Module 4

Multi-Resolution Analysis

Lesson 12

Multi-resolution Analysis: Discrete Wavelet Transforms

Instructional Objectives

At the end of this lesson, the students should be able to:

1. Define Discrete Wavelet Transforms (DWT) and its inverse.
2. Compute DWT and inverse DWT through subband coding and decoding.
3. Extend the DWT concepts to two dimensions.
4. Apply DWT and inverse DWT on images.
5. Explain the block diagram of a DWT-based still image compression and coding system.

12.0 Introduction

In lesson-10, we had presented the basic theory of wavelets and defined the scaling and the wavelet functions for approximation and detailed analysis respectively. In lesson-11, we presented the concepts of subband coding and decoding. In this lesson, we shall attempt to relate the scaling and wavelet functions to subband coding. In lesson-10, while discussing wavelet series, we had seen how to express a continuous function in terms of the scaling and the wavelet functions, which were also continuous and only their scaling and translations were discrete. In practice, we are required to deal with discrete signals, as in digital images and hence, the concepts of wavelet series get modified into its discrete form, in terms of the Discrete Wavelet Transforms (DWT). After introducing its definition and its mathematical forms in the present lesson, we shall present a multi-scale filter-bank structure to compute the DWT at two or more scales. It is shown that the filter-bank structure resembles the analysis-synthesis filter bank structures for subband coding. In this lesson, we shall also see how to apply DWT on two-dimensional signals, i.e. images. In recent years, DWT applied on images have attracted considerable attention, after the most recent still image compression standard JPEG-2000 have adopted DWT as the transform. The DWT coefficients across the subbands bear a definite spatial relationship, which we need to utilize for space-frequency localization and encoding of DWT coefficients.

12.1 Discrete Wavelet Transforms and its inverse

Let us refer back to Section-10.4 on wavelet series. There, we had expressed a continuous one-dimensional signal f(x) in terms of a series expansion consisting of continuous scaling and wavelet functions, as given in equation (10.14). Now, instead of considering a continuous signal, if we consider a discrete sequence $s(n)$, defined for $n = 0,1,\cdots\cdots$, the resulting coefficients in the series expansion are called the *Discrete Wavelet Transforms (DWT)* of $s(n)$. The coefficients of series

expansions shown in equations (10.15) and (10.16) get modified for discrete signal to obtain the DWT coefficients, given by

$$W_{\varphi}(j_0,k)=\frac{1}{\sqrt{M}}\sum_{n}s(n)\varphi_{j_0,k}(n) \quad \text{.....(12.1)}$$

$$W_{\psi}(j,k)=\frac{1}{\sqrt{M}}\sum_{n}s(n)\psi_{j,k}(n) \quad \text{.....(12.2)}$$

where, $j \geq j_0$ and $s(n)$, $\varphi_{j_0,k}(n)$ and $\psi_{j_0,k}(n)$ are functions of discrete variables $n=0,1,\cdots\cdots,M-1$. Equation (12.1) computes the approximation coefficients and equation (12.2) computes the detail coefficients. The corresponding Inverse Discrete Wavelet Transform (IDWT) to express the discrete signal in terms of the wavelet coefficients can be written as

$$s(n)=\frac{1}{\sqrt{M}}\sum_{k}W_{\varphi}(j_0,k)\varphi_{j_0,k}(n)+\sum_{j=0}^{\infty}\sum_{k}W_{\varphi}(j,k)\psi_{j,k}(n) \quad \text{..... (12.3)}$$

Normally, we let $j_0=0$ and select M to be a power of 2 $\left(M=2^J\right)$, so that the summations are performed over $j=0,1,\cdots\cdots,J-1$ and $k=0,1,2,\cdots\cdots 2^{j-1}$.

