Chapter 2

Background

2.1 Basics of Discrete Wavelet Transform (DWT)

In image compression system, the popular and effective ways are based on Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT). The JPEG, MPEG1, and MPEG2 standards are based on DCT, while the JPEG2000 and MPEG4 standards are based on DWT. Both DCT and DWT have their own advantages and disadvantages because of their inborn properties. (For example, one of the main disadvantages of DCT is the so-called "blocking artifact.") Many researches have shown that for higher compression ratios, the compression performance of DWT is better than that of DCT, and the visual picture quality is also better even the PSNR values are the same.

Wavelet transform performs image analysis by multi-resolution. Any decomposition of an image into wavelets involves a pair of waveforms: one to represent the high frequencies corresponding to the detailed parts of an image (wavelet function ψ) and one for the image's low frequencies or smooth parts (scaling function ϕ).

A 2-D image can be derived from 1-D DWT. So the scaling function for 2-D DWT can be obtained by multiplying two 1-D scaling functions:

$$\phi(x, y) = \phi(x)\phi(y).$$

Also, the wavelet functions for 2-D DWT can be obtained by multiplying two wavelet functions or one wavelet and one scaling function for 1-D. This shows, for a 2-D image, there are three wavelet functions in three directions:

Horizontal: $\psi_h(x, y) = \phi(x)\psi(y)$.

Vertical: $\psi_v(x, y) = \psi(x)\phi(y)$.

Diagonal: $\psi_d(x, y) = \psi(x)\psi(y)$.

The 2-D scaling function ϕ for multi-resolution approximation can be obtained as: $\varphi_{j,m,n}(x, y) = 2^{\frac{j}{2}}\varphi(2^j x - m, 2^j y - n)$

The 2-D wavelet functions ψ are given as:

$$\psi^i_{j,m,n}(x, y) = 2^{\frac{j}{2}}\psi^i(2^j x - m, 2^j y - n), \quad i = \{H, V, D\}$$

As a result, each resolution has three types of detailed images: horizontal (HL), vertical (LH) and diagonal (HH).

The results in four different subbands (LL, HL, LH, and HH) in the decomposition corresponding to four types of transformed coefficients: $W_\varphi(j_0, m, n)$, $W^H_\psi(j, m, n)$, $W^V_\psi(j, m, n)$, and $W^D_\psi(j, m, n)$. $W_\varphi(j_0, m, n)$ coefficients are the approximation part of the image and correspond to the LL subband. $W^H_\psi(j, m, n)$ coefficients contain the horizontal details and correspond to the HL subband. $W^V_\psi(j, m, n)$ coefficients contain the vertical details and correspond to the LH subband. $W^D_\psi(j, m, n)$ coefficients represent the diagonal details in the image and constitute the HH subband. In short, the results of wavelet transform are sets of wavelet coefficients shown below, which measure the contribution of the wavelets at these locations and scales.

$$W_\varphi(j_0, m, n) = \frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x, y)\varphi_{j_0,m,n}(x, y)$$

$$W^i_\psi(j, m, n) = \frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x, y)\psi^i_{j,m,n}(x, y), \quad i = \{H, V, D\}$$

Then, the decomposition can be repeated on the average part of an image (LL). Fig. 2-1(a) shows the sketch map.

Two-dimensional DWT is implemented by convolving an input N×N image with a pair of low-pass and high-pass filters and down-sampling by 2. Recursive decompositions of an image on the average part using DWT represent an image different scales or subbands. Fig. 2-1(b) shows the sketch map.

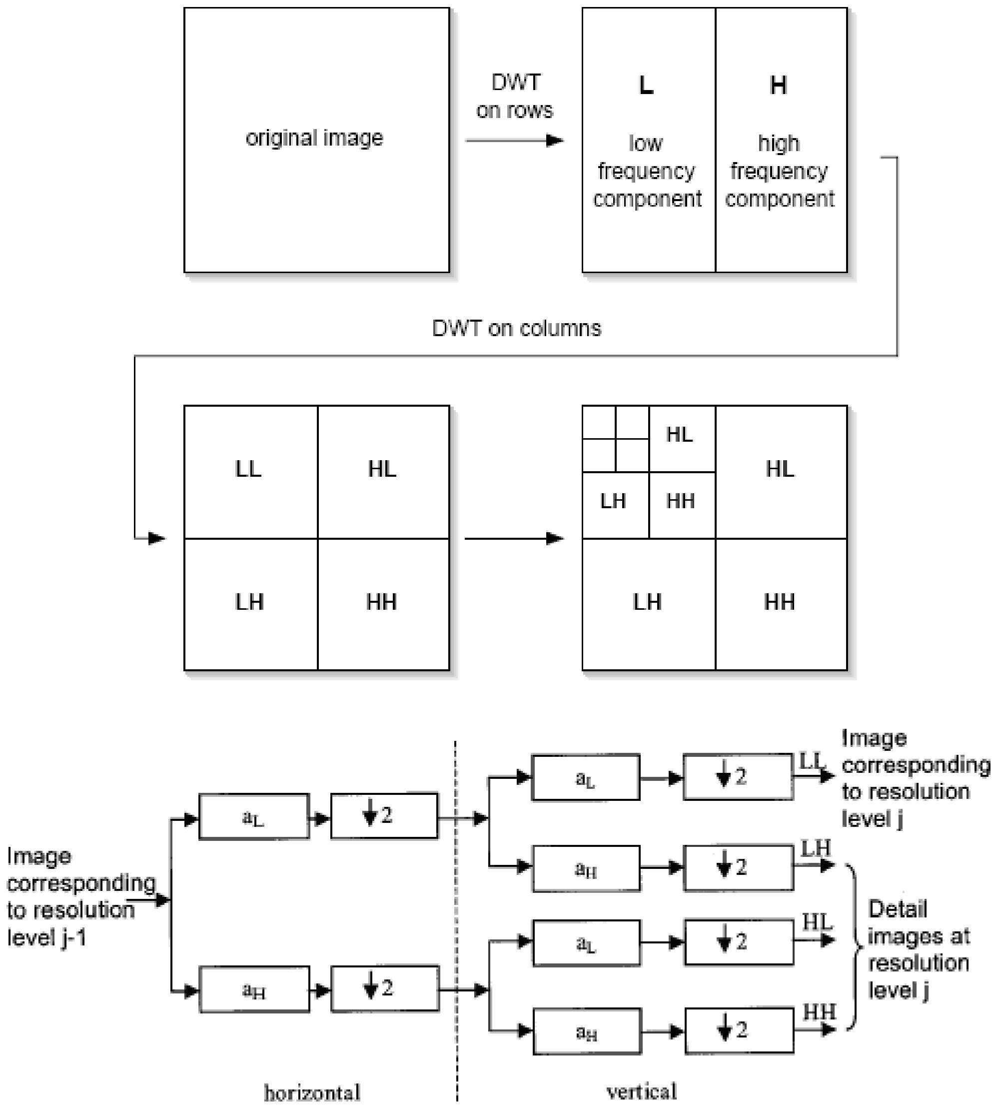

Fig. 2-1. (a) Structure of wavelet decomposition (3-level). (b) 2-D DWT.

The synthesis stage performs up-sampling and filtering in the reverse order to reconstruct the image. The equation of 2-D inverse discrete wavelet transform (IDWT) is given as:

$$f(x,y)=\frac{1}{\sqrt{MN}}\sum_m\sum_n W_\varphi(j_0,m,n)\varphi_{j_0,m,n}(x,y)+\frac{1}{\sqrt{MN}}\sum_{i=H,V,D}\sum_{j=j_0}^{\infty}\sum_m\sum_n W_\psi^i(j,m,n)\psi_{j,m,n}^i(x,y)$$

2.2 Wavelets

There are many wavelet types. They can be distinguished in five types.

1. Orthogonal wavelets with FIR filters:

   These wavelets can be defined through the filter order *N*. Families of such wavelets include Haar, Daubechies, Coiflet, and Symlets.

