Elementary Theory of Wavelets and Filter Bank

Pinal Choksi

Assistant Professor, Science & Humanity Department, Vadodara Institute of Engineering, Kotambi, Vadodara, Gujarat, India

Abstract-- Many of us are interested in detecting the irregularity in very small region of a transient signal which cannot be detected by necked eyes. That can be possible by discretizing the region in which the irregularity lie. There are two ways to discretize the signal. In theory one can use Fourier transform on the signal and cut in M-pieces. Where as in practice discretization can be obtained by applying multiple number of filters (Filter Bank).

This paper will give knowledge of wavelets and filter banks and later on connection between them. These are rapidly developing topics in real time. The technique of filter banks (for discrete signals) and wavelets (to represent functions) are used throughout signal and image processing for compression, denoising, enhancement, motion estimation and pattern recognition. New wavelets continue to be constructed for new applications.

To understand this process we need to have a brief knowledge about the basic concepts like Fourier Transform, Short-Time Fourier Transform, Wavelet Transform, MRA, Filter and Filter bank.

Fourier Transform:

A Fourier transform of f (t) $\in$ L^1 (R) is given as

$$F(w) = F\big(f(t)\big) = \int_{-\infty}^{\infty} e^{-iwt} f(t)dt$$

$$= < f\,(t),\ e^{iwt}>$$

It is also hold for f (t) $\in$ L^2 (R)

Fourier transform is a mathematical description of the relationship between function of time and corresponding function of frequency.

Fourier transform widely used in science and engineering that convert a signal which is intensity vs time in to a spectrum which is an amplitude vs frequency.

Limitation of Fourier Transform:

The following are limitations of Fourier Transform.

1. It can be computed for only one frequency at a time.
2. A small change in the signal would affect the entire frequency spectrum of a signal.
3. The formula does not give any information about frequencies which develops with respect to time.

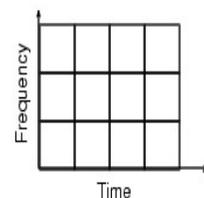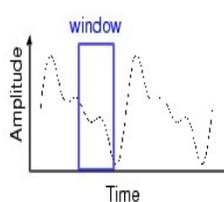Therefore, to overcome above limitations we switch on to Short-Time Fourier Transform.

Short-Time Fourier Transform (STFT):

A STFT is used to get approximate information from both time and frequency domain simultaneously. It is defined as

$$STwf(to, w) = \int_{-\infty}^{\infty} f(t)\overline{w(t - to)}e^{-iwt}dt$$

$$= <f\ (t),\ w\ (t\text{-}to)\ e^{iwt}>$$

Now we shall show how to discretize the signal by using STFT.

The STFT represents a sort of compromise between the time and frequency based views signal. It provides some information about both when and at what frequencies a signal event occurs. However, you can

only obtain this information with limited precision, and that is determined by the size of the window.

Limitation of STFT:

The limitation is that once you choose a particular size for the time window, that window is the same for all frequencies. Many signals require a more flexible approach one where we can vary the window size to determine more accurately either time or frequency, but that is possible in wavelet transform.

Wavelet:

History of wavelet is in 1983, Jean Morlet a French Geophysicist working in the oil company first introduce the concept of a wavelet. He developed a new time frequency analysis calling "Wavelets of constant shape."

In 1984, Grossman and Morlet said the mathematical foundation for the wavelet transform using the tools like Hilbert space, frame etc.

In 1986 Daubechies, Grossman, Meyer developed the theory regarding wavelets. This was the beginning of the wavelet or rather golden era of development of wavelet concept.

We all are interested in knowing what wavelet is. How it is useful. Basically small waves are called wavelets.

Wavelet has small concentrated burst of finite energy in the time domain (this condition makes the wavelet little) and it exhibits some oscillation in time (this conditions makes wavelet wavy), hence the word wavelet.

Thus, wavelets have oscillatory behavior in some interval and then decays rapidly to zero, outside the interval.

One major advantage afforded by wavelets is the ability to perform local analysis that is, to analyze a localized area of a larger signal.

