WAVELETS AND MULTIRATE DIGITAL SIGNAL PROCESSING

Lecture 1: Introduction

Prof.V.M.Gadre, EE, IIT Bombay

1 Introduction

This lecture introduces the subject of wavelets and multirate digital signal processing. It provides an inspiration to understand this subject at a greater depth.

The subject of wavelets follows a basic exposition to subjects like signal analysis, system theory and digital signal processing. It could be considered as an advanced course on signal processing. However, this does not imply that the concepts introduced in this course are difficult to understand. In fact, the concepts are easier than a basic course on signal analysis. The basic course had introduced the idea of abstraction i.e. abstraction of signals, systems, transforms, analysis in different domains etc. On the other hand, wavelets bring us closer to reality. In this sense, the course is very easy. In a basic course, we assume that the signals last forever. For example, while calculating the Fourier transform, we represent any signal in terms of basis functions and these basis functions last from $t = -\infty$ to $t = +\infty$, where t denotes time. However, no signal in this world can last forever. Thus, we should deal with signals in finite domains.

Example 1

In fact, we understand finite domains very well, if that finite domain is the natural domain. To reflect more on this, consider an example of a piece of an audio signal which is finite in time. Here, time is the **natural domain**. From a signal processing perspective, we wish to find the **content** in that audio signal, by enhancing some parts of that signal and suppress others. We may even be interested in characterizing the system. But for doing all these things, we deal with finite time signal and the abstraction of everlasting signals is unnecessary.

Example 2

Consider another example, in which we have a picture of a 'face'. In this case, the **natural domain** is space and it is 2-Dimensional. The face has various features like eyebrows, forehead, nose, lips etc. Suppose we wish to isolate a particular feature, say lips. This requires localization in the spatial domain. Here, again, the amount of data is finite.

Example 3

Another example which explains localization is a piece of audio in which a number of notes are sung. It may be called as a 'raga' in Indian tradition and the notes may be called as the components of the raga. Now, we aim to make a system that takes the rendition of this 'raga' and identifies the notes that compose it. To achieve this, we need to segment the signal in time. For example, the first note may be played for 1 second, the second note for 0.5 seconds and so on. This demands segmentation in time. Moreover, we should also understand that all the notes are not of fixed lengths. The length of the time segment is also important. But more important is to understand the concept of 'Notes' in the signal processing context.

In the basic course, we were exposed to the idea of frequency domain. We know that signals have embedded inside them, a collection of sine waves. These sine waves are continuous for continuous time signals while they are sampled for discrete sequences. Thus most reasonable signals can be thought of as a collection of sine waves. In principle, if the signal is not periodic, we require an infinity of sine waves whose frequency ranges from 0 to ∞. For periodic signals, we have a discrete set of sine waves, possibly finite or possibly infinite. Thus, a different domain is more useful to analyze the signal. Now if we query about the 'Notes' in the raga, it is equivalent to asking the frequency content in the audio piece *i.e.* what points on the frequency axis are occupied by this note? What are the locations where the transform is prominent? This is one of the most fundamental inspirations to study wavelets. Before continuing with this example, it is worthwhile to introduce the term **'wavelets'**.

1.1 Wavelets

Fourier transforms deal with sine waves. Sine waves have many nice properties. Firstly they occur naturally. For example, an electrical engineer recognizes sine wave as naturally emerging from an electricity generation system. Secondly, sine waves are the most analytic, the smoothest possible periodic functions. They also have the power to express many other waveforms *i.e.* they form a very good basis. Addition of two sine waves of the same frequency but with possibly different amplitudes and phases, gives a sine wave of the same frequency with possibly different amplitude and phase. A sine wave on differentiation or integration is a sine wave of the same frequency. Any linear combination of all these operations on a sine wave results in a sine wave of the same frequency. But the biggest drawback of sine waves is that they need to last forever. If the sine wave is truncated (a one sided sine wave, for example), the response to this signal by a system, in general, is different from the response which would be obtained if the signal would be a sine wave from $t = -\infty$ to $t = +\infty$. There would be transients which are not periodic. All the beautiful properties mentioned above, are no longer valid. So, if we need to apply the principles studied in a basic course, we need something unrealistic i.e. a sine wave which lasts forever. To be **more realistic** in our demands, it is **appropriate to deal with wavelets** rather than waves. **Wavelets are waves that last for a finite time, or more appropriately, they are waves that are not predominant forever**. They may be significant in a certain region of time and insignificant elsewhere or they might exist only for finite time duration. For example, a sine wave that exists only between $t = 0 \; and \; t = 1 \; msec$ is, in principle, a wavelet (though not a very good one), a wave that doesn't last forever.