2. Biorthogonal wavelets with FIR filters:

   These wavelets can be defined through the two scaling filters N_d and N_r, for decomposition and reconstruction, respectively. The BiorSplines wavelet family is a family of this type.

3. Orthogonal wavelets without FIR filter, but with scale function:

   These wavelets can be defined through the definition of the wavelet function and the scaling function. The Meyer wavelet family is a family of this type.

4. Wavelets without FIR filter and without scale function:

   These wavelets can be defined through the definition of the wavelet function. Families of such wavelets include Morlet and Mexican_hat.

5. Complex wavelets without FIR filter and without scale function:

   These wavelets can be defined through the definition of the wavelet function. Families of such wavelets include Complex Gaussian and

Shannon.

We choose four often-used wavelet families in our paper. They are Haar Wavelet (HW) family, Daubechies Wavelet (DW) family, Coiflet Wavelet (CW) family, and Biorthogonal Wavelet (BW) family. Discussion of wavelets begins with Haar wavelet. It is the first and simplest one. Haar wavelet is discontinuous, and resembles a step function. It represents the same wavelet as Daubechies when N is equal to 1 (DW1). Daubechies invented compactly support orthogonal wavelets, thus made discrete wavelet analysis practicable. Coiflet wavelet has $2N$ moments equal to 0 and the scaling function has $2N$-1 moments equal to 0. The two functions have a support of length $6N$-1. Biorthogonal wavelet exhibits the property of linear phase, which is needed for signal and image reconstruction. It uses two wavelets, one for decomposition (on the left side) and the other for reconstruction (on the right side) instead of the same single one. Fig. 2-2 shows some wavelet functions of each wavelet family.