12.2 Computing DWT and IDWT through subband analysis and synthesis

We are now going to show that DWT and subband decomposition have a strong relationship and that it is possible to compute DWT through subband decomposition. To show this relationship, we first substitute the definition of wavelet functions $\psi_{j,k}(n)$, given in equation (10.7) of section-10.3 (lesson-10) into the DWT equation (12.2) and obtain

$$W_{\psi}(j,k)=\frac{1}{\sqrt{M}}\sum_{n}s(n)2^{j/2}\psi\left(2^j n-k\right) \quad \text{..... (12.4)}$$

We rewrite the expression for wavelet function in terms of weighted sum of shifted, double-resolution scaling function, given in equation (10.11) as

$$\psi(n)=\sum_{r}h_{\psi}(r)\sqrt{2}\varphi(2n-r) \quad \text{..... (12.5)}$$

where, the real variable x is replaced by the integer variable n and r is the shift. If we scale n in equation (12.5) by 2^j, translate it by k and let $m=2k+r$, we obtain

$$\psi\left(2^{j}n-k\right)=\sum_{r}h_{\psi}(r)\sqrt{2}\varphi\left(2\left(2^{j}n-k\right)-r\right)$$
$$=\sum_{m}h_{\psi}(m-2k)\sqrt{2}\varphi\left(2^{j+1}n-m\right) \quad \text{..... (12.6)}$$

Substituting the above expression in equation (12.4), we obtain

$$W_{\psi}(j,k)=\frac{1}{\sqrt{M}}\sum_{n}s(n)2^{j/2}\left[\sum_{m}h_{\psi}(m-2k)\sqrt{2}\varphi\left(2^{j+1}n-m\right)\right] \quad \text{..... (12.7)}$$

Interchanging the order of summation, we obtain

$$W_{\psi}(j,k)=\sum_{m}h_{\psi}(m-2k)\left[\frac{1}{\sqrt{M}}\sum_{n}s(n)2^{(j+1)/2}\varphi\left(2^{j+1}n-m\right)\right] \quad \text{..... (12.8)}$$

The term written within the square bracket is identified as the DWT equation for the approximation coefficient, given in equation (12.1) with *j* replaced by (*j* + 1) and equation (12.8) can be rewritten as

$$W_{\psi}(j,k)=\sum_{m}h_{\psi}(m-2k)W_{\varphi}(j+1,k) \quad \text{..... (12.9)}$$

Similarly, if we start with the DWT equation (12.1), use the definition of scaling functions, given in equation (10.2) and apply the weighted summation expressions of equation (10.4), it is possible to obtain in a similar manner as above,

$$W_{\varphi}(j,k)=\sum_{m}h_{\varphi}(m-2k)W_{\varphi}(j+1,k) \quad \text{..... (12.10)}$$

Equations (12.9) and (12.10) show a relationship between the DWT coefficients at adjacent scales. Both $W_{\psi}(j,k)$ and $W_{\varphi}(j,k)$ are obtained by convolving the scale (j+1) approximation coefficient $W_{\varphi}(j+1,k)$ with $h_{\psi}(-n)$ and $h_{\varphi}(-n)$ respectively and then subsampling the convolved output by a factor of 2. This is depicted in a block diagram form in fig.12.1.

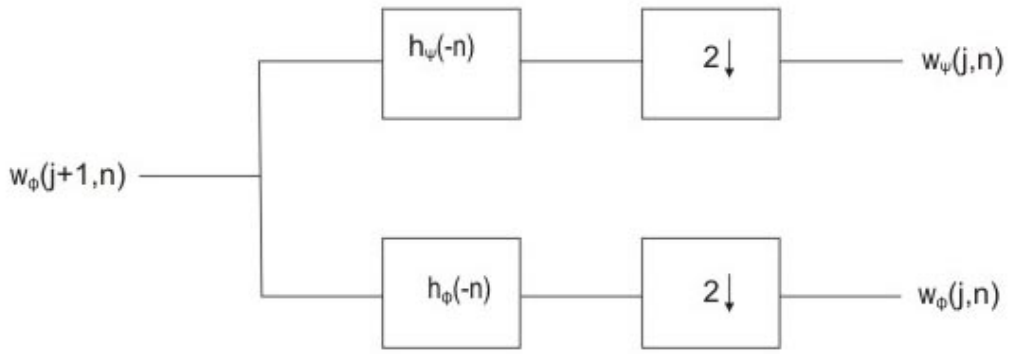

Fig12.1: *Analysis of DWT coefficients (one-stage)*

The block diagram looks identical to that of the analysis filter block diagram for subband coding, as shown in fig.11.1 with $h_0(n) = h_\varphi(-n)$ and $h_1(n) = h_\psi(-n)$. Computation of DWT through the above approach can be iterated to obtain further analysis of the approximation coefficients, as depicted in fig.12.2 for the two highest scales of the transform, where the samples of the original signal itself is assumed to be the highest scale coefficient $W_\varphi(J,n)$.

