

IMAGE FORMATION

Two-dimensional (2-D) images can be acquired in several ways. Using optical instrumentation such as a camera or a microscope, a 2-D image of a three-dimensional (3-D) object can be obtained through optical lens geometry. Through medical imaging modalities, 2-D and 3-D images of an organ can be obtained using transmission, emission, reflectance, diffraction, or nuclear resonance methods as outlined in the previous chapter. Most recent advances in medical imaging allow the acquisition of even 4-D time-varying image sequences of 3-D organs such as a beating heart. For example, magnetic resonance imaging (MRI) can be used to obtain time-based image sequences of a beating heart. The sophisticated medical imaging modalities such as 3-D computed tomography (CT) and MRI require multidimensional reconstruction methods to create images. Nevertheless, the basic principles of image formation are utilized throughout the image reconstruction, processing, and display processes. This chapter provides some of the basic principles of image formation that are universally applicable, from general purpose optical camera-based 2-D imaging to more sophisticated multidimensional medical imaging systems.

An analog image is described by the spatial distribution of brightness or gray levels that reflect a distribution of detected energy. The image can be displayed using a medium such as paper or film. A photograph on paper may show a black and white image with gray levels representing grades of brightness from black to white. While black and white images require only one gray level or intensity variable, the color images are produced by using multiple variables. For examples, three basic colors, red, green, and blue (RGB), could be used as three variables for representing color images. When combined together, the RGB intensities can produce a selected color at a spatial location in the image. Thus, a true color image can be represented by three RGB component images with the corresponding gray-level or intensity variables (1–3). Other color image representation schemes can be derived from the basic RGB color representation (1–3). For example, a commonly used color representation scheme is intensity I , saturation S , and hue H , which can be expressed as (3)

$$\begin{aligned}
I &= \frac{R+G+B}{3} \\
S &= 1 - \frac{3}{(R+G+B)}(\min(R, G, B)) \\
H &= \begin{cases} \theta & \text{if } B \leq G \\ 2\pi - \theta & \text{if } B > G \end{cases} \\
\text{with } \theta &= \cos^{-1} \left\{ \frac{\frac{1}{2}[(R-G) + (R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\}
\end{aligned} \tag{2.1}$$

where angle θ is measured with respect to the red axis in the *ISH* coordinate space and R , G , and B are, respectively, the red, green, and blue values in the RGB coordinate space.

The advantage of the *ISH* color representation scheme is that it provides the overall intensity component, which is the average of the RGB values. The RGB color images of skin lesions have been converted into an *ISH* coordinate system for multichannel intensity and texture-based segmentation (4).

It should be noted that a black and white image can be converted into a pseudo-color image by transforming the original gray levels of the black and white image into the specific components of a color representation scheme (1–3). One simple method of pseudo-color transformation is intensity slicing in which the entire range of gray levels of the black and white image is divided into three subranges. The three subranges are then assigned to the RGB input of the color monitor for pseudo-color visualization. Such a pseudo-color representation often helps visualization of structures in medical images (3, 5, 6).

For a digital image representation suitable for processing on the computer, the image has to be discrete in both spatial and intensity (gray level) domains. A discrete spatial location of finite size with a discrete gray-level value is called a pixel. For example, an image of 1024×1024 pixels may be displayed in 8-bit gray-level resolution. This means that each pixel in the image may have any value from 0 to 255 (i.e., a total of 256 gray levels). The pixel dimensions would depend on the spatial sampling. An analog image may be converted into a digital image using image scanners or digitizers. On the other hand, digital images may be acquired directly using an array of discrete detectors or sensors as found in digital cameras or CT medical imaging systems. The pixel dimensions may be carefully correlated with the actual physical dimensions of objects in the real world. For example, if a real-world object is sampled with 1 mm spatial sampling in each direction for image formation, a 2-D image with 1024×1024 pixel resolution may represent the actual spatial dimension of $10.24 \text{ cm} \times 10.24 \text{ cm}$.

2.1. IMAGE COORDINATE SYSTEM

In the process of image formation, the object coordinates are mapped into image coordinates. If the image formation system is linear, a simple transformation with

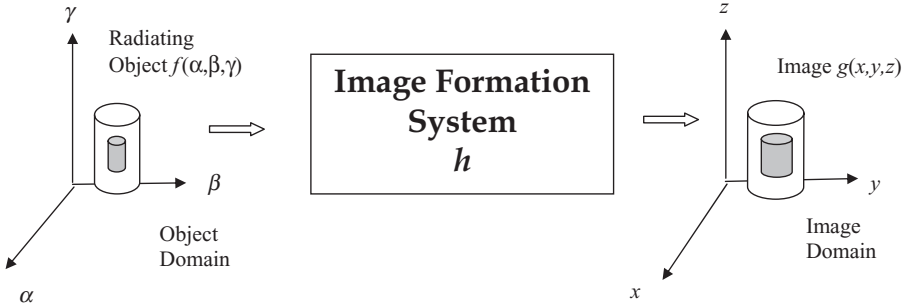


Figure 2.1 The object and image domain coordinate systems.

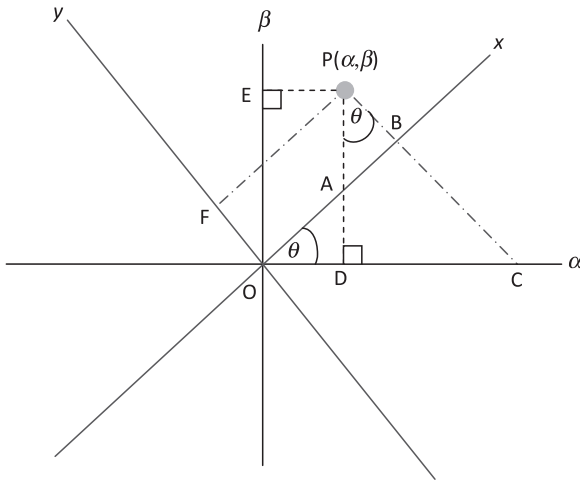


Figure 2.2 A pixel P is projected on the original (α, β) and (x, y) coordinate system rotated by θ .

proper translation, scaling, and rotation can define the correspondence between the coordinate systems of the object and image domains. Such transformation operations may be used to restore real-world features and to establish correspondence among images for registration and interpretation. Figure 2.1 shows the object and image coordinate systems such that the object coordinate system (α, β, γ) is mapped onto the image coordinate system (x, y, z) as output from an image formation system.

2.1.1 2-D Image Rotation

Let us consider rotating a 2-D image $f(\alpha, \beta)$ by an angle θ . Let us assume that (α, β) coordinate system is rotated by angle θ to (x, y) coordinates. Figure 2.2 shows a pixel P in image $f(\alpha, \beta)$ projected on the (x, y) coordinate system.

The pixel P can be projected on (α, β) and (x, y) coordinate systems as

$$\begin{aligned} \alpha &= OD; & \beta &= PD \\ x &= OB; & y &= PB. \end{aligned} \quad (2.2)$$

Considering ΔPDC and ΔOBC , the following relationships can be observed:

$$\begin{aligned} x &= OB = OC \cos \theta \\ x &= (OD + DC) \cos \theta \\ x &= (\alpha + \beta \tan \theta) \cos \theta \\ x &= \alpha \cos \theta + \beta \sin \theta. \end{aligned} \quad (2.3)$$

and

$$\begin{aligned} y &= PA \cos \theta \\ y &= (PD - AD) \cos \theta \\ y &= (\beta - \alpha \tan \theta) \cos \theta \\ y &= \beta \cos \theta - \alpha \sin \theta. \end{aligned} \quad (2.4)$$

Thus, a 2-D rotation transformation for output image $g(x, y)$ obtained by rotating $f(\alpha, \beta)$ is given by

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}. \quad (2.5)$$

2.1.2 3-D Image Rotation and Translation Transformation

If an object $f(\alpha, \beta, \gamma)$ is mapped into an image $g(x, y, z)$ through a linear image formation system h , a transformation involving translation and rotation may be established to define object and image coordinate systems as

$$\mathbf{G} = \mathbf{R}(\mathbf{F} - \mathbf{T}) \quad (2.6)$$

where \mathbf{G} and \mathbf{F} are, respectively, image and object domain coordinate systems denoted as column vectors and \mathbf{R} and \mathbf{T} are, respectively, rotation and translation matrices.

Translation is simply a vector subtraction operation. Scaling is a vector multiplication operation. For rotation in 3-D, three rotations about three axes can be defined in a sequence to define the complete rotation transformation. It can be shown that a rotation transformation can be defined by the following three operations:

1. Rotation of $\mathbf{G}_1(\alpha, \beta, \gamma)$ about β by an angle θ such that $\mathbf{G}_1(\zeta, \tau, \sigma) = \mathbf{R}_\theta \mathbf{G}_1(\alpha, \beta, \gamma)$

$$\text{where } R_\theta = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \quad (2.7)$$

2. Rotation of $\mathbf{G}_1(\zeta, \tau, \sigma)$ about α by an angle ϕ such that $\mathbf{G}_2(\omega, \varepsilon, \nu) = \mathbf{R}_\phi \mathbf{G}_1(\zeta, \tau, \sigma)$

$$\text{where } R_\phi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{bmatrix} \quad (2.8)$$

3. Rotation of $\mathbf{G}_2(\omega, \varepsilon, \nu)$ about γ by an angle ψ such that $\mathbf{F}(x, y, z) = \mathbf{R}_\psi \mathbf{G}_2(\omega, \varepsilon, \nu)$

$$\text{where } R_\psi = \begin{bmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.9)$$

It should be noted that the sequence of rotational operations is important because these operations are not commutative.

2.2. LINEAR SYSTEMS

It is usually desirable for an image formation system to behave like a linear, spatially invariant system. In other words, the response of the imaging system should be consistent, scalable, and independent of the spatial position of the object being imaged. A system is said to be linear if it follows two properties: scaling and superposition. In mathematical representation, it can be expressed as

$$h\{aI_1(x, y, z) + bI_2(x, y, z)\} = ah\{I_1(x, y, z)\} + bh\{I_2(x, y, z)\} \quad (2.10)$$

where a and b are scalar multiplication factors, and $I_1(x, y, z)$ and $I_2(x, y, z)$ are two inputs to the system represented by the response function h .

It should be noted that in real-world situations, it is difficult to find a perfectly linear image formation system. For example, the response of photographic film or X-ray detectors cannot be linear over the entire operating range. Nevertheless, under constrained conditions and limited exposures, the response can be practically linear. Also, a nonlinear system can be modeled with piecewise linear properties under specific operating considerations.

2.3. POINT SOURCE AND IMPULSE FUNCTIONS

The image formation process requires an energy source that is used on the object through an imaging operation, for example, reflection or transmission. As the response of the imaging operation is recorded on a film or array of detectors, an image can be modeled in terms of responses of spatially distributed point sources. As shown later in this chapter, if image formation is considered to be linear, the image can be described in terms of integration of responses of spatially distributed point sources. Thus, an image can be characterized through responses of point sources, also known as point spread functions (PSF).

A point source can be modeled through an impulse or Dirac delta function. To understand the properties of an impulse function, let us first define a 2-D rectangular function with a unit magnitude over a unit area (Fig. 2.3) as

$$\text{rect}(x, y) = \begin{cases} 1 & \text{for } |x| \leq \frac{1}{2}, |y| \leq \frac{1}{2} \\ 0 & \text{otherwise} \end{cases} \quad (2.11)$$

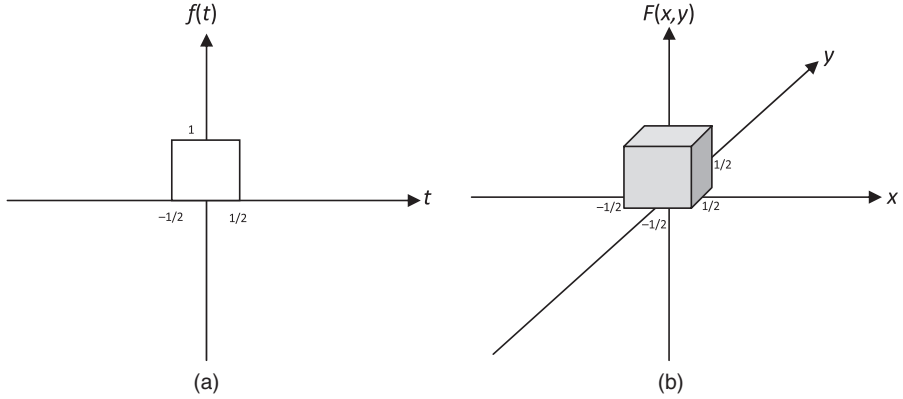


Figure 2.3 (a) One-dimensional unit rectangular function; (b) two-dimensional unit rectangular function.

