WAVELETS AND MULTIRATE DIGITAL SIGNAL PROCESSING

Lecture 37: Modulation Analysis & 3-band filter bank, Application

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

1 Previous Lecture

In the previous lecture we have discussed one way to analyze the general system with analysis and synthesis filters, namely the approach of polyphase decomposition. Essentially polyphase decomposition is a time decomposition approach where we recognize all sequence in question whether the input sequence, the output sequence, filter impulse response, could be thought of comprising of as many subsequence as the number by which the sequence is decimated and interpolated *i.e.,* downsampling factor and upsampling factor. For example, if we have downsampling and upsampling by 2, we think of odd and even number point on all sequence of interest. Based on this decomposition we identify relation between output polyphase and input polyphase component through filter polyphase component.
Naturally this is difficult to do in time domain, hence we use Z-domain. We also noted that condition of perfect reconstruction amounts to a condition on product of polyphase matrix correspond to analysis and synthesis side.

2 Final step in polyphase approach

$$\begin{bmatrix} \text{Output Polyphase} \\ \text{Vector} \end{bmatrix} = \begin{bmatrix} \text{Synthesis} \\ \text{polyphase} \\ \text{Matrix} \end{bmatrix} \begin{bmatrix} \text{Analysis} \\ \text{Polyphase} \\ \text{Matrix} \end{bmatrix} \begin{bmatrix} \text{Input} \\ \text{Polyphase} \\ \text{vector} \end{bmatrix}$$

All the matrices are of order M, which is the factor of downsampling and upsampling. Product of synthesis and analysis polyphase matrices of order M is equal to square matrix of size $M \times M$.
Question to be answer is that what should this matrix be for perfect reconstruction.
For perfect reconstruction we require:

$$Y(Z) = C_0 Z^{-D} X(Z) \tag{1}$$

If we decompose $Y(Z)$ and $X(Z)$ we get,

$$Y(Z) = \sum_{K=0}^{M-1} Z^{-K}\, Y_{K,M}(Z^M) \tag{2}$$

$$X(Z) = \sum_{K=0}^{M-1} Z^{-K}\, X_{K,M}(Z^M) \tag{3}$$

Therefore,

$$\sum_{K=0}^{M-1} Y_{K,M}(Z^M)\, Z^{-K} = C_0 Z^{-D} \sum_{K=0}^{M-1} X_{K,M}(Z^M) Z^{-K} = C_0 \sum_{K=0}^{M-1} Z^{-(D+K)} X_{K,M}(Z^M) \tag{4}$$

We need to separate $D+K$: D is fixed for all K, $D+K$ essentially carries out a rearrangement of polyphase component.
For *e.g.*, $M = 3$, $D = 5$.

K	K+D	
0	5	≡2
1	6	≡0
2	7	≡1

Therefore 0^{th} number polyphase component of input is mapped to 2^{nd} number polyphase component of the output. Similarly, 1^{st} and 2^{nd} number polyphase component of input are mapped to 0^{th} number and 1^{st} number polyphase component of the output respectively.
What we have is the cyclic rearrangement of the polyphase component and therefore, if we look at product matrix what we require is that every row and column must have only one non zero entry *i.e.*, if we take every row and every column there is exactly one non zero entry and that is identical. The constant factor in each of the entry is C_0 and the additional delay depends on D.

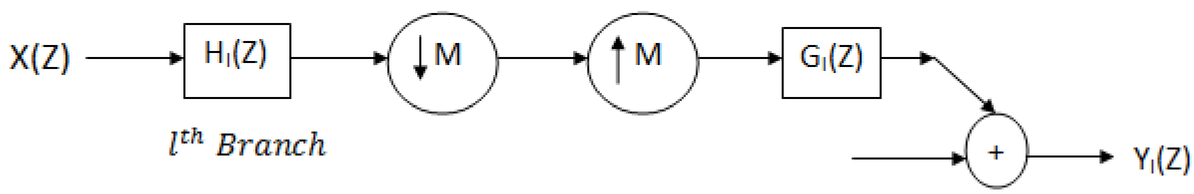
2.1 Summary of the method

For perfect reconstruction,
synthesis matrix $\times$ *Analysis matrix* = *(following form) Each row and column has exactly one entry of the form* $C_0 Z^{-L}$.
where, L depends on D.
For *e.g.*, $D = 5$, $M = 3$
Polyphase matrices written in Z^3 , then $L = 3$.
Polyphase matrices written in Z , then $L = 1$.

