Chapter 7
Spatial Domain Image Transforms

7.1 Introduction

In the material presented in Chap. 5 we looked at a number of geometric processing operations that involve the spatial properties of an image. Averaging over adjacent groups of pixels to reduce noise, and looking at local spatial gradients to enhance edge and line features, are examples. There are, however, more sophisticated approaches for processing the spatial domain properties of an image. The most recognisable is probably the Fourier transformation which we consider in this chapter, allowing us to understand what are called the spatial frequency components of an image. Once we can use the Fourier transformation we will see that it offers a powerful method for generating the sorts of operation we did with templates in Chap. 5.

There are other spatial transformation techniques as well. Some we will mention in passing that find application in image compression in the video and television industry. More recent techniques, such as the wavelet transform, have emerged as important image processing tools in their own right, including in remote sensing. The wavelet transform is introduced later in this chapter.

We commence with some necessary mathematical background that leads, in the first instance, to the principles of sampling theory. That is sampling, not in the statistical sense that we use when testing map accuracy in Chap. 11, but in the role of digitising a continuous landscape to produce an image composed of discrete pixels. This material is based on two mathematical fields that the earth science reader may not have encountered in the past: the first is complex numbers and the second is integral calculus. We will work our way through that material as carefully as possible but for those readers not needing a background on transformation methods, this chapter can be passed over without affecting an understanding of the remainder of the book.
7.2 Special Functions

Three special functions are important in understanding the development of sampling theory and the transformations treated here. We will consider them as functions of time, because that is the way most are presented in classical texts, but they will be interpreted as functions of position, in either one or two dimensions as required, later in the chapter.

7.2.1 The Complex Exponential Function

Several of the functions we meet here involve imaginary numbers which arise when we try to take the square root of a negative number. The most basic is the square root of \(-1\). Although that may seem to be an unusual concept, it is an enormously valuable mathematical construct in developing transformations. It is sufficient here to consider \(\sqrt{-1}\) as a special symbol rather than try to understand the logical implications of taking the square root of something that is negative. It is given the symbol \(j\) in the engineering literature, but is represented by \(i\) in the mathematical literature.

By definition, a complex exponential function \(^{1}\) that is periodic with time is

\[
g(t) = e^{j\omega t}
\]  

(7.1)

This is best looked at as a vector that rotates in an anticlockwise direction in a two-dimensional plane (called an Argand diagram) described by real and imaginary number axes as shown in Fig. 7.1. The concept of an imaginary number is developed below. If the exponent in (7.1) were negative the vector would rotate in the clockwise direction. As the vector rotates from its position at time zero its projection onto the axis of real numbers plots out the cosine function whereas its projection onto the axis of imaginary numbers plots out the sine function. One complete rotation of the vector, covering \(360^\circ\) or \(2\pi\) radians, takes place in a time \(t = T\) defined by \(\omega T = 2\pi\). \(T = 2\pi/\omega\) is called the period of the function, measured in seconds, and \(\omega\) is called its radian frequency, with units of radians per second. Note, for example, if the period is 10 ms, the radian frequency would be \(200\pi = 628\) rad s\(^{-1}\). Often we describe frequency \(f\) in hertz rather than as \(\omega\) in radians per second. The two measures are related by

\[
\omega = 2\pi f
\]  

(7.2)

The complex exponential expression in (7.1) can be written\(^2\)

\(^{1}\) We use the symbol \(g\) here for a general function, rather than the more usual \(f\), to avoid confusion with the symbol universally used for frequency.

\(^2\) This is known as Euler’s theorem.
or, if the sign of the exponent is reversed,

\[ e^{-j\omega t} = \cos \omega t - j \sin \omega t \]  

Both (7.3a) and (7.3b) can be written in the form \( g(t) = a(t) + jb(t) \) in which \( a(t) \) is called the real part of \( g(t) \) and \( b(t) \) is called the imaginary part.

The expressions in (7.3a) and (7.3b) can be demonstrated from Fig. 7.1 if the rotating vector is represented by two Cartesian coordinates with unit vectors \((1,j)\) in the horizontal and vertical directions respectively. In this way the symbol \( j \) is nothing other than a vector that points in the vertical direction. From (7.3a) we can see that

\[ \cos \omega t = Re\{e^{j\omega t}\} \]  

\[ \sin \omega t = Im\{e^{j\omega t}\} \]  

in which \( Re \) and \( Im \) are operators that pick out the real (horizontal) or imaginary (vertical) parts of the complex exponential. Alternatively, from (7.3a) and (7.3b) we can see that

\[ \cos \omega t = \frac{1}{2}(e^{j\omega t} + e^{-j\omega t}) \]  

\[ \sin \omega t = \frac{1}{2j}(e^{j\omega t} - e^{-j\omega t}) \]
7.2.2 The Impulse or Delta Function

An important function for understanding the properties of sampled signals, including digital image data, is the impulse function. It is also referred to as the Dirac delta function, or simply the delta function. It is spike-like, of infinite amplitude and infinitesimal duration. It cannot be defined explicitly. Instead, it is described by a limiting operation in the following manner.

Consider the rectangular pulse of duration $\alpha$ and amplitude $1/\alpha$ shown in Fig. 7.2. Note that the area under the curve is 1. If we let the value of $\alpha$ go to 0 then the function grows in amplitude and tends to an infinitesimal width. We define the delta function by that limiting operation. As a formal definition, the best we can do then is to say

$$\delta(t) = 0 \text{ for } t \neq 0$$  \hspace{1cm} (7.6a)

and\(^3\)

$$\int_{-\infty}^{\infty} \delta(t) dt = 1$$  \hspace{1cm} (7.6b)

This turns out to be sufficient for most purposes in engineering and science.

Equations (7.6a, 7.6b) define a delta function at the origin; an impulse located at time $t_0$ is defined by

$$\delta(t - t_0) = 0 \text{ for } t \neq t_0$$  \hspace{1cm} (7.6c)

and

$$\int_{-\infty}^{\infty} \delta(t - t_0) dt = 1$$  \hspace{1cm} (7.6d)

---

\(^3\) Recall that the integral of a function over its range is equal to the area under its curve.
If we take the product of a delta function with another function the result, from (7.6c), is
\[ \frac{d}{C_0} t - t_0(g(t)) = \frac{d}{C_0} t - t_0(g(0)) \] (7.7)

From this we see
\[ \int_{-\infty}^{\infty} \delta(t - t_0)g(t)dt = \int_{-\infty}^{\infty} \delta(t - t_0)g(t_0)dt = g(t_0) \int_{-\infty}^{\infty} \delta(t - t_0)dt = g(t_0) \] (7.8)

This is known as the sifting property of the impulse.

### 7.2.3 The Heaviside Step Function

The Heaviside or unit step function is shown in Fig. 7.3 and is defined by
\[ u(t - t_0) = \begin{cases} 1 & \text{for } t \geq t_0 \\ 0 & \text{for } t < t_0 \end{cases} \] (7.9)

The step and delta functions are related by
\[ \delta(t - t_0) = \frac{du(t - t_0)}{dt} \]

### 7.3 The Fourier Series

We now come to a very important concept in the analysis of functions and signals, which we will apply later to images. If a function of time is periodic, in the sense that it repeats itself with some regular interval such as the square waveform shown...
in Fig. 7.4, then it can be written as the sum of sinusoidal signals or the sum of complex exponential signals, called a Fourier series. We write a periodic function as \( g(t) = g(t + T) \) where, again, \( T \) is its period. In the terminology of the complex exponential the Fourier series of the function \( g(t) \) is written

\[
g(t) = \sum_{n=-\infty}^{\infty} G_n e^{j\omega_0 n t} \quad (7.10a)
\]

in which \( \omega_0 = 2\pi / T \), and \( n \) is an integer. The coefficients \( G_n \) tell us how much of each sinusoidal frequency component is present in the composition of \( g(t) \). Notice however that there are coefficients with positive and negative indices corresponding to positive and negative frequency components. That can be understood by noting in (7.5a) and (7.5b) that the pure trigonometric functions are composed of exponentials with positive and negative exponents. The two sided summation in (7.10a) recognises that property. One might ask why the trigonometric functions themselves are not used in developing the Fourier series. The fact is they can be and are; it is just that the exponential form is more convenient mathematically and has become the standard expression in engineering and science.

