1 Fourier Transform

In essence, the world is just elements and compounds.

1.1 Introduction

Fourier transform has played a key role in image processing for many years, and it continues to be a topic of interest in theory as well as application. The fundamental principle behind Fourier transform is that a pattern can be treated as a signal, and as such, it can be represented by elementary components of the signal. If we can define elementary components to represent or approximate a pattern under analysis, we can determine how significant an elementary component in a given pattern. The elementary components found in the signal can be used to describe the given pattern. Fourier transform is useful for pattern analysis and description because different patterns can be distinguished by the transformed spectra (Fig. 1.1) [1], while similar patterns will have similar transformed spectra even they are affected by noise and other variations. It can be observed that the spectrum of Fig. 1.1a clearly shows patterns of both horizontal and vertical directions, while that in Fig. 1.1b shows patterns of random fashion. They demonstrate the power of Fourier transform in image analysis and understanding.

1.2 Fourier Series

1.2.1 Sinusoids

To understand how Fourier transform works, it has to start with understanding how sinusoids work. It is important to understand the relationship between frequency and period. Figure 1.2 shows a sine wave and its harmonic waves. It shows how the change of variable scaling or horizontal stretching affects the sine wave's frequency and periods.

D. Zhang, *Fundamentals of Image Data Mining*, Texts in Computer Science, https://doi.org/10.1007/978-3-030-17989-2_1

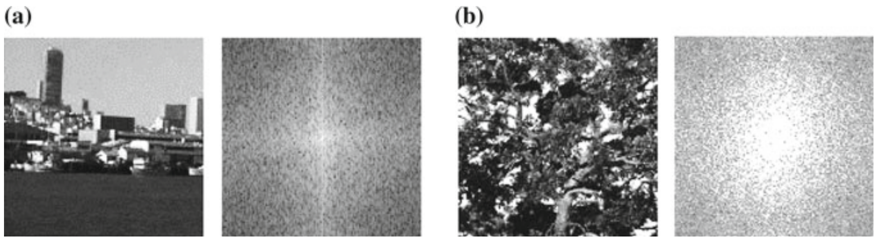

Fig. 1.1 Fourier spectra of different images. **a** A scenic image at the left and its Fourier spectrum at the right; **b** a tree image in the left and its Fourier spectra in the right. The brighter the pixel, the higher magnitude of the spectrum

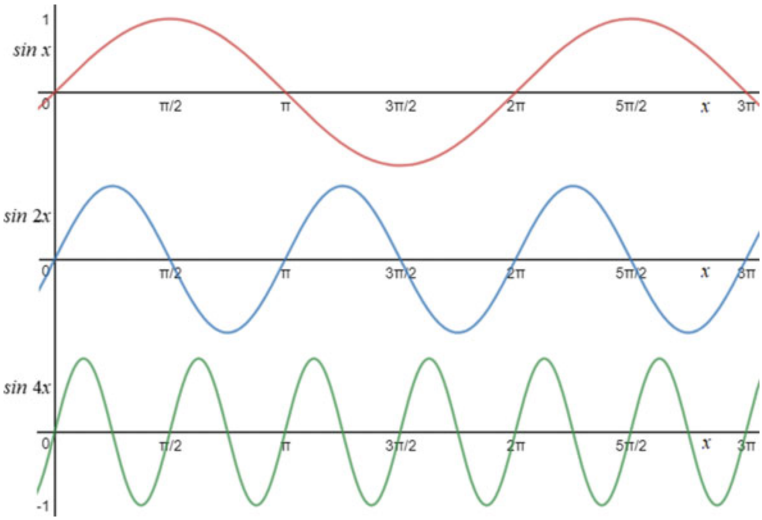

Fig. 1.2 Three sine waves sin(nx) with different periods and frequencies

Sine waves	Frequency	Periods
$\sin x$	$\frac{1}{2\pi}$	2π
$\sin 2x$	$\frac{1}{\pi}$	π
$\sin 4x$	$\frac{2}{\pi}$	$\frac{\pi}{2}$
$\ldots$		
$\sin nx$	$\frac{n}{2\pi}$	$\frac{2\pi}{n}$

As can be seen, as the variable scaling factor n increases, the period of sin(nx) becomes shorter and the frequency becomes higher. For example, the period of sin (x) is 2π, while the periods of sin($2x$) and sin($4x$) are π and $1/2\pi$, respectively. Consequently, the frequencies of the three sine waves are $1/2\pi$, $1/\pi$, and $2/\pi$, respectively. Similarly, the period and frequency of sin(nx) are $2\pi/n$ and $n/2\pi$, respectively.

The sine waves expressed this way do not have an easy interpretation of the frequency and periods, because both of them are in angular terms. Now let's replace n with $2\pi n$ and change the sine function from $\sin(nx)$ to $\sin(2\pi nx)$, see what will happen.

Sine waves	Frequency	Periods
$\sin(2\pi x)$	1	1
$\sin(4\pi x)$	2	$\frac{1}{2}$
$\sin(8\pi x)$	4	$\frac{1}{4}$
...		
$\sin(2\pi nx)$	n	$\frac{1}{n}$

Now both the periods and frequencies are easier to understand. For example, the period of $\sin(2\pi x)$ is 1, while the periods of $\sin(4\pi x)$ and $\sin(8\pi x)$ are 1/2 and 1/4, respectively. Consequently, the frequencies of the three sine waves are 1, 2, and 4, respectively (Fig. 1.3). Therefore, the period and frequency of $\sin(2\pi nx)$ are $1/n$ and n, respectively. This is much easier to interpret.