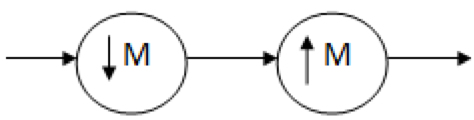
Think about: If we have to obtain perfect reconstruction but if the analysis synthesis system together becomes linear shift invariant system, that mean one can equivalently treat output as a result of input been acted upon by a LSI system with a certain transfer function. What can we say about the entries of this product matrix when this LSI invariance is present in over all analysis synthesis structure? Can we attribute a certain structure to this product matrix?

3 Modulation Approach

It is frequency domain approach. Consider one of the branch:



We wish to establish a relation across this branch. All these branches with different $G_l(Z)$ will come together in summation to form $Y_0(Z)$. Now look at



This is essentially a multiplication by a periodic sequence $P_M[n]$

$$P_M[n] = 1, \qquad n\ is\ a\ multiple\ of\ M$$
$$= 0, \qquad else$$

Now, in modulation approach we think of the process of downsampling followed by upsampling as modulation by sequence and that sequence is broken into its component sequence each of which is exponential.
Consider one period *i.e.*, $P_M[n]$ restricted to $0, 1, ...(M-1)$. Obtain its DFT.

$$\tilde{P}_M[k] = \sum_{n=0}^{M-1} P_M[n] W_M^{-nk} \quad ; W_M = e^{j\frac{2\pi}{M}} \tag{5}$$

This $\tilde{P}_M[k]$ is dot product of one period of the sequence with the exponential.

$$\tilde{P}_M[k] = 1, \quad k = 0, 1, ...(M-1) \tag{6}$$

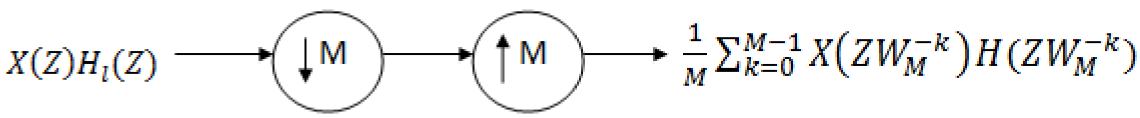
Take Inverse Fourier Transform.

$$P_M[k] = \frac{1}{M} \sum_{k=0}^{M-1} 1.W_M^{nk} \quad ; \forall n \tag{7}$$

Now,

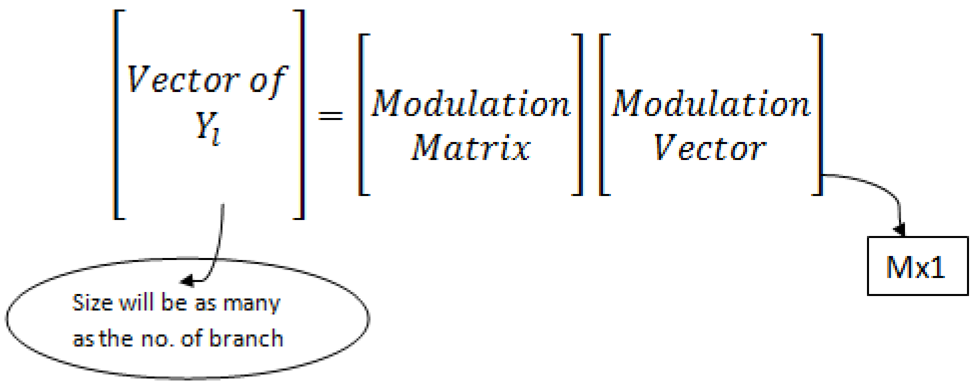
$$X(Z) \rightarrow [H_l(Z)] \rightarrow X(Z)H_l(Z) \overset{Modulated\ by}{\rightarrow} \frac{1}{M} \sum_{k=0}^{M-1} W_M^{nk} \tag{8}$$

When we modulate a sequence by α^n, $Z \leftarrow Z\alpha^{-1}$ in the Z-transform. If we use this property repeatedly and note that Z-transform is the linear operator



So we have M modulates of input being acted upon by corresponding M modulates of Analysis filter.

$$Y_l(Z) = G_l(Z) \sum_{k=0}^{M-1} X(ZW_M^{-k}) H_l(ZW_M^{-k}) \tag{9}$$



l^{th} row of modulation matrix =

$$G_l(Z)[H_l(ZW_M^{-0}).H_l(ZW_M^{-1}).H_l(ZW_M^{-2})......H_l(ZW_M^{-(M-1)})] \tag{10}$$