Consider a sinusoidal signal with a small discontinuity one so tiny as to be barely visible. Such a signal easily could be generated in the real world, perhaps by a power fluctuation or a noisy switch.
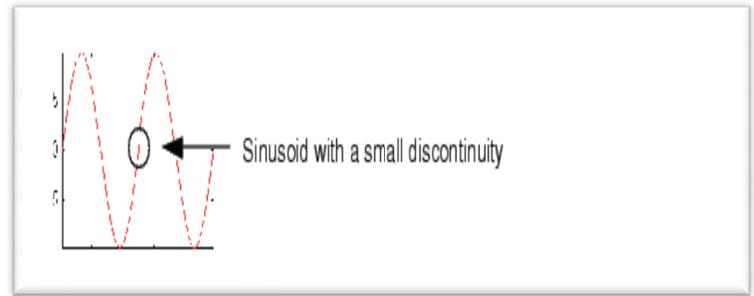

A plot of the Fourier coefficients of this signal shows nothing particularly interesting; a flat spectrum with two peaks representing a signal frequency. However, a plot of wavelet coefficients clearly shows the exact location in time of the discontinuity.

Fourier Coefficients   Wavelet Coefficients
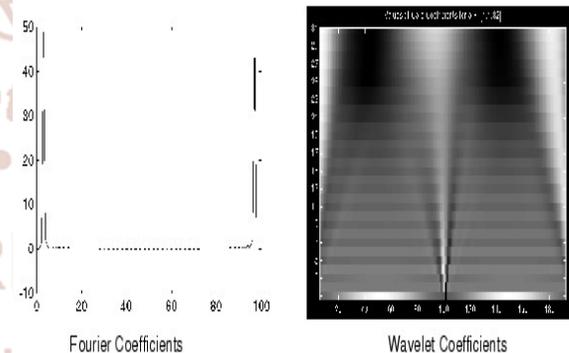
Wavelet analysis is capable of revealing aspects of data that other signal analysis techniques aspects like trends, breakdown points, discontinuities in higher derivatives and self-similarity. Furthermore because it affords a different view of data than those presented by traditional techniques, wavelet analysis can often compress or de-noise a signal without appreciable degradation.

Indeed, in their brief history within the signal processing field, wavelets have already proven themselves to be an indispensable addition to the analyst's collection of tools. Hence it is very popular in field of analysis.

Equate wavelets with sine waves, which are the basis of Fourier analysis. Sinusoids do not have limited duration they extend from minus to plus infinity. And where sinusoids are smooth and predictable, wavelets tend to be irregular and asymmetric.


Sine Wave   Wavelet (db10)
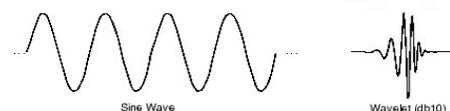
Fourier analysis consists of breaking up a signal into sine waves of various frequencies. Similarly, wavelet analysis is the breaking up of a signal into shifted and scaled versions of the original (or mother) wavelet.

As in above image of sine waves and wavelets, seen that intuitively that signals with sharp changes might be better analyzed with an irregular wavelet than with a smooth sinusoid, just as some foods are better handled with a fork than a spoon.

Wavelet Transform:

Wavelet transform represents a windowing technique with variable-sized regions. Wavelet transform allows the use of long time intervals where we want more precise low-frequency information and shorter regions where we want high frequency information.

Here what this looks like in contrast with the time-based, frequency-based and STFT views of a signal:
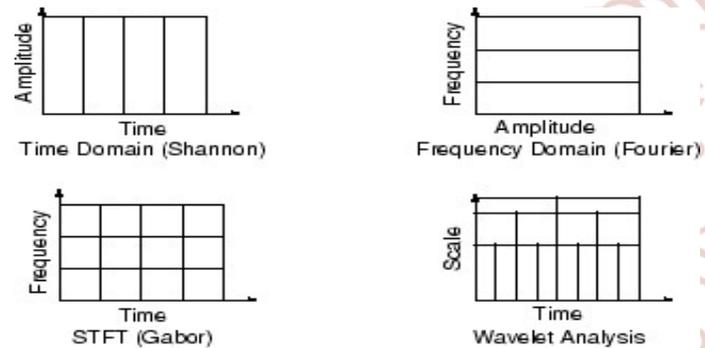
As in above image noticed that wavelet transform does not use a time-frequency region, but rather s time-scale region.