The expansion coefficients \( G_n \) are, in general, complex numbers. Finding their values is the most significant part of using the Fourier series. They are given by

\[
G_n = \frac{1}{T} \int_{-T/2}^{T/2} g(t) e^{-j\omega_0 n t} dt \quad (7.10b)
\]

To understand their importance consider the Fourier series of the square waveform in Fig. 7.4. Over the range \((-T/2, T/2)\) covered by the integral the square waveform is zero, except between \((-T/4, T/4)\) over which it is unity. Therefore Eq. (7.10b) becomes

---

Fig. 7.4 A square waveform with period \( T \)
The first few values of this last expression for positive and negative values of $n$ are

\[ n = -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5 \]

\[ G_n = \frac{1}{T} \int_{-T/4}^{T/4} e^{-jn\omega_0 t} dt = \frac{1}{n\pi} \sin \frac{n\pi}{2} \]

These are shown plotted in Fig. 7.5 in two forms: in the second, they are represented by their amplitudes and phase angles. Any complex number can be written in either of two forms—the Cartesian and polar forms. Which one to use is dictated by the application at any given time. The Cartesian form $a + jb$ can be converted to the polar form $Re^{j\theta}$ where $R = \sqrt{a^2 + b^2}$ and $\theta = \tan^{-1}(b/a)$. In our work the polar form, in which the number has an amplitude $R$ and phase $\theta$, is most common.
The graphs shown in Fig. 7.5 are frequently referred to as spectra, as indicated; generically the set of $G_n$ is also referred to as the spectrum of the function $g(t)$. Note that an angle (phase) of $\pm 180$ is equivalent to $-1$ in complex numbers. These results tell us that the square wave of Fig. 7.4 can be made up from 0.500 parts of a constant, plus 0.318 parts of a pure sinusoid with the same fundamental frequency as the square wave, minus 0.106 parts of a sine wave with three times the frequency (said to be the third harmonic), plus 0.064 parts of a sine wave at five times the frequency (fifth harmonic), and so on. The square wave has no even harmonic components. Again, remember that the two components for the same frequency either side of the origin are the amplitudes of the positive and negative exponential components that constitute a sinusoid as in (7.5a) and (7.5b).

7.4 The Fourier Transform

The Fourier series describes a periodic function as the sum of a discrete number of sinusoids, expressed in complex exponentials at integral multiples of the fundamental frequency. For functions that are non-periodic, called aperiodic, decomposition into sets of sinusoids is still possible. However, whereas the Fourier series yields a set of distinct and countable components, the spectrum of a non-periodic function can consist of an infinite set of sinusoids at every conceivable frequency. To find that composition we use the Fourier transform, or Fourier integral, defined by

$$G(x) = \int_{-\infty}^{\infty} g(t)e^{-j\omega t}dt$$

(7.11a)

This is the equivalent expression to (7.10b). The major difference is that there is no sense of period in this equation; in addition, the frequency term in the exponent is now a variable $\omega$ rather than a discrete set $\{ n\omega_0 \}$ as was the case for the Fourier series. Writing the transform as a function of the continuous frequency variable indicates the likelihood that it exists at all possible frequencies.

If we know the Fourier transform, or spectrum, of a function then the function can be reconstructed according to

$$g(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\omega)e^{j\omega t}d\omega$$

(7.11b)

This is the equivalent of the Fourier series expression of (7.10a).

To demonstrate the application of the Fourier transform consider the single unit pulse shown in Fig. 7.6a. From (7.11a) this is seen to be

---

5 There are a number of conventions for defining the Fourier integral, largely to do with where the $2\pi$ is placed between (7.11a) and (7.11b).
which is shown plotted in Fig. 7.6b. Again, note that the frequency axis accommodates both positive and negative frequencies, which together compose sinusoids. Note also that every possible frequency exists in the spectrum of the pulse, apart from a set where the spectrum crosses the frequency axis. The spectrum has negative as well as positive values. When negative the phase spectrum has a value of \(-180^\circ\).

An important Fourier transform is that of an impulse function. From the sifting property of the impulse and the fact that 
\[ e^0 = 1, \]
the transform is
\[ G(\omega) = \int_{-\infty}^{\infty} \delta(t)e^{-j\omega t}dt = 1, \]
This tells us that the impulse is composed of equal amounts of every possible frequency! It also suggests that functions which change rapidly in time will have large numbers of frequency components. It is interesting to consider the Fourier transform of a constant \(c\). That can be shown to be
\[ G(\omega) = \int_{-\infty}^{\infty} ce^{-j\omega t}dt = 2\pi c\delta(\omega) \]
In other words, as expected, the spectrum of a constant exists only at the origin in the frequency domain.6 This expression can be derived based on the properties of the complex exponential; it can be verified by working backwards through (7.11b). While it might seem strange having the impulse function appear as a multiplier for

---

6 If the transform at (7.11a) and (7.11b) had used the \(2\pi\) as a denominator in (7.11a) then the \(2\pi\) would not appear in the transform of a constant. If the transforms were defined in terms of frequency \(f = \omega/2\pi\), the \(2\pi\) factor would not appear at all. Some authors use that form.
the constant, we can safely interpret that as just a reminder that the constant exists at $\omega = 0$ and not as something that changes the amplitude of the constant.

The Fourier transform of a periodic function, normally expressed using the Fourier series, is important. We can obtain it by substituting (7.10a) into (7.11a) to give

$$G(\omega) = \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} G_n e^{j\omega_0 t} e^{-j\omega t} dt$$

i.e.

$$G(\omega) = \sum_{n=-\infty}^{\infty} G_n \int_{-\infty}^{\infty} e^{j(n\omega_0 - \omega)t} dt$$

which becomes

$$G(\omega) = 2\pi \sum_{n=-\infty}^{\infty} G_n \delta(\omega - n\omega_0) \quad (7.12)$$

This last expression uses the property for the Fourier transform of a constant (in this case unity). It tells us that the only frequencies that can exist in the Fourier transform of a periodic function are those which are integral multiples of the fundamental frequency of the periodic waveform.

The last example is important because it tells us that we do not need to work with both Fourier series and Fourier transforms. Because the Fourier transform also applies to periodic functions, it is sufficient in the following to focus on the transform alone.

### 7.5 The Discrete Fourier Transform

Because our interest is in digital imagery, which is simply a two-dimensional version of the one-dimensional functions we have been considering up to now, it is important to move away from continuous functions of time (or any other independent variable) and instead look at the situation when the independent variable is discrete, and the dependent variable consists of a set of samples.

Suppose we have a set of $K$ samples over a time interval $T$ of the continuous function $g(t)$. Call these samples $g(k), k = 0 \ldots K - 1$. The individual samples occur at the times $t_k = k\Delta t$ where $\Delta t$ is the spacing between samples. Note that $T = K\Delta t$. Consider now how (7.11a) needs to be modified to handle the set of samples rather than a continuous function of time. First, obviously the function itself is replaced by the samples. Secondly the integral over time is replaced by the sum over the samples and the infinitesimal time increment $dt$ in the integral is replaced by the sampling time increment $\Delta t$. The time variable $t$ is replaced by $k\Delta t = \frac{k T}{K}, k = 0 \ldots K - 1$. So far this gives as a discrete form of (7.11a)
We now have to consider how to treat the frequency variable $\omega$. We are developing this discrete form of the Fourier transformation so that it can handle digitised data and so that it can be processed by computer. Therefore, the frequency domain also has to be digitised by replacing $\omega$ by the frequency samples $\omega = r\Delta\omega$, $r = 0 \ldots K - 1$. We have deliberately chosen the number of samples in the frequency domain to be the same as the number in the time domain for convenience. What value now do we give to the increment in frequency $\Delta\omega$? That is not easily answered until we have treated sampling theory later in this chapter, so for the present note that it will be $2\pi/T$ and thus is directly related to the time over which the original function has been sampled. With this treatment of the frequency variable the last expression for the discrete version of the Fourier transform now becomes

$$G(r) = \Delta t \sum_{k=0}^{K-1} \gamma(k)e^{-j2\pi rk} r = 0 \ldots K - 1$$

It is common to define

$$W = e^{-j2\pi/K}$$

so that the last expression is written

$$G(r) = \Delta t \sum_{k=0}^{K-1} \gamma(k)W^{rk} r = 0 \ldots K - 1$$ (7.14a)

Equation (7.14a) is known as the \textit{discrete Fourier transform} (DFT). In a similar manner we can derive a \textit{discrete inverse Fourier transform} (IDFT) to allow the original sequence $\gamma(k)$ to be reconstructed from the frequency samples $G(r)$. That is given by

$$\gamma(k) = \frac{1}{T} \sum_{r=0}^{K-1} G(r)W^{-rk} k = 0 \ldots K - 1$$ (7.14b)