A family of delta functions can be derived from rectangular functions as

$$\begin{aligned} \partial_n(x, y) &= n^2 \text{rect}(nx, ny) \\ &= \begin{cases} n^2 & \text{for } |x| \leq \frac{1}{2n}, |y| \leq \frac{1}{2n} \\ 0 & \text{otherwise} \end{cases} \quad \text{for } n = 1, 2. \end{aligned} \quad (2.12)$$

It can be seen from Equation 2.12 that as n increases, the base of the impulse reduces and magnitude increases. For $n \rightarrow \infty$, the delta function is zero everywhere except at $(0, 0)$, where its magnitude becomes infinite. This is also called Dirac delta or impulse function. However, the basic constraint remains the same as

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \partial(x, y) dx dy = 1. \quad (2.13)$$

The Dirac delta or impulse function can be used in modeling point sources as it has good sifting properties:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \partial(x, y) dx dy = f(0, 0)$$

and

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \partial(x - \alpha, y - \beta) dx dy = f(\alpha, \beta). \quad (2.14)$$

The above property can be used to sample a multidimensional function. For example, a 2-D array of impulse functions, called comb function, as shown in Figure 2.4, can be used to sample an image, and can be expressed as

$$\text{comb}(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \partial(x - m, y - n). \quad (2.15)$$

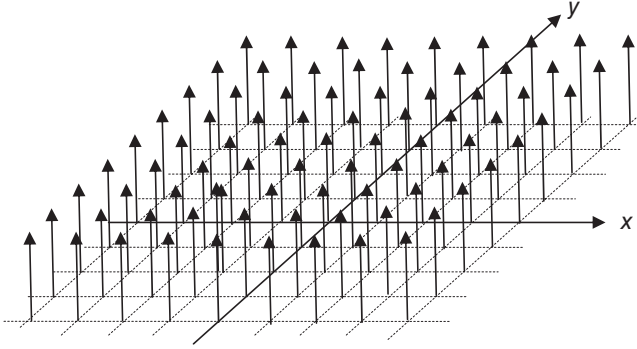


Figure 2.4 A schematic diagram of two-dimensional comb (x, y) function. The vertical arrows represent approaching infinite magnitude in the limiting case as $n \rightarrow \infty$ (see text for details).

2.4. PROBABILITY AND RANDOM VARIABLE FUNCTIONS

In any statistical data or experiment, a set of outcomes can be defined as $\Omega = \{\omega_i\}$; $i = 1, \dots, N$ with probability p_i to each outcome ω_i such that

$$p_i \geq 0 \quad \text{and} \quad \sum_{i=1}^N p_i = 1. \quad (2.16)$$

Let us define a random variable \mathbf{f} as a function of the outcome as $\mathbf{f}(\omega_i)$ that is assigned an arbitrary but fixed number z such that for any z , the set $\{\mathbf{f} \leq z\}$ is an event. The probability of the event $\{\mathbf{f} \leq z\}$ is called the distribution function of the random variable \mathbf{f} and may be denoted as $P_f(z)$. Properties of the distribution function $P_f(z)$ includes (13, 18)

- a. $P_f(-\infty) = 0$; $P_f(\infty) = 1$
- b. $P_f(z_1) \leq P_f(z_2)$ if $z_1 \leq z_2$
- c. Let us denote the probability of random variable \mathbf{f} within the range $z_1 < \mathbf{f} \leq z_2$ by $Pr\{z_1 < \mathbf{f} \leq z_2\}$ then

$$Pr\{z_1 < \mathbf{f} \leq z_2\} = P_f(z_2) - P_f(z_1)$$

- d. $p_f(z) = \frac{dP_f(z)}{dz}$ where $P_f(z)$ is called the probability density function of random variable \mathbf{f} .

It can be noted that

$$Pr\{z_1 < \mathbf{f} \leq z_2\} = \int_{z_1}^{z_2} p_f(z) dz. \quad (2.17)$$

The expected or mean value $E(\mathbf{f})$ of a random variable can be defined as

$$E(\mathbf{f}) = \int_{-\infty}^{\infty} zp_f(z) dz = \mu_f. \quad (2.18)$$

The variance σ_f^2 of \mathbf{f} can be defined as

$$\sigma_f^2 = E\{(\mathbf{f} - \mu_f)^2\} = \int_{-\infty}^{\infty} (z - \mu_f)^2 p_f(z) dz. \quad (2.19)$$

A commonly used probability density function in statistical estimation-based image reconstruction, processing, and analysis operations is normal or Gaussian density function given as

$$p_f(z) = \frac{1}{\sigma_f \sqrt{2\pi}} \exp \frac{-(z - \mu_f)^2}{2\sigma_f^2}. \quad (2.20)$$

2.4.1 Conditional and Joint Probability Density Functions

Conditional probability of an event A when B has occurred is defined by

$$Pr(A/B) = \frac{Pr(A \cap B)}{Pr(B)} = \frac{Pr(AB)}{Pr(B)}. \quad (2.21)$$

If the two events A and B are said to be independent, then

$$Pr(A/B) = \frac{Pr(AB)}{Pr(B)} = \frac{Pr(A)Pr(B)}{Pr(B)} = Pr(A). \quad (2.22)$$

If events A and B are not independent, then the Equation 2.21 can be expressed as

$$Pr(AB) = Pr(A/B)Pr(B). \quad (2.23)$$

Similarly, the conditional probability of event B given that event A has occurred can be given as

$$Pr(B/A) = \frac{Pr(BA)}{Pr(A)} = \frac{Pr(AB)}{Pr(A)}$$

or

$$Pr(AB) = Pr(B/A)Pr(A). \quad (2.24)$$

From Equations 2.23 and 2.24, it is implied that

$$Pr(A/B)Pr(B) = Pr(B/A)Pr(A). \quad (2.25)$$

Equation 2.25 yields to Bayes' theorem that the conditional probability of determining event A given B can be expressed in the form of a priori knowledge of class probabilities (18) as

$$Pr(A/B) = \frac{Pr(B/A)}{Pr(B)} Pr(A). \quad (2.26)$$

A generalized version of Bayes' rule can be used to improve a priori knowledge of A , and the new information gives out the posteriori probability of A given B . In the generalized version, with a set of events A_i ; $i = 1, \dots, n$ as

$$Pr(A_i / B) = \frac{Pr(B / A_i)Pr(A_i)}{Pr(B)} = \frac{Pr(B / A_i)Pr(A_i)}{\sum_{i=1}^n Pr(B / A_i)Pr(A_i)}. \quad (2.27)$$

The conditional distribution function $P_f(z/B)$ of a random variable \mathbf{f} while B has occurred, and its expected value, are defined as

$$P_f(z / B) = \frac{Prob(\{\mathbf{f} \leq z\} \cap B)}{Prob(B)} \quad (2.28)$$

$$E\{\mathbf{f} / B\} = \int_{-\infty}^{\infty} zp_f(z / B)dz. \quad (2.29)$$

The set $\Omega = \{\omega_i\}$; $i = 1, \dots, N$ of all events can be characterized by the individual probability distribution functions of respective random variables as

$$p_{fi}(z_i) = Prob\{\mathbf{f}_i \leq z_i\}; \quad i = 1, 2, \dots, N. \quad (2.30)$$

The joint probability density function of the random variables $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N$ is defined as

$$p_{f1, f2, \dots, fN}(z_1, z_2, \dots, z_N) = \frac{\partial^N P_{f1, f2, \dots, fN}(z_1, z_2, \dots, z_N)}{\partial z_1 \partial z_2 \dots \partial z_N} \quad (2.31)$$

where $P_{f1, f2, \dots, fN}(z_1, z_2, \dots, z_N) = Prob\{\mathbf{f}_1 \leq z_1, \mathbf{f}_2 \leq z_2, \dots, \mathbf{f}_N \leq z_N\}$.

2.4.2 Independent and Orthogonal Random Variables

The random variables $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N$ are independent if their joint density function can be expressed as the product of individual density functions as

$$p_{f1, f2, \dots, fN}(z_1, z_2, \dots, z_N) = p_{f1}(z_1)p_{f2}(z_2) \dots p_{fN}(z_N). \quad (2.32)$$

It follows that if random variables $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N$ are uncorrelated, their expected values can be explicitly expressed as

$$E\{\mathbf{f}_i \mathbf{f}_j\} = E\{\mathbf{f}_i\}E\{\mathbf{f}_j\} \quad \text{for all } i, j \text{ and } i \neq j. \quad (2.33)$$

The above expression becomes 0 when the random variables $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N$ are orthogonal with

$$E\{\mathbf{f}_i \mathbf{f}_j\} = 0 \quad \text{for all } i, j \text{ and } i \neq j. \quad (2.34)$$

Orthogonality is an important property that allows one to design appropriate basis functions for series expansion and several transformations such as wavelet transform for image processing and compression tasks.

In statistical estimation-based image reconstruction and restoration tasks, functions of the random variable are often used. As functions of random variables are also random variables, their joint probability distributions are related to Jacobian transformation. Let us assume random variables $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_N$ are functions of random variables $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N$ defined as (13, 18)

$$\begin{aligned}
\mathbf{g}_1 &= \Psi_1(\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N) \\
\mathbf{g}_2 &= \Psi_2(\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N) \\
&\dots \\
\mathbf{g}_N &= \Psi_N(\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N).
\end{aligned} \tag{2.35}$$

Following the above representation, the joint density function of the random variables $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_N$ can be expressed as

$$P_{g_1, g_2, \dots, g_N}(x_1, x_2, \dots, x_N) = \frac{\partial^N P_{g_1, g_2, \dots, g_N}(x_1, x_2, \dots, x_N)}{\partial z_1 \partial z_2 \dots \partial z_N} \tag{2.36}$$

where $P_{g_1, g_2, \dots, g_N}(x_1, x_2, \dots, x_N) = \text{Prob} \{ \mathbf{g}_1 \leq x_1, \mathbf{g}_2 \leq x_2, \dots, \mathbf{g}_N \leq x_N \}$.

The probability distribution functions for random variables $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_N$ and $\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N$ are thus related by

$$P_{g_1, g_2, \dots, g_N}(x_1, x_2, \dots, x_N) = \frac{P_{f_1, f_2, \dots, f_N}(z_1, z_2, \dots, z_N)}{|J(z_1, z_2, \dots, z_N)|} \tag{2.37}$$

where Jacobian transformation $J(z_1, z_2, \dots, z_N) =$

$$\begin{vmatrix}
\frac{\partial \Psi_1}{\partial z_1} & \frac{\partial \Psi_1}{\partial z_2} & \dots & \frac{\partial \Psi_1}{\partial z_N} \\
\frac{\partial \Psi_2}{\partial z_1} & \frac{\partial \Psi_2}{\partial z_2} & \dots & \frac{\partial \Psi_2}{\partial z_N} \\
\dots & \dots & \dots & \dots \\
\frac{\partial \Psi_N}{\partial z_1} & \frac{\partial \Psi_N}{\partial z_2} & \dots & \frac{\partial \Psi_N}{\partial z_N}
\end{vmatrix}.$$

2.5. IMAGE FORMATION

In general, image formation is a neighborhood process (7). One can assume that a radiant energy, such as a light source to illuminate an object, is represented by the function $f(\alpha, \beta, \gamma)$. For example, in case of photographic imaging, the object function $f(\alpha, \beta, \gamma)$ may be a distribution of the light reflected by the object. An image represented by the function $g(x, y, z)$ is obtained through an image formation system that transports the radiant energy from the object to the image plane (7–13). If the image plane is 2-D, such as in photographic imaging using a film camera, the projection of the 3-D object is acquired on a 2-D plane. If the image formation system is 3-D, a 3-D image is acquired and then can be manipulated to obtain 2-D views. For example, in X-ray transmission-based CT, a 3-D image can be reconstructed and used to compute any cross-sectional view as shown in Figure 2.5.

The image formation process can be similarly defined for other forms of imaging such as emission imaging, as shown in Figure 2.6. The only difference is that the object itself radiates the energy in emission imaging and therefore does not require any external radiation energy source. Although the imaging formation process can be described in a similar way, the radiant energy distribution has to be modeled differently in emission imaging.

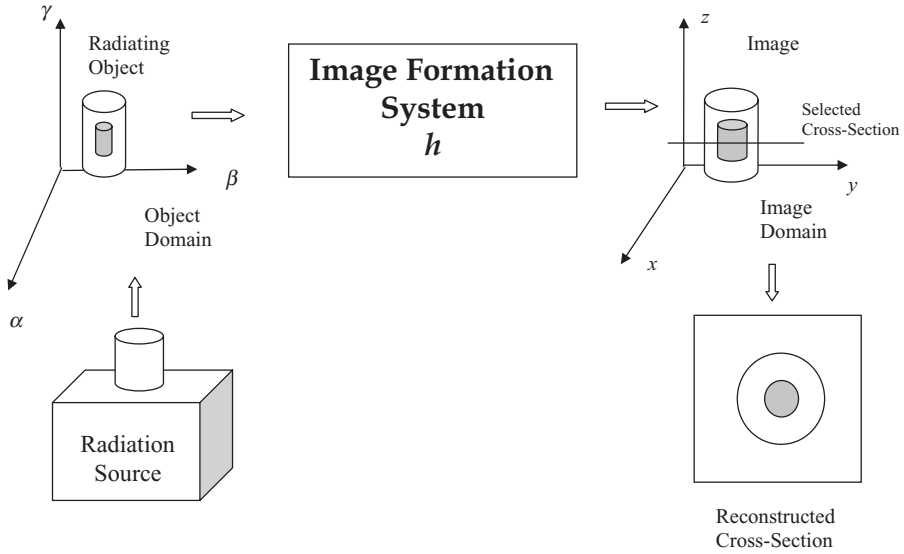


Figure 2.5 A schematic block diagram of a general 3-D image formation system requiring an external radiation source (such as light) for photographic imaging, or X ray for transmission imaging.