Example 3 revisited

Now, we go back to the example of audio clip. The audio clip comprises of many notes of varying time lengths and hence requires time segmentation. On the other hand, identifying notes in a particular time segment, involves segmentation in the frequency domain. Thus we are asking for a **simultaneous localization** in time and frequency domain. But the **uncertainty principle** in nature puts constraints on this simultaneous localization beyond a point. In signal processing, we call it **uncertainty in time and frequency domain**. Thus, the resolution in the time domain is increased at a compromise in the resolution in the frequency domain. It could be even intuitively argued that, for a shorter audio clip, it is more difficult to identify a note in that time segment than a note played for a longer period of time. But, what is not intuitive from this discussion is that we cannot go down to identifying one particular frequency precisely. So, if we wish to come down to a point on the time axis, we need to spread all over the frequency axis and vice versa. This is the stronger version of this principle. However the weaker version is more subtle. Even if we select a time region on the time axis and ask for the region of frequencies which are predominant in that time region, even then there is a restriction on the simultaneous length or measure of the time and frequency regions. In fact, the more we focus in time, the less we focus in frequency. This could be best explained by examples.

1.2 A few examples

Example 1

Consider a mobile communication system in which a bit stream is transmitted at 1Mbps. If the bit interval is uniform, the time interval for 1 bit is 1 μsec. This indicates segmentation in time. Now, consider that there are two mobile operators operating in a given region. All the users in the region are using mobile from any of the two operators. To avoid interference, these operators should be separated in some domain. They cannot be separated in the time domain since the users of different operators can use the mobiles simultaneously. So the separation could be in frequency domain. Thus, every operator is allocated a particular bandwidth *i.e.* a region in the frequency domain so that the users of that operator can operate in that frequency region only. This is indicative of segmentation in frequency. Thus, in a mobile communication case, there is a desire to localize in time and frequency simultaneously *i.e.* transmitting a bit in a time interval of 1 μsec, indicating a localization in time and transmitting in a particular frequency region only indicating localization in frequency. Thus, this is an example of simultaneous localization in time and frequency.

Example 2

Consider another example of a biomedical signal, say ECG signal (electro-cardiographic waveform) in which various features of the ECG signals are analysed. There are various segments in a typical ECG signal which are often indexed by letters say P, Q etc. These segments are of unequal length.. In fact, biomedical engineers often talk about what are called as evoked potentials. They give stimulus to a bio-medical system and evoke a response and the waveform corresponding to that response is called an evoked potential. An evoked potential typically has fast-varying parts in the response and slow-varying parts in the response. Obviously, the slower parts of the response are predominantly located in the lower ranges of the frequency region while the quicker parts of the response are predominantly located in the higher ranges of the frequency. To isolate the quicker parts of the response, is it sufficient to pass the signal through a conventional high pass filter? Here arises the time frequency conflict. Indeed, it is not sufficient. In fact we need a different perspective on filtering. We need to identify, in different parts of the time axis, which regions of the frequency axis are predominant and then identify different parts of the frequency axis that need to be emphasized in different time ranges. This is yet another example of time frequency conflict. In a basic course, we understand the domains very well because we keep them apart. But one normally needs to consider the two domains together and when we try to do so, there is a fundamental conflict.