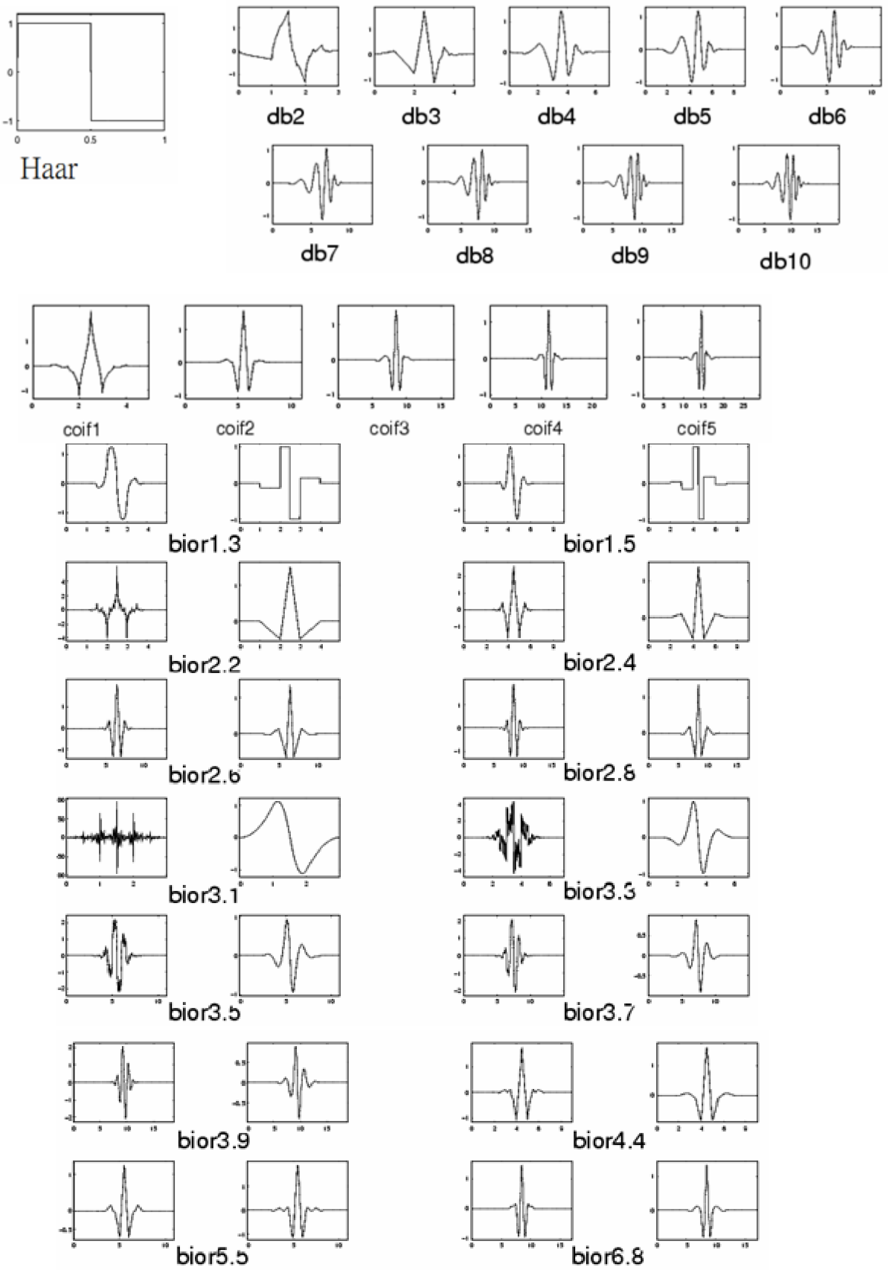

Fig. 2-2. Wavelet functions of each wavelet family.

Important properties of wavelet families are compact support (lead to efficient implementation), symmetry (useful in avoiding dephasing in image processing), orthogonality (allow fast algorithm), regularity and degree of smoothness (related to filter order or filter length). The properties of each wavelet family are shown in Table. 2-1.

Family	Orthogonal	Compact Support	Symmetry
Haar	●	●	●
Daubechies	●	●	Far from
Coiflet	●	●	Near from
Biorthogonal		●	●

Table. 2-1. Properties of each wavelet family.

HW, DW, and CW are parameterized by filter order (N) that determines filter length (L). BW uses filters with similar or dissimilar orders for decomposition (N_d) and reconstruction (N_r). Although the filter length L is determined by filter order N, relationship between them is different in different wavelet families. For example, filter length for DW family is $2N$, and for CW family is $6N$. HW is the special case of DW with $N = 1$ and $L = 2$ (DW1). Filter lengths of BW are approximately $\{\max(2N_d, 2N_r) + 2\}$, but effective lengths are different for LPF and HPF used for decomposition and reconstruction. Table. 2-2 shows the details. And Table. 2-3 shows filter coefficients of some wavelets.

<table>
<tr><th><u>Wavelet Family</u></th><th><u>Filter Order N</u></th><th><u>Filter Length L</u></th></tr>
<tr><td>Haar</td><td>1</td><td>2</td></tr>
<tr><td>Daubechies</td><td>A positive integer.</td><td>2N</td></tr>
<tr><td>Coiflet</td><td>1, 2, 3, 4, 5</td><td>6N</td></tr>
<tr><td>Biorthogonal</td><td>
<table>
<tr><th>N_r</th><th>N_d</th></tr>
<tr><td>1</td><td>1, 3, 5</td></tr>
<tr><td>2</td><td>2, 4, 6, 8</td></tr>
<tr><td>3</td><td>1, 3, 5, 7, 9</td></tr>
<tr><td>4</td><td>4</td></tr>
<tr><td>5</td><td>5</td></tr>
<tr><td>6</td><td>8</td></tr>
</table>
</td><td>
<table>
<tr><th>(N_r, N_d)</th><th>L_d</th><th>L_r</th></tr>
<tr><td>1, 1</td><td>2</td><td>2</td></tr>
<tr><td>1, 3</td><td>6</td><td>2</td></tr>
<tr><td>1, 5</td><td>10</td><td>2</td></tr>
<tr><td>2, 2</td><td>5</td><td>3</td></tr>
<tr><td>2, 4</td><td>9</td><td>3</td></tr>
<tr><td>2, 6</td><td>13</td><td>3</td></tr>
<tr><td>2, 8</td><td>17</td><td>3</td></tr>
<tr><td>3, 1</td><td>4</td><td>4</td></tr>
<tr><td>3, 3</td><td>8</td><td>4</td></tr>
<tr><td>3, 5</td><td>12</td><td>4</td></tr>
<tr><td>3, 7</td><td>16</td><td>4</td></tr>
<tr><td>3, 9</td><td>20</td><td>4</td></tr>
<tr><td>4, 4</td><td>9</td><td>7</td></tr>
<tr><td>5, 5</td><td>9</td><td>11</td></tr>
<tr><td>6, 8</td><td>17</td><td>11</td></tr>
</table>
</td></tr>
</table>