A more general form of sine function is expressed as $\sin(2\pi n/L)$, which has a period of L/n and frequency of n/L. This is extremely helpful to analyze signals with arbitrary periodicity and frequencies.

Sine waves	Frequency	Periods
$\sin\left(\frac{2\pi x}{L}\right)$	$\frac{1}{L}$	L
$\sin\left(\frac{4\pi x}{L}\right)$	$\frac{2}{L}$	$\frac{L}{2}$
$\sin\left(\frac{8\pi x}{L}\right)$	$\frac{4}{L}$	$\frac{L}{4}$
...		
$\sin\left(\frac{2\pi nx}{L}\right)$	$\frac{n}{L}$	$\frac{L}{n}$

1.2.2 Fourier Series

One of the most important and interesting discoveries in mathematics is that any math function can be approximated with a *series of sinusoids* (sine and cosine waves), called *Fourier series*. Now consider a signal function $f(x)$ with period L,

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos\frac{2\pi nx}{L} + b_n \sin\frac{2\pi nx}{L}) \tag{1.1}$$

To determine the Fourier coefficients a_n and b_n, we multiply both sides of the above equation with either $\sin\left(\frac{2\pi nx}{L}\right)$ or $\cos\left(\frac{2\pi nx}{L}\right)$ and do the integral in $[-L/2, L/2]$.

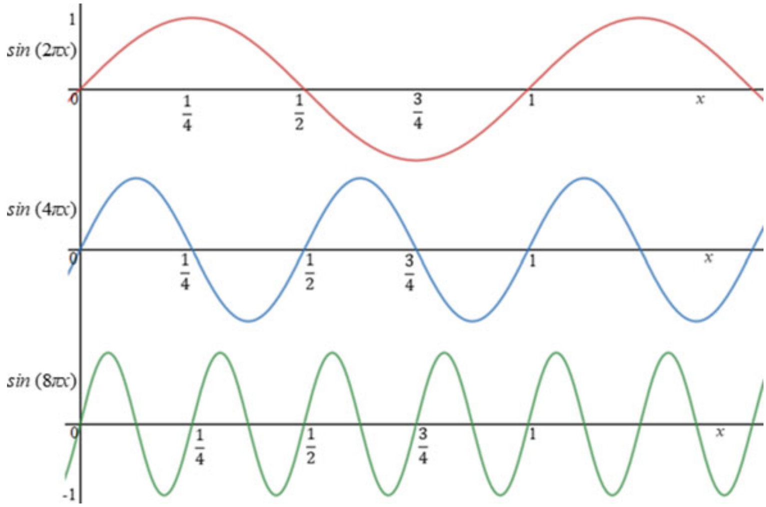

Fig. 1.3 Three sine waves $\sin(2\pi nx)$ with different periods and frequencies

It can be shown that sine and cosine waves have the following convenient properties:

$$\int_{-L/2}^{L/2} \cos\frac{2\pi\, nx}{L}\cos\frac{2\pi\, mx}{L}\,dx = \begin{cases} L/2 & for\, n = m \\ 0 & for\, n \neq m \end{cases} \tag{1.2}$$

$$\int_{-L/2}^{L/2} \sin\frac{2\pi\, nx}{L}\sin\frac{2\pi\, mx}{L}\,dx = \begin{cases} L/2 & for\, n = m \\ 0 & for\, n \neq m \end{cases} \tag{1.3}$$

$$\int_{-L/2}^{L/2} \sin\frac{2\pi\, nx}{L}\,\cos\frac{2\pi\, mx}{L}\,dx = 0 \tag{1.4}$$

$$\int_{-L/2}^{L/2} \sin\frac{2\pi nx}{L}\,dx = 0 \tag{1.5}$$

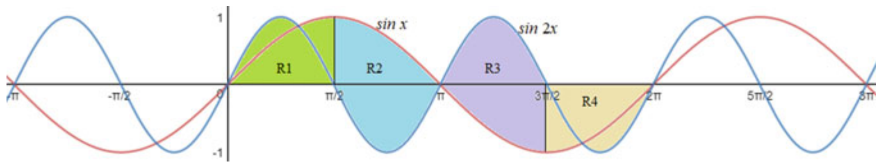

Fig. 1.4 Illustration of $\int \sin x \sin(2x)dx = 0$

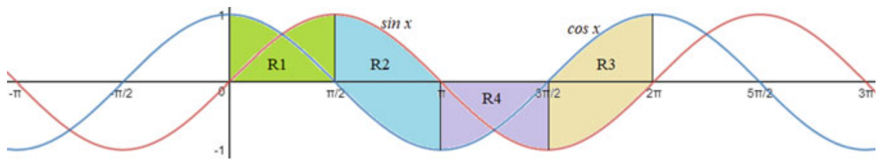

Fig. 1.5 Illustration of $\int \sin x \cos x dx = 0$

$$\int_{-L/2}^{L/2} \cos \frac{2\pi nx}{L} dx = 0 \tag{1.6}$$

To prove the above properties, we only need to demonstrate that $\int \sin(nx) \sin(mx)dx = 0 \;\; (m \neq n)$ and $\int \sin(nx) \cos(mx)dx = 0$; the others are obvious due to the symmetry of sine and cosine waves. Without loss of generality, we just need to show $\int \sin x \sin(2x)dx = 0$ and $\int \sin x \cos x \, dx = 0$.