2 Brief outline of the course

This course will start with a particular tool to analyze signals, *i.e.* the **Haar Multiresolution Analysis**. Haar, a mathematician, proposed a dual of the idea of Fourier analysis. In Fourier analysis, we represent even discontinuous or non-smooth waveforms into a linear combination of extremely smooth functions namely the sine waves. Haar proposed the idea of taking smooth functions and convert them into a linear combination of effectively discontinuous functions. For example, data in a digital communication system, *e.g.* an audio signal, image, or a video, is transmitted it with a large level of discontinuity. To record a digital audio, we first sample the signal then we digitize the signal, and then record it. All these are highly discontinuous operations. In fact we are not only, forcibly, introducing discontinuity in time but also in amplitude. Thus, representing a smooth audio signal into a discontinuous bit stream is very beneficial to digital communication system. In fact a digital recording is even better than an analog recording. Therefore, the first few lectures will look at one whole angle of wavelets and multirate digital signal processing based on the principles that Haar propounded. Understanding Haar multiresolution analysis(MRA) in depth leads to better understanding of many of principles of wavelets and multirate processing, specifically the two band processing. After Haar MRA, the sequence would be

- The Daubechies family of multiresolution analysis.(Daubechies is the name of the mathematician who proposed this family of multiresolution analysis).
- The uncertainty principle, fundamentally and in terms of its implications.
- The Continuous wavelet transform(CWT). In the Haar multiresolution analysis, we have a certain discretization in the variables associated with the wavelet transform. In the continuous wavelet transform, the variables become continuous.
- Some of the generalisations of the ideas like **'wave packet transform'**, variants of wavelet transforms.
- In the last phase of the course, we shall look at some of the important applications, where wavelets and multirate digital processing provide great advantages.

3 Multirate digital signal processing

Let us look at some of the developments in the subject of multirate signal processing. The connection with wavelets will also be seen. Consider the biomedical example discussed previously. The biomedical signal has fast-varying parts and slow-varying parts of the response. The slow-varying parts of the response are likely to last for a longer period of time while the quicker parts of the response are likely to last for a shorter period of time. So, apart from considering localization in time and also in frequency, the localization required for higher frequencies and lower frequencies is also important. Lower frequencies in response have lower time resolution. Resolution means the ability to resolve, the ability to be able to identify specific components. How can the frequency axis be narrowed down? More appropriately, how much do we need to narrow down? It is often desired (not always) that the higher frequency components be compromised on frequency resolution but not on time resolution. So things that are transient, demand time resolution and things that occupy lower frequency ranges, demand frequency resolution. Hence for increasing frequencies, more frequency resolution is needed opposed to time resolution. This brings the idea of multirate processing. So, if we have higher frequencies, we should use smaller sampling interval and vice versa, in a discrete processing system. This leads to more efficient processing operations. In an evoked potential response, frequent sampling is not required for the lower frequency components. It increases data without any advantage. On the other hand, while handling quicker components, if our sampling rate is less or inadequate, 'aliasing' is introduced. Hence, same sampling rate is not used for all the frequency components. Unlike a basic course, we deal here with sequences that are obtained with different sampling rates in the same system. This brings the idea of multirate digital signal processing.

3.1 Filter banks

Going further, the idea of filter banks is important as opposed to filters. In a biomedical example, to separate components, many different operators have to be simultaneously used. So a system of filters is needed, which has certain individual characteristics as well as collective characteristics. Thus, analysis as well as synthesis is required. In addition to that, we also require localization. So a bank of filters, refers to a set of filters which either have a common input or a common point of output. This concept of a bank of filters, and in fact, two bank filters namely analysis and synthesis filter bank, taken together, is very central to multirate signal processing. So a two band filter bank will be studied in this course at a greater depth. The concept of two band filter bank is of great importance to be able to construct wavelets. In fact from the Haar multiresolution analysis example, we will see that an intimate relationship exists between the Haar wavelet and a two band Haar filter bank.