3 Wavelet Transform

To understand the realworld zoom in zoom out.

3.1 Discrete Wavelet Transform

Gabor transform can be written as

$$G_{\sigma_x \sigma_y}(\tau_1, \tau_2) = \sum_{x=0}^{N} \sum_{y=0}^{N} f(x, y) g^*_{\sigma_x \sigma_y}(x - \tau_1, y - \tau_2) \tag{3.1}$$

where * means complex conjugate and the Gabor function $g_{\sigma_x, \sigma_y}(x, y)$ is given as

$$g_{\sigma_x \sigma_y}(x, y) = \frac{1}{2\pi\sigma_x \sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] e^{j2\pi\omega x} \tag{3.2}$$

The group of Gabor functions $g_{\sigma_x, \sigma_y}(x, y)$ are windowed waveform functions, called wavelets. But the Gabor wavelets are not orthogonal, which means there is a correlation between different Gabor wavelets. This correlation results in redundancy in the extracted wavelet features computed from images or signals. The FWHM approach in Sect. 2.3 causes loss of frequency. This is undesirable for image or signal representation. The window size is also an issue similar to that of STFT. These issues can be overcome by using orthogonal wavelets with varying window size.

The general form of a 2D orthogonal wavelet can be formulated as follows:

$$\psi_{a_1 a_2 b_1 b_2}(x, y) = \frac{1}{\sqrt{a_1 a_2}} \psi\left(\frac{x - b_1}{a_1}, \frac{y - b_2}{a_2}\right) \tag{3.3}$$

where a_1, a_2 are the scale parameters and b_1, b_2 are the position parameters. Similar to a 2D FT, a 2D wavelet also has the property of separability:

D. Zhang, *Fundamentals of Image Data Mining*, Texts in Computer Science, https://doi.org/10.1007/978-3-030-17989-2_3

$$\psi_{a_1a_2b_1b_2}(x,y) = \frac{1}{\sqrt{a_1}}\psi\left(\frac{x-b_1}{a_1}\right)\frac{1}{\sqrt{a_2}}\psi\left(\frac{y-b_2}{a_2}\right) \tag{3.4}$$

The *Discrete Wavelet Transform* (DWT) on a function or image $f(x, y)$ is given as

$$W^s(k,l) = \frac{1}{s}\sum_n\sum_m f(m,n)\psi\left(\frac{m-k}{s}\right)\psi\left(\frac{n-l}{s}\right) \tag{3.5}$$

where (k, l) is the position of the wavelet and s is the scale. If we compare a wavelet with a magnifying glass, the position vector (k, l) represents the location of the magnifying glass and the scale s represents the distance between the magnifying glass and the image. By adjusting the position and scale, the wavelet can analyze an image in the same way as we analyze an image using a magnifying glass.

3.2 Multiresolution Analysis

As explained in Sect. 1.4, the window size (time) and the frequency band are inversely proportional. That is, when the window size is halved, the frequency band captured by the window is twice higher. The frequency bandwidth is typically arranged in octave, that is, the next bandwidth is twice the width of the previous one. The inverse relationship between window size and frequency of a DWT can be demonstrated using a 1D signal with a Nyquist sampling rate of 1,024 Hz. As can be seen in the following table, as the wavelet decomposition level (scale) and the window size increases, the bandwidth becomes narrower and narrower until it reduces to a single point, which is equivalent to a FT frequency.

Decomposition level	Window size	Frequency band
1	2	512–1,023
2	4	256–511
3	8	128–255
4	16	64–127
5	32	32–63
6	64	16–31
7	128	8–15
8	256	4–7
9	512	2–3
10	1,024	1

The above table can be illustrated using a time–frequency plane. Figure 3.1 shows the time–frequency planes of both DWT and STFT side by side. The wavelet time–frequency plane is shown on Fig. 3.1 Left. In contrast, STFT uses one window size for all frequencies as shown in Fig. 3.1 Right. Compared with STFT, a