If we substitute (7.14a) into (7.14b) we see they do in fact constitute a transform pair. To do so we need different indices for $k$; we will use $l$ instead of $k$ in (7.14b) so that

$$\gamma(l) = \frac{1}{T} \sum_{r=0}^{K-1} G(r)W^{-rl} = \frac{1}{T} \sum_{r=0}^{K-1} \Delta t \sum_{k=0}^{K-1} \gamma(k)W^{rk}W^{-rl}$$

i.e.

$$\gamma(l) = \frac{1}{K} \sum_{k=0}^{K-1} \gamma(k) \sum_{r=0}^{K-1} W^{r(k-l)}$$
From the properties of the complex exponential function the second sum is zero for \( k \neq l \) and is \( K \) when \( k = l \), so that the right hand side becomes \( \gamma(l) \) as required. An interesting by-product of this analysis has been that the \( \Delta t \) and \( T \) divide to leave \( K \), the number of samples. As a result, the transform pair in (7.14a) and (7.14b) can be written in the simpler form, used in software that computes the discrete Fourier transform:

\[
G(r) = \sum_{k=0}^{K-1} \gamma(k)W^{rk} \quad r = 0 \ldots K - 1
\]  

\[
\gamma(k) = \frac{1}{K} \sum_{r=0}^{K-1} G(r)W^{-rk} \quad k = 0 \ldots K - 1
\]

These last two expressions are particularly simple. All they involve are the sets of samples to be transformed (or inverse transformed) and the complex constants \( W \), which can be computed beforehand.

### 7.5.1 Properties of the Discrete Fourier Transform

Three properties of the DFT and IDFT are important.

**Linearity**

Both the DFT and IDFT are linear operations. Thus, if the set \( G_1(r) \) is the DFT of the sequence \( \gamma_1(k) \) and the set \( G_2(r) \) is the DFT of the sequence \( \gamma_2(k) \) then, for any constants, \( a \) and \( b \), \( aG_1(r) + bG_2(r) \) is the DFT of \( a\gamma_1(k) + b\gamma_2(k) \).

**Periodicity**

From (7.13) \( W^{\pm mkK} = 1 \) where \( m \) and \( k \) are integers, so that

\[
G(r + mK) = \sum_{k=0}^{K-1} \gamma(k)W^{(r+mK)k} = G(r)
\]

Similarly

\[
\gamma(k + mK) = \frac{1}{K} \sum_{r=0}^{K-1} G(r)W^{-(r+mK)} = \gamma(k)
\]

Thus both the sequence of samples and the set of transformed samples are periodic with period \( K \). This has two important implications. First, to generate the Fourier series of a periodic function it is only necessary to sample it over a single period. Secondly, the spectrum of an aperiodic function will be that of a periodic repetition of that function over the sampling duration—in other words it is made to look periodic by the limited time sampling. Therefore, for a limited time function such
as the rectangular pulse shown in Fig. 7.6, it is necessary to sample the signal beyond the arguments for which it is non-zero so it looks approximately aperiodic.

**Symmetry**
Put \( r' = K - r \) in (7.15a) to give

\[
G(r') = \sum_{k=0}^{K-1} \gamma(k) W^{(K-r)k}
\]

Since \( W^{KK} = 1 \) then \( G(K - r) = G^*(r) \) where the * represents the complex conjugate. This implies that the amplitude spectrum is symmetric about \( K/2 \) and the phase spectrum is antisymmetric.

### 7.5.2 Computing the Discrete Fourier Transform

Evaluating the \( K \) values of \( G(r) \) from the \( K \) values of \( \gamma(k) \) in (7.15a) requires \( K^2 \) multiplications and \( K^2 \) additions, assuming that the values of \( W^{rk} \) have been calculated beforehand. Since those numbers are complex, the multiplications and additions required to evaluate the Fourier transform are also complex. It is the multiplications that are the problem; complex multiplications require significant computing resources, so that transforms involving of the order of 1,000 samples can take significant time. Fortunately, a fast algorithm, called the *fast Fourier transform* (FFT), is available.\(^7\) It only requires \( K^2 \log_2 K \) complex multiplications, which is a substantial reduction in computational demand. The implementation of the DFT in software uses the FFT algorithm. The only penalty in using this method is that the number of samples taken of the function to be transformed, and the number of samples in the transform, each have to be a power of two.

### 7.6 Convolution

#### 7.6.1 The Convolution Integral

Before proceeding to look at the Fourier transform of an image it is of value to appreciate the operation called convolution. It was introduced in Sect. 5.8 in the context of geometric enhancement of imagery. We now look at it in more detail because of its importance in understanding both sampled data and the spatial processing of images. As with the development of the Fourier transform we commence with its application to one dimensional, continuous functions of time. We will then

---

modify it to apply to samples of functions. After having considered the Fourier transform of an image we will look at the two-dimensional version of convolution.

Convolution is defined in terms of a pair of functions. For \( g_1(t) \) and \( g_2(t) \) the result is

\[
y(t) = g_1(t) \ast g_2(t) = \int_{-\infty}^{\infty} g_1(\tau)g_2(t - \tau)d\tau
\]

(7.17)
in which the symbol \( \ast \) indicates convolution. The operation commutative, i.e. \( g_1(t) \ast g_2(t) = g_2(t) \ast g_1(t) \), a fact sometimes used when evaluating the integral.

We can understand the convolution operation if we break it down into the following four steps, which are illustrated in Fig. 7.7:

i. Folding: form \( g_2(-\tau) \) by taking the mirror image of \( g_2(\tau) \) about the vertical axis

ii. Shifting: form \( g_2(t - \tau) \) by shifting \( g_2(-\tau) \) by the variable amount \( t \)

iii. Multiplication: form the product \( g_1(\tau)g_2(t - \tau) \)

iv. Integration: compute the area under the product

### 7.6.2 Convolution with an Impulse

Convolution of a function with an impulse is important in understanding sampling. The delta function sifting theorem in (7.8) gives

\[
y(t) = g(t) \ast \delta(t - t_0) = \int_{-\infty}^{\infty} g(\tau)\delta(t - \tau - t_0)d\tau = g(t - t_0)
\]

Thus, the result is to shift the function to a new origin. Clearly, for \( t_0 = 0 \), \( y(t) = g(t) \).

### 7.6.3 The Convolution Theorem

This theorem can be verified using the definitions of convolution and the Fourier transform.\(^8\) It has two forms:

\[
\text{If } y(t) = g_1(t) \ast g_2(t) \quad \text{then } Y(\omega) = G_1(\omega)G_2(\omega) \quad (7.18a)
\]

\[
\text{If } Y(\omega) = G_1(\omega) \ast G_2(\omega) \quad \text{then } y(t) = 2\pi g_1(t)g_2(t) \quad (7.18b)
\]

---

7.6 Convolution

7.6.4 Discrete Convolution

Just as we can modify the Fourier transform formula to handle discrete samples rather than continuous functions we can do the same with convolution. Suppose $\gamma_1(k)$ and $\gamma_2(k)$ are sampled versions of the functions $g_1(t)$ and $g_2(t)$ then (7.17) can be written in discrete form as

$$y(k) = \sum_n \gamma_1(n)\gamma_2(k - n) \quad \text{for all } k$$

(7.19)
The integral has been replaced by the sum operation. Strictly $d\tau$ should be replaced by the time increment between the samples, but it is usually left out in the discrete form. To evaluate (7.19) the set of samples $\gamma_2(k)$ is first reversed in order to form $\gamma_2(-n)$ and then slid past $\gamma_1(n)$ as shown in Fig. 7.8, taking products where samples coincide and then summing the results.

7.7 Sampling Theory

Discrete time functions, such as the sequence of samples that we considered when developing the discrete Fourier transform, and digital images, can be considered to be the result of sampling the corresponding continuous functions or scenes on a regular basis. The periodic sequence of impulses, spaced $\Delta t$ apart, sometimes called a Dirac comb,

\[ \mathcal{D}(t) = \sum_{k=-\infty}^{\infty} \delta(t - k\Delta t) \]  

(7.20)

can be used to extract a uniform set of samples from a function $g(t)$ by forming the product

\[ \mathcal{D}(t)g(t) \]

(7.21)
From (7.7) this is seen to be a sequence of samples of value \( g(kT)\delta(t - k\Delta t) \), which we represent by \( \gamma(k) \). Despite the undefined magnitude of the delta function we will be content in this treatment to regard the product simply as a sample of the function \( \gamma(t) \), so that (7.21) can be interpreted as a set of uniformly spaced samples of \( \gamma(t) \).