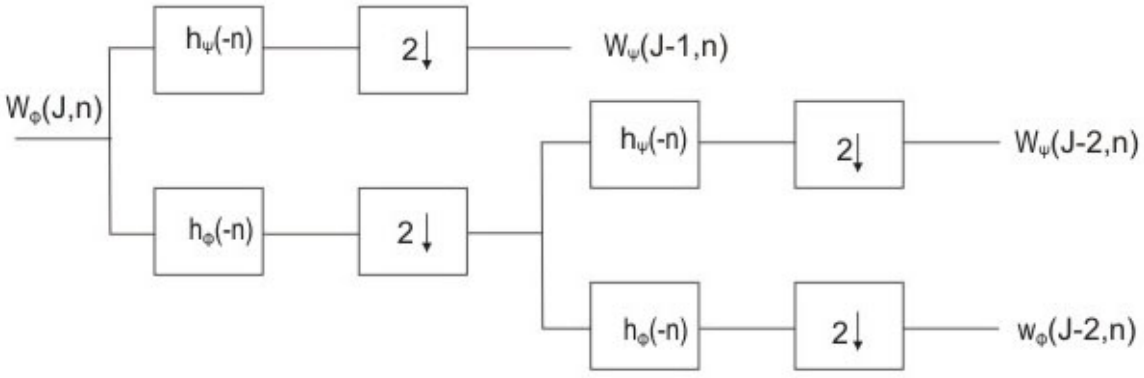

Fig12.2 Analysis of DWT filters (two stage)

The corresponding frequency splitting characteristics for the analysis filter outputs is shown in fig.12.3.

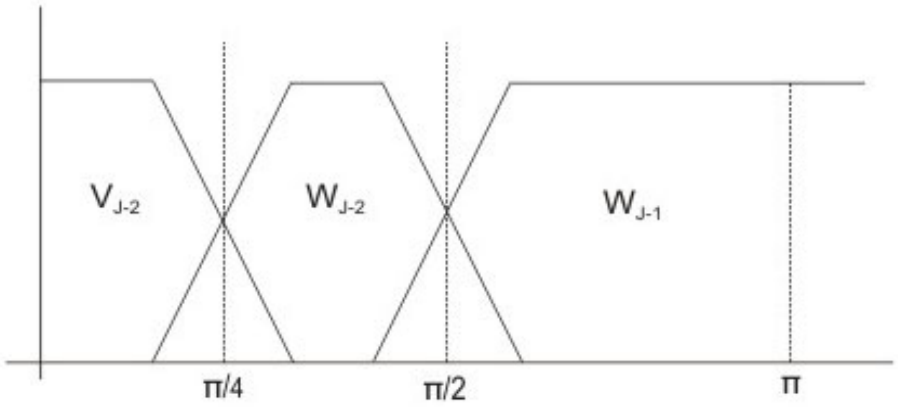

Fig12.3 Frequency Splitting of analysis filter bank

In this figure, V_J corresponds to the scaling space in which the original signal s (n) resides. V_{J-1} and W_{J-1} correspond to the low-pass (approximation) and the high-pass (details) analysis space of the first filter bank. The space V_{J-1} is further analyzed into V_{J-2} and W_{J-2}.The two-stage filter bank shown in fig.12.2 can be extended to any number of scales.

This approach therefore implements DWT in an efficient way through subband decomposition and is also known as the Fast Wavelet Transforms (FWT). An equally efficient approach exists to implement the Inverse Discrete Wavelet Transforms (IDWT) through synthesis of the subband signals, as shown in fig.12.4.