The Continuous Wavelet Transform (CWT):

If $\psi \in L^2$ (R) satisfies the "admissibility" condition

$$C\varphi = \int_{-\infty}^{\infty} f(t)\overline{\varphi\left(\frac{t-b}{a}\right)} dt \; ; f \in L^2 (R)$$

$$= \int_{-\infty}^{\infty} f(t)\overline{\psi b, a\,(t)} \; dt$$

$$= \; <f(t), \; \psi_{b,a} \; (t) \; >$$

where $\psi_{b,a}\,(t) = \dfrac{\varphi\frac{(t-b)}{a}}{\sqrt{|a|}}$

Where $a, b \in R$ with $a \neq 0$ as parameter a,b gives effect of magnification or contraction and translation respectively. $\psi_{b,a}$ (t) are called baby wavelets.

Wavelets are very very small and having irregular behavior in very small interval. Therefore area covered under the curve is approximately zero.

i.e. $\int_{-\infty}^{\infty} \varphi(t)dt = 0$; $\varphi \; \epsilon \quad L^2 \quad (R) \quad$ where $\varphi(t)$ is the basic or mother wavelet.

Scaling:

Wavelet analysis produces a time-scale view of a signal, and now what is exactly mean by scale.

Scaling a wavelet simply means stretching (or compressing) it. To go beyond colloquial descriptions such as "stretching" that is introduce the scale factor, often denoted by the letter ᵃ. In sinusoids, the effect of the scale factor is very easy to see:
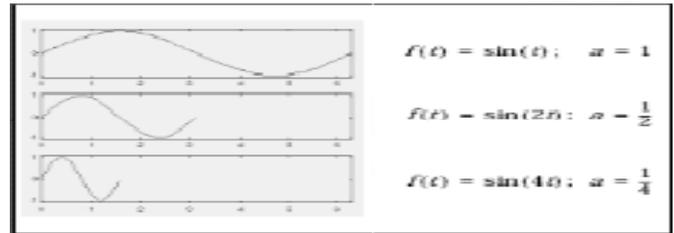
The scale factor works exactly the same with wavelets. The smaller the scale factor, the more "compressed" the wavelet.
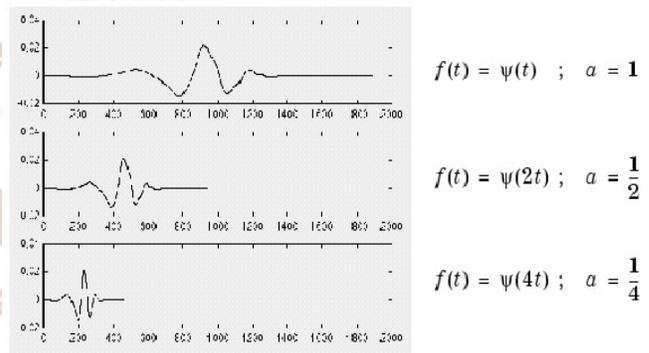
It is clear from the diagrams that, for a sinusoidal sin (wt), the scale factor ᵃ is inversely propotional to the radian frequencyω. Similarly, with wavelet analysis, the scale is inversely proportional to the frequency of the signal.

Shifting:

Shifting a wavelet simply means delaying its onset. Mathematically, delaying a function f(t) by k is represented by f(t-k).

Discrete Wavelet Transform: (DWT)

The continuous wavelet transform has unique advantage as its widely width is control by scale 'ᵃ'. However computational load of continuous wavelet transform is very heavy in order to capture all characteristics. To minimize excess of computation in continuous wavelet transform mathematician has developed discrete wavelet transform which can be obtained from the process known as multi resolution analysis.