It is important now to know the Fourier transform of the samples in (7.21). We will find that by calling on the convolution theorem in (7.18b), which requires the Fourier transforms of \( g(t) \) and \( \mathfrak{D}(t) \): Since \( \mathfrak{D}(t) \) is periodic we can work via the Fourier series coefficient formula of (7.10b), which shows

\[
\mathfrak{D}_n = \frac{1}{\Delta t} \int_{-\Delta t/2}^{\Delta t/2} \delta(t)e^{-j\omega_0 t}dt = \frac{1}{\Delta t}
\]

so that from (7.12) the Fourier transform of the periodic sequence of impulses in (7.20) is

\[
\mathfrak{D}(\omega) = \frac{2\pi}{\Delta t} \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_s)
\]

(7.22)

in which \( \omega_s = 2\pi/\Delta t \). Thus, the Fourier transform of the periodic sequence of impulses spaced \( \Delta t \) apart in the time domain is itself a periodic sequence of impulses in the frequency domain spaced apart \( \frac{2\pi}{\Delta t} \) rad s\(^{-1}\), or \( \frac{\Delta t}{2}\) Hz.

Suppose \( g(t) \) has the Fourier transform, or spectrum, \( G(\omega) \); then from (7.18b) and (7.22) the spectrum of the samples of \( G(\omega) \) represented by (7.21) is given by convolving \( G(\omega) \) with (7.22). Convolution with an impulse shifts a function to a new origin centred on the impulse so that the outcome of this operation is a periodic repetition of the spectrum \( G(\omega) \)—the spectrum of the unsampled function—as shown in Fig. 7.9. The repetition period in the frequency domain is determined by the rate at which the time function is sampled—it is the inverse of the time domain sampling rate. Note that if the sampling rate is high then the repeated segments of the spectrum of the original function are well separated. If the sampling rate is low then those segments are closer together.

Imagine that the frequency components of the original function are limited to frequencies below \( \mathcal{B} \text{Hz} \) (or \( \frac{2\pi\mathcal{B}}{\Delta t} \) rad s\(^{-1}\)). \( \mathcal{B} \) is called the bandwidth of \( g(t) \). Not all real non-periodic functions have a limited bandwidth. The single pulse of Fig. 7.6 is an example. However, it suits our purpose here to assume that there is a limit \( \mathcal{B} \) to the frequency composition of those functions of interest to us. That allows us to introduce a particularly important concept in the sampling of signals and images. It is clear that if adjacent spectral segments in Fig. 7.9 are to remain separated then we require

\[
\text{sampling rate} = \frac{1}{\Delta t} > 2\mathcal{B}
\]

(7.23)

In other words, the rate at which the function \( g(t) \) is sampled must exceed twice its bandwidth. Should that not be the case then the segments of the spectrum will overlap as shown in Fig. 7.9d, causing a form of distortion called \textit{aliasing}. 

7.7 Sampling Theory 219
The sampling rate of $2B$ in (7.23) is called the Nyquist rate. Equation (7.23) is often called the sampling theorem.

Return now to the concept of the discrete Fourier transform. On the basis of the sampling material just presented we know that the spectrum of a band limited sampled signal is a periodic repetition of the spectrum of the original unsampled signal. We now need to represent the spectrum by a set of samples so that we can handle the data digitally. Effectively that means sampling the spectrum shown in Fig. 7.10a by a periodic sequence of impulses (a sampling comb) in the frequency domain. For this purpose consider an infinite periodic sequence of impulses in frequency spaced $Df$ or $Dω/2π$ apart as seen in Fig. 7.10b. Using (7.20) and (7.22), but going from the frequency to the time domain, it can be shown that the inverse transform of the frequency domain sampling sequence is a set of impulses in the time domain, spaced $T = 1/Df$ apart.

Multiplication of the spectrum of the sampled time function in Fig. 7.10a by the frequency domain sampling function in Fig. 7.10b produces the sampled spectrum of Fig. 7.10c. By the convolution theorem of (7.18a), that is equivalent to convolving the periodic sequence of impulses in the time domain shown in Fig. 7.10b by the original time function samples in Fig. 7.10a to produce the periodic repetition of the samples of the time function shown in Fig. 7.10c. The repetition period of the group of samples is $T$. It is convenient if the number of
samples used to represent the spectrum is the same as the number of samples taken of the time function. Let that number be \( K \), consistent with the development in Sect. 7.5. Since the time domain samples are spaced \( \Delta t \) apart, the duration of sampling is \( K \Delta t \). By the manner in which we have drawn Fig. 7.10, we have, for convenience, synchronised the total sampling period of the time function \( T \) with the repetition period of the inverse Fourier transform of the frequency domain sampling comb in Fig. 7.10b. As a result, \( \Delta f = 1/T = 1/K\Delta t \), the inverse of the sampling duration. Similarly, the total unambiguous bandwidth in the frequency domain is \( K \Delta f = 1/\Delta t \) which covers just one segment of the spectrum.

7.8 The Discrete Fourier Transform of an Image

7.8.1 The Transformation Equations

The previous sections have treated functions with a single independent variable. That could have been time, or position along a line of image data. We now turn our attention to functions with two independent variables to allow Fourier transforms of images to be computed and understood. Despite this implicit increase in complexity we will find that we can take full advantage of the material of the previous sections to help in our understanding. Let
\[
\phi(k, l) \quad k, l = 0 \ldots K - 1 \tag{7.24}
\]

be the brightness of a pixel at location \(k, l\) in an image of \(K \times K\) pixels. The set of image pixels is a digital sample of the scene recorded by the remote sensing imaging instrument. Therefore, the Fourier transform of the scene is given by the discrete Fourier transform of the set of pixel brightnesses. Building on the material of Sect. 7.5 it can be seen that the discrete Fourier transform of an image is given by

\[
\Phi(r, s) = \sum_{k=0}^{K-1} \sum_{l=0}^{K-1} \phi(k, l) W^{rk} W^{sl} \tag{7.25a}
\]

An image can be reconstructed from its transform according to

\[
\phi(k, l) = \frac{1}{K^2} \sum_{r=0}^{K-1} \sum_{s=0}^{K-1} \Phi(r, s) W^{-rk} W^{-sl} \tag{7.25b}
\]

### 7.8.2 Evaluating the Fourier Transform of an Image

We can rewrite (7.25a) as

\[
\Phi(r, s) = \sum_{k=0}^{K-1} W^{rk} \sum_{l=0}^{K-1} \phi(k, l) W^{sl} \tag{7.26}
\]

The right hand sum will be recognised as the one-dimensional Fourier transform of the \(k\)th row of pixels in the image, which we write as

\[
\Phi(k, s) = \sum_{l=0}^{K-1} \phi(k, l) W^{sl} \quad k = 0 \ldots K - 1 \tag{7.27}
\]

Thus, the first step is to Fourier transform each row of the image. We then replace the row by its transform. The transformed pixels are now addressed by the spatial frequency index \(s\) across the row rather than the positional index \(l\). Using (7.27) in (7.26) we have

\[
\Phi(r, s) = \sum_{k=0}^{K-1} \Phi(k, s) W^{rk} \quad s = 0 \ldots K - 1 \tag{7.28}
\]

which is the one-dimensional discrete Fourier transform of the \(s\)th column of the image, after the row transforms of (7.27) have been computed. Therefore, to compute the two dimensional Fourier transform of an image, it is only necessary to transform each row individually to generate an intermediate image, and then transform that result by column to yield the final transform. Both the row and column
transformations would be carried out using the fast Fourier transform algorithm, which requires $K^2 \log_2 K$ complex multiplications to transform the complete image.

### 7.8.3 The Concept of Spatial Frequency

Entries in the Fourier transformed image $U(r, s)$ represent the composition of the original image in terms of spatial frequency components, vertically and horizontally. Spatial frequency is the image analogue of the time frequency of signal. A sinusoidal function with a high frequency alternates rapidly, whereas a low-frequency function changes slowly with time. Similarly, an image with high spatial frequency in, say, the horizontal direction shows frequent changes of brightness with position horizontally. A picture of a crowd of people would be a particular example, whereas a head and shoulders view of a person reading the news on television is likely to be characterised mainly by a low spatial frequencies. Typically, an image is composed of a collection of both horizontal and vertical components with different spatial frequencies of differing strengths. They are what the discrete Fourier transform describes.

The upper left-hand pixel in $U(r, s)$—$U(0, 0)$—is the average brightness value of the image. In engineering this would sometimes be called the DC value. That is the component of the spectrum with zero frequency variation in both directions. Thereafter, pixels in $U(r, s)$ both horizontally and vertically represent components with frequencies that increment by $1/K$ where the original image is of size $K \times K$.