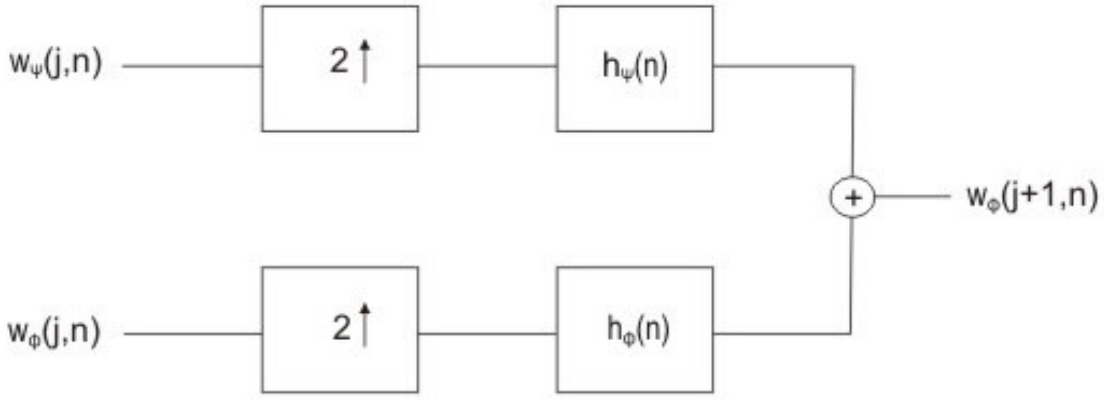

Fig12.4 Synthesis of subband signals (one-stage)

Here, the approximation coefficient $W_\varphi(j+1,n)$ at a higher scale is synthesized from the approximation and detail coefficients $W_\varphi(j,n)$ and $W_\psi(j,n)$ respectively at the next lower scales. IDWT computation through synthesis filter banks is thus the mirror of DWT computation through analysis filter banks.

12.3 Two-dimensional DWT

The concepts of one-dimensional DWT and its implementation through subband coding can be easily extended to two-dimensional signals for digital images. In case of subband analysis of images, we require extraction of its approximate forms in both horizontal and vertical directions, details in horizontal direction alone (detection of horizontal edges), details in vertical direction alone (detection of vertical edges) and details in both horizontal and vertical directions (detection of diagonal edges). This analysis of 2-D signals require the use of following two-dimensional filter functions through the multiplication of separable scaling and wavelet functions in n_1 (horizontal) and n_2 (vertical) directions, as defined below:

$$\varphi(n_1,n_2)=\varphi(n_1)\varphi(n_2) \quad \text{..... (12.11)}$$

$$\psi^H(n_1,n_2)=\psi(n_1)\varphi(n_2) \quad \text{..... (12.12)}$$

$$\psi^V(n_1,n_2)=\varphi(n_1)\psi(n_2) \quad \text{..... (12.13)}$$

$$\psi^D(n_1,n_2)=\psi(n_1)\psi(n_2) \quad \text{..... (12.14)}$$

In the above equations, $\varphi(n_1,n_2)$, $\psi^H(n_1,n_2)$, $\psi^V(n_1,n_2)$ and $\psi^D(n_1,n_2)$ represent the approximated signal, signal with horizontal details, signal with vertical details and signals with diagonal details respectively. The 2-D analysis filter implemented through separable scaling and wavelet functions is shown in fig.12.5.

