations become less noisy as more high frequency information is filtered out. Thus this provides a simple method to de-noise the signal. Fig. 6(b) shows the de-noised signal using level-5 approximation and Daubechies db3 wavelet. In comparison to the original signal, it is much cleaner and the reflected pulse can be clearly seen. The measured reflection time corresponds to twice the cable length. This indicates the fault is at the cable termination rather than inside the cable and thus agrees with the ultrasonic detector finding.
One disadvantage of the above method is that the fast changing features of the original signal is lost. Note the smoothing effect on the wavefront in Fig. 6(b). This would reduce the accuracy of the measurement of the time delay between the first pulse and its reflection. An elegant alternative to overcome this problem is the technique called thresholding whereby the details are discarded only if the magnitudes exceed a certain limit. The procedure is to examine the details vectors of the wavelet decomposition, select the appropriate threshold coefficients and reconstruct the new details signals. The toolbox provides two calling functions: one to calculate the default threshold parameters and the other to perform the actual de-noising. Applying these functions, the result is shown in Fig. 6(c). It retains well the sharp detail of the original but is somewhat noisier. It may be possible to improve the result by trying other thresholds [11].

Another application of wavelet transform in high voltage processes:
- power quality disturbance recognition,
- fault detection in electromotors,
- fault diagnosis of nonlinear analog circuits,
- monitoring of partial discharges in high voltage cables,
- signal figures compression,
- and many more.

Conclusion
Most of basic wavelet theory has been done. The mathematics has been worked out in excruciating detail and wavelet theory is now in the refinement stage. The refinement stage involves generalizations and extensions of wavelets, such as extending wavelet packet techniques.

The future of wavelets lies in the as-yet uncharted territory of applications. Wavelet techniques have not been thoroughly worked out in applications such as practical data analysis, where for example discretely sampled time-series data might need to be analyzed. Such applications offer exciting avenues for exploration.

This work was supported by Science and Technology Assistance Agency under the contact number APVV-20-006005 and VEGA 1/3142/06 projects.

REFERENCES

Author: Ing. Dušan Medved, Technical University of Košice, Department of Electric Power Engineering, Mäsiarska 74, 04210 Košice, Slovakia, E-mail: Dusan.Medved@tuke.sk;