In most cases we would know the scale of the image, in other words the distance on the ground covered by the $K$ pixels across the lines and down the columns. That allows us to define the spatial frequency increment in terms of metres$^{-1}$. For example, if an image covered $5 \times 5$ km, then the spatial frequency increment in both directions is $2 \times 10^{-4}$ m$^{-1}$.

### 7.8.4 Displaying the DFT of an Image

In Fig. 7.9 we saw that the one dimensional discrete Fourier transformation of a function is periodic with period $K$. The same is true of the discrete Fourier transform of an image. The $K \times K$ pixels of $U(r, s)$ can be viewed as one period of an infinitely periodic two-dimensional array in the manner depicted in Fig. 7.11. We also saw that the amplitude spectrum of the one dimensional DFT is symmetric about $K/2$. Similarly $U(r, s)$ is symmetric about its centre. Therefore, no new amplitude information is shown by displaying transform pixels horizontally and vertically beyond $K/2$. Rather than ignore them (since their accompanying phase is important) the display is adjusted as shown in Fig. 7.11 to bring $U(0, 0)$ to the centre. In that manner the pixel at the centre of the Fourier transform array represents the image average brightness value. Pixels away from the centre represent the
proportions of image components with increasing spatial frequency. That is the usual manner for presenting two-dimensional image transforms. Examples of spectra displayed in this manner are given in Fig. 7.12. To make visible those components with smaller amplitudes the logarithmic scaling \( D(r,s) = \log |\Phi(r,s)| \) is sometimes used

\[
D(r,s) = \log |\Phi(r,s)|
\]  

(7.29)

### 7.9 Image Processing Using the Fourier Transform

Having a knowledge of the discrete Fourier transform of an image allows us to develop more general geometric processing operations than those treated in Chap. 5. In preparation for this, note that the high spatial frequency content of an image is associated with frequent changes of brightness with position. Edges, lines and some types of noise are examples of high spatial frequency data. In contrast, gradual changes of brightness with position account for the low frequency components of the spatial spectrum. Since ranges of spatial frequency are identifiable with regions in the spectrum we can understand how the spectrum of an image can be altered to produce different geometric enhancements of the image itself. For example, if regions near the centre of the spectrum are removed, leaving behind only the high frequencies, and the image is then reconstructed from the modified spectrum, a version containing only edges and line-like features will be produced. On the other hand, if the high frequency components are removed, leaving behind only the region near the centre of the spectrum, the reconstructed image will appear smoothed, since edges, lines and other high-frequency detail will have been removed.

Modification of the two-dimensional discrete image spectrum in the manner just described can be expressed as the product of the image spectrum \( \Phi(r,s) \) and a filter function \( H(r,s) \) to generate the new spectrum:

---

To implement simple sharpening or smoothing as described above $H(r,s)$ would be set to 0 for those frequency components to be removed and 1 of those frequency components to be retained. Equation (7.30) also allows more complicated filtering operations to be carried out. For example, a specific band of spatial frequencies could be excluded if they corresponded to some form of periodic noise, such as the line striping sometimes observed with line scanner data. Also, the entries in $H(r,s)$ can be different from 0 or 1, allowing more sophisticated changes to the spectrum of an image. Figure 7.13 shows the effect of applying ideal (sharp cut-off) filters to the image segment in Fig. 7.12. The filter cut off values are shown superimposed over the log amplitude spectrum of the image by circles.

The low pass filtered images are those generated by retaining only those frequency components inside the circles, whereas the high pass filtered versions are made up from those spectral components outside the filter circles. Even though the filters are shown for convenience over the amplitude spectra they are applied to the full complex Fourier transform of the original image. Modification of the spatial

$$Y(r,s) = H(r,s)\Phi(r,s)$$  \hspace{1cm} (7.30)
(geometric) features of an image in the frequency domain in this manner involves three steps. First, the image has to be Fourier transformed to produce a spectrum. Secondly, the spectrum is modified according to (7.30). Finally the image is reconstructed from the modified spectrum using the inverse discrete Fourier transform. Together, these three operations require $2K^2\log_2 K + K^2$ multiplications.

### 7.10 Convolution in two Dimensions

The convolution theorem for functions given in Sect. 7.6.3 has a two dimensional counterpart, again in two forms:

1. **Equation (7.31a)**
   \[
   y(k, l) = \Phi(k, l) \ast h(k, l) \quad \text{then} \quad Y(r, s) = \Phi(r, s)H(r, s) \quad (7.31a)
   \]

2. **Equation (7.31b)**
   \[
   Y(r, s) = \Phi(r, s) \ast H(r, s) \quad \text{then} \quad y(k, l) = \phi(k, l)h(k, l) \quad (7.31b)
   \]

Unlike (7.18b) there is no $2\pi$ scaling factor here since the spatial frequency variables $r, s$ are equivalent to frequency $f$ in hertz and not frequency $\omega$ in radians per second.

The operation in (7.31a) and (7.31b) is the discrete version of convolution shown in (5.13) and described in Sect. 5.8. Equation (7.31a) shows that any of the geometric enhancement operations that can be implemented by modifying the image spectrum can also be carried out by performing a convolution between the image itself and the inverse Fourier transform of the filter function $H(r, s)$. Conversely, operations such as simple mean value filtering described in Sect. 5.3.1 can be implemented in the spatial frequency domain; that needs the Fourier transform of the template. The template has to have the same dimensions as the

---

**Fig. 7.13** Examples of low pass and high pass spatial filtering based on the Fourier transform of the original image; in this case filters with sharp cut-offs have been used.
image but with values of 0 everywhere except for the set of pixels that are used to implement the prescribed operation.

### 7.11 Other Fourier Transforms

If (7.3b) is substituted in (7.11a) we have

\[ G(x) = \int_{-\infty}^{\infty} g(t)(\cos \omega t - j \sin \omega t) \, dt \]

which can be separated into

\[ G(x) = \int_{-\infty}^{\infty} g(t) \cos \omega t \, dt \] (7.32a)

and

\[ G(x) = \int_{-\infty}^{\infty} g(t) \sin \omega t \, dt \] (7.32b)

the first of which is called a Fourier cosine transform, and the second of which is called a Fourier sine transform. They are applied to even and odd functions respectively, in which case the integrals usually go from 0 to \( \infty \). There is a discrete version of the Fourier cosine transform, called the DCT or discrete cosine transform, which is given by discretising (7.32a) in the same manner we discretised (7.11a) to generate the DFT. The DCT finds widespread application in video compression for the television industry.\(^{10}\)

### 7.12 Leakage and Window Functions

In Sect. 7.7 we noted that a sampled function can be regarded as the unsampled version multiplied by an infinite periodic sequence of impulses. The spectrum of the set of samples produced is the spectrum of the original function convolved with the spectrum of the sequence of impulses; we saw that in Fig. 7.9.

In practice it is not possible to take an infinite number of samples of the function. Instead, sampling is commenced at a given time and terminated after some period \( \tau \), as illustrated in Fig. 7.14. This can be represented by a long pulse of unit amplitude and duration \( \tau \)—a sampling window—that multiplies the infinite sequence of samples. The spectrum of the finite set of samples is, as a consequence, modified by being

convolved by the spectrum of the sampling window, again as illustrated in Fig. 7.14. Since the sampling window is a rectangular pulse its Fourier transform is as shown in Fig. 7.6. Because the pulse is long compared with the sampling interval, the spectrum shown in Fig. 7.6 is compressed and looks like a finite amplitude impulse, thus approximating well the situation with an infinite sampling comb. However when finite in length, its side lobes cause problems during convolution with the spectrum of the sequence of samples, causing distortion, as depicted in Fig. 7.14.

To minimise that form of distortion, which is referred to as leakage, the rectangular sampling window is replaced by a function which avoids the sharp turn on and turn off with time that characterises the rectangular function. There are several candidates for these so-called window functions,\(^{11}\) perhaps the most common of which is the raised cosine or Hanning window:

\[
w(t) = 0.5 - 0.5 \cos \left( \frac{2\pi t}{\tau} \right)
\]

(7.33)

This has smaller side lobes in its spectrum than the simple rectangular pulse and, as a result, leads to less leakage distortion.

---

If the function being sampled is periodic, and the samples are taken over one or several full periods, leakage will not occur. Otherwise it is always a matter for consideration, and window functions generally need to be used.

7.13 The Wavelet Transform

7.13.1 Background

In principle, the Fourier transform can be used to represent any signal by a collection, sometimes infinite, of sinusoidal functions. Likewise, the two-dimensional spatial Fourier transform can be used to model the distribution of brightness values in an image by using a collection of two dimensional sinusoidal basis functions.