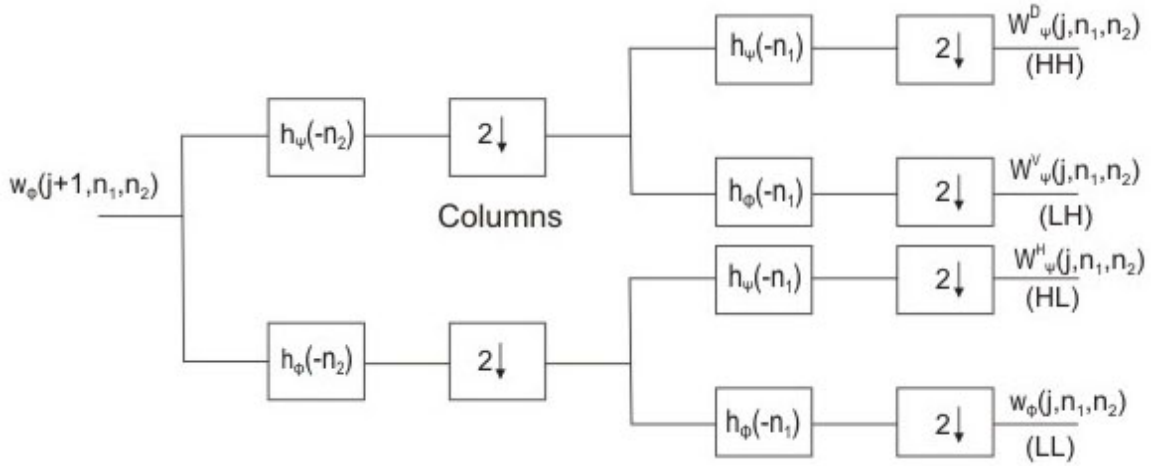

Fig12.5 2-D analysis filtering through separable scaling and wavelet functions.

The filtering in each direction follows subsampling by a factor of two, so that each of the subbands corresponding to the filter outputs contain one-fourth of the number of samples, as compared to the original 2-D signal. The output of the analysis filter banks is the Discrete Wavelet Transformed (DWT) coefficients. Fig.12.6 shows the decomposed subbands of the DWT. The bands $\varphi(n_1, n_2)$, $\psi^H(n_1, n_2)$, $\psi^V(n_1, n_2)$ and $\psi^D(n_1, n_2)$ are also referred to as LL, LH, HL and HH respectively, where the first letter represents whether it is low-pass (L) or high-pass (H) filtered along the columns (vertical direction) and the second letter represents whether the low-pass or high-pass filtering is applied along the rows (horizontal direction). It is possible to iteratively apply the 2-D subband decompositions as above on any of the subbands. Commonly, it is the LL subband (the approximated signal) that requires analysis for further details. Like this, if the LL subband is iteratively decomposed for analysis, the resulting subband partitioning is called the dyadic partitioning. Click on fig.12.6 repeatedly to get the second, third, fourth and so on levels of dyadic partitions. It may be noted that every level of decomposition subsamples the newly created subbands by a factor of two along the rows and columns (that is, by a factor of four) as compared to the previous level of decomposition. Note however that the total number of DWT coefficients considering all the subbands always remains same as that of the total number of pixels in the image. As we go further up in the levels of decomposition, we suffer a loss of resolution in the newly created subbands. This is to say that the first level of decomposition extracts the finest resolution of details; the subbands created in the second level of decomposition extract coarser details than the first one and so on.

Fig.12.6 Dyadic Partitioning of subbands for images

It may be noted that the process of analysis filtering is lossless. It is therefore possible to have a perfect reconstruction of the original 2-D signal (image) by a reverse process of synthesis filtering, as shown in fig.12.7, which is just the mirror of fig.12.5.