Figure 1.4 illustrates $\int \sin x \sin(2x)dx$ in one period. We divide the *Sum of Product* (SoP) of the two functions within one period into four regions: R1–R4, marked with different colors. It is easy to observe that the SoP magnitude of the four regions is exactly the same; however, the signs of the four corresponding SoPs are +, −, −, and +, respectively, resulting in the total SoP of the period as 0. Applying this to all the other periods, it can be shown $\int \sin x \sin(2x)dx = 0$ on the entire x axis.

Figure 1.5 illustrates $\int \sin x \cos x \, dx$ in one period. Similar to the above, we divide the Sum of Product (SoP) of the two functions in the single period into four regions: R1–R4, marked with different colors. Again, it is easy to observe that the SoP magnitude of the four regions are exactly the same, however, the signs of the 4 corresponding SoPs are +, −, + and −, respectively, resulting in the total SoP of the period as 0. Applying this to all other periods, it can be shown $\int \sin x \cos x \, dx = 0$ on the entire x axis.

By making use of the integral identities and orthogonality of (1.2)–(1.6), the Fourier coefficients are obtained as follows:

$$\left.\begin{aligned} a_0 &= \frac{1}{L}\int_{-L/2}^{L/2} f(x)dx \\ a_n &= \frac{2}{L}\int_{-L/2}^{L/2} f(x)\cos\frac{2\pi nx}{L}dx \\ b_n &= \frac{2}{L}\int_{-L/2}^{L/2} f(x)\sin\frac{2\pi nx}{L}dx \end{aligned}\right\} \tag{1.7}$$

$n = 1, 2, \ldots$

1.2.3 Complex Fourier Series

Using the Euler formula

$$e^{ix} = \cos x + i \sin x \tag{1.8}$$

where $i = \sqrt{-1}$. It is easy to work out

$$\cos(x) = \frac{1}{2}\left(e^{ix} + e^{-ix}\right) \tag{1.9}$$

$$\sin(x) = \frac{1}{2i}\left(e^{ix} - e^{-ix}\right) = -\frac{i}{2}\left(e^{ix} - e^{-ix}\right) \tag{1.10}$$

Now, by replacing the sinusoids in the Fourier series of (1.1) with the above two equations, we obtain the complex Fourier series:

$$\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n\cos\frac{2\pi nx}{L} + b_n\sin\frac{2\pi nx}{L}) \\ &= \frac{a_0}{2} + \frac{1}{2}\sum_{n=1}^{\infty} a_n(e^{i\frac{2\pi nx}{L}} + e^{-i\frac{2\pi nx}{L}}) - \frac{i}{2}\sum_{n=1}^{\infty} b_n(e^{i\frac{2\pi nx}{L}} - e^{-i\frac{2\pi nx}{L}}) \\ &= \frac{a_0}{2} + \frac{1}{2}\sum_{n=1}^{\infty}(a_n - ib_n)e^{i\frac{2\pi nx}{L}} + \frac{1}{2}\sum_{n=1}^{\infty}(a_n + ib_n)e^{-i\frac{2\pi nx}{L}} \end{aligned} \tag{1.11}$$

which can be written as the complex Fourier series

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{i2\pi nx/L} \tag{1.12}$$

The exponential form of orthogonality is as follows:

$$\int_{-\frac{L}{2}}^{\frac{L}{2}} e^{-\frac{i2\pi mx}{L}} e^{\frac{i2\pi nx}{L}} = \begin{cases} L & for\, m = n \\ 0 & otherwise \end{cases} \tag{1.13}$$

Now by multiplying both sides of the Fourier series (1.12) with $e^{-i2\pi nx/L}$ and do integral in [0, L], we obtain complex Fourier coefficients:

$$c_n = \frac{1}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} f(x)\, e^{-i\frac{2\pi nx}{L}} dx, n = 0, \pm 1, \pm 2, \ldots \tag{1.14}$$

1.3 Fourier Transform

Equation (1.14) indicates that the coefficients of the Fourier series are determined by $f(x)$, while (1.12) indicates that $f(x)$ can be reconstructed from Fourier coefficients c_n. Therefore, the Fourier series establish a unique correspondence between $f(x)$ and its Fourier coefficients. Now, consider the integral of (1.14):

$$Lc_n = \int_{-\frac{L}{2}}^{\frac{L}{2}} f(x) e^{-j\frac{2\pi nx}{L}} dx \tag{1.15}$$

where $j = \sqrt{-1}$. If we let $L \to \infty$, n/L becomes continuous and $n/L \to u$, (1.15) becomes

$$F(u) = \int_{-\infty}^{\infty} f(x) \exp(-j2\pi ux)\, dx \tag{1.16}$$

Now, by substituting (1.6) into (1.12) and replacing the sum with an integral by using $n/L \to u$ and $1/L \to du$, (1.12) becomes

$$f(x) = \int_{-\infty}^{\infty} F(u) \exp(j2\pi ux)\, du \tag{1.17}$$

The $F(u)$ of (1.16) is called the forward *Fourier transform* or FT, and (1.17) is called the *inverse Fourier transform* or FT^{-1}.