Many of the image features of interest to us occur over a short distance, including edges and lines. Also, when dealing with functions of time, we are sometimes interested in representing short time signals rather than those that last for a long period. As an illustration, an organ playing a single, pure tone generates a signal that is well-modelled by simple sinusoids. In contrast, when a single note is played on a piano we have an approximately sinusoidal signal, at the frequency of the key played, which lasts for just a short time. We can still find the Fourier transform representation of the piano note—its spectrum—but there are other ways to represent such a short time signals, just as there are other ways of representing or modelling image features that change over a short distance. The wavelet transformation is generally more useful in such situations than the Fourier transform. It is based on the definition of a wavelet, which is a wavelike signal that is limited in time (or space, in the spatial domain). The theory of the wavelet transformation is quite detailed, especially when treated comprehensively. Here we provide a simple introduction in which the mathematical detail is kept to a minimum and some concepts are simplified, so that its common usage in image processing can be understood. It finds most application in image compression and coding, and in the detection of localised features such as edges and lines.

7.13.2 Orthogonal Functions and Inner Products

The Fourier series and transform expansions of (7.10a) and (7.11b) are special cases of the more general representation of a function by a sum of other functions, expressible as

---

\[ g(t) = \sum_n a_n \psi_n(t) \quad (7.34) \]

The \( \psi_n(t) \) are called basis functions and the \( a_n \) are expansion coefficients. We saw how to find the expansion coefficients for complex exponential basis functions in (7.10b). To do so depends on a property of the basis functions called orthogonality, which means:

\[ \int \psi_m(t) \psi_n(t) dt = 1 \text{ for } m = n, \text{ and zero otherwise} \quad (7.35) \]

The range of the integral depends on the actual basis functions themselves. In (7.10b) the range extends over one period of the sinusoidal basis functions. If (7.35) holds, then from (7.34) we can see

\[ \int g(t) \psi_m dt = \int \sum_n a_n \psi_n(t) \psi_m dt = \sum_n a_n \int \psi_n(t) \psi_m dt = a_m \]

which gives us the procedure for calculating values for the expansion coefficients. That is seen in explicitly (7.10b), and in the Fourier transform formula of (7.11a).

It is fundamental to many functional representations of the form of (7.34) that an orthogonal basis set is chosen so that the expansion coefficients are easily established. In the general theory of the wavelet transform, in which we will seek to represent practical functions and images by basis functions that exist over only a limited interval, the same is essentially true.

Operations like that in (7.35) are called inner products and are written symbolically as

\[ \langle \psi_m(t), \psi_n(t) \rangle = \int \psi_m(t) \psi_n(t) dt \]

### 7.13.3 Wavelets as Basis Functions

What sorts of wavelet basis functions should we use in practice? Whatever functions we choose they have to be able to model events that occur at different positions in time, or space when we look at images, and to accommodate events that, while being localised, can occur over different ranges of time or position. To achieve that the wavelet basis functions generally have to have two indices (one for location and one for spread) so that a function can be represented, or expanded, as
Figure 7.15 shows such a set of functions. We see that a fundamental function is translated and scaled so that it can cover instances at different times and over different durations.

Rather than define translations and scalings arbitrarily we restrict attention to binary scalings, in which the range is shrunk by progressive factors of two, and to dyadic translations, in which the shift amount is an integral multiple of the binary scaling factor. That means that the set of wavelet functions we are dealing with are built up from a basis function \( \psi(t) \), sometimes called the mother wavelet or generating wavelet, such that all other wavelets are defined by

\[
g(t) = \sum_{j,k} c_{j,k} \psi_{j,k}(t)
\]

in which \( j \) is the scaling factor and \( k \) is the integral multiple of the scaling by which the shift occurs. The factor \( 2^{j/2} \) is included so that the integral of the squared amplitude of the wavelet is unity, one of the requirements of a wavelet basis function. The relationship in (7.37) is applied in Fig. 7.15, although the amplitude scaling is omitted for clarity. Note that, apart from being integers, \( j \) and \( k \) are arbitrary at this stage. Putting (7.37) in (7.36) we have
\[ g(t) = \sum_{j,k} c_{j,k} 2^{j/2} \psi(2^j t - k) \]  

(7.38)

The set of coefficients \( c_{j,k} \) is sometimes called the wavelet transform of \( g(t) \), or its wavelet spectrum.

### 7.13.4 Dyadic Wavelets with Compact Support

It is possible to restrict further the set of wavelets of interest by establishing a relationship between \( j \) and \( k \). If we constrain our attention to so-called compact functions \( g(t) \) that are zero outside the interval \([0,1]\) we can use a single index \( n \) to describe the set of basis functions, where

\[ n = 2^j + k \]

The basis functions, while still having the general form in (7.37), can then be indexed by the single integer \( n \), in which case the wavelet expansion, or representation, of the time restricted signal is

\[ g(t) = \sum_{n=0}^{\infty} c_n \psi_n(t) \]  

(7.39)

The expansion coefficients are given by

\[ c_n = \int_{-\infty}^{\infty} g(t) \psi_n(t) \, dt \]  

(7.40)

We do not pursue this version explicitly any further here.

### 7.13.5 Choosing the Wavelets

Not all finite time (or space) functions can be used as wavelets; it is only those that satisfy the so-called admissibility criterion that can be employed in wavelet expansions.\(^\text{13}\) Fortunately, for most of the work of interest to remote sensing image processing a simpler approach is possible, based on the concept of filter banks, which avoids the need specifically to treat a range of candidate wavelet families. Although originally developed separately, the filter bank and wavelet approaches are related. Rather than continue with the theoretical development of continuous wavelets as such, we will now focus on filter banks. As the name suggests a filter

The bank is made up of a set of filters that respond to different frequency ranges in a signal or image. Each of the filters is a finite impulse response (FIR) digital filter. A background in that material, while helpful, is not required for the following development.

### 7.13.6 Filter Banks

#### 7.13.6.1 Sub Band Filtering, and Downsampling

Suppose we want to represent the signal $g(t)$ by a wavelet model. The first step in the filter bank approach is to separate the signal into its low and high frequency components by passing it through two filters as illustrated in Fig. 7.16; one extracts the low frequencies from the signal and the other the high frequencies. In this case the filters are chosen so that they divide the spectrum of the signal into two halves, as indicated, which makes the following development simple. In practice other arrangements are also used.

We can represent the filtering operation diagrammatically as in Fig. 7.17, which shows the filters as blocks. The output of each filter block is given by the convolution between the input signal and the function of time—the filter function—that describes the operation of the filter. The filter function is also called its impulse response. That name comes about because if an impulse is fed into the filter then the output (the response) is the filter function, as seen from Sect. 7.6.2.

---

14 See also Sect. 5.8 and Fig. 5.13.
If we use the convolution theorem of (7.18a) we can represent the output of the filter in the frequency domain as the product of the Fourier transform of the input signal and the Fourier transform of the filter impulse response, called the filter’s transfer function.

We now assume that the input signal has been sampled and is represented by the sequence $c_k(t)$:

When dealing with sampled functions we must also represent the impulse response of the filter by its sampled counterpart. Call this $h_0(k)$ for the low pass filter and $h_1(k)$ for the high pass filter, as seen in Fig. 7.17.

The results of sending the sampled input signal through the two filters are

$$y_0(k) = \sum_{n} \gamma(n) h_0(k - n)$$  \hspace{1cm} (7.41a)

and

$$y_1(k) = \sum_{n} \gamma(n) h_1(k - n)$$  \hspace{1cm} (7.41b)

For simplicity we have left out limits on the summations, but they are understood to cover the full range of samples. Assume that we have sampled $g(t)$ at or just above the Nyquist limit of (7.23), which says that the sampling rate should exceed twice the bandwidth of the signal if there is to be no (aliasing) distortion, and so that the signal can be fully recovered from the samples. Such an assumption of near Nyquist rate sampling helps simplify the following development; it can be shown that the results hold in general just so long as sampling is above the Nyquist rate.

We now make an interesting observation about the sequence of samples in (7.41a); $y_0(k)$ is sampled at the same rate as $g(t)$ and yet, because it is the output of
a low pass filter with a cut-off frequency that is half of the original signal bandwidth, the Nyquist limit is halved. In other words we have twice the sampling rate necessary to avoid aliasing distortion in $y_0(k)$. We could therefore halve the rate and still avoid distortion; with that reduced rate $y_0(k)$ can still be reconstructed if necessary. We halve the sampling rate by dropping out every second sample in $y_0(k)$, leaving a sequence that is half the size of the original. Although not as obvious, we can also halve the sampling rate of $y_1(k)$ by dropping every second sample.\(^{15}\) The process of dropping one sample in two, in each of $y_0(k)$ and $y_1(k)$, is called downsampling.