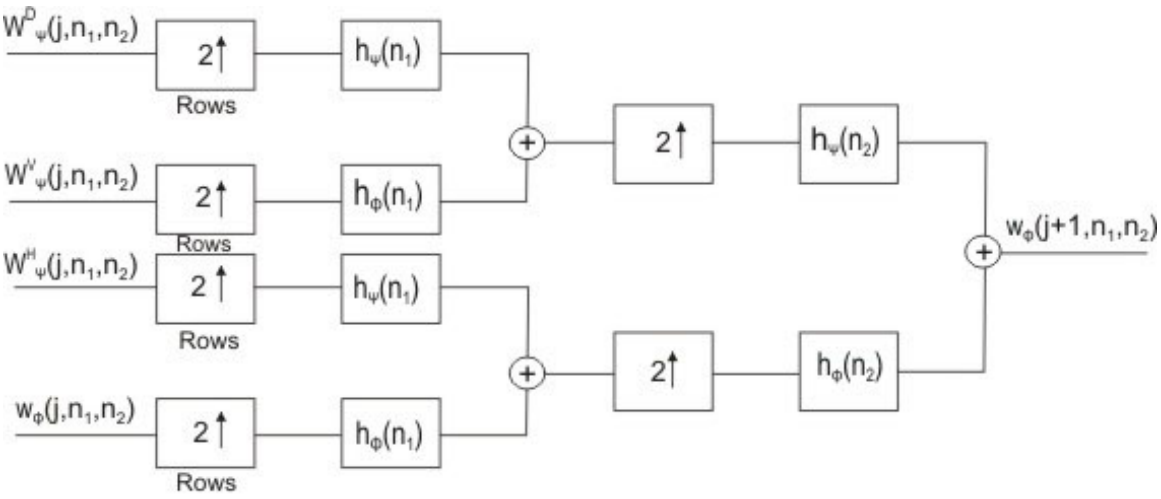

Fig12.7 2-D synthesis filtering for image econstruction

It may be noted that the synthesis filter banks along the rows and columns are associated with an upsampling by a factor of two so that the reconstructed image can be shown at the original resolution. The synthesis filter banks therefore perform the IDWT, which is also lossless, like the DWT.

12.4 Applying DWT and IDWT on images

Although the previous section introduced the idea of performing two-dimensional DWT through an extension of one-dimensional DWT concepts, the exact implementation on digital images is discussed in this section. Based on the definitions of one-dimensional DWT, given in equation (12.1) and (12.2) and the four directionally sensitive separable scaling and wavelet functions defined in equations (12.11) to (12.14), the two-dimensional DWT of an image function $s(n_1,n_2)$ of size $N_1 \times N_2$ may be expressed as

$$W_\varphi(j_0,k_1,k_2) = \frac{1}{\sqrt{N_1 N_2}} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} s(n_1,n_2) \varphi_{j_0,k_1,k_2}(n_1,n_2) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(12.15)}$$

$$W_\psi^i(j_0,k_1,k_2) = \frac{1}{\sqrt{N_1 N_2}} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} s(n_1,n_2) \psi^i{}_{j_0,k_1,k_2}(n_1,n_2) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(12.16)}$$

where $i = \{H,V,D\}$ indicate the directional index of the wavelet function. As in one-dimensional case, j_0 represents any starting scale, which may be treated as $j_0 = 0$. Given the equations (12.15) and (12.16) for two-dimensional DWT, the image function $s(n_1,n_2)$ is obtained through the 2-D IDWT, as given below

$$s(n_1,n_2)=\frac{1}{\sqrt{N_1N_2}}\sum_{k_1}\sum_{k_2}W_\varphi(j_0,k_1,k_2)\varphi_{j_0,k_1,k_2}(n_1,n_2)$$

$$+\frac{1}{\sqrt{N_1N_2}}\sum_{i=H,V,D}\sum_{j=0}^{\infty}\sum_{k_1}\sum_{k_2}W_\psi^i(j,k_1,k_2)\psi^i{}_{j,k_1,k_2}(n_1,n_2) \quad \text{.....(12.17)}$$

The 2-D scaling and wavelet functions used in equations (12.15) to (12.17) can be realized through separable, one-dimensional FIR digital filters of impulse responses $h_\varphi(-n)$ and $h_\psi(-n)$. Various forms of filters are used, for example Haar filters having $h_\varphi(n)=\left\{\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right\}$ and $h_\psi(n)=\left\{\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right\}$.

If the 2-D analysis filter bank, shown in fig.12.5 with Haar filter coefficients is applied on the digital image shown in fig.12.8, what results is a Discrete Wavelet Transformed (DWT) image with first level of decomposition, as shown in fig.12.9.

Fig.12.8 Original Lena Image

Fig12.9 DWT applied on Lena Image after one level of decomposition

It may be noted that the LL, LH, HL and HH subband partitions indicate the approximated image and the images with horizontal edges, vertical edges and diagonal edges respectively. Decomposing the LL subband further to have a second level of DWT decomposition results in fig.12.10.

Fig12.10 DWT applied on Lena Image after two levels of decomposition

The pixels within each subband represent the DWT coefficients corresponding to the subband. By applying IDWT on the DWT coefficients, we obtain the reconstructed image, as shown in fig.12.11.