Table. 2-2. Filter order and filter length of each wavelet family.

	DW1	DW2	DW5	CW2	BW2.2		BW4.4	
k	**$a_L(k)$**	**$a_L(k)$**	**$a_L(k)$**	**$a_L(k)$**	**$a_L(k)$**	**$s_L(k)$**	**$a_L(k)$**	**$s_L(k)$**
0	0.7071	-0.1294	0.0033	-0.0007	0	0	0	0
1	0.7071	0.2241	-0.0126	-0.0018	-0.1768	0.3536	0.0378	-0.0645
2		0.8365	-0.0062	0.0056	0.3536	0.7071	-0.0238	-0.0407
3		0.4830	0.0776	0.0237	1.0607	0.3536	-0.1106	0.4181
4			-0.0322	-0.0594	0.3536	0	0.3774	0.7885
5			-0.2423	-0.0765	-0.1768	0	0.8527	0.4181
6			0.1384	0.4170			0.3774	-0.0407
7			0.7243	0.8127			-0.1106	-0.0645
8			0.6038	0.3861			-0.0238	0
9			0.1601	-0.0674			0.0378	0
10				-0.0415				
11				0.0164				

Table. 2-3. Filter coefficients of some wavelets.

2.3 Image Compression Schemes

The goal of image compression is to represent an image as accurately as possible using the fewest number of bits. There are two kinds of image compression schemes: lossless and lossy. The main difference between them is whether or not the original data can be recovered completely in a compression system.

In a lossless compression scheme, every bit of image that is originally in the image is hold after the image is decompressed, so all of the information is restored. The two important and popular lossless compression algorithms are Huffman coding and Lempel-Ziv coding.

On the other hand, in a lossy compression scheme, typically there is some distortion between the original image and the decompressed image. It permanently reduces certain information, so a part of the original information cannot be reconstructed when the image is decompressed (even the distortion may not be noticed by human eyes). Fig. 2-3 shows the fundamental components in a lossy image compression scheme. Compression is obtained by applying a transform to decorrelate the image data, quantizing the resulting transform coefficients, and entropy coding the quantized values.

A lossy image compression scheme is typically comprised of three major parts. First, a transform is performed. Commonly used transforms are the Fourier Transform, Discrete Cosine Transform (DCT), and Discrete Wavelet Transform (DWT). Take DWT for example, a wavelet bank decomposes the image into wavelet coefficients. Second, the coefficients are quantized. A quantizer simply reduces the number of bits needed to store the transformed coefficients by reducing the precision of those values. It is a lossy and irreversible procedure. Quantization can be performed on each individual coefficient, which is known as Scalar Quantization (SQ). Quantization can also be performed on a group of coefficients together, and this is known as Vector Quantization (VQ). Third, the entropy encoder encodes the quantized coefficients into a bit stream to give better overall compression. It uses a model to accurately determine the probabilities for each quantized value and produces an appropriate code based on these probabilities so that the resultant output code stream will be smaller than the input stream. The commonly used entropy encoders are the Huffman encoder and the arithmetic encoder. Entropy encoding is a lossless and reversible procedure. To recover the original data, we can apply the above procedures in a reverse order.