The next step is to low and high pass filter $y_0(k)$. In this case the cut-off frequency of the filter $h_0(k)$ is set to $B/4$ because the bandwidth of $y_0(k)$ is $B/2$. Similarly the band pass of $h_1(k)$ is from $B/4$ to $B/2$. Thus both filters now have a bandwidth that is half that used in the first stage; that affects the impulse response as we show below in Fig. 7.19.

Again, the result is downsampled by dropping every second sample in the resultant signals. We continue in that fashion, by filtering the low frequency signals at each step, using filters with bandwidths half of their previous values, and downsampling, until the point is reached where the resulting low frequency signal is a single number. That process is shown in Fig. 7.18 for a three stage filter bank. It has decomposed the sequence $\gamma(k)$ into the set \{y_1(k), y_{01}(k), y_{001}(k), y_{000}(k)\} as a result of the chosen $h_0(k)$ and $h_1(k)$. The latter play the part of basis functions while the \{y_1(k), y_{01}(k), y_{001}(k), y_{000}(k)\} act as expansion coefficients. That is the wavelet spectrum of $\gamma(k)$ on the $h_0(k)$, $h_1(k)$ basis. At first it seems odd having two apparently different basis functions. Indeed there are not, as we will see soon in Sect. 7.13.6.3.

\[^{15}\] See Castleman, \textit{loc. cit.}, Sect. 14.4.3.2. As an alternative explanation for those with a signal processing background, the ability to drop every second sample in $y_1(k)$ can be appreciated if we recognise that because the spectrum below $B/2$ is unfilled, the high pass signal can be frequency downshifted (translated) to the origin without distortion, whereupon its highest frequency component is then at $B/2$ and not $B$. The downshifted version can then also be represented without distortion at the downsampled rate.
The high pass filters $h_1(k)$ constitute the wavelets, with the first—that with a bandwidth of $B/2$—being the mother wavelet. As we move from left to right in Fig. 7.18 the impulse responses broaden (dilate) as expected of a set of wavelet basis functions, even though that is opposite to the scaling from coarse to fine in the family shown in Fig. 7.15.

The impulse responses of the low pass filters, $h_0(k)$, are called the scaling vectors and their outputs are the scaling functions.\(^\text{16}\) To see the changes in impulse response all we have to do is find the time equivalent versions of the low pass and high pass filters, using (7.11b). For a low pass filter with cut off frequency $B/N$, with $N = 2^j, j = 1, 2, \ldots$ it can be shown that the amplitude of the continuous time impulse response is given by

$$g(t) = \frac{B}{N\pi} |\text{sinc}(Bt/N)|$$

which is shown plotted in Fig. 7.19. It can be demonstrated that the magnitude of the impulse response for the equivalent high pass filters is the same. Time dilation by factors of two is evident in the figure as the bandwidth is decreased by those same factors.

\(^{16}\) They are also referred to as the approximations to the original signal; they are a successive set of reduced resolutions of the original and are sometimes said to form an image pyramid. They can be employed in multiresolution analysis.
7.13.6.2 Reconstruction from the Wavelets, and Upsampling

In principle we can reconstruct the filtered signal after a single stage by choosing two new filters in the manner seen in Fig. 7.20, shown in terms of transfer functions rather than impulse responses since that leads quickly to an important result. The signals are also expressed in their frequency domain (Fourier transformed) versions. By using the frequency domain representation we can avoid convolution.

As in Fig. 7.17 the outputs of the left hand filters are

\[ Y_0(f) = H_0(f)G(f) \quad \text{and} \quad Y_1(f) = H_1(f)G(f) \]

The output from the right hand summing device is

\[ \Psi(f) = R_0(f)Y_0(f) + R_1(f)Y_1(f) \]

which, when substituting for \( Y_0(f) \) and \( Y_1(f) \), gives

\[ \begin{align*}
\Psi(f) &= R_0(f)H_0(f)G(f) + R_1(f)H_1(f)G(f) \\
&= [R_0(f)H_0(f) + R_1(f)H_1(f)]G(f)
\end{align*} \]

Now if

\[ R_0(f)H_0(f) + R_1(f)H_1(f) = 1 \] (7.42a)

then

\[ \Psi(f) = G(f) \]

Equation (7.42a) shows a general relationship that must hold for perfect reconstruction. However, since the filtered samples were downsampled, they must be upsampled before the reconstruction filters can generate the correct result.\(^{17}\) That is done by inserting zeros between each successive sample in \( y_0(k) \) and \( y_1(k) \). An outcome of doing that is that (7.42a) generalises to

\[ R_0(f)H_0(f) + R_1(f)H_1(f) = 2 \] (7.42b)

The reconstruction segment of Fig. 7.20 can be cascaded as was the analysis block in Fig. 7.18 to create a synthesis or reconstruction filter bank.

---

It would be desirable if the same filter bank could be used in reverse order to reconstruct the original signal from the transformed version. That turns out to be possible when the impulse response samples are symmetric, which is the simple case we treat in this section. In general the samples have to be reversed when reconstructing.\(^{18}\) The special case is shown in Fig. 7.21. As noted above, since, in Fig. 7.18, the sequences have been downsampled after each stage of filtering we have to upsample during reconstruction to restore each successive set of sequences to the right number of samples.

**7.13.6.3 Relationship Between the Low and High Pass Filters**

We now consider the interdependence of \(h_0(k)\) and \(h_1(k)\). Clearly, they are intimately related because one is the complement of the other in the frequency domain, as observed in Fig. 7.16. What does that mean for the samples themselves? If we view the high pass filter in the frequency domain as a frequency shifted version of the low pass filter we can employ the frequency shift property of the Fourier transform to show the \(h_0(k), h_1(k)\) relationship. To use that theorem we note first that the spectrum of the high pass filter is essentially the same as that of the low pass filter, but shifted along the frequency axis by an amount \(B\). That is expressed as

\[
H_1(f) = H_0(f - B)
\]

From the frequency shift theorem\(^{19}\)

\[
h_1(t) = e^{j2\pi Bt}h_0(t)
\]

We now have to express this in discrete form. The filter bandwidth \(B\) in a digital filter is set by the time between samples. A higher sampling rate leads to a greater bandwidth, and vice versa. If we write the continuous time variable as an integral

\[18\] See P.S. Addison, *The Illustrated Wavelet Handbook*, IOP Publishing, Bristol, 2002, Fig. 3.18, and Strang and Nguyen, *loc. cit.*

\[19\] See Brigham, *loc. cit.*
multiple of the sampling interval \( t = k\Delta t \) then \( B = 1/2\Delta t \) so that, in discrete form, the last expression becomes

\[ h_1(k) = e^{jk\pi}h_0(k) \]

The complex exponential with an exponent that is an integral multiple of \( \pi \) is \( \pm 1 \). That tells us that the impulse response of the high pass filter is the same as that of the low pass filter, except that every odd numbered sample has its sign reversed. That applies for the case of the ideal low and high pass filters. More generally, it also requires a reversal of the order of the samples.\(^{20}\)

### 7.13.7 Choice of Wavelets

In the filter bank development of Sect. 7.13.6 we have presented a decomposition and synthesis methodology that emulates wavelet analysis. It is based on the specification of the impulse responses of the low and high pass filters, which we now need to be more specific about because they describe the wavelets that are the basis functions for a given situation.

Although (7.38) is acceptable as a general expression of a wavelet expansion, it is more convenient, when comparing it to the filter bank approach, if we re-express the decomposition as the combination of an expansion in wavelets \( \psi(t) \) plus a companion scaling function \( \varphi(t) \) in the form

\[ g(t) = A\varphi(t) + \sum_{j,k} c_{j,k} 2^{j/2}\psi(2^jt - k) \]  

(7.43)

The scaling function satisfies the scaling or dilation equation

\[ \varphi(t) = \sum_k h_0(k)\varphi(2t - k) \]  

(7.44)

while the mother wavelet satisfies the wavelet equation

\[ \psi(t) = \sum_k h_1(k)\varphi(2t - k) \]  

(7.45)

As an illustration of these suppose the scaling function is the constant between 0 and 1 shown in Fig. 7.22 and the filter impulse response is a simple two sample average for the low pass filter \( [h_0(0) = 1, h_0(1) = 1] \) and a simple two sample difference \( [h_1(0) = 1, h_1(1) = -1] \) for the high pass filter. Then (7.44) and (7.45) become, respectively

\[ \varphi(t) = \varphi(2t) + \varphi(2t - 1) \]  

(7.46a)

\(^{20}\) See Strang and Nguyen, loc. cit.
Fig. 7.22 The Haar scaling function and wavelets generated with the scaling, dilation and recurrence relations—this is the first subset of 8 Haar wavelets; also shown at the top is the scaling function on a half time scale; note that the $2^{-1/2}$ amplitude scaling in (7.37) has not been included.