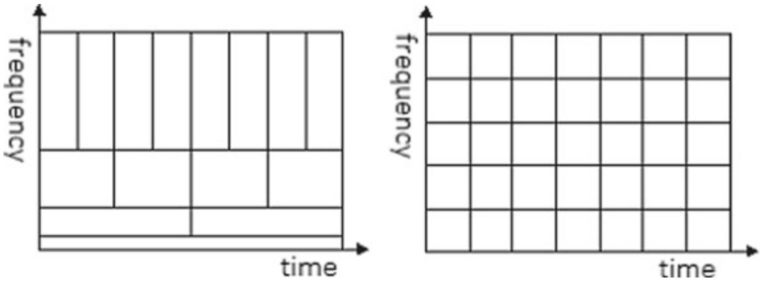


Fig. 3.1 Different frequency tiling of spectral plane. Left: the wavelet spectrum; Right: the STFT spectrum

DWT yields very high resolution at lower frequencies while sacrifices resolution at higher frequencies. Therefore, a DWT is a multiresolution tool.

Wavelets analyze and represent a signal with multiresolution. This is extremely useful, because lower resolution represents a summary and higher resolution represents fine details of a signal, both are essential in signal analysis and representation. The multiresolution representation is done through repeating rounds of scaling (low pass, L) and wavelet transform (high pass H) on a signal. The scaling captures the low frequency information of the signal and the wavelet captures the high frequency information of the signal. At each round of the wavelet transform, a low-resolution signal (low frequency L) and a fine details signal (high frequency H) are obtained, both are half the size of the original signal. Since the information in the fine details signal is usually scarce, most of the information in the original signal is captured by the lower resolution version, this achieves great efficiency of signal representation. To recover the signal, the inverse wavelet transform is applied, which is done through repeating rounds of expansion.

3.3 Fast Wavelet Transform

3.3.1 DTW Decomposition Tree

For a 2D image, the rows and columns are treated as 1D signals. Due to the separability property of DWT, the two passes at each round of the DWT are done at the rows and the columns separately. The 2D digital wavelet transform on an image is illustrated in Figs. 3.2, 3.3 and 3.4.

- **Horizontal transform**. At the first step of level 1 decomposition, each row of the image is scaled (weighted average) and wavelet transformed (weighted difference).

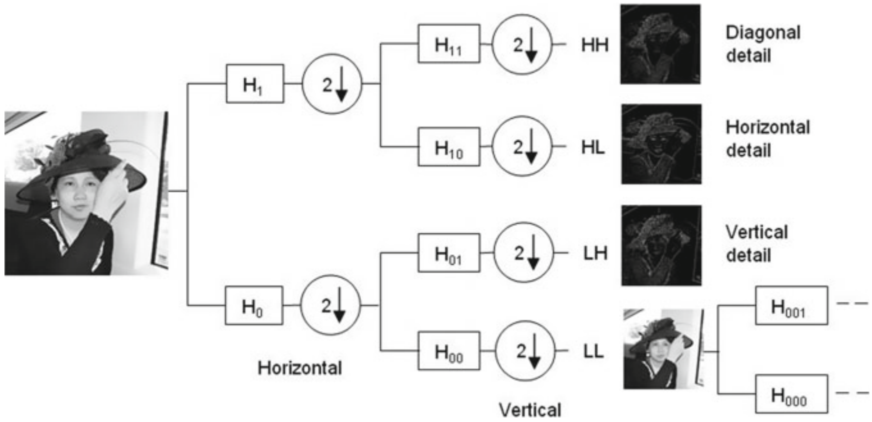


Fig. 3.2 The 2D DWT decomposition tree for a lady image

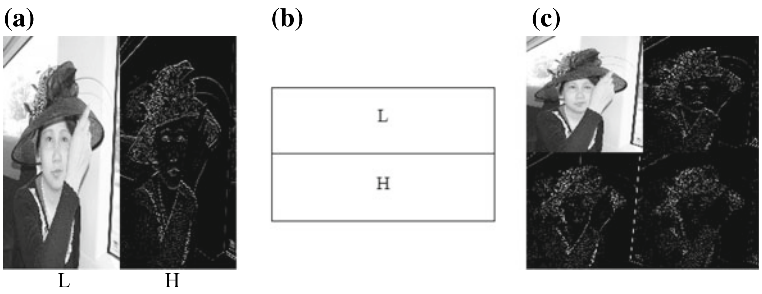


Fig. 3.3 Illustration of 2D DWT process on a lady image. **a** Horizontal transform; **b** vertical transform; **c** spectrum of level 1 2D DWT decomposition