For perfect reconstruction we first want alias cancellation.
Alias cancellation means, no contribution from $X(ZW_M^{-k})$, $k \neq 0$.
Essentially we ask for: First column of modulation matrix is the only Non-Zero column. This is very stringent requirement. It is sufficient but not necessary. A more general condition is: *Sum of columns in the modulation matrix = 0* $\forall k \neq 0$.
Example: $M = 3$ and 3 channels.

$$\begin{bmatrix} G_0(Z) & 0 & 0 \\ 0 & G_1(Z) & 0 \\ 0 & 0 & G_2(Z) \end{bmatrix} = \begin{bmatrix} H_0(Z) & H_0(ZW_3^{-1}) & H_0(ZW_3^{-2}) \\ H_1(Z) & H_1(ZW_3^{-1}) & H_1(ZW_3^{-2}) \\ H_2(Z) & H_2(ZW_3^{-1}) & H_2(ZW_3^{-2}) \end{bmatrix}$$

What we need is this:

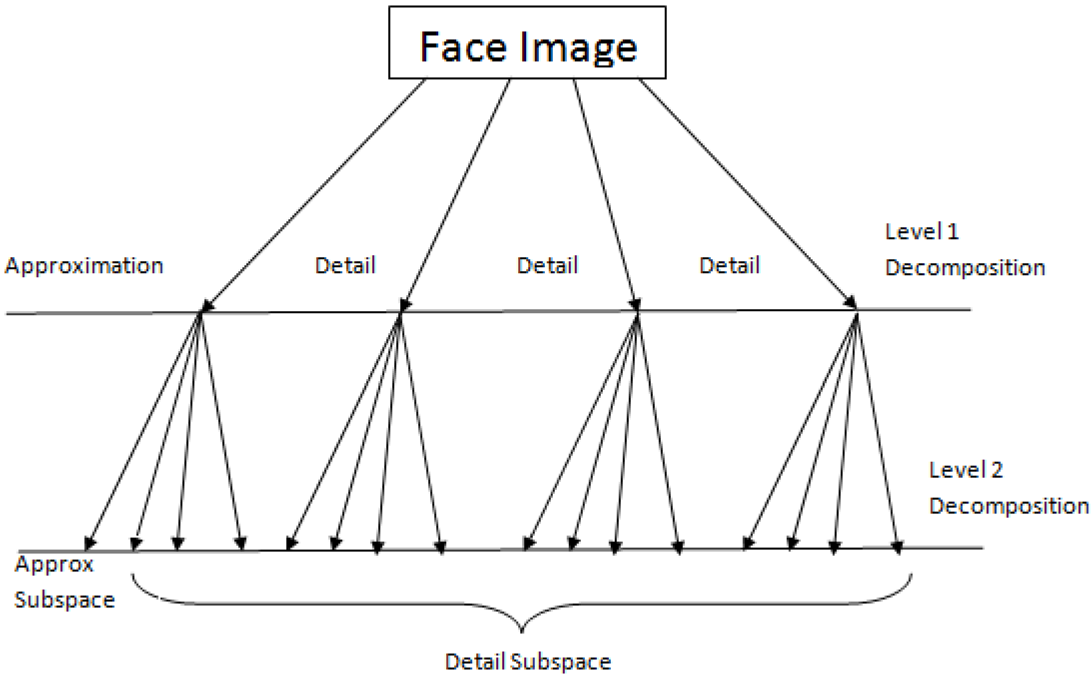
$$\sum_{l=0}^{2} G_l(Z)H_l(ZW_3^{-k}) = 0, \quad k = 1, 2 \tag{11}$$

Think about: Consider the ideal 3-band 3-channel filter bank where analysis and synthesis filters are each triple band filter. So low pass filter is ideal filter with passband from 0 to $\frac{\pi}{3}$, the middle filter is bandpass filter with passband from $\frac{\pi}{3}$ to $\frac{2\pi}{3}$ and last is ideal high pass filter with passband from $\frac{2\pi}{3}$ to π. Work out modulation terms explicitly.