1.4 Discrete Fourier Transform

1.4.1 DFT

Discrete Fourier Transform (DFT) is particularly useful for digital pattern analysis, because digital patterns exist in discrete form. To define DFT from Fourier series, $f(x)$ is first discretized into N samples in $[0, L]$:

$$f(0), f(\Delta x), f(2\Delta x), \ldots, f((N-1)\Delta x) \tag{1.18}$$

where Δx is the sample step in spatial domain and $L = N\Delta x$, and then $f(x)$ can be expressed as

$$f(k) = f(k\Delta x), k = 0, 1, 2, \ldots, N-1 \tag{1.19}$$

Now consider the Fourier coefficients (1.14):

$$\begin{aligned} c_n &= \frac{1}{L}\int_{-L/2}^{L/2} f(x)e^{-j\frac{2\pi nx}{L}}dx \\ &= \frac{1}{L}\int_{0}^{L} f(x)e^{-j\frac{2\pi nx}{L}}dx \end{aligned} \tag{1.20}$$

By substituting $L = N\Delta x$, $f(x) = f(k)$, $x = k\Delta x$, and $dx = \Delta x$ into the above equation, it yields

$$\begin{aligned} c_n &= \frac{\Delta x}{N\Delta x}\sum_{k=0}^{N-1} f(k)e^{-j\frac{2\pi nk\Delta x}{N\Delta x}} \\ &= \frac{1}{N}\sum_{k=0}^{N-1} f(k)e^{-j\frac{2\pi nk}{N}} n = 0, 1, 2, \ldots, N-1 \end{aligned} \tag{1.21}$$

Therefore, the DFT of $f(x)$ is given as

$$F(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\exp(-j2\pi ux/N) \quad u = 0, 1, 2, \ldots, N-1 \tag{1.22}$$

By substituting (1.22) into (1.12), the inverse DFT is obtained as

$$f(x) = \sum_{u=0}^{N-1} F(u)\exp(j2\pi ux/N) \quad x = 0, 1, 2, \ldots, N-1 \tag{1.23}$$

1.4.2 Uncertainty Principle

Assume $f(x)$ is a signal in a time period of $\Delta T = [0, L]$, the sampling step Δu in frequency domain and the sampling step Δx in spatial domain are related by the following expression:

$$\Delta u = \frac{1}{\Delta T} = \frac{1}{N\Delta x} \tag{1.24}$$

Basically, (1.24) tells that the frequency sampling step is inversely proportional to the spatial sampling step. This is known as the *uncertainty principle*, which means that increasing spatial resolution (reduce Δx) reduces the frequency resolution and vice versa. In other words, higher spatial resolution and higher frequency resolution cannot be achieved simultaneously. This is the key reason behind the multiresolution analysis such as wavelets which will be discussed later on in Chap. 3.

Since the Δx depends on the sampling rate f_s, and the relationship between Δx and f_s is given by $\Delta x = 1/f_s$, the above inequality becomes

$$\Delta u \geq \frac{f_s}{N} \tag{1.25}$$

and the uth frequency is given by

$$f_u = u \cdot \frac{f_s}{N} \tag{1.26}$$

It should be noted that the uth frequency computed from (1.22) is not the actual frequency; instead, the uth frequency is the uth *bin* of frequency. In other words, u is the bin number, and the actual frequency is given by (1.26): $f_u = u \cdot \Delta u$, and Δu is called the bin size of DFT. If $f_s = N$, $\Delta u = 1$, this is often the assumption in DFT. However, this is not always the case. When $f_s \gg N$, $\Delta u \gg 1$, this will be demonstrated in Sect. 2.2. Equation (1.25) is another form of the *uncertainty principle*. Given a sampling rate, in order to increase the frequency resolution (reduce Δu), it has to increase the sample or window size N, which reduces the spatial resolution. This is called the *trade-off* between spatial resolution and frequency resolution.

It should also be noted that any window size of a DFT is relative according to (1.25). Specifically, a window size is relative to the sampling rate or sampling frequency. A window of N samples is smaller in a signal with faster sampling rate than that in a signal with slower sampling rate. For example, in a signal with 44,000 Hz sampling rate, a window of 128 samples has a duration of 128/44,000 = 0.0029 s. However, in a signal with 22,000 Hz sampling rate, the duration of a 128 samples window is 128/22,000 = 0.0058 s, which is twice the size as that in the first signal. This indicates that a bin number (u) of a DFT computed from windows of different sizes or different signals means different frequencies.

The inverse relationship between frequency resolution and spatial resolution (window size) can be demonstrated using the following example. Suppose there are two sine waves with very small frequency difference [2]:

$$\begin{array}{ll} \text{Sine wave one:} & \sin(2\pi \times 0.05x) \\ \text{Sine wave one:} & \sin(2\pi \times 0.0501x) \end{array}$$

In this case, Δu = 0.0001. If we plot the two sine waves in one graph (Fig. 1.6), one in red and the other in blue, the two signals do not show a difference in the first 100 samples, which means a small window cannot discern the difference of the two signals. However, if we show the two signals in a very large window (5,000 samples), at the end of the window, they are 180° out of phase. This is because the periods of the two sine waves are 20 and 19.96, respectively. Assume a sampling frequency of 1 Hz. For a 100 samples window, the difference between the two signals is just 5.01 − 5 = 0.01 period, which is almost indiscernible. However, with a 5,000 samples window, the difference between the two signals is 250.5 − 250 = 0.5 period, which is more than sufficient to distinguish the two signals. It is more convenient to explain this case using (1.25), because the smallest frequency difference can be detected in a 100 samples window is Δu = 1/100 = 0.01, while in