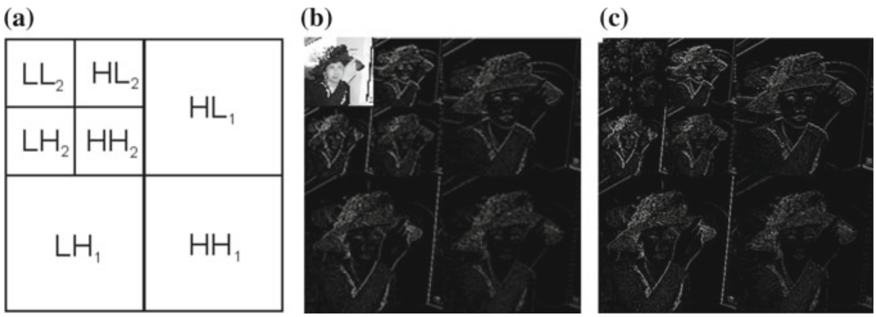


Fig. 3.4 Two levels of 2D wavelet decomposition. **a** The spectrum plane of two levels of 2D DWT; **b** the spectrum of two levels of 2D DWT on the lady image; **c** the complete decomposition of the lady image

- The result from the first step are two half images, one with scaling coefficients (L) and the other with wavelet coefficients (H), both are half width of the image row (Fig. 3.3a).

- **Vertical transform**. In the next, the scaling (L) and wavelet transform (H) are applied on each column of the two half images from the previous step (Fig. 3.3b).
  - The result from the second step is four quarter sized images, which are: summary (LL, top left), vertical details (LH, bottom left), horizontal details (HL, top right), and diagonal details (HH, bottom right) (Fig. 3.3c).

- **Level 2 decomposition**. The above steps are repeated on the LL image for the next round of DWT (Fig. 3.4b).

- **DWT spectrum**. The level 1 decomposition process can be repeated until the summary image can no longer be decomposed further, the final spectrum of the wavelet transform on the lady image is shown in Fig. 3.4c. It can be observed that the DWT spectrum captures the essence of the image while discarding all the redundant details. This is very useful for image analysis.

This process can be summarized in mathematical terms. Suppose the scaling and wavelet functions are ϕ, ψ, respectively. At each level of decomposition, the following 4 quarter sized images are resulted from the DWT by using (3.5): average/summary image $\phi(x, y)$, horizontal difference/details image $\psi^{\mathrm{H}}(x, y)$, vertical difference/details image $\psi^{\mathrm{V}}(x, y)$, and diagonal difference/details image $\psi^{\mathrm{D}}(x, y)$.

$$\phi(x,y) = \phi(x)\,\phi(y) \quad \to \mathrm{LL} \tag{3.6}$$

$$\psi^{\mathrm{H}}(x,y) = \phi(x)\,\psi(y) \quad \to \mathrm{HL} \tag{3.7}$$

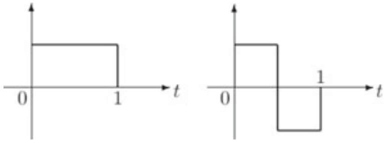
$$\psi^{\mathrm{V}}(x,y) = \psi(x)\,\phi(y) \quad \to \mathrm{LH} \tag{3.8}$$

$$\psi^{\mathrm{D}}(x,y) = \psi(x)\,\psi(y) \quad \to \mathrm{HH} \tag{3.9}$$

3.3.2 1D Haar Wavelet Transform

The DWT can be demonstrated using Haar wavelet, the scaling function and wavelet function of Haar transform are shown in Fig. 3.5. Given the unique characteristic or shape of the Haar transform functions, the scaling and wavelet transform of Haar wavelet become simply the *average* and *difference* (or *details*).