4 Application: Face recognition using wavelet packet analysis

It is a application from image processing. Face recognition is one of the biometric authentication problem. It incorporates change in signal so faces can be changed abruptly hence it has limited application in biometric authentication. It could be better used in activity analysis. Let say we are looking at Surveillance application in which we have to map a particular region in which no person is allowed. In that case if we use face detection and recognition, we can get a very good hand over a particular region. So we can say if somebody enters at a particular time and what he is doing. This is activity tracking and abnormality detection.
The basic difference between face detection and face recognition is that in face detection we look of similar features in all the faces while in face recognition we look for features which are different across the faces. In this application we deal with face recognition, so we assume that some region of face which contains image is given.
A typical face contains nose, mouth, *etc.* These are in focus when we look in different scales in different subspaces because they contain different frequencies as well as facial regions. So we need to go for a decorrelation in spatial as well as frequency domain. Therefore this is the signal where we need to look into time as well as frequency localization and that is why we use wavelets as they give better representation in terms of time and frequency. With wavelet packet analysis we not decompose the low frequency band but also the high frequency band. this gives richer representation of the face image. So there are two things here, we are decorrelating face in time as well as frequency and we are getting better face representation. So given an image we need to decompose it using wave packet analysis and then we can represent it by features

and these features are well decorrelated so we can do good classification.
The only problem stands is with the dimensionality. If we decompose this image into multiple subbands (say 10-15 subbands) , then the data we have for classification is huge which is difficult to manage. So we go for *moment's based approach* and we take only 1^{st} and 2^{nd} moment's mean and variance of feature and then we can classify.



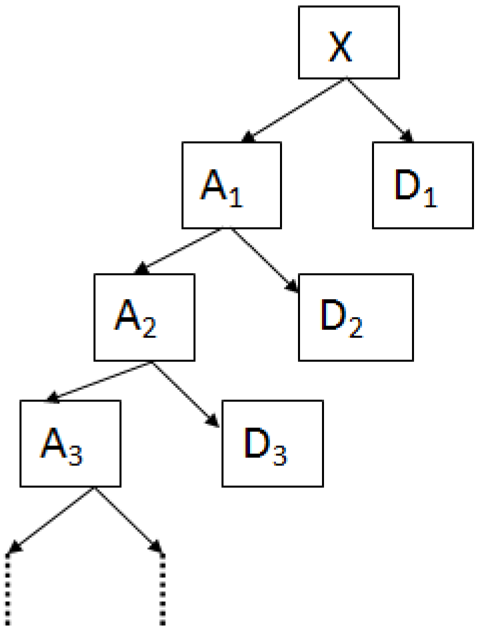
Two-Level Wavelet packet decomposition

Face image is decomposed in approximation and detail subspace in level 1. In level 2 decomposition we can decompose approximation and details of level 1 further and we will get around 16 subspaces in which 1 is approximation subspace and 15 are detail subspace. Now we need to generate features from this, we will use mean and variance only for features representation. When we utilize euclidean distance we generally have features which are not mean and variance together because we are now looking at probability distribution function rather then individual values. So we need a distance matrix which takes care of mean as well as variance so normal euclidean distance will not perform better in that case.
The detailed approach is discussed in subsequent chapter.

5 Application of wavelets in Data Mining

Data mining is ensemble of tool which we use to deal with huge amount of data efficiently for our purpose.
Data mining involves two thing. One is storing the data with help of efficient data structure and second is the retrieval of data from data structure for own purpose. When we have to store and retrieve the data, the user basically interacts with the data structure by the help of some responses and queries. In practice when we have large amount of time series data, the queries that are generally encountered are not point queries but are spread over some larger duration of time. For *e.g.*, if data is stock prices of a company for a particular period of time then queries which we encounter are generally like on which month the stock prices had a rising

trend or on which week stock prices has deep dip.
In order to get this data we need to do some post processing on the raw data and we get the output of this query. This way we interact with the data structure and we get the output.
Wavelet transform on a signal gives us the information at various levels of abstraction and with various translates and various scales. Raw data can be thought of a sequence or signal *i.e.*, daily stock price as a signal than depending upon the response or queries we can say that we are interested on various translates in the data such as we are interested in the data 2 months back or 6 months back. This 2 months or 6 months are different translates. Also we are interested in a data at different scales such as we are interested in a weekly data 2 months back, therefore 2 months become a translates and week becomes a scale. Wavelet can be used efficiently to deal with such problems. In this case we need very little post processing to analyze this type of data.



This is trend and surprise abstraction tree (TSA tree). Here 'X' is original sequence, 'A' and 'D' are are Discrete wavelet transform of the parent node (X).
As we go down this tree, at each level we increase our abstraction by one level and we need a trends or surprise at any level, say at 2^{nd} level, first we need to extract out A_2 node and do post processing to get trend information corresponding to this level and similarly extract D_2 node and post process to get the surprise information corresponding to that level. The next part is to implement this in memory efficient manner. For this we need to store the leaf nodes (*i.e.*, D_1 , D_2 ....D_n A_{n+1}) and we can extract any other node.
Node dropping and coefficient dropping which talks about further compressing the leaf node information to get a better memory efficient implementation of this method.