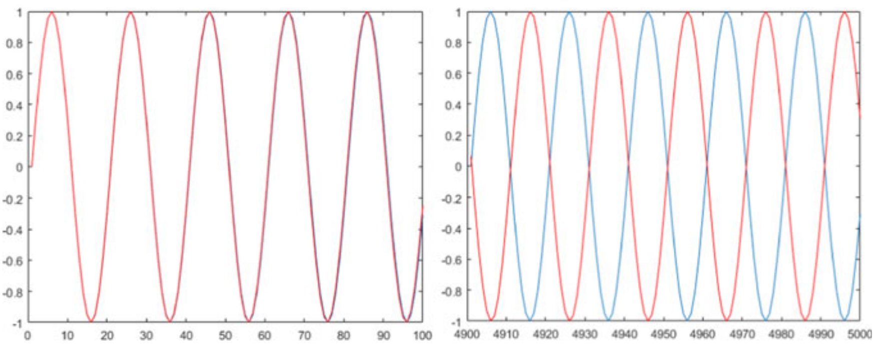

Fig. 1.6 Inverse relationship between spatial and frequency resolution. Left: the first 100 samples of the two sine waves; right: the last 100 samples of the two sine waves

a 5,000 samples window, $\Delta u = 1/5{,}000 = 0.0002$, which is able to distinguish the two signals.

It demonstrates that a smaller window gives poor frequency resolution, while a larger window gives higher frequency resolution. This is because the larger the window, the more samples, and the more low frequencies can be computed.

1.4.3 Nyquist Theorem

Because frequency is measured by the number of cycles in a period of time, and the smallest cycle consists of two samples, for a signal of size N, only $N/2$ frequencies can be computed from the DFT. This is called the *Nyquist theorem*.

Another way to express the *Nyquist theorem* is that in order to reconstruct/recover a signal appropriately ("appropriately" means recover the "essence" or low frequency while ignoring the "nuance" or high frequency), the sampling rate of the signal must be at least twice the highest frequency in the signal. Figure 1.7 demonstrates this fact. The figure shows three signals of 1 s length. The top signal is a sine wave with 1 cycle/period (frequency = 1) which can be recovered or reconstructed appropriately by at least two samples (marked with red dots). The middle signal is a sine wave with two cycles/periods in a second (frequency = 2); it needs at least four samples to recover the signal appropriately. The

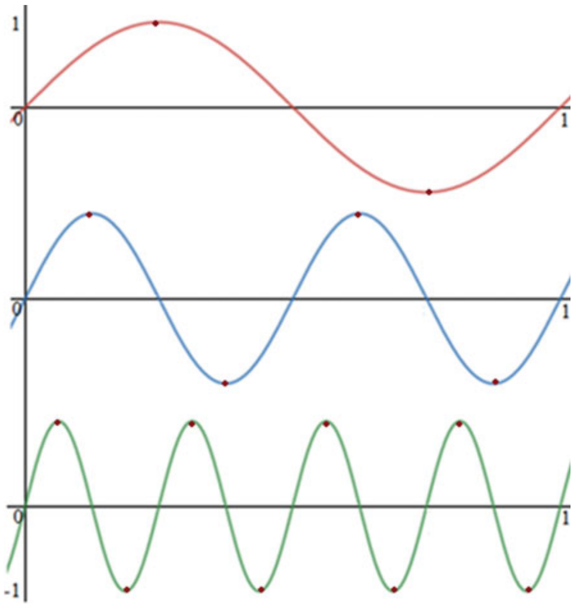

Fig. 1.7 Illustrations of different sampling rates for three signals of the same time length

bottom signal is a sine wave with 4 cycles/periods in a second (frequency = 4); it needs at least eight samples to recover the signal appropriately, and so on so forth.

In the case of images, frequency is related to structure size, and small structures are known to have high frequency. Because the smallest structure in an image requires 2 pixels to discern, the highest frequency which can be captured in an image is 1/2 pixels.

1.5 2D Fourier Transform

For a two-variable function $f(x, y)$ defined in $0 \leq x, y < N$, its Fourier transform pair is given by

$$F(u, v) = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x, y)\exp[-j2\pi(ux + vy)/N] \tag{1.27}$$

for $u, v = 0, 1, 2, \ldots, N - 1$, and $j = \sqrt{-1}$.

$$f(x, y) = \frac{1}{N}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1} F(u, v)\exp[j2\pi(ux + vy)/N] \tag{1.28}$$

for $x, y = 0, 1, 2, \ldots, N - 1$.

Although the number of $F(u, v)$ resulted from Fourier transform is usually large, the number of significant $F(u, v)$ (or $F(u)$) (large magnitude) is usually small. This is because the higher frequencies only represent the finest pattern details which are not so useful in many applications. This means that a meaningful approximation of original pattern $f(x, y)$ (or $f(x)$) can be constructed from a small number of $F(u, v)$ (or $F(u)$). This forms the basis of Fourier signal processing and Fourier pattern analysis.

1.6 Properties of 2D Fourier Transform

Fourier transform has the following important properties which are useful for image analysis.

1.6.1 Separability

The discrete Fourier transform can be expressed in the separable form

$$
\begin{aligned}
F(u,v) &= \frac{1}{\sqrt{N}}\sum_{x=0}^{N-1}\left[\frac{1}{\sqrt{N}}\sum_{y=0}^{N-1} f(x,y)\exp\left(-\frac{j2\pi vy}{N}\right)\right]\exp\left(-\frac{j2\pi ux}{N}\right) \\
&= \frac{1}{\sqrt{N}}\sum_{x=0}^{N-1} F(x,v)\exp\left(-\frac{j2\pi ux}{N}\right) \\
&= FT_x\{FT_y[f(x,y)]\}
\end{aligned} \tag{1.29}
$$

where FT_x and FT_y are the 1D FTs on row and column, respectively.

The advantage of the separability is that $F(u, v)$ can be obtained in two steps by successive applications of 1D FT which can be computed using the *Fast Fourier Transform* (FFT).

1.6.2 Translation

The translation property of the Fourier transform is given by

$$FT[f(x-x_0, y-y_0)] = F(u,v)\cdot\exp[-j2\pi(ux_0+vy_0)/N] \tag{1.30}$$

It indicates that a shift in spatial domain results in a phase change in frequency domain. That means the magnitude of Fourier transform is invariant to translation. This is a desirable feature, because, in many applications, the phase information is discarded which leaves the FT features invariant to translation.

1.6.3 Rotation

To find the relationship between a rotated function $f(x, y)$ and its spectrum, let's assume the function $f(x, y)$ is rotated by an angle θ, and the function after the rotation is $f(x', y')$. Then the relationship between two corresponding points of the two functions is as follows:

$$x' = x\cos\theta + y\sin\theta \tag{1.31}$$

$$y' = y\cos\theta - x\sin\theta \tag{1.32}$$

$$x = x'\cos\theta - y'\sin\theta \tag{1.33}$$

$$y = x'\sin\theta + y'\cos\theta \tag{1.34}$$

By substituting (1.33) and (1.34) into (1.27), we have

$$\begin{aligned} F(u', v') &= \frac{1}{N}\sum_{x'=0}^{N-1}\sum_{y'=0}^{N-1} f(x', y') \exp\left[-j2\pi\left(\frac{ux'\cos\theta - uy'\sin\theta + vx'\sin\theta + vy'\cos\theta}{N}\right)\right] \\ &= \frac{1}{N}\sum_{x'=0}^{N-1}\sum_{y'=0}^{N-1} f(x', y') \exp\left[-j2\pi x'\left(\frac{u\cos\theta + v\sin\theta}{N}\right)\right] \exp\left[-j2\pi y'\left(\frac{v\cos\theta - u\sin\theta}{N}\right)\right] \\ &= \frac{1}{N}\sum_{x'=0}^{N-1}\sum_{y'=0}^{N-1} f(x', y') \exp\left[-j2\pi\left(\frac{x'u' + y'v'}{N}\right)\right] \end{aligned} \tag{1.35}$$

where

$$u' = u\cos\theta + v\sin\theta \tag{1.36}$$

$$v' = v\cos\theta - u\sin\theta \tag{1.37}$$

Therefore, rotating $f(x, y)$ by an angle of θ in spatial domain rotates $F(u, v)$ by the same angle in frequency domain.

The rotation property can be proved more conveniently by considering $f(x, y)$ and FT in either complex domain or polar space. A point (x, y) in complex domain can be expressed as

$$z = x + jy \tag{1.38}$$

By using Euler's formula, it is simple to shown that

$$\begin{aligned} ze^{-j\theta} &= (x + jy)\cdot(\cos\theta - j\sin\theta) \\ &= x\cos\theta + jy\cos\theta - jx\sin\theta + y\sin\theta \\ &= (x\cos\theta + y\sin\theta) + j(y\cos\theta - x\sin\theta) \\ &= x' + jy' \end{aligned} \tag{1.39}$$

Equation (1.39) shows that a point z rotated by an angle θ is equivalent to z times $e^{-j\theta}$. In other words, the following is true:

$$f(x', y') = f(x, y)\cdot e^{-j\theta} \tag{1.40}$$

Equation (1.40) is a more concise and convenient rotation formula than (1.31) and (1.32). By substituting (1.40) into (1.27), we obtain the FT of the rotated function $f(x', y')$:

$$\begin{aligned} F(u',v') &= \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x',y')\exp\left[-j2\pi\left(\frac{ux+vy}{N}\right)\right] \\ &= \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)e^{-j\theta}\exp\left[-j2\pi\left(\frac{ux+vy}{N}\right)\right] \\ &= F(u,v)\cdot e^{-j\theta} \end{aligned} \tag{1.41}$$

Therefore, we obtain the same result as shown in (1.35).

If we consider both $f(x, y)$ and $F(u, v)$ in polar space, they can be expressed as $f(r, \theta)$ and $F(\rho, \phi)$, respectively, where

$$x = r\cos\theta, y = r\sin\theta; \quad u = \rho\cos\phi, v = \rho\sin\phi \tag{1.42}$$

(r, θ) is the polar coordinates in image plane and (ρ, ϕ) is the polar coordinates in frequency plane. The differentials of x and y are

$$\left.\begin{aligned} dx &= \cos\theta\, dr - r\sin\theta\, d\theta \\ dy &= \sin\theta\, dr - r\cos\theta\, d\theta \end{aligned}\right\} \tag{1.43}$$

The Jacobian of (1.43) is r. Therefore, by substituting both (1.42) and (1.43) into 2D continuous FT, the 2D FT in polar space is given by the following equations:

$$\begin{aligned} F(\rho,\phi) &= \int_0^{\infty}\int_0^{2\pi} f(r,\theta)e^{-j2\pi(rcos\theta\rho cos\phi + rsin\theta\rho sin\phi)} rdrd\theta \\ &= \int_0^{\infty}\int_0^{2\pi} f(r,\theta)e^{-j2\pi r\rho\cos(\theta-\phi)} rdrd\theta \end{aligned} \tag{1.44}$$

Suppose $f(r, \theta)$ is rotated for an angle of θ_0 to $f(r, \theta + \theta_0)$. Let $\theta' = \theta + \theta_0$, then

$$\theta = \theta' - \theta_0 \quad \text{and} \quad dq = d\theta' \tag{1.45}$$

Now, in (1.44), by substituting $f(r, \theta)$ with $f(r, \theta')$ and substituting θ with (1.45), we obtain

$$F(\rho,\phi') = \int_0^{\infty}\int_0^{2\pi} f(r,\theta')e^{-j2\pi r\rho\cos[\theta'-(\phi+\theta_0)]} rdrd\theta' \tag{1.46}$$

Equation (1.46) means

$$FT[f(r, \theta + \theta_0)] = F(r, \phi + \theta_0) \tag{1.47}$$

Again, this yields the same result as (1.35) and (1.41). Equation (1.47) also tells that in polar domain, the rotation of an image causes a translation or shift on its FT spectrum. This property is useful for feature normalization.

1.6.4 Scaling

For two scalars a and b, the scale property of Fourier transform is given by

$$FT[f(ax, by)] = \frac{1}{ab} F\left(\frac{u}{a}, \frac{v}{b}\right) \tag{1.48}$$

It indicates the scaling of $f(x, y)$ with a and b in x and y directions in spatial domain (time domain in 1D case) causes inverse scaling of magnitude of $F(u, v)$ in frequency domain. That means, if you stretch $f(x, y)$ in spatial domain, you shrink $F(u, v)$ in frequency domain and vice versa. This proves the *uncertainty principle* from another perspective. In general terms, enlarging an object in an image gives rise to lower frequencies in spectral domain while shrinking an object in an image gives rise to higher frequencies in spectral domain. This property is useful in dealing with image scaling.

1.6.5 Convolution Theorem

The *convolution theorem* states that the FT of a convolution between two functions is equal to the product of two FTs. Specifically, given two function f and g, the following are true:

$$FT[f * g] = FT[f] \cdot FT[g] \tag{1.49}$$

$$f * g = FT^{-1}\{FT[f] \cdot FT[g]\} \tag{1.50}$$

where $f * g$ means convolution. Because of the separability property of 2D FT, we only need to prove the 1D case.

$$\begin{aligned} FT[f * g] &= \sum_n \sum_m f(m) g(n-m) e^{-\frac{j2\pi nu}{N}} \\ &= \sum_m f(m) \sum_n g(n-m) e^{-\frac{j2\pi nu}{N}} \\ &= \sum_m f(m) FT[\mathrm{g}] e^{-\frac{j2\pi mu}{N}} (translation\, property) \\ &= FT[g] \sum_n f(m) e^{-\frac{j2\pi mu}{N}} \\ &= FT[f] \cdot FT[g] \end{aligned} \tag{1.51}$$

Convolution theorem shows that convolution in spatial domain can be done by an FT (FFT in practice) and a product. This is a useful feature because both FFT and product are much more efficient than spatial convolution.

1.7 Techniques of Computing FT Spectrum

The magnitude image of a Fourier transform is called an FT spectrum. The intensity of an FT spectrum has a very large dynamic range; it is impossible to display this large range in a gray level image. For example, the dynamic range of spectral values of the Lena image [3] is [0, 31, 744], and Fig. 1.8a shows the FT spectrum without scaling. It can be seen that the spectral image reveals little information about the input image. Conventional *thresholding* (Fig. 1.8b) and *linear scaling* (Fig. 1.8c) do not work well for such a large range of values.

The common practice of displaying an FT spectrum is to do a *logarithmic* transformation of the spectral values to bring down the large spectral values to well within the display range of 255 and raise the lower spectral values in the meantime. However, the logarithm transformed spectrum does not have sufficient contrast between lower frequency and higher frequency spectral values as shown in Fig. 1.8d. A more effective way to display FT spectrum is to apply a logarithm transform on the spectral values followed by a linear scaling to map the spectral magnitudes to [0, 255] using (1.52):

$$F'(u, v) = 255 \times \frac{\log(1 + |F(u, v)|)}{\log[1 + \max(|F(u, v)|)]} \tag{1.52}$$

Figure 1.8e shows the FT spectrum using (1.52). It can be seen from Fig. 1.8e that there are three directional features in the FT spectrum: horizontal, vertical, and diagonal. The strong horizontal feature is due to the vertical pole on the left-hand