Fig12.11 Reconstructed Lena image

In this case, the reconstruction is lossless, since the filter coefficients are exactly represented and the DWT coefficients are not quantized. However, in practice, we have to transmit the images in limited bandwidth situations, necessitating compression and hence, the DWT coefficients are to be quantized and efficiently encoded. We shall study such encoding schemes in the next lessons.

12.5 DWT based still image compression system

It may be noted from the results of DWT applied over the images that mostly the LL subband at the highest level of decomposition has significant information content and all other subbands have less significant content. This is expected, since all natural images are rich in low frequency information content, as compared to the high-frequency content and the results demonstrate excellent energy compaction properties of DWT. This property can be very effectively utilized to achieve image compression.

Fig.12.12 shows the block diagram of an image compression and coding system based on DWT. The DWT coefficients are suitably quantized and the quantized coefficients are grouped within the subband or across the subbands, depending upon the coding scheme employed. The grouping basically exploits the self-similarity of DWT across different subbands. The quantization and subband coefficient grouping are integrated in some of the coding schemes. Finally, the quantized and grouped DWT coefficients are encoded through a variable length coder, before generating the encoded bitstream. Fig.12.13 shows the block diagram of the decoder, which does the reversal of the encoder operations to reconstruct the image. Again, exact reconstruction is never possible, since the quantization is employed. However, it has been observed that at very low bit-rate, DWT shows better reconstruction results as compared to block-based DCT at comparable bit-rate and does not suffer from blocking artifacts as in DCT.

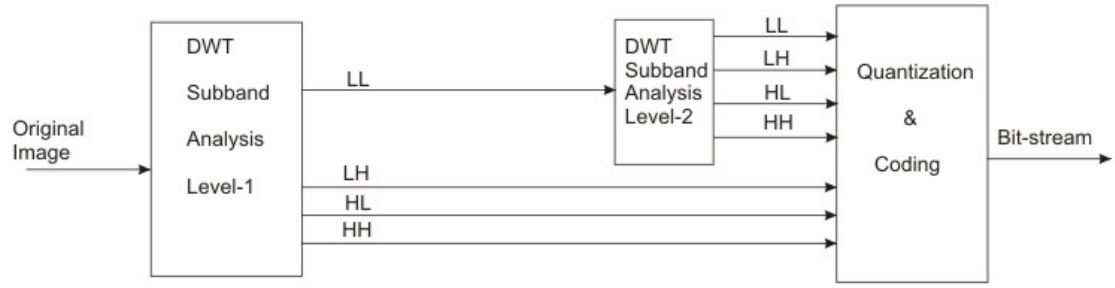

Fig . 12.12 A two-level DWT-based image compression system

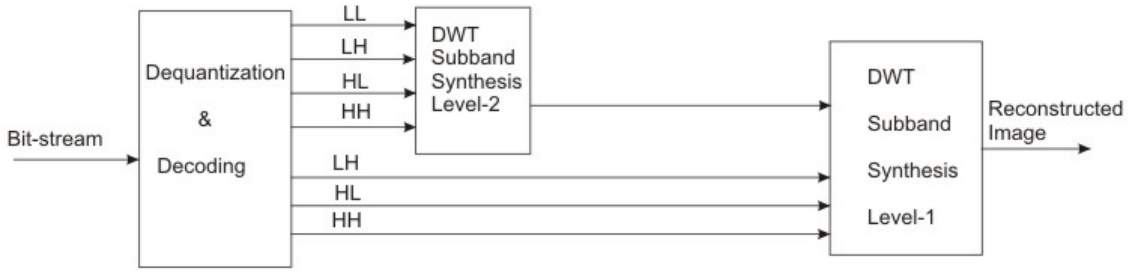

Fig . 12.13 A two-level image decoding system based on DWT

12.6 Conclusion

The previous lessons had introduced us to the theory of wavelets and subband coding. In this lesson, we presented the theory of discrete wavelet transforms and showed how it can be implemented for images through subband decomposition. DWT based still image compression system has been explained and it is seen to improve the coding performance by eliminating the blocking artifacts. The advantages of DWT has resulted in its acceptance as the coding technique in JPEG-2000, the latest standard in still image compression.