Suppose x and y are two neighboring points, the scaling coefficient and wavelet coefficient are given by s and d, respectively:



(a) $\varphi(t)$ (b) $\psi(t)$

Fig. 3.5 Harr scaling function and wavelet function

$$s = (x + y)/2 \quad \text{and} \quad d = (x - y)/2 \tag{3.10}$$

The inverse Haar transform is then given by *addition* and *subtraction*:

$$x = s + d \qquad \text{and} \quad y = s - d \tag{3.11}$$

- Given an even length discrete signal of $(a_0, a_1, \ldots, a_{2n}, a_{2n+1})$.
- It is first organized into pairs $((a_0, a_1), \ldots, (a_{2n}, a_{2n+1}))$.
- By applying the first round of Haar transform, the coefficients of the transform are given by $((s_0, s_1, \ldots, s_n), (d_0, d_1, \ldots, d_n))$.
- The second round of Haar transform can be performed on the sequence of s, and so on so forth.

For example, suppose [11, 9, 5, 7] is a 4-point digital signal, the following demonstrates the process of a Haar wavelet transform on the signal.

Resolution	Averages	Details
4	[11, 9, 5, 7]	
2	[10, 6]	[1, −1]
1	[8]	[2]

Therefore, the Haar wavelet transform of [11, 9, 5, 7] is given by [8, 2, 1, −1]. As can be seen, after the wavelet transform, the first value captures the most significant information while the last two values are very small. This is helpful in signal processing and analysis, because more attention can be given to the most significant information.

The wavelet transform can be performed more efficiently by using matrix multiplication. The following is an example of 4 × 4 Haar wavelet transform matrix.

$$H_4 = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 2 & -2 & 0 & 0 \\ 0 & 0 & 2 & -2 \end{bmatrix} \tag{3.12}$$

For the above 4-point signal, using the Haar wavelet transform matrix, the transform coefficients are given by

$$h_4 = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 2 & -2 & 0 & 0 \\ 0 & 0 & 2 & -2 \end{bmatrix}\begin{bmatrix} 11 \\ 9 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \\ 1 \\ -1 \end{bmatrix} \tag{3.13}$$

3.3.3 2D Haar Wavelet Transform

For a 2D image, this is done on both the rows and columns separately.

Suppose the following is a 4 × 4 image I:

102	56	68	152
24	62	46	32
52	92	72	84
76	60	92	60

Step 1. *Horizontal scaling* of image I (horizontal pairwise average, L):

79	110
43	39
72	78
68	76

Step 2. *Horizontal wavelet transform* of I (horizontal pairwise difference, H):

23	−42
−19	7
−20	−6
8	16

The image I_c after horizontal transform by combining the results from the above two steps:

79	110	23	−42
43	39	−19	7
72	78	−20	−6
68	76	8	16

Step 3. *Vertical scaling* of I_c (vertical pairwise average, LL, HL)

61	74.5	2	−17.5
70	77	−6	5

Step 4. *Vertical wavelet transform* of I_c (vertical pairwise difference, LH, HH)

18	35.5	21	−24.5
2	1	−14	−11

The image after the first round of Haar wavelet transform by combining the results from Step 3 and 4:

LL	61	74.5	2	−17.5	HL
	70	77	−6	5	
LH	18	35.5	21	−24.5	HH
	2	1	−14	−11	

The second round of Haar DWT repeats the Steps 1–4 on the LL band and this process can be continued until required levels of decomposition is achieved. Similar to the 1D case, the first quarter of the wavelet transformed image contains the most significant information.

3.3.4 Application of DWT on Image

Figure 3.6 demonstrates the complete process of computing the Haar wavelet transform on the lady image using the DWT decomposition tree of Fig. 3.2. At each next level of the decomposition, the image is halved, therefore, the DWT is very efficient and fast.

3.4 Summary

Wavelets are an extension or an improvement to windowed FT in two aspects: *orthogonality* and *multiresolution*. Orthogonality means that there is no redundancy between DWT channels. Multiresolution means to analyze an image by zooming in and zooming out, which is like studying a map with a magnifying glass. This is achieved by adapting image resolution to wavelet size/scale.

The contrast between wavelets and windowed FT can be easily understood in frequency plane as shown in Fig. 3.1. Basically, with DWT, we retain higher resolution at very low-frequency band at the cost of losing resolution at high-frequency band. This is sensible because low-frequency information is much more important than high frequency information to human perception.