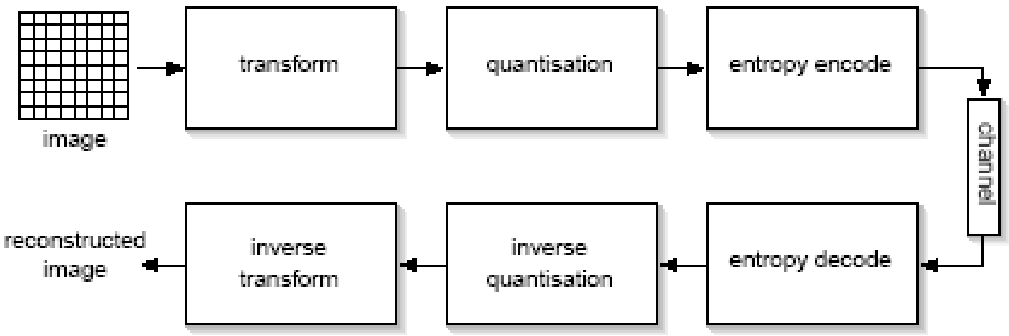

Fig. 2-3. The fundamental parts in a lossy image compression scheme.

2.4 Image Quality Evaluations

In a compression system, there are objective and subjective image quality evaluations.

A standard objective measure of image quality is reconstruction error. Suppose that an input image element block $\{x(n)\}$, n=0, 1… *N*-1 is reproduced as $\{y(n)\}$, n=0, 1… *N*-1. The reconstruction error $r(n) = x(n) - y(n)$.

The variances of $x(n)$, $y(n)$ and $r(n)$ are σ_x^2, σ_y^2 and σ_r^2, respectively. In the special case that mean is zero, variances are simply equal to respective mean-square values over appropriate sequence length *M*.

$$\sigma_z^2 = \frac{1}{M}\sum_{n=1}^{M} z^2(n), \qquad z = x, y \text{ or } r.$$

Signal-to-noise ratio (SNR) is defined as the ratio between signal variance and reconstruction error variance [mean-square error (MSE)]. It is usually expressed in decibels (dB).

$$\text{SNR(dB)} = 10 \log_{10}\left(\frac{\sigma_x^2}{\sigma_r^2}\right) = 10 \log_{10}\left(\frac{\sigma_x^2}{\text{MSE}}\right)$$

When the input is an R-bits image, σ_x^2 can be replaced by $(2^R - 1)^2$. For example, the peak SNR (PSNR) of an 8-bits image can be defined as:

$$\text{PSNR} = 10 \log_{10}\left(\frac{255^2}{\text{MSE}}\right)$$

Typical PSNR values range between 20 and 40. Generally when PSNR is 40 dB or greater, the original and reconstructed images are virtually indistinguishable by human observers. The actual value of PSNR is not meaningful, but the comparison between two values for different reconstructed images gives one measure of picture quality.

Subjective image quality evaluations are difficult to experiment because the results vary depending on the test conditions and cost a lot of time. And the standard objective measure has been recognized as inadequate because of its low correlation with human visual perception. In many applications, it is very important to choose an image compression system that gives the best subjective quality, but the quality has to be evaluated objectively. So a perception-based, quantitative distortion measure, called the Picture Quality Scale (PQS), was developed for evaluating the quality of compressed images [16].

PQS is a perception-based objective evaluation. It is constructed by regression with Mean Opinion Score (MOS) that is a perception-based subjective evaluation and is a five-level grading scale. [Table. 2-4] PQS can be simply expressed as a linear combination of uncorrelated principle distortion measures Zj, that is,

$$\text{PQS} = b_0 + \sum_{j=1}^{J} b_j Z_j$$

Where $\{bj\}$ are the partial regression coefficients obtained by multiple linear regression of $\{Zj\}$ against the MOS.

For images with high quality, it is possible to obtain values of PQS lager than 5. Also for images with low quality scale, PQS can be negative values (meaningless results).

Grading Scales	Impairment
5	Imperceptible
4	Perceptible, but not annoying
3	Slightly annoying
2	Annoying
1	Very annoying

Table. 2-4. The scales of the MOS.