Multi Resolution Analysis: (MRA)

MRA decomposes space L^2 (R) in a set of approximate subspace $\{V_j\}$ $j \in Z$

$V_j = \ldots\ldots \oplus W_{j-2} \oplus W_{j-1}$

Where { V_j } is generated by some scaling function $\Phi \in (R)$ and {W_j } is generated by some wavelet $\psi \in L^2(R)$

Which satisfies the following axioms:

1. {0} $\subset \cdots \ldots \ldots V_{-2} \subset V_{-1} \subset \cdots \ldots \ldots \subset L^2(R)$ therefore, the sequence of subspaces V_j is nested.
2. $\cap_{j \in Z} V_j = \{0\}$
3. Close $\cup_{j \in Z} V_j = L^2(R)$ i.e. closure of union of V_j gives whole of $L^2(R)$
4. $V_{j+1} = V_j \oplus W_j$
5. $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1, \ j \in Z}$
6. There exist a scaling function $\Phi(x) \in V_0$ such that {$\Phi_{j,k}$} is an orthonormal basis that spans V_j.
7. F(x) $\in V_0 \Leftrightarrow f(2x) \in V_{j+1, \ j \in Z}$

Filter:

Filter is an operator which maps one signal to another signal.

Filter has basically two properties.

➢ Linearity:

If H is a filter? Operator.

$$H((\alpha f + \beta g) = \alpha H(f) + \beta H(g)$$

Where α, β are constants and f, g are two functions $(f, g \in I)$

➢ Time invariant or Shift invariant:

A filter H is called shift invariant or time invariant if it communicates with the shift operator HD = DH. This means that delaying the filtered signal is the same as filtering the delayed signal.

▪ When filter is operated on an impulse say δ then the effect generated is known as impulse response. The impulse response is given by filter coefficients.

h = (h_n) = H δ …………..(1)

▪ The fourier transform of the impulse response is known as frequency response of the filter.

H(e^{iw}) = $\sum_n h_n e^{-inw}$ ($by \ using \ z - transform \ and \ taking \ z = e^{iw}$)

But frequency response is the combination of amplitude and phase angle so,

H(e^{iw}) = $\left|\sum_n h_n e^{-inw}\right| e^{i\phi(w)}$

= |H(eiw) |$e^{i\phi(w)}$

We can describe filter in three different way that is in time domain, z-domain and in term of linear algebra.

Filter Bank:

Suppose discrete or an analog signal 's' which is limited with bandwidth Ω. We can write signal as the sum of M signals. Each of which have bandwidth Ω/M. In this way, a wide band signal can be split into M signals of smaller band and transmitted over a channel with smaller bandwidth. The receiver can reconstruct the original signal. In theory we can compute, Fourier transform of signal 's', cut this in M pieces and back transform. In practice, we are applying M filters to the signal and each of these filters generates one signal with limited bandwidth. This is called M channel filter bank.

Now, suppose we are applying 5 filters to the signal and each of these filters generates one of the 5 signals with small bandwidth. This is called 5-channel filter bank.

Now M=2

Suppose a discrete signal's'. We can apply a low pass filter H and a high pass filter G. which splits the signal's' in two parts. The part which contains the low frequencies which gives a low resolution idea of the signal and the other part which contains the high frequencies which gives the detail information. This is called 2-channel filter bank. It splits frequency band in two sub bands.

It is possible to stretch the half bands again to the full bandwidth by down-sampling. Down sampling means signals passing through filter in which we delete odd samples and keep only even number of samples. Down sampling is denoted by (↓)

If s=s_n is given signal.

S' = ↓ s if s_n ' = s_{2n}

In general, s' = (↓ M)s if s_n' =s_{nM} (M is number of filters).

In z-domain,

S'= ↓ s ↔ $s'(z^2) = \frac{G(z)+H(-z)}{2}$

(by applying low pass filter & high pass filter G(z) = H(-z))

$$\leftrightarrow S'(z) = \frac{G(z^{1/2}) + H(-z^{1/2})}{2}$$

$$\leftrightarrow S'(z) = \frac{S(z^{1/2}) + S(-z^{1/2})}{2}$$

In the frequency domain,

$$S'(e^{iw}) = \frac{S(e^{iw/2}) + S(e^{-i\ /2})}{2}$$

This shows that if bandwidth of S is π then band width of S' is 2π.