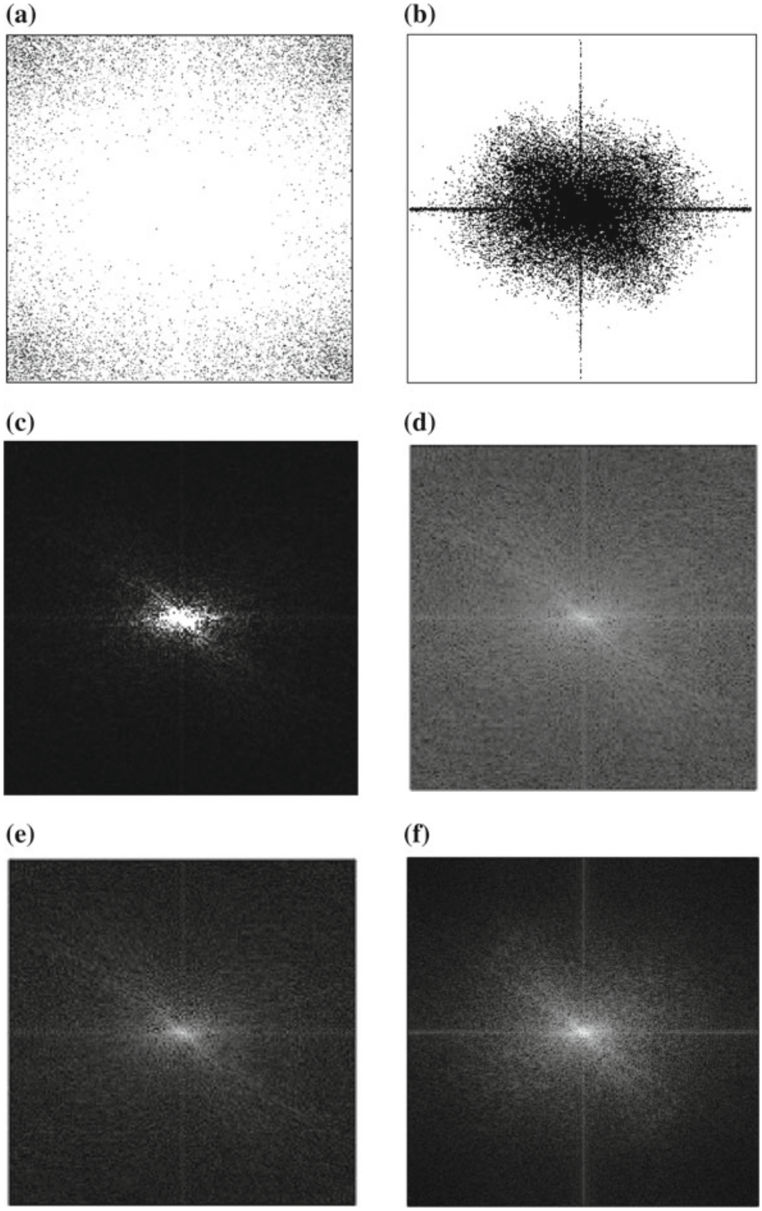

Fig. 1.8 FT spectra with different methods. **a** FT spectrum without scaling; **b** FT spectrum with thresholding value 10; **c** FT spectrum with linear scaling; **d** FT spectrum with log transform; **e** FT spectrum with both log and linear transform; and **f** FT spectrum with enhanced contrast from (**e**)

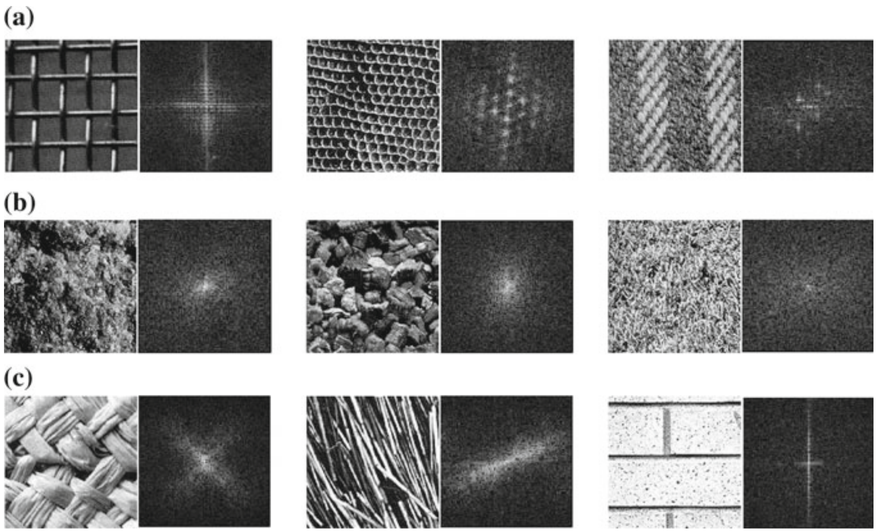

Fig. 1.9 FT spectra of different types of patterns. **a** Regular patterns and their FT spectra on the right; **b** random patterns and their FT spectra on the right; and **c** directional patterns and their FT spectra on the right

side of the input image, while the vertical and diagonal features are due to the rim of the hat and the black arch on the right-hand side of the input image.

The logarithmic transformation, however, enhances the low magnitude values, while compressing high magnitude values into a relatively small pixel range. Therefore, if an image contains some important high magnitude information, this may lead to loss of information. An alternative solution to further increase the spectral contrast is to decrease the compression rate by scaling down the spectrum image intensity before applying the logarithmic transform. This is because the logarithmic function has a less degree of compression at places close to the origin. Figure 1.8f shows the FT spectrum with enhanced contrast, which is equivalent to highlighting the low-frequency area with a spotlight.

The FT spectrum reveals key information about an image if displayed properly. Figure 1.9 shows three different types of homogenous patterns and their FT spectra on the right-hand side of the patterns. It can be seen that the FT spectra have generally accurately captured the three types of texture features: regularity, randomness, and directionality. This is the primary motive for the development of short-time FT and wavelets, which attempt to capture local and changing patterns.