![Diagram of Haar wavelets and scaling function](image)

\[
\psi(t) = \varphi(2t) - \varphi(2t - 1) \tag{7.46b}
\]

which are also shown plotted in Fig. 7.22. The recurrence relationship in (7.37) then allows others in the set to be generated. These are the Haar wavelets. When used as the basis for a filter bank implementation of the wavelet transform the expressions in (7.46a) and (7.46b) need to be scaled by $1/\sqrt{2}$ to give the correct square integral for the wavelets.

Haar wavelets are generated by the simple sum and difference filters above. A variety of other types is available, many of which are examples of the Daubechies wavelets\(^{21}\) that can also be implemented readily as filter banks. The family of Daubechies wavelets includes the Haar wavelet as its simplest case.

\(^{21}\) See Strang and Nguyen, *loc. cit.*, and Addison, *loc. cit.*, Fig. 3.15.
7.14 The Wavelet Transform of an Image

The application of the discrete wavelet transformation to imagery is similar to the manner in which the discrete Fourier transform is applied. First, the rows of the image are transformed using the first stage process in Fig. 7.18. Every second column is then discarded, which corresponds to downsampling the row transformed data. Next the columns are transformed and every second row discarded. As illustrated in Fig. 7.23 that leads to four new images:

- a version that has been low pass filtered in both row and column, and which is referred to as the approximation image;
- a version that has been high pass filtered in the horizontal direction, thereby emphasising the horizontal edge information;
- a version that has been high pass filtered in the vertical direction, thereby emphasising the vertical edge information;
- a version that has been high pass filtered in both the vertical and horizontal directions, thereby emphasising diagonal edge information.

Because of the downsampling operations by column and row, the dimensions of those transformed images are half those of the original, giving us the same number of pixels overall. This is demonstrated with the small 2 × 2 image embedded in Fig. 7.23, which does not require downsampling, using the operations in (7.46a) and (7.46b). The single stage of Fig. 7.18 can be extended by transforming the low resolution image to produce yet a new low frequency approximation and high pass detail images, such as seen in Fig. 7.24 for two stages. As with the one dimensional wavelet transform, image reconstruction can be carried out by using the filter bank in reverse if the impulse response is symmetric.

7.15 Applications of the Wavelet Transform in Remote Sensing Image Analysis

Since the original image, such as the swan in the example of Fig. 7.24, can be completely reconstructed from any stage of the discrete wavelet transform using a reconstructing filter bank, it is clear that any level of decomposition contains all the original information content. Specifically, the low pass approximation image and the three high pass detail images in that figure, among them hold all the information. Interestingly, the detail images are much less complex than the approximation and can be compressed using a number of standard techniques without significant information loss. That allows the image data to be stored or transmitted using far fewer bits of digital data than the original. Compression

22 MATLAB® contains a large number of wavelet bases that can be used in signal and image processing.
based on the discrete wavelet transform is one of its principal applications in image processing.\textsuperscript{23}

The wavelet transform can also be applied to decompose a single pixel spectrum into a more compact form as a feature reduction tool prior to classification.  

7.16 Bibliography on Spatial Domain Image Transforms

Many texts in image processing for engineering and science cover much of the material presented in this chapter. The difficulty, however, is that the level of mathematical detail is sometimes quite high. Books that are comprehensive but not overly complex are


Introductory treatments can also be found in books on remote sensing image processing including


Material on the Fourier transform and the fast Fourier transform is given in


which remains one of the best treatment is available.


provides a detailed treatment of the Fourier transform and convolution for those with a higher level of mathematics expertise, as does the standard text in the field


Wavelets are covered in books that range from detailed mathematical treatments through to those which focus on applications. Castleman (above) has a good but difficult to read section. An idiosyncratic, but very comprehensive and readable account, is given in


---

A good overview of the application of wavelets in a wide range of physical situations including medicine, finance, fluid flow, geophysics and mechanical engineering will be found in


while a very good discussion on wavelet applications in astronomical image processing is given in


This book also includes a helpful section on the use of the Fourier transform for image filtering. It shows the use of wavelets for multiresolution image analysis and how, by filtering specific wavelet components, features at particular levels of detail can be enhanced. Other texts that could be consulted are


the last of which has a focus on SAR imagery and medical imaging.

### 7.17 Problems

7.1. Using (7.35) demonstrate that the complex exponentials $e^{im\omega t}$, where $m$ is an integer, are an orthogonal set.

7.2. Verify the results of Sect. 7.6.2 using a simple sketch.

7.3. Using the Fourier transform of an impulse and the convolution theorem, verify the result of Sect. 7.6.2 mathematically.

7.4. Using (7.15a) compute the discrete Fourier transform of a square wave using 2, 4 and 8 samples per period respectively.

7.5. Compute the discrete Fourier transform of a unit pulse of width $2a$. Use 2, 4 and 8 samples over a time interval equal to $8a$. Compare the results to those obtained in problem 7.4.

7.6. Image smoothing can be undertaken by computing averages over a square or rectangular window or by filtering in the spatial frequency domain. Consider just a single line of image data. Determine the corresponding spatial frequency domain filter function for a simple three pixel averaging filter to be used on that line of data. That requires the calculation of the discrete Fourier transform of a unit pulse.

7.7. As in problem 7.6 consider a single line of image data. One way of applying a low pass filter to that data is to choose an ideal filter function in the spatial frequency domain that has a sharp cut-off, such as that shown in Fig. 7.16. Determine the corresponding function in the image domain by calculating
the inverse Fourier transform of the ideal filter. Taking into account the
discrete pixel nature of the image, approximate the inverse transform by an
appropriate one dimensional template.

7.8. Are window functions required if a periodic signal is sampled over an
integral number of periods?

7.9. In Fig. 7.9 suppose the function \( g(t) \) is a sinusoid of frequency \( B \) Hz. Its
spectrum will consist of two impulses, one at \( +B \) Hz and the other at \(-B\)Hz.
Produce the spectrum of the sinusoid obtained by taking three samples every
two periods. Suppose the sinusoid is to be reconstructed from the samples by
feeding them through a low pass filter that will pass all frequency compo-
nents up to \( 1/2\Delta t \), where \( \Delta t \) is the sampling interval, and will exclude all
other frequencies. Describe the nature of the reconstructed signal; this will
give an appreciation of aliasing distortion.

7.10. The periodic sequence of impulses in Fig. 7.9 is an idealised sampling
function. In practice it is not possible to generate infinitesimally short
samples of a function; rather, the samples will have a finite, although short,
duration. That can be modelled mathematically by replacing the infinite
sequence of impulses by a periodic sequence of finite-width pulses. One way
of representing that periodic sequence of pulses is as the convolution of a
single pulse and a periodic sequence of impulses. Using that model, describe
what modifications are required in Fig. 7.9 to account for samples of finite
duration.

7.11. Explain the appearance of the spectrum shown in Fig. 7.12 for the sequence
of uniformly spaced vertical lines.

7.12. By examining Fig. 7.24 how many pixels are there in the collected images
after the first and second stages of the discrete wavelet transformation? What
would be the case without downsampling?

7.13. The histogram of an image before wavelet transformation generally will
have many filled bins. The same would be expected of the approximation
images at each stage of the transformation. Qualitatively, what might the
general shape of the histograms of the detail images look like in, say, Fig. 7.24? The simple example illustrated in Fig. 7.23 may help in
answering this question.

7.14. The time function \( g(t) \), with Fourier transform \( G(f) \), is acceptable as a basic
wavelet if it satisfies the admissibility criterion

\[
\int_{-\infty}^{\infty} \frac{|G(f)|^2}{|f|} df < \infty
\]

Show that the wavelets associated with the high pass filter in Fig. 7.16
satisfy the criterion.

7.15. Find the Fourier spectrum of a short, time-limited pure sinusoid, such as
might occur when hitting a single note on a piano. In time that can be
modelled by a sinusoidal signal that commences at a given time and stops a
short time later. You may find this easier to handle using the convolution theorem, rather than calculating the Fourier transform from scratch.

7.16. The square waveform of Fig. 7.4 can be generated by convolving a unit pulse, such as that shown in Fig. 7.6a, with a periodic sequence of impulses. Using that model, and the convolution theorem, verify the spectrum of Fig. 7.5 from the spectrum of Fig. 7.6b.