For the high pass band $\pi/2 \leq |w| < \pi$ we have to shift the spectrum first to low pass band

$|w| < \pi/2$ which corresponds to adding π to w. This means that for the shifted spectrum. The frequency response is given by

$$\sum s_n e^{-i\ (w+\pi)} = \sum s_n e^{-i}\ e^{-in\pi}$$
$$= \sum (-1)^n s_n e^{-inw}$$

On the analysis side of a two channel filter bank. We have the application of the two filters $\widehat{H}*$ and $\widehat{G}*$ which are both followed by down sample operation. Here we use the notation $\widehat{H}*$ to denote that the transfer function of this filter is $\widehat{H}* = \sum_k h_k z^k$

Similarly for $\widehat{G}*$ to denote that transfer function of this filter is $\widehat{G}* = \sum_k g_k z^k$ $\widehat{H}*(z)$ is the transform of the time reversed sequence h^* where $h = h_k$

On the synthesis side, of the filter bank one finds the mirror image of the analysis side.

First the signals are up sampled. This means that between every two samples a zero is introduced. This is denoted as

$$s' = \uparrow s \leftrightarrow s'_{2n} = s_n \ and \ s'_{2n-1} = 0 \leftrightarrow S'(z)$$
$$= S(z^2)$$

After the up sampling some filters H and G are applied and two result signals Y and X are added. The synthesis side should up to the analysis computing such that the resulting signal S' is again equal to the original signal S.

The operation of the filter bank can also be written in terms of (infinite) matrices. Filtering by low pass filter $\widehat{H}*$ means multiplication with the infinite toeplitz matrix with entries the impulse response $\widehat{h}* = (\widehat{h}_{-n})$ i.e. the Fourier coefficient of $\widehat{H}*$.

Down sampling means that we skip every other row (the odd ones).

Thus the operation in upper branch of the analysis part gives LP (z) as the result of multiplication with the adjoint of the matrix.

$$\widehat{H} = \begin{bmatrix} . & & & & & \\ \dots & \widehat{h5} & \widehat{h3} & \widehat{h1} & \widehat{h}_{-1} & \dots \\ \dots & \dots & \widehat{h_4} & \widehat{h_2} & \widehat{h_0} & \dots \\ \dots & \dots & \dots & \widehat{h_3} & \widehat{h_1} & \dots \end{bmatrix} \text{-------------(*)}$$

Similarly, for the lower branch. Thus the vector p of the samples of LP (z) and the vector q of the samples of HP (z) are obtained from the samples s of S(z) as follows

$$\begin{bmatrix} p \\ q \end{bmatrix} = \begin{bmatrix} \widehat{H^*} \\ \widehat{G^*} \end{bmatrix} s = \widetilde{K^*} s$$

On the synthesis side, the up sampling followed by filtering with H means that we multiply with toeplitz matrix whose entries are the impulse response coefficients (i.e. the Fourier coefficients of H(z)) and in which every other column is deleted. That is the matrix H which is defined like $\widehat{H}$ but without the tildes. That is as follows

$$H = \begin{bmatrix} . & & & & \\ \dots & h_5 & h_3 & h_1 & \dots \\ \dots & \dots & h_4 & h_2 & h_0 & \dots \\ \dots & \dots & \dots & h_3 & h_1 & \dots \\ & & & & & . \\ & & & & & \dots \end{bmatrix}$$

The matrix G can be defined similarly for the other branch on the synthesis side.

The samples $\hat{s}$ of the result $\hat{S}(z)$ are then computed from the samples of LP (z) and HP (z) by

$$\hat{s} = H_p + G_q$$
$$= [H\ G] \begin{bmatrix} p \\ q \end{bmatrix}$$
$$\equiv K \begin{bmatrix} p \\ q \end{bmatrix}$$

We shall have $\hat{s} = s$ if $\widetilde{K}*K = I$

The recursive application of a two channel filter bank leads to an M channel filter bank.

If the 2-channel filter banks split in equal band-widths then the bandwidths of the end channels will not be the same.

Wavelet Decomposition and Reconstruction:

Consider the general structure of multiresolution analysis and wavelets as in

$V_j = \ldots\ldots \oplus W_{j-2} \oplus W_{j-1}$, j$\in Z$

Where $\{V_j\}$ is generated by scaling function$\Phi \in L^2(R)$ and $\{W_j\}$ is generated by some wavelet $\psi \in L^2(R)$.

By third property of MRA we know every function f in $L^2(R)$ can be approximated closely as is desired by $f_N \in V_N$, for some $N \in Z$

As, $V_j = V_{j-1} \oplus W_{j-1}$, for any j$\in Z$

f_N has a unique decomposition $f_N = f_{N-1} \oplus g_{N-1}$ where $f_{N-1} \in V_{N-1}$ and $g_{N-1} \in W_{N-1}$

by repeating this process, we have

$f_N = g_{N-1} + g_{N-2} + \cdots + g_{N-M} + f_{N-M}$ --------------- ---- (6)

Where $f_j \in V_j$ and $g_j \in W_j$ and M is so chosen that f_{N-M} is sufficiently "blurred". The decomposition in (6) is unique is called "wavelet decomposition" ; and the "blur" is measured in terms of the variation(or frequency or number of cycles per unit length) of f_{N-M}.

A less efficient "stopping criterion" is to require $\|f_{N-M}\|$ to be smaller than some threshold.

Now an algorithm approach for expressing f_N as a direct sum of its components $g_{N-1}, g_{N-2}, \ldots, g_{N-M}$ and f_{N-M} and recovering f_N from these components

Since both the scaling function $\Phi \in V_0$ and the wavelet $\psi \in W_0$ are in V_1 and since V_1 is generated by $\Phi_{1,k}(x) = 2^{\frac{1}{2}}\Phi(2x - k); k \in Z$

There exists two sequences $\{p_k\}$ and $\{q_k\} \in l^2$ such that

$\qquad \Phi(x) = \sum_k p_k \Phi(2x - k)$ ----------------- (7)

$\qquad \varphi(x) = \sum_k q_k \Phi(2x - k)$ ----------------- (8)
For all x $\in R$

The formulas (7) and (8) are called the "two scale relation " of the scaling function and wavelet respectively. Since both $\Phi(2x)$ and $\Phi(2x - 1)$ are in V_1 and $V_1 = V_0 \oplus W_0$

There are four l^2 sequences which are denoted by $\{a_{-2k}\}, \{b_{-2k}\}, \{a1_{-2k}\}, \{b1_{-2k}\}, k \in Z$

Such that

$\Phi(2x) = \sum_k [a_{-2k}\Phi(x - k) + b_{-2k}\varphi(x - k)]$-------- ----- (9)

$\Phi(2x - 1) = \sum_k [a_{1-2k}\Phi(x - k) + b_{1-2k}\varphi(x - k]$---- (10)

This two formulas (9) and (10) can combine into a single formula:

$\Phi(2x - l) = \sum_k [a_{l-2k}\Phi(x - k) + b_{l-2k}\varphi(x - k]$, $l \in Z$ ---- (11)

Which is called 'decomposition relation' of Φ and φ.

Now two pairs sequences $(\{p_k\}, \{q_k\})$ and $(\{a_k\}, \{b_k\})$ all of which are unique due to direct sum relationship. $V_1 = V_0 \oplus W_0$

These sequences are used to formulate the following reconstruction and decomposition algorithm. Here $\{p_k\}$ and $\{q_k\}$ are called reconstruction sequences, while $\{a_k\}$ and $\{b_k\}$ are called decomposition sequences.

CONCLUSION

The lifting scheme which is an efficient computational scheme to compute wavelet transform can be derived directly from the polyphase matrix; it can be applied in much more general situation than classical wavelet filter banks. Therefore, we introduce it via an alternative approach to wavelets (subdivision schemes) which will give a framework in which it is easier to consider more general situations.

The signal is split into a low pass and a band pass part by the filters $\widetilde{H}$ and $\widetilde{G}$ respectively. This corresponds to computing the V_N and the W_N part of a signal in V_{N+1}. The results are then sub sampled, to remove redundancy. Next a (primal) lifting step is executed with the filter. At the synthesis side the same operations are undone in opposite order to obtain original signal.

References

[1] Charles K. Chui: An introduction to wavelets: MRA, Wavelet Decompositions and Reconstructions. Academic Press San Diego New York Boston 1992

[2] Anthony Teolis: Computational Signal Processing with Wavelets: what is wavelet? Printed and bound by Hamilton printing New York.

[3] Adhemar Bultheel: Wavelets with applications in signal and image processing: Filters and Filter Bank.

[4] Richard R. Goldberg: Fourier Transform: The Fourier Transform on L.Cambridge at the University Press 1970.