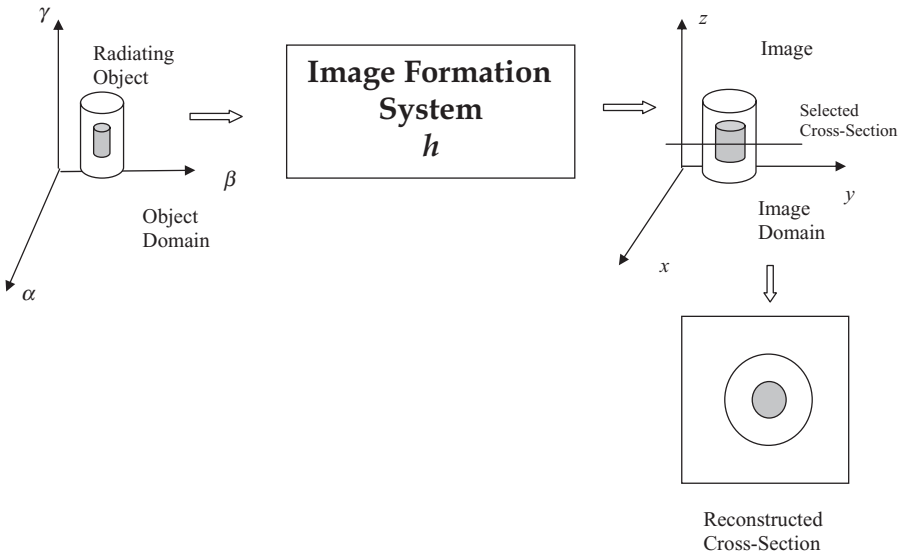


Figure 2.6 A schematic block diagram of an emission-based image formation system commonly used for fluorescence or nuclear medicine imaging modalities.

Whether the imaging process is 2- or 3-D, the image is formed through integration of a function of radiation energy distribution over the entire spatial extent from both object and image coordinates. This function is known as the response function of the image formation system. The parameters describing the response

function could be very different and can be modeled differently across imaging systems. For example, X-ray and PET imaging modalities are characterized by different response functions involving different parameters. Diagnostic X rays attenuate while passing through the body because of the absorption properties of the medium within the body. On the other hand, PET imaging is influenced by absorption as well as scattering coefficients of the object being imaged.

Let us represent a generalized response function of an image formation system as h . A general image formation process follows the principles of non-negativity and superposition. Thus,

$$\begin{aligned} f(\alpha, \beta, \gamma) &\geq 0 \\ g(x, y, z) &\geq 0. \end{aligned} \quad (2.38)$$

Considering two spatial points $f_1(\alpha, \beta, \gamma)$ and $f_2(\alpha, \beta, \gamma)$ in the object domain and $g_1(x, y, z)$ and $g_2(x, y, z)$ in the image domain, the superposition principle follows as

$$g_1(x, y, z) + g_2(x, y, z) = h(x, y, z, \alpha, \beta, \gamma, f_1(\alpha, \beta, \gamma)) + h(x, y, z, \alpha, \beta, \gamma, f_2(\alpha, \beta, \gamma)). \quad (2.39)$$

where $h(x, y, z, \alpha, \beta, \gamma, f_i(\alpha, \beta, \gamma))$ for $i = 1, 2$ is the general response function that depends on the object and image coordinates and the spatial point $f_i(\alpha, \beta, \gamma)$ of the object distribution function.

If the image formation system is considered to be linear, the response function does not depend on the object distribution and hence

$$h(x, y, z, \alpha, \beta, \gamma, f(\alpha, \beta, \gamma)) = h(x, y, z, \alpha, \beta, \gamma) f(\alpha, \beta, \gamma). \quad (2.40)$$

Using the additive property of radiating energy distribution to form an image, one can write

$$g(x, y, z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x, y, z, \alpha, \beta, \gamma, f(\alpha, \beta, \gamma)) d\alpha d\beta d\gamma. \quad (2.41)$$

If the image formation system is assumed to be linear, the image expression becomes

$$g(x, y, z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x, y, z, \alpha, \beta, \gamma) f(\alpha, \beta, \gamma) d\alpha d\beta d\gamma. \quad (2.42)$$

As described above, the response function $h(x, y, z, \alpha, \beta, \gamma)$ is called the point spread function of the image formation system. The PSF depends on the spatial extent of the object and image coordinate systems. The expression $h(x, y, z, \alpha, \beta, \gamma)$ is the generalized version of the PSF described for the linear image formation system that can be further characterized as a spatially invariant (SI) or spatially variant (SV) system. If a linear image formation system is such that the PSF is uniform across the entire spatial extent of the object and image coordinates, the system is called a linear spatially invariant (LSI) system. In such a case, the image formation can be expressed as

$$g(x, y, z) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x - \alpha, y - \beta, z - \gamma) f(\alpha, \beta, \gamma) d\alpha d\beta d\gamma. \quad (2.43)$$

In other words, for an LSI image formation system, the image is represented as the convolution of the object is radiant energy distribution and the PSF of the image

formation system. It should be noted that the PSF is basically a degrading function that causes a blur in the image and can be compared to the unit impulse response, a common term used in signal processing.

A 3-D LSI imaging system is very difficult to implement in real practice. Consider a photographic camera with a lens that has a fixed focal length. Such a camera will only focus on object points in a selected x - y plane in the object domain on a fixed x - y plane in the image domain. Depending on the depth of focus of the lens, all points out of the focused object plane (with different z -values in Fig. 2.1) will have different PSF and magnification to cause different amounts of blur in the 2-D image plane. Furthermore, the curvature of the lens will cause spatially variant PSF even within the focused object plane because of spherical aberrations. However, the spherical aberrations can be removed or minimized if high-quality lenses with appropriate aperture setting are used in the camera. Thus, a high-quality photographic camera can produce a sharp, spatially invariant image of the focused object plane. In this case, the above convolution equation for a fixed z value can be expressed as

$$g(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x - \alpha, y - \beta) f(\alpha, \beta) d\alpha d\beta \quad (2.44)$$

or

$$g = h \otimes f.$$

where \otimes denotes a convolution operator.

As shown above, an image can be expressed as the convolution of the object distribution with the PSF for an LSI imaging system without any noise consideration. The noise in the detection instrumentation and data acquisition system can be modeled as additive and signal independent. This is a simplification but true for electronic instrumentation for linear imaging system. With additive noise assumption, the image formation process can be expressed as

$$g(x, y) = f(x, y) \otimes h(x, y) + n(x, y). \quad (2.45)$$

In most medical imaging systems, the instrumentation is designed in such a way so as to provide maximum linearity and spatially invariant properties of image formation. The advantage is rather obvious. Since it is easier to experimentally evaluate the PSF of an imaging system, the acquired image $g(x, y, z)$ can be updated or restored through image processing methods. General causes of blurry PSF include the finite size of the source, parallax and geometrical effects, radiation scatter, and nonlinear responses of the detector systems. The quality of medical images can be improved through image restoration and enhancement methods such as inverse filtering and Weiner deconvolution. These methods use the knowledge of the PSF of the imaging scanner in frequency filtering techniques that are described in Chapter 6.

2.5.1 PSF and Spatial Resolution

The PSF of an imaging system can be obtained experimentally by imaging a point source (defined above) and measuring the blur or spread of the point source. Usually, the PSF as the image of a point source (Fig. 2.7) is well approximated by a Gaussian

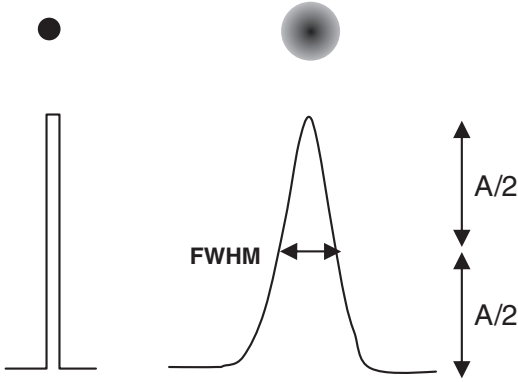


Figure 2.7 A point source (top left) and its point spread function (PSF) for an imaging system (top right); their one-dimensional intensity profiles are respectively shown in the second row.

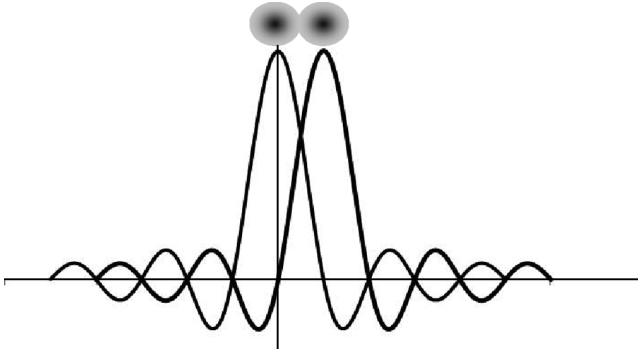


Figure 2.8 Two imaging point sources resolved through Rayleigh criterion applied to the respective sinc-based PSFs.

function as shown by the one-dimensional (1-D) intensity profile at the bottom right of Figure 2.7. If the PSF is arbitrarily approximated as a Gaussian function, the resolution of the image is described through full width at half maximum (FWHM) parameter of the Gaussian distribution. For example, in 1-D, the Gaussian distribution for PSF with mean μ and standard deviation σ can be represented as

$$h(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (2.46)$$

The FWHM in the above case is expressed as

$$\text{FWHM} = 2\sqrt{2\ln 2}\sigma \cong 2.36\sigma. \quad (2.47)$$

With an FWHM-based resolution criterion, two point sources are said to be resolved if they are separated by a distance greater than the FWHM. In case the PSF is described as a sinc function (see Fig. 2.8), a Rayleigh criterion can be used that



Figure 2.9 An X-ray imaging phantom with holes of different sizes.

requires the peak intensity of one PSF to at least coincide or be farther away from the first zero-crossing point of the second PSF in order to resolve two point sources.

For testing the spatial resolution performance of a medical imaging system, a physical phantom is used with a number of tubes of different diameters embedded in a cylindrical enclosure. For example, Figure 2.9 shows a spatial resolution phantom made out of aluminum with holes of different diameters. Such a phantom with holes acting as point sources can be used for X-ray imaging systems. There are varieties of phantoms that are used to meet specific needs of testing spatial resolution performance of a given imaging modality. Some phantoms are made out of an enclosure with circular tubes and hollow objects of different shapes and diameters. Based on imaging modality such as SPECT, PET, and MRI, the phantom is filled with a specific material (e.g., radioactive material for SPECT imaging system) for testing spatial resolution capabilities as well as contrast performance.

In addition to PSF, other measures are also often used in characterizing the spatial resolution performance of the imaging system. Such measures include line spread function (LSF) and modulation transfer function (MTF). The LSF is often used in X-ray imaging that is evaluated experimentally using a phantom with a grid of thin lead bars, and also defined as

$$LSF(y) = \int_{-\infty}^{\infty} PSF(x, y) dx. \quad (2.48)$$

An MTF can be defined in the frequency domain that is characterized by the frequency response of the PSF as

$$MTF(\omega_x, \omega_y, \omega_z) = \iiint PSF(x, y, z) e^{-j2\pi\omega_x x} e^{-j2\pi\omega_y y} e^{-j2\pi\omega_z z} dx dy dz \quad (2.49)$$

where ω_x , ω_y , ω_z are respectively spatial frequencies along x -, y -, and z -directions in the frequency domain (see Section 2.7 for Fourier transform).

2.5.2 Signal-to-Noise Ratio

The signal-to-noise ratio (SNR) is defined as the power ratio of original signal to noise. While the original signal can be regarded as useful information that is represented by the mean or expected value of the signal, noise can be considered variations in the background represented as standard deviation. Let us represent power

of signal and noise as P_{signal} and P_{noise} respectively with A_{signal} and A_{noise} as amplitudes. SNR is then expressed as

$$\text{SNR} = \frac{P_{\text{signal}}}{P_{\text{noise}}} = \left(\frac{A_{\text{signal}}}{A_{\text{noise}}} \right)^2$$

$$\text{SNR(dB)} = 10 \log \frac{P_{\text{signal}}}{P_{\text{noise}}} = 20 \log \left(\frac{A_{\text{signal}}}{A_{\text{noise}}} \right). \quad (2.50)$$

For an image about which we have no other knowledge about the noise in its formation, the SNR can be represented by the ratio of mean and standard deviation or root-mean-square of all gray-level values of all pixels as

$$\text{SNR} = \frac{\bar{g}}{\sqrt{\frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N [g(x, y) - \bar{g}]^2}} \quad (2.51)$$

where $\bar{g} = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N g(x, y)$.

With additive noise assumption as used in Equation 2.45, image formation equation for an LSI imaging system can be expressed in frequency domain as

$$G(u, v) = F(u, v)H(u, v) + N(u, v) \quad (2.52)$$

where $G(u, v)$, $F(u, v)$ and $H(u, v)$ are, respectively, Fourier transform of acquired image, original object distribution, and PSF with u and v as spatial frequencies and $N(u, v)$ is the noise spectrum.

Considering that the acquired image represents the signal, and the noise spectrum can be estimated, the SNR can be computed through the ratio of signal and noise power spectra as

$$\text{SNR} = \frac{\sum_{u=0}^{P-1} \sum_{v=0}^{Q-1} |G(u, v)|^2}{\sum_{u=0}^{P-1} \sum_{v=0}^{Q-1} |N(u, v)|^2}. \quad (2.53)$$

Alternatively, with the knowledge of $f(x, y)$ (as is the case when a phantom is used for imaging), noise can be considered as the difference between the acquired image $g(x, y)$ and $f(x, y)$. The SNR then can be computed as

$$\text{SNR} = \frac{\sum_{x=1}^M \sum_{y=1}^N [g(x, y)]^2}{\sum_{x=1}^M \sum_{y=1}^N [f(x, y) - g(x, y)]^2}$$

$$= \frac{\sum_{x=1}^M \sum_{y=1}^N [g(x, y)]^2}{\sum_{x=1}^M \sum_{y=1}^N [n(x, y)]^2}. \quad (2.54)$$

2.5.3 Contrast-to-Noise Ratio

Medical images are often characterized by contrast-to-noise ratio (CNR) as a measure of image quality to discriminate an object of interest from its surrounding. In the context of medical imaging, CNR provides a measure of the capability to visualize physiological structures, lesions, or abnormalities in the image.

Contrast is a localized feature in the image that is described by the difference between the gray-level value of an object and its background. While there are several ways to define contrast feature, it is usually normalized in the 0–1 range as

$$C(x, y) = \frac{|f_{obj} - f_{back}|}{f_{obj} + f_{back}}$$

or

$$C(x, y) = \frac{|f_{obj} - f_{back}|}{\max\{f_{obj}, f_{back}\}}$$

or

$$C(x, y) = \frac{|f_{obj} - f_{back}|}{f_{ref}}. \quad (2.55)$$

where f_{obj} and f_{back} are, respectively, the average gray-level values or signal intensities in the object and background, and f_{ref} is a predetermined gray-level value or signal intensity in the image.

As described in the SNR section above, noise can be defined in different ways to determine CNR. Considering the image formation as expressed in Equation 2.52, the CNR can be expressed as

$$CNR = \frac{|f_{obj} - f_{back}|}{\sigma(x, y)}$$

or

$$CNR = \frac{\sum_{\text{for all objects}} |f_{obj} - f_{back}|^2}{[\sigma(x, y)]^2}. \quad (2.56)$$

where $\sigma(x, y)$ is the standard deviation of the noise in the image $f(x, y)$.

2.6. PIN-HOLE IMAGING

A pin-hole imaging method was originally developed for general-purpose photography. It is the simplest form of a camera for taking pictures on film or an array of sensors placed in the image plane. The pin-hole imaging method is used in many biomedical imaging systems including the nuclear medicine imaging modalities: SPECT and PET. The radiation from the object plane enters into the image plane through a pin hole. The plane containing the pin hole is called the focal plane. Figure 2.10 shows a schematic diagram of a pin-hole imaging system.

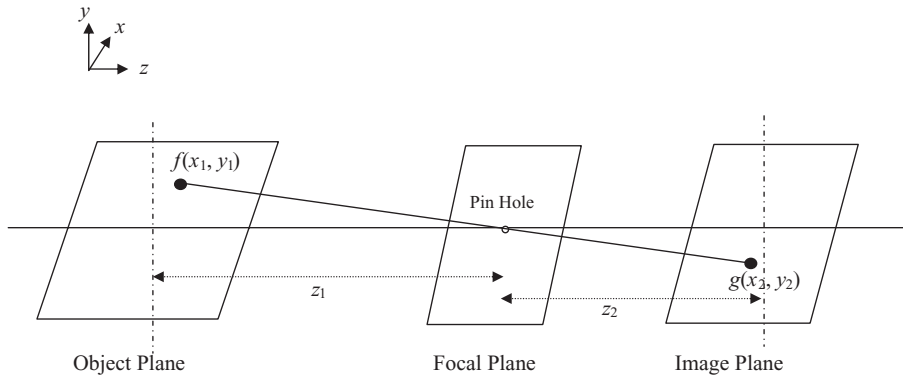


Figure 2.10 A schematic diagram of a pin-hole imaging camera.

If a point in the object plane is considered to have $(x_1, y_1, -z_1)$ coordinates mapped into the image plane as (x_2, y_2, z_2) coordinates, then

$$x_2 = -\frac{z_2 x_1}{z_1} \quad \text{and} \quad y_2 = -\frac{z_2 y_1}{z_1}. \quad (2.57)$$

It can be noted that such a pin-hole imaging system is inherently spatially variant providing a magnification factor, $M = -\frac{z_2}{z_1}$. However, if the collection efficiency of

the radiant energy in the image plane is considered to be a constant throughout the image plane independent of the location of the source of the radiant energy point in the object plane, the imaging can be considered an LSI process. Generalizing the object plane with the 2-D coordinate system (α, β) and the corresponding image plane with the coordinate system (x, y) , the general response function can include the magnification factor M so that the image formation equation can be expressed as

$$h(x, y; \alpha, \beta) = h(x - M\alpha, y - M\beta)$$

and

$$g(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x - M\alpha, y - M\beta) f(\alpha, \beta) d\alpha d\beta. \quad (2.58)$$

2.7. FOURIER TRANSFORM

The Fourier transform plays a very significant role in medical imaging and image analysis. The Fourier transform is a linear transform that provides information about the frequency spectrum of the signal, and is used in image processing for image enhancement, restoration, filtering, and feature extraction to help image interpretation and characterization. It is also used in image reconstruction methods for medical imaging systems. For example, the Fourier transform is used in image reconstruction in MRI. The Fourier transform can be applied to a signal to obtain frequency

information if the function representing the signal is absolutely integrable over the entire domain and has only a finite number of discontinuities and a finite number of maxima and minima in any finite region. In addition, the function must not have any infinite discontinuities.

A 2-D Fourier transform, FT, of an image $g(x, y)$ is defined as*

$$G(u, v) = FT\{g(x, y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) e^{-j2\pi(ux+vy)} dx dy \quad (2.59)$$

where u and v are the spatial frequencies in the x - and y -dimensions.

Once an image is transformed into the frequency domain, the degradation related to the noise and undesired frequencies can be filtered out. The filtered frequency information can then be used to recover the restored image through an inverse Fourier transform. Since the Fourier transform is a linear transform, the inverse transform can be used to obtain the original image from spatial frequency information if no filtering is performed in the frequency domain. A 2-D inverse Fourier transform is defined as

$$g(x, y) = FT^{-1}\{G(u, v)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(u, v) e^{j2\pi(ux+vy)} du dv. \quad (2.60)$$

The Fourier transform provides a number of useful properties for signal and image processing applications. Some of the important properties are briefly described below.

1. Linearity: Fourier transform, FT, is a linear transform.

$$FT\{ag(x, y) + bh(x, y)\} = aFT\{g(x, y)\} + bFT\{h(x, y)\} \quad (2.61)$$

2. Scaling: It provides proportional scaling.

$$FT\{g(ax, by)\} = \frac{1}{ab} G\left(\frac{u}{a}, \frac{v}{b}\right) \quad (2.62)$$

3. Translation: Translation of a function in the spatial domain introduces a linear phase shift in the frequency domain.

$$FT\{g(x-a, y-b)\} = G(u, v) e^{-j2\pi(ua+vb)} \quad (2.63)$$

4. Convolution: The convolution of two functions in spatial domain is represented by multiplication of their respective spectra in the frequency domain.

$$FT\left\{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\alpha, \beta) h(x-\alpha, y-\beta) d\alpha d\beta\right\} = G(u, v) H(u, v). \quad (2.64)$$

5. Cross-correlation: The cross-correlation operation of two functions in the spatial domain is represented by the multiplication of the conjugate of one spectrum with another spectrum in the frequency domain.

*Using general notation, lowercase functions such as $g(x, y)$ are represented in the spatial domain with the spatial coordinate system (x, y) while upper case functions such as $G(u, v)$ are represented in the Fourier or frequency domain with the frequency coordinate system (u, v) . Also, a and b have been used to represent constants for scaling, translation, and shifting operations.

$$FT \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g^*(\alpha, \beta) h(x + \alpha, y + \beta) d\alpha d\beta \right\} = G^*(u, v) H(u, v). \quad (2.65)$$

6. Auto-correlation: The Fourier transform of the auto-correlation of a function (such as a stochastic random signal) in the spatial domain is equal to the square of the absolute value of its Fourier transform, which is called the power spectrum.

$$FT \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g^*(\alpha, \beta) g(x + \alpha, y + \beta) d\alpha d\beta \right\} = G^*(u, v) G(u, v) = |G(u, v)|^2 \quad (2.66)$$

7. Parseval's theorem: The total energy is conserved in both spatial and frequency domains such that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) g^*(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(u, v) G^*(u, v) du dv. \quad (2.67)$$

8. Separability: If $g(x, y)$ is separable in x - and y -dimensions, the Fourier transform of $g(x, y)$ will also be separable.

$$g(x, y) = g_x(x) g_y(y).$$

then

$$FT\{g(x, y)\} = FT_x\{g_x(x)\} FT_y\{g_y(y)\}. \quad (2.68)$$

Figure 2.11a shows a vertical stripe image of 128×128 pixels generated from a sinusoid signal with a time period of 8 pixels. The image of the Fourier transform of the vertical stripe image is shown in Figure 2.11b. Two impulses corresponding to the frequency of the sinusoidal stripe signal can be seen in the Fourier transform image. Also, the vertical orientation of the stripes in the spatial domain and the horizontal orientation of the impulses in the Fourier domain should be noted. The effect of rotation in the spatial and frequency domains can be seen in Figure 2.12.

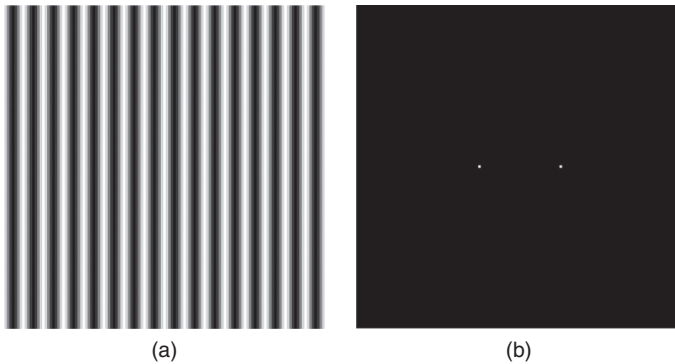


Figure 2.11 (a) A vertical stripe image generated from a sinusoidal waveform of a period of 8 pixels and (b) the magnitude image of its Fourier transform.

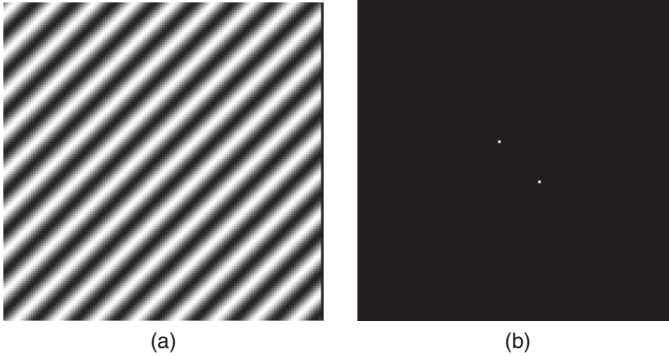


Figure 2.12 (a) A rotated stripe image and (b) magnitude image of its Fourier transform.

2.7.1 Sinc Function

The Fourier transform of a rectangular function is a sinc function. This has a specific significance in image processing because many image processing operations involve rectangular functions directly or indirectly. For example, a sampling period may be characterized by a rectangular window function, signal averaging is done over a rectangular window, and finite bandwidth is a specific case of the spectrum limited by a rectangular window function. Furthermore, a pixel in a digital image has a square shape, a specific instance of a rectangular function. For simplicity, let us consider a rectangular function in one dimension as

$$f(x) = \begin{cases} A & \text{for } |x| \leq \frac{w}{2} \\ 0 & \text{otherwise} \end{cases} \quad (2.69)$$

Let $F(\omega)$ be the Fourier transform of $f(x)$ with ω as spatial frequency:

$$\begin{aligned} F(\omega) &= \int_{-\infty}^{\infty} f(x) e^{-j2\pi\omega x} dx = \int_{-w/2}^{w/2} A e^{-j2\pi\omega x} dx \\ &= \frac{-A}{j2\pi\omega} \left[e^{-j2\pi\omega x} \right]_{-w/2}^{w/2} \\ &= \frac{A}{j2\pi\omega} \left[e^{j\pi\omega w} - e^{-j\pi\omega w} \right] \\ &= A w \frac{\sin(\pi\omega w)}{(\pi\omega w)} \\ &= A w \operatorname{sinc}(z) \end{aligned} \quad (2.70)$$

where $\operatorname{sinc}(z) = \frac{\sin(\pi z)}{(\pi z)}$ and $z = \omega w$.

As shown in Figure 2.13, the $\operatorname{sinc}(z)$ function has high dc value with the oscillating lobes that go down in amplitude as frequency increases. The oscillatory

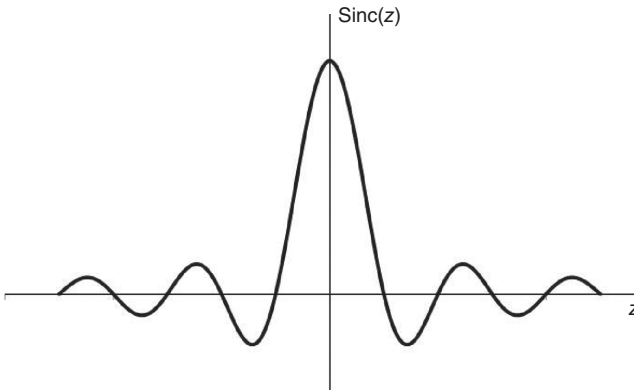


Figure 2.13 A sinc function: Fourier transform of a rectangular function.

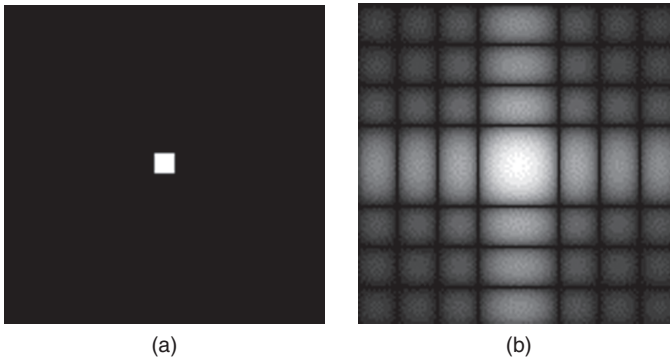


Figure 2.14 (a) An image with a square region at the center and (b) the logarithmic magnitude image of its Fourier transform.

nature of the sinc function creates dramatic effects in image reconstruction and processing. Figure 2.14 shows an image of a square region and its Fourier transform. The sinc-based oscillation pattern along both horizontal and vertical axes in the Fourier domain can be seen in Figure 2.14.

2.8. RADON TRANSFORM

The Radon transform defines projections of an object mapping the spatial domain of the object to its projection space. For example, the 2-D Radon transform of a 2-D object defines 1-D line integrals in the projection space, which is also called the Radon domain (14–17). In his paper published in 1917 (14), Radon provided a mathematical proof that an object can be reconstructed uniquely from its infinite continuous projections. In the early 1970s, Hounsfield and Cormack independently used the mathematical foundation established by Radon for reconstructing images of a human organ from its projections (15, 16). The projections were acquired through an X-ray CT imaging scanner (15).

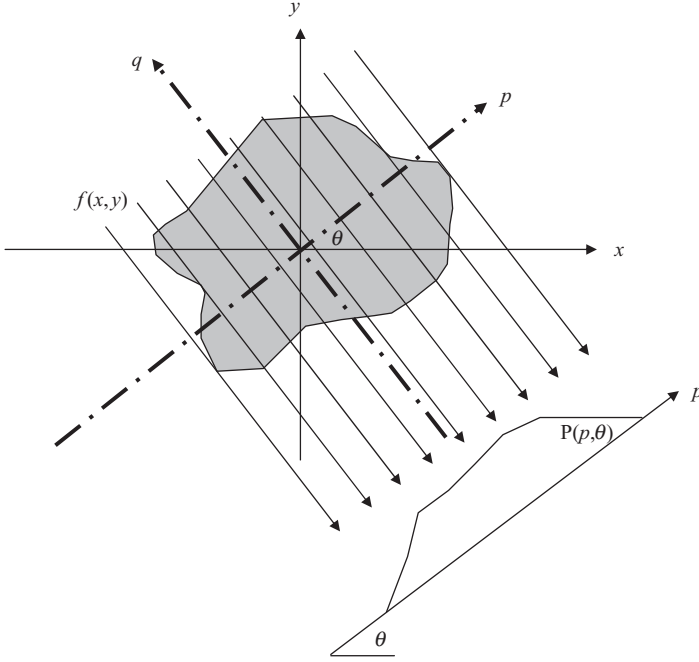


Figure 2.15 Line integral projection $P(p, \theta)$ of the two-dimensional Radon transform.

Let us define a 2-D object function $f(x, y)$ and its Radon transform by $\mathbf{R}\{f(x, y)\}$. Let us use the rectangular coordinate system (x, y) in the spatial domain. The Radon transform is defined by the projection $P(p, \theta)$ in the polar coordinate system as

$$P(p, \theta) = \mathbf{R}\{f(x, y)\} = \int_L f(x, y) dl$$

where the line integral \int_L is defined along the path L such that

$$x \cos \theta + y \sin \theta = p. \quad (2.71)$$

Figure 2.15 shows a line integral that is computed along the parallel arrow lines that are sampled along the p axis and are defined by the angle θ . A set of line integrals or projections can be obtained for different θ angles.

The polar coordinate system (p, θ) can be converted into rectangular coordinates in the Radon domain by using a rotated coordinate system (p, q) as

$$\begin{aligned} x \cos \theta + y \sin \theta &= p \\ -x \sin \theta + y \cos \theta &= q. \end{aligned} \quad (2.72)$$

The above implies

$$\mathbf{R}\{f(x, y)\} = \int_{-\infty}^{\infty} f(p \cos \theta - q \sin \theta, p \sin \theta + q \cos \theta) dq. \quad (2.73)$$

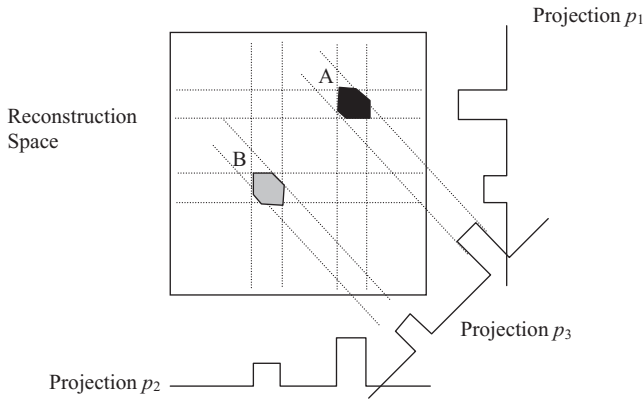


Figure 2.16 A schematic diagram for reconstructing images from projections. Three projections are back-projected to reconstruct objects A and B.

A higher-dimensional Radon transform can be defined in a similar way. For example, the projection space for a 3-D Radon transform would be defined by 2-D planes instead of lines.

The significance of using the Radon transform for computing projections in medical imaging is that an image of a human organ can be reconstructed by back-projecting the projections acquired through the imaging scanner. Figure 2.16 shows a schematic diagram of an image reconstruction process, called the backprojection method. Three simulated projections of two objects A and B are back-projected into the reconstruction space. Each projection has two segments of values corresponding to the objects A and B. When the projections are back-projected, the areas of higher values represent two reconstructed objects yielded by the intersection of back-projected data. It should be noted that the reconstructed objects may have geometrical or aliasing artifacts because of the limited number of projections used in the imaging and reconstruction processes. To improve the geometrical shape and accuracy of the reconstructed objects, a large number of projections should be used. The use of the Radon transform in image reconstruction from projection for applications in CT is described in Chapter 8.

2.9. SAMPLING

Computerized image processing and analysis requires that the image information be in a digital format. For example, a 2-D image $f_d[x, y]$ may be represented by a 2-D matrix of discrete numbers denoting the brightness levels placed uniformly in two dimensions. A picture element with a brightness level is called a pixel. The pixels are positioned in the image space based on the spatial sampling rate. Let us assume that a 2-D image signal $f_a(x, y)$ exists in the analog domain. This signal needs to be sampled or discretized in order to create a digital image. One simple example is a digital camera with a charge-coupled device (CCD) array that provides a digital

image through a 2-D array of light detectors. The physical size of each detector and its position provides the sampling of data points in the digital image. Thus, for a given object size and the discrete number of detectors, a sampling spatial frequency in the object space can be determined. The important question is whether the spatial sampling frequency is adequate to capture the fine details of the object. The sampling theorem provides the mathematical foundation of the Nyquist criterion to determine the optimal sampling rate for discretization of an analog signal without the loss of any frequency information. The Nyquist criterion states that to avoid any loss of information or aliasing artifact, an analog signal must be sampled with a sampling frequency that is at least twice the maximum frequency present in the original signal.

The sampling function is basically a series of delta functions represented in the same dimensions as of the signal. For example, when sampling a 1-D signal, a sampling function is defined as a series of 1-D delta functions. For a 2-D image, a 2-D distribution of delta functions is defined as

$$s(x, y) = \text{comb}(x, y) = \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \delta(x - j_1\Delta x, y - j_2\Delta y) \quad (2.74)$$

where Δx and Δy are, respectively, the spacing of data points to be sampled in x - and y -directions.

Figure 2.17 shows an image of a 2-D distribution of Gaussian impulses (or comb function) and its Fourier transform.

The sampled version of the image $f_d[x, y]$ is given by

$$f_d[x, y] = f_a(x, y)s(x, y) = \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} f_a(j_1\Delta x, j_2\Delta y)\delta(x - j_1\Delta x, y - j_2\Delta y). \quad (2.75)$$

Let us consider, ω_x and ω_y to be the spatial frequencies in x - and y -directions, respectively, and $F_s(\omega_x, \omega_y)$ to be the Fourier transform of the sampled image $f_d[x, y]$.

Using the convolution theorem, it can be shown that

$$F_s(\omega_x, \omega_y) = \frac{1}{\Delta x \Delta y} \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} F_a(\omega_x - j_1\omega_{xs}, \omega_y - j_2\omega_{ys}) \quad (2.76)$$

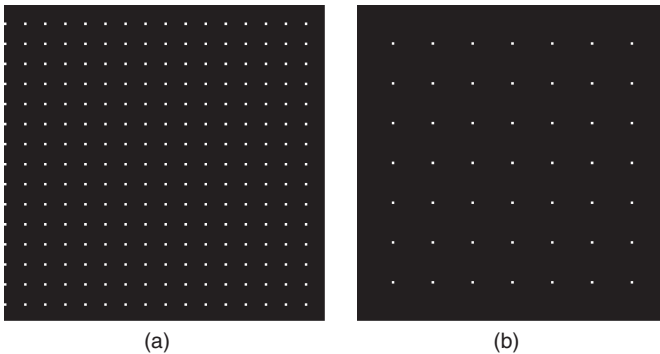


Figure 2.17 (a) A 2-D distribution of Gaussian impulses (or comb function) in the spatial domain and (b) its representation in the Fourier domain.

where $F_a(\omega_x, \omega_y)$ is the Fourier transform of the analog image $f_a(x, y)$ and ω_{xs} and ω_{ys} represent the Fourier domain sampling spatial frequencies such that

$$\omega_{xs} = 2\pi/\Delta x \quad \text{and} \quad \omega_{ys} = 2\pi/\Delta y. \quad (2.77)$$

For a good discrete representation, the analog signal should be sampled in such a way that the information from the sampled signal can be recovered without any loss or artifact. In other words, the sampling process must not cause any loss of frequency information.

It can be easily shown that the multiplication operation of $f_a(x, y)$ with the sampling signal $s(x, y)$ would create a convolution operation in the Fourier domain resulting in multiple copies of the sampled image spectrum with the $2\pi/\Delta x$ and $2\pi/\Delta y$ spacing in ω_x and ω_y directions, respectively.

Let us assume that the image as the 2-D signal in the spatial domain is bandlimited, for example, the Fourier spectrum $F_a(\omega_x, \omega_y)$ is zero outside the maximum frequency components ω_{xmax} and ω_{ymax} in x - and y -directions as shown in Figure 2.18a.

To recover the signal without any loss, the multiple copies of the image spectrum must not overlap. If this condition is not met, the overlapped region of image spectrum will create aliasing and the original signal or image cannot be recovered by any filtering operation. Thus, to avoid overlapping of image spectra, it is necessary that

$$\begin{aligned} \omega_{xs} &\geq \omega_{xmax} = 2\pi f_{xmax} \\ \omega_{ys} &\geq \omega_{ymax} = 2\pi f_{ymax} \end{aligned} \quad (2.78)$$

where f_{xmax} and f_{ymax} are the maximum spatial frequencies available in the image in x - and y -directions, respectively.

Figure 2.18b shows a good sampling of the bandlimited image signal shown in Figure 2.18a. Since the sampling frequency is higher than the Nyquist rate, multiple copies of the sampled spectrum do not overlap. Such a sampling does not cause any loss of information.

If the above condition is not satisfied, the undersampling will cause the overlapping of the high-frequency regions of the copies of image spectrum in the frequency domain. Such a sampled image will show an aliasing artifact, making it difficult to visualize the high-frequency information such as edges. Figure 2.18c shows undersampling of the bandlimited image signal shown in Figure 2.18a. Since the sampling frequency is lower than the Nyquist rate, multiple copies of the sampled spectrum overlap, causing the loss of high-frequency information in the overlapped regions. Due to the loss of information, the overlapping of high-frequency regions results in aliasing artifacts in the image.

To remove the aliasing artifact, a low-pass filter is required in the Fourier domain, prior to sampling, to remove the potential overlapped region of image spectra. After filtering the high-frequency information from the analog image, an inverse Fourier transform can recover the image in the spatial domain, but the recovered image will suffer from the lack of high-frequency information. Therefore, it is important to sample the image acquisition space adequately (preferably at a higher rate than the Nyquist rate) in medical imaging by placing an appropriate array

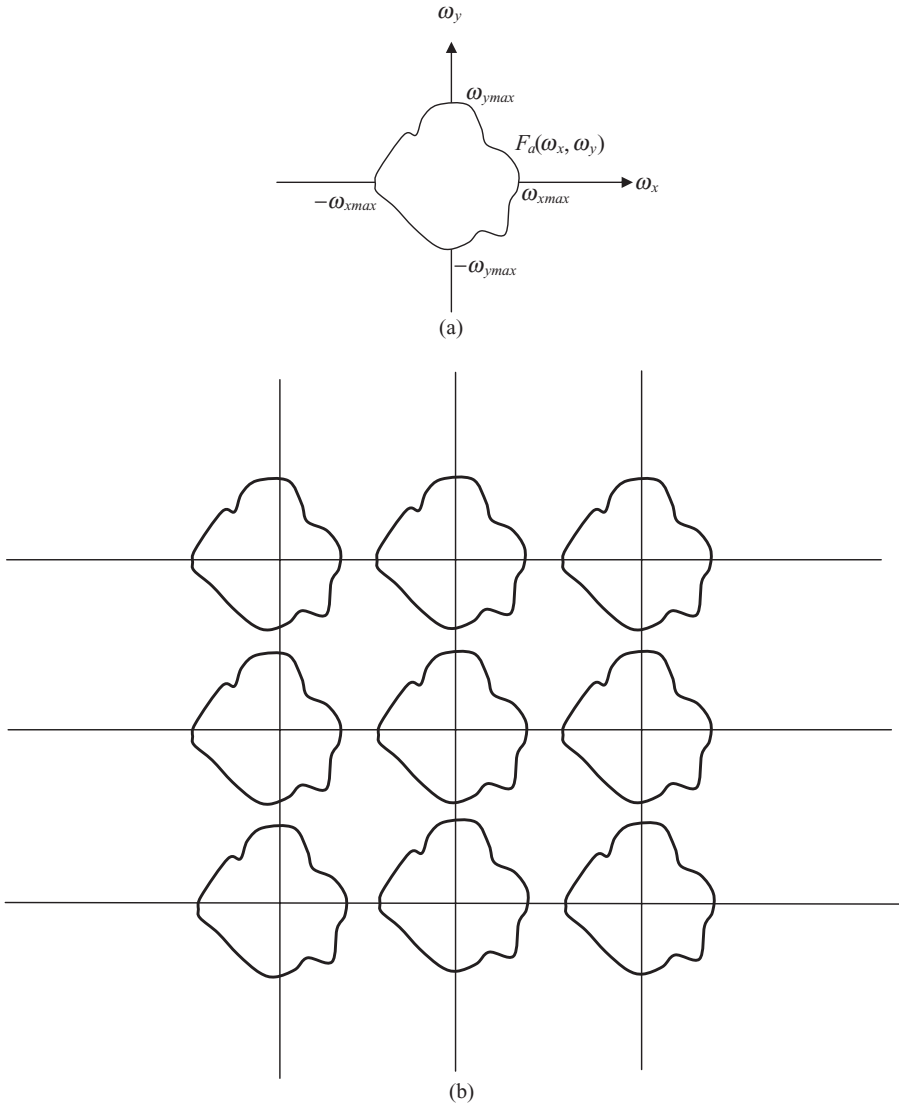
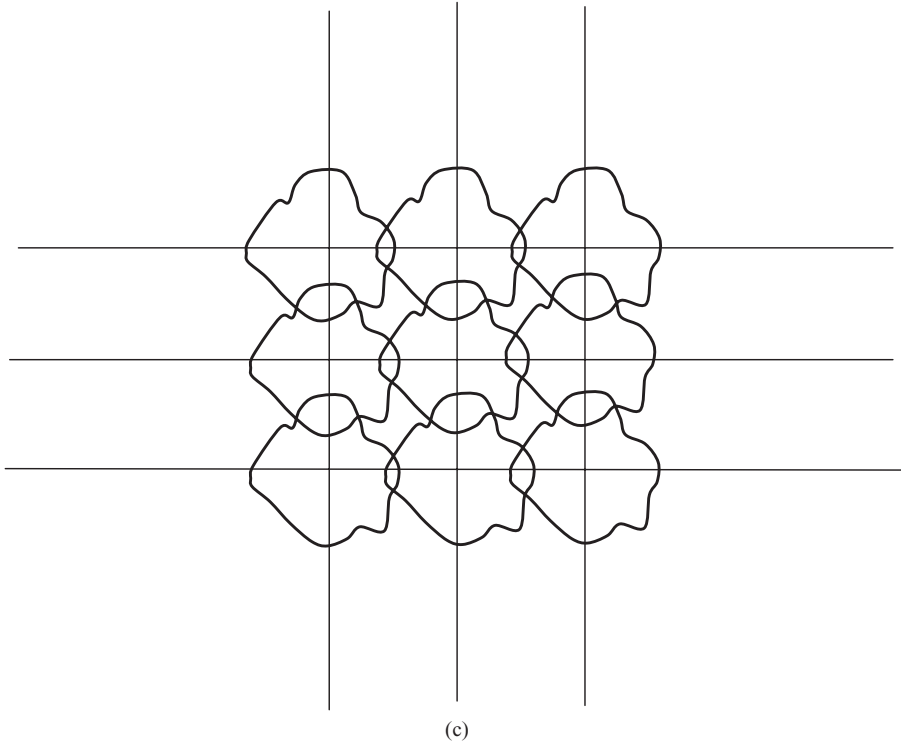


Figure 2.18 (a) The Fourier spectrum of a bandlimited two-dimensional signal; (b) sampling of the bandlimited signal with a sampling frequency higher than the Nyquist rate; (c) undersampling of the bandlimited signal with a sampling frequency lower than the Nyquist rate.

of detectors so that the acquired images do not suffer from aliasing artifacts. Figure 2.19 shows the effect of aliasing. An aliasing sinusoidal signal of a period of 1.8 pixels is shown in Figure 2.19a. The logarithmic magnitude of the Fourier transform of the aliasing signal shown in Figure 2.19a is shown in Figure 2.19b. Figure 2.19c shows the central cross section of the logarithmic magnitude of the Fourier transform, along the horizontal direction, of the image shown in Figure 2.11a, which is

Figure 2.18 *Continued*

sampled satisfying the Nyquist criterion. The two peaks of the Fourier transform can be seen without any distortion in Figure 2.19c. Figure 2.19d shows the central cross-section along the horizontal direction of the logarithmic magnitude of the Fourier transform of the aliasing pattern shown in Figure 2.19a showing the distortion of the Fourier spectrum.

2.10. DISCRETE FOURIER TRANSFORM

The discrete Fourier transform (DFT), $F(u, v)$ of an image $f(x, y)$ is defined as

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi \left(\frac{xu}{M} + \frac{yv}{N} \right)} \quad (2.79)$$

where u and v are frequency coordinates in the Fourier domain such that $u = 0, 1, \dots, M-1$ and $v = 0, 1, \dots, N-1$.

The inverse DFT in two dimensions can be defined as

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi \left(\frac{xu}{M} + \frac{yv}{N} \right)}. \quad (2.80)$$

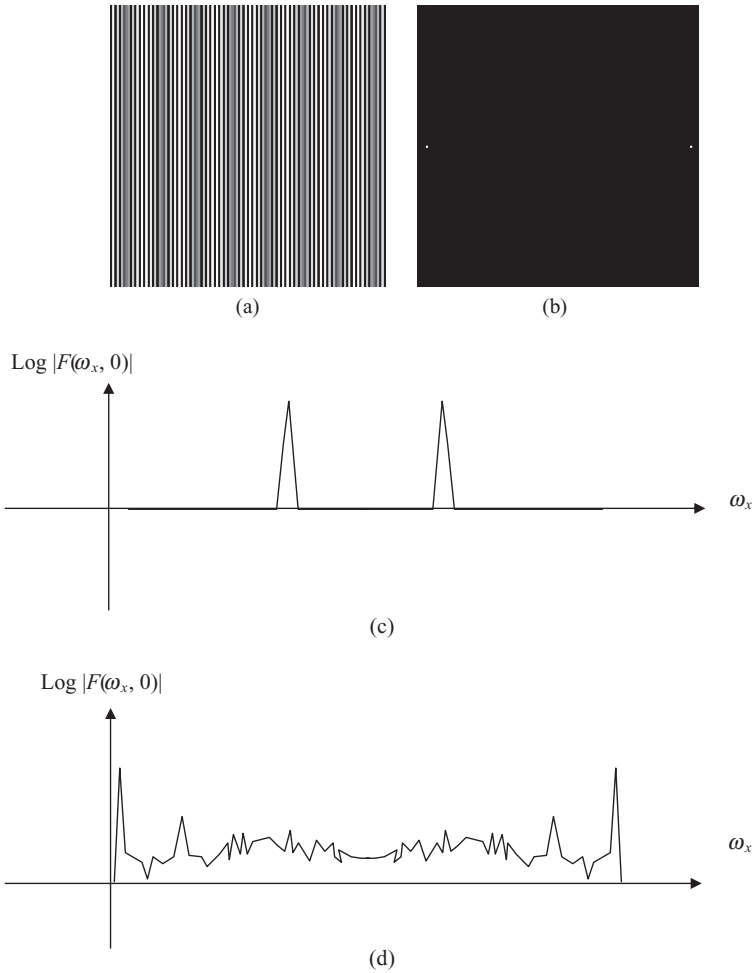


Figure 2.19 (a) A vertical stripe image generated from a sinusoidal waveform of a period of 1.8 pixels; (b) the logarithmic magnitude image of its Fourier transform; (c) the central cross-section of the logarithmic magnitude of the Fourier transform, along the horizontal direction, of the signal shown in Figure 2.8a, which is sampled, satisfying the Nyquist criterion; and (d) the central cross-section along the horizontal direction of the logarithmic magnitude of the Fourier transform of the aliasing pattern shown in Figure 2.15a.

The properties of the Fourier transform as described above are valid for the DFT as well. However, the numerical implementation of the DFT and fast Fourier transform (FFT) may make some approximations and finite computations. The errors in computations and approximations may cause artifacts in the image spectrum and may not allow an implementation of Fourier transform to be exactly reversible.

2.11. WAVELET TRANSFORM

The wavelet transform is a method for the complete frequency localization of a signal. The Fourier transform provides information about the frequency components present in the signal. However, Fourier analysis suffers from the drawback of the loss of localization or time information when transforming information from the time domain to the frequency domain. When the frequency representation of a signal is analyzed, it is impossible to tell when a particular frequency event took place. If the signal properties do not change much over time, this drawback may be ignored. However, signals change with interesting properties over time or space. An electrocardiogram signal changes over the time marker with respect to heartbeat events. Similarly, in the context of 2- and 3-D images, a signal or a property represented by the image (such as blood flow) may change over the sampled data points in space. Thus, Fourier analysis, in general, does not provide a specific event (frequency) localized information with respect to time (in time series signals) or space (in images). This drawback of Fourier transfer can be somewhat addressed with the use of short-time Fourier Transform (STFT) (19–22). This technique computes the Fourier transform to analyze only a small section of the signal at a time. As a matter of fact, STFT maps a signal into separate functions of time and frequency. The STFT provides some information about frequency localization with respect to a selected window. However, this information is obtained with limited precision determined by the size of the window. A major shortcoming of STFT is that the window size is fixed for all frequencies, once a particular size for the time window is chosen. In real applications, signals may require a variable window size in order to accurately determine event localization with respect to frequency and time or space.

Wavelet transform may use long sampling intervals where low-frequency information is needed, and shorter sampling intervals where high-frequency information is available. The major advantage of wavelet transform is its ability to perform multiresolution analysis for event localization with respect to all frequency components in data over time or space. Thus, wavelet analysis is capable of revealing aspects of data that other signal analysis techniques miss, such as breakdown points, and discontinuities in higher derivatives (19–22).

Wavelet transform theory uses two major concepts: scaling and shifting. Scaling, through dilation or compression, provides the capability to analyze a signal over different windows (sampling periods) in the data, while shifting, through delay or advancement, provides translation of the wavelet kernel over the entire signal. Daubechies wavelets (20) are compactly orthonormal wavelets that make discrete wavelet analysis practicable. Wavelet analysis has been used in numerous applications in statistics, time series analysis, and image processing (21–25).

The wavelet transform is based on the principle of linear series expansion of a signal using a set of orthonormal basis function. Through linear series expansion, a signal $f(t)$ can be uniquely decomposed as a weighted combination of orthonormal basis functions as

$$f(t) = \sum_n a_n \varphi_n(t) \quad (2.81)$$

where n is an integer index with $n \in \mathbb{Z}$ (\mathbb{Z} is a set of all integers), a_k are weights, and $\varphi_n(t)$ are orthonormal basis functions.

It follows that $f(t) \in V$ as the closed span of the expansion set $\{\varphi_n(t)\}$ denoted by V and expressed as

$$V = \overline{\text{Span}_n \{\varphi_n(t)\}}. \quad (2.82)$$

The weights or coefficients of expansions can then be expressed as

$$a_n = \langle \tilde{\varphi}_n(t), f(t) \rangle \quad (2.83)$$

where $\{\tilde{\varphi}_n(t)\}$ is set of dual basis function, and $\langle \rangle$ represents the inner product.

Using the orthonormality condition of the set of basis functions, Equation 9.46 can be expressed as

$$a_n = \langle \varphi_n(t), f(t) \rangle$$

$$\text{with } \langle \varphi_j(t), \varphi_k(t) \rangle = \delta_{jk} = \begin{cases} 0 & j \neq k \\ 1 & j = k \end{cases}. \quad (2.84)$$

The wavelet transform utilizes the series expansion of a signal using a set of orthonormal basis functions that are generated by scaling and translation of the mother wavelet $\psi(t)$ and the scaling function $\phi(t)$. The wavelet transform decomposes the signal as a linear combination of weighted basis functions to provide frequency localization with respect to the sampling parameter such as time or space. The multiresolution approach of the wavelet transform establishes a basic framework of the localization and representation of different frequencies at different scales.

If the sampling parameter is assumed to be time, a scaling function $\phi(t)$ in time t can be defined as

$$\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k) \quad (2.85)$$

where j is a scaling parameter, and k is a translation parameter, and $j, k \in \mathbb{Z}$, set of all integers, and $\{\phi_{j,k}(t)\}$ spans over the subspace $V_j \in L^2(\mathbf{R})$ where \mathbf{R} is the space of all real numbers.

The scaling and translation generates a family of functions using the following “dilation” equations:

$$\phi(t) = \sqrt{2} \sum_{n \in \mathbb{Z}} h_\phi(n) \phi(2t - n) \quad (2.86)$$

where $h_\phi(n)$ is a set of filter (generally called as low-pass filter) coefficients.

To induce a multiresolution analysis in $L^2(\mathbf{R})$, it is necessary to have a nested chain of closed subspaces defined as

$$\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots \subset L^2(\mathbf{R}). \quad (2.87)$$

Let us define a function $\psi(t)$ as the “mother wavelet” such that its translated and dilated versions form an orthonormal basis of $L^2(\mathbf{R})$ to be expressed as

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k). \quad (2.88)$$

The wavelet basis induces an orthogonal decomposition of $L^2(\mathbf{R})$, that is, one can write

$$\cdots \subset W_{-1} \subset W_0 \subset W_1 \subset W_2 \subset \cdots \subset L^2(\mathbf{R}) \quad (2.89)$$

where W_j is a subspace spanned by $\psi(2^j t - k)$ for all integers $j, k \in \mathbf{Z}$.

This basic requirement of a multiresolution analysis is satisfied by the nesting of the spanned subspaces similar to scaling functions, that is, $\psi(t)$ can be expressed as a weighted sum of the shifted $\psi(2t-n)$ as

$$\psi(t) = \sqrt{2} \sum_{n \in \mathbf{Z}} h_\psi(n) \psi(2t - n) \quad (2.90)$$

where $h_\psi(n)$ is a set of filter (generally called high-pass filter) coefficients.

The wavelet-spanned multiresolution subspace satisfies the relation

$$\begin{aligned} V_{j+1} &= V_j \oplus W_j \\ &= V_{j-1} \oplus W_{j-1} \oplus W_j \\ &= V_{j-2} \oplus W_{j-2} \oplus W_{j-1} \oplus W_j \end{aligned}$$

where \oplus denotes the union operation of subspaces, and

$$L^2(\mathbf{R}) = \cdots \oplus W_{-2} \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus W_2 \oplus \cdots \quad (2.91)$$

Figure 2.20 shows the wavelet multiresolution space showing the distribution of the original $L^2(\mathbf{R})$ space into subspaces through wavelet decomposition. Three levels of decomposition are shown. In the first level of decomposition, the original space V_{j+1} is divided into V_j and W_j spectral subspaces as the output of the low-pass and high-pass filters, respectively. These subspaces are then further decomposed into smaller spectral bands. It can be seen from the hatched spaces that for a complete spectral representation of the original signal in V_{j+1} space, it is sufficient to include the last low-pass filter space V_{j-2} and all high-pass filter spaces W_{j-2} , W_{j-1} and W_j . However, other combinations involving different spectral bands are possible.

Wavelet functions span the orthogonal complement spaces. The orthogonality requires the scaling and wavelet filter coefficients to be related through the following:

$$h_\psi(n) = (-1)^n h_\phi(1-n). \quad (2.92)$$

Let $x[n]$ be an arbitrary square summable sequence representing a signal in the time or space domain such that

$$x[n] \in L^2(\mathbf{R}). \quad (2.93)$$

V_{j+1}							
V_j				W_j			
V_{j-1}	W_{j-1}			WV_{j-1}	WW_{j-1}		
V_{j-2}	W_{j-2}	WV_{j-2}	WW_{j-2}	WVV_{j-2}	WWV_{j-2}	WWW_{j-2}	WWW_{j-2}

Figure 2.20 The wavelet multiresolution space with three levels of decomposition.

The series expansion of a discrete signal $x[n]$ using a set of orthonormal basis functions $\phi_k(n)$ is given by

$$x[n] = \sum_{k \in \mathbb{Z}} \langle \phi_k[l], x[l] \rangle \phi_k[n] = \sum_{k \in \mathbb{Z}} X[k] \phi_k[n] \quad (2.94)$$

where $X[k] = \langle \phi_k[l], x[l] \rangle = \sum_l \phi_k^*(l) x[l]$.

where $X[k]$ is the transform of $x[n]$. All basis functions must satisfy the orthonormality condition, that is,

$$\langle \phi_k[n], \phi_l[n] \rangle = \delta[k - l]$$

with

$$\|x\|^2 = \|X\|^2. \quad (2.95)$$

The series expansion is considered complete if every signal can be expressed using the expression in Equation 2.95. Similarly, using a set of bi-orthogonal basis functions, the series expansion of the signal $x[n]$ can be expressed as

$$\begin{aligned} x[n] &= \sum_{k \in \mathbb{Z}} \langle \phi_k[l], x[l] \rangle \tilde{\phi}_k[n] = \sum_{k \in \mathbb{Z}} \tilde{X}[k] \tilde{\phi}_k[n] \\ &= \sum_{k \in \mathbb{Z}} \langle \tilde{\phi}_k[l], x[l] \rangle \phi_k[n] = \sum_{k \in \mathbb{Z}} X[k] \phi_k[n] \end{aligned} \quad (2.96)$$

where $\tilde{X}[k] = \langle \phi_k[l], x[l] \rangle$ and $X[k] = \langle \tilde{\phi}_k[l], x[l] \rangle$ and $\langle \phi_k[n], \tilde{\phi}_l[n] \rangle = \delta[k - l]$.

Using the quadrature-mirror filter theory, the orthonormal bases $\phi_k[n]$ can be expressed as low-pass and high-pass filters for decomposition and reconstruction of a signal. It can be shown that a discrete signal $x[n]$ can be decomposed into $X[k]$ as

$$x[n] = \sum_{k \in \mathbb{Z}} \langle \phi_k[l], x[l] \rangle \phi_k[n] = \sum_{k \in \mathbb{Z}} X[k] \phi_k[n]$$

where

$$\begin{aligned} \phi_{2k}[n] &= h_0[2k - n] = g_0[n - 2k] \\ \phi_{2k+1}[n] &= h_1[2k - n] = g_1[n - 2k] \end{aligned}$$

and

$$\begin{aligned} X[2k] &= \langle h_0[2k - l], x[l] \rangle \\ X[2k + 1] &= \langle h_1[2k - l], x[l] \rangle. \end{aligned} \quad (2.97)$$

In Equation 2.97, h_0 and h_1 are, respectively, the low-pass and high-pass filters for signal decomposition or analysis, and g_0 and g_1 are, respectively, the low-pass and high-pass filters for signal reconstruction or synthesis. A perfect reconstruction of the signal can be obtained if the orthonormal bases are used in decomposition and reconstruction stages as

$$\begin{aligned} x[n] &= \sum_{k \in \mathbb{Z}} X[2k] \phi_{2k}[n] + \sum_{k \in \mathbb{Z}} X[2k + 1] \phi_{2k+1}[n] \\ &= \sum_{k \in \mathbb{Z}} X[2k] g_0[n - 2k] + \sum_{k \in \mathbb{Z}} X[2k + 1] g_1[n - 2k]. \end{aligned} \quad (2.98)$$

As described above, the scaling function provides low-pass filter coefficients and the wavelet function provides the high-pass filter coefficients. A multiresolution signal representation can be constructed based on the differences of information available at two successive resolutions 2^j and 2^{j-1} . Such a representation can be computed by decomposing a signal using the wavelet transform. First, the signal is filtered using the scaling function, a low-pass filter. The filtered signal is then sub-sampled by keeping one out of every two samples. The result of low-pass filtering and subsampling is called the scale transformation. If the signal has the resolution 2^j , the scale transformation provides the reduced resolution 2^{j-1} . The difference of information between resolutions 2^j and 2^{j-1} is called the “detail” signal at resolution 2^j . The detail signal is obtained by filtering the signal with the wavelet, a high-pass filter, and subsampling by a factor of two. Thus, wavelet-based decomposition of a signal $x[n]$ is obtained using a low-pass filter $h_0[k]$ (obtained from the scaling function) and a high-pass filter $h_1[k]$ (obtained from the wavelet function). The signal decomposition at the j^{th} stage can thus be generalized as

$$\begin{aligned} X^{(j)}[2k] &= \langle h_0^{(j)}[2^j k - l], x[l] \rangle \\ X^{(j)}[2k+1] &= \langle h_1^{(j)}[2^j k - l], x[l] \rangle. \end{aligned} \quad (2.99)$$

The signal can be reconstructed back from the decomposed coefficients using the reconstruction filters as

$$x[n] = \sum_{j=1}^J \sum_{k \in \mathbb{Z}} X^{(j)}[2k+1] g_1^{(j)}[n-2^j k] + \sum_{k \in \mathbb{Z}} X^{(j)}[2k] g_0^{(j)}[n-2^j k]. \quad (2.100)$$

The signal decomposition and reconstruction stages implementing Equations 2.99 and 2.100 are schematically shown in Figure 2.21, where H_0 and H_1 are shown,

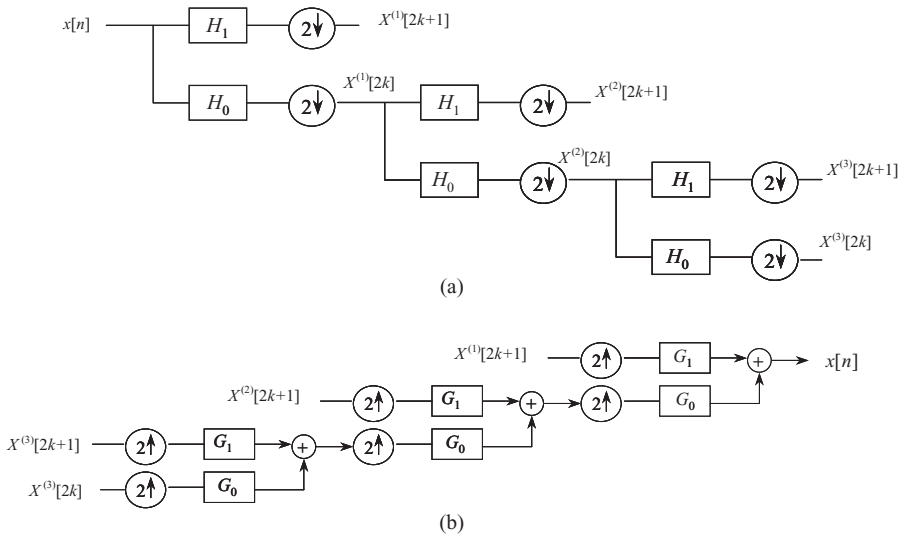


Figure 2.21 (a) A multiresolution signal decomposition using wavelet transform and (b) the reconstruction of the signal from wavelet transform coefficients.

respectively, as low-pass and high-pass filters for signal decomposition, and G_0 and G_1 are shown, respectively, as low-pass and high-pass filters for signal reconstruction.

In order to apply wavelet transform for image processing, the above method for 1-D wavelet transform for signal decomposition is applied to two individual 1-D vectors extracted from the 2-D image: the first vector reading pixels along the rows of the image (horizontal sampling), and the second vector reading pixels along the columns of the image (vertical sampling). The image (say at resolution 2^{i+1}) is first low-pass and high-pass filtered along the rows (Fig. 2.22). The result of each filtering process is subsampled. Next, the subsampled results are low-pass and high-pass filtered along each column. The results of these filtering processes are again subsampled. The combination of filtering and subsampling processes essentially provides the band-pass information. The frequency band denoted by A_j in Figure 2.22 is referred to as the low-low-frequency band. It contains the scaled low-frequency information. The frequency bands labeled D_j^1 , D_j^2 , and D_j^3 denote the detail information. They are referred to as low-high, high-low, and high-high-frequency bands, respectively. Thus, an image is decomposed into four quadrants: low-low, low-high, high-low, and high-high-frequency bands. This scheme can be iteratively applied to an image for further decomposition into narrower frequency bands; that is, each frequency band can be further decomposed into four narrower bands. However, for wavelet decomposition as shown in Figure 2.20, only low-low band is further decomposed into four quadrants of the next lower resolution. A two-level wavelet decomposition-based quadrant structure is shown in Figure 2.21b. Since each level of decomposition reduces the resolution by a factor of two, the length of the filter limits the number of levels of decomposition.

Several wavelets satisfying the mathematical constraints of providing orthonormal bases for signal decomposition have been developed by Daubechies (20). Least asymmetric wavelets computed with different support widths have been designed for signal and image processing applications (20–22). Figure 2.23 shows the scaling and wavelet functions with the corresponding decomposition and reconstruction low-pass and high-pass filters for Daubechies's db4 wavelet. Each filter is limited to eight coefficients in length. Figure 2.24 shows an X-ray mammographic image and its complete three-level decomposition using the db4 wavelet-based low-pass and high-pass decomposition filters shown in Figure 2.23. Figure 2.25 shows a denoised image of the X-ray mammogram shown in Figure 2.23 after soft thresholding the coefficients in the wavelet decomposition domain and then reconstructing the image using the reconstruction filters (shown in Figure 2.23). The wavelet description and operations can be found in the MATLAB wavelet toolbox. Image processing operations using wavelets are further described in Chapter 9.

Wavelet transform has been widely used in signal and image processing applications as well as in data mining and information processing applications. Wavelet transform provides many favorable properties, such as hierarchical and multiresolution decomposition structure, linear time and space relationships, and a wide variety of basis functions to extract specific spectral features for analysis (20–27).

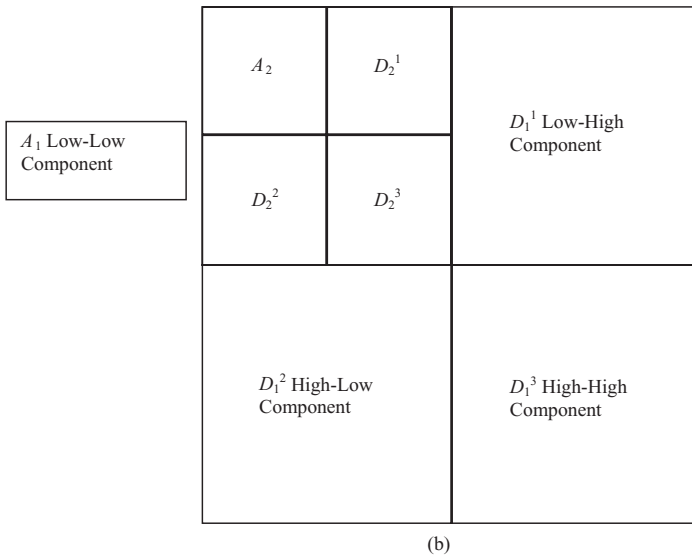
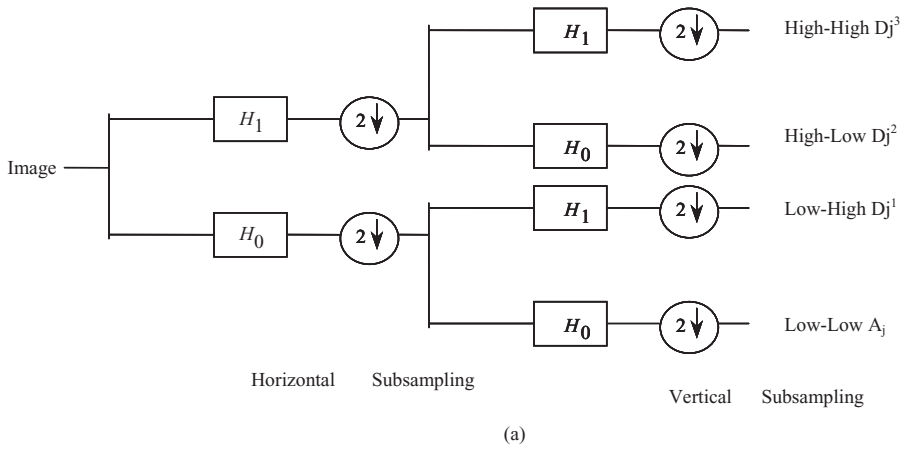


Figure 2.22 (a) Multiresolution decomposition of an image using the wavelet transform. (b) Wavelet transform-based image decomposition: the original resolution image ($N \times N$) is decomposed into four low-low A_1 , low-high D_1^1 , high-low D_1^2 , and high-high D_1^3 images, each of which is subsampled to resolution $\left(\frac{N}{2} \times \frac{N}{2}\right)$. The low-low image is further decomposed into four images of $\left(\frac{N}{4} \times \frac{N}{4}\right)$ resolution each in the second level of decomposition. For a full decomposition, each of the “Detail” component can also be decomposed into four sub-images with $\left(\frac{N}{4} \times \frac{N}{4}\right)$ resolution each.

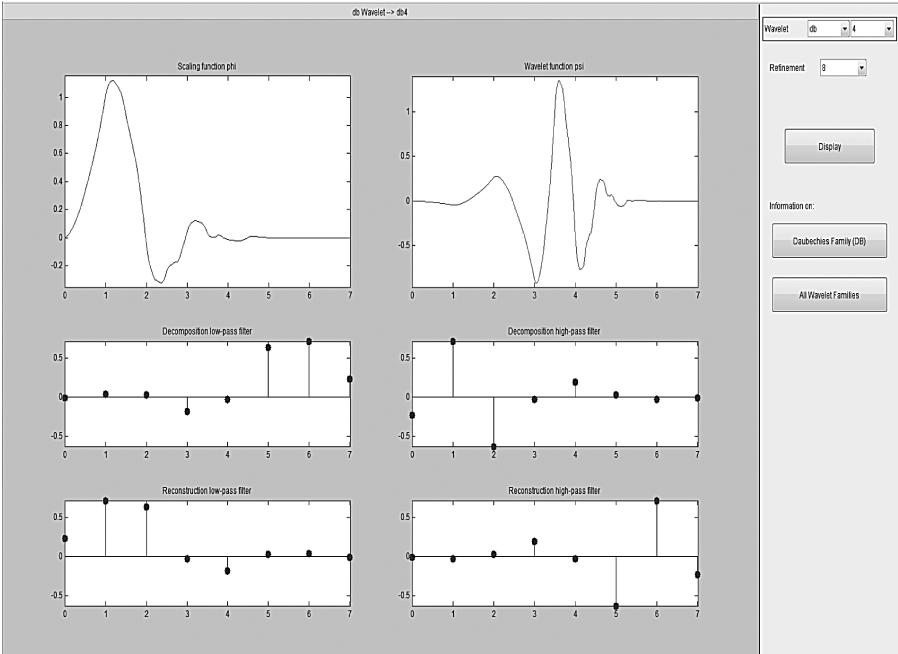


Figure 2.23 Scaling and wavelet functions with corresponding decomposition and reconstruction low-pass and high-pass filters for Daubechies’s db4 wavelet. The image is produced using the MATLAB wavelet toolbox GUI.

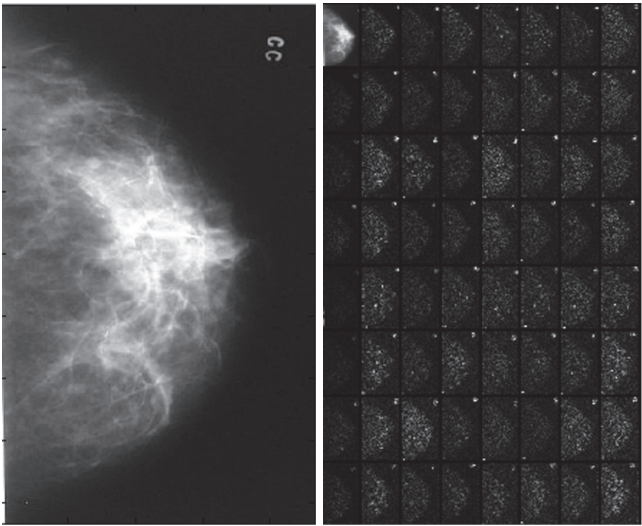


Figure 2.24 An image of X-ray mammogram (left) and its complete decomposition using Daubechies’s db4 wavelet.

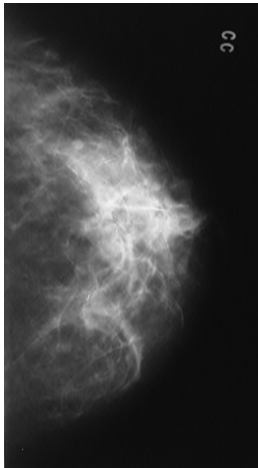


Figure 2.25 A denoised image of the X-ray mammogram (shown in Fig. 2.23) reconstructed after soft thresholding coefficients in the wavelet decomposition domain.

2.12. EXERCISES

- 2.1. What is the difference between a linear and a nonlinear imaging system? What problems would you expect if an imaging system is nonlinear and spatially variant?
- 2.2. Why it is important to design instrumentation in medical imaging that can provide spatially invariant point spread function (PSF)? What is the main advantage in dealing with linear spatially invariant medical imaging systems?
- 2.3. Prove Parseval's theorem in two dimensions.
- 2.4. Extend the Radon transform to three dimensions. Develop the mathematical expressions defining the projection domain.
- 2.5. Prove and explain the convolution and correlation properties of two-dimensional discrete Fourier transform (DFT).
- 2.6. Show that the frequency spectrum of a circularly symmetric function is also circularly symmetric.
- 2.7. Show that the Fourier transform of a unit impulse function located at the origin is a constant.
- 2.8. Show that the Fourier transform of a spatially distributed impulse function as $\text{comb}(x, y)$ is a $\text{comb}(u, v)$ function in the frequency domain.
- 2.9. Show that Fourier transform of a sampled function is an infinite periodic sequence of copies of Fourier transform of the original continuous function.
- 2.10. For the image formation as represented by Equation 2.44, show that $S_{gg}(u, v) = S_{ff}(u, v)|H(u, v)|^2$ where $S_{ff}(u, v)$ and $S_{gg}(u, v)$ are, respectively, the spectral densities of the random fields represented by $f(x, y)$ and $g(x, y)$; and $H(u, v)$ is the Fourier transform of point spread function, $h(x, y)$.

- 2.11.** Show that the Fourier transform of 1-D rectangular function of amplitude A and duration t is a sinc function given by $\left(At \frac{\sin(\pi ut)}{\pi ut} \right)$.
- 2.12.** Define the Nyquist criterion for sampling a 2-D signal. What should be the maximum sampling period to avoid aliasing if the highest frequency in the signal is 27 KHz?
- 2.13.** A medical imaging scanner has a detector system that cannot acquire the data faster than 1.0 KHz with a spatial resolution of $1 \text{ mm} \times 1 \text{ mm}$. Let us assume that the object is dynamically changing its contrast with a period 0.1 ms and needs to see pathology with circular dimension of about 1 mm diameter on average. Can this imaging system provide sufficient information through its images? Explain your answer.
- 2.14.** Display an image in MATLAB. Mark five landmarks in the image and note their positions. Create a rotation transform to rotate the image by 30 degrees along the horizontal axis and 45 degrees along the vertical axis. Implement the transformation in MATLAB. Find the landmarks and note their positions. Are all landmarks rotated by exactly the same vector? If not, why not?
- 2.15.** Apply the reverse rotation transformation to the rotated image in Problem 2.3. Subtract this image from the original image in the MATLAB environment. Does the resulting image contain all zeros? If not, explain the reason.
- 2.16.** Write a program in MATLAB to implement and display 2-D Radon transform. Obtain eight projections of a chest radiograph covering the angular span of 180 degrees around the image. Display the line-integral projections.
- 2.17.** Create a phantom image by assigning 255 gray-level value to the 2×2 pixel area in the center of a 256×256 pixel image. All other pixels are assigned zero value. Use 2-D FFT routine in MATLAB and display the Fourier spectrum image.
- 2.18.** Use the chest radiograph image and apply the 2-D FFT routine in MATLAB. Display the Fourier spectrum image. Now take 2-D inverse Fourier transform of the spectrum image to recover the spatial domain image. Subtract this image from the original image. Does the image contain all zero values?
- 2.19.** Can Fourier transform provide localization of spatial frequencies in an image? Elaborate your answer with an example.
- 2.20.** Explain the wavelet transform with respect to scaling and wavelet functions.
- 2.21.** How is the wavelet transform different from the short-time Fourier transform?
- 2.22.** Why must the bases in wavelet decomposition-based series expansion satisfy the condition of orthonormality?
- 2.23.** Describe a three-level wavelet processing with decomposition filters. Using the output of the decomposition filters, describe a method for loss-less perfect reconstruction.

- 2.24.** Use a brain image for wavelet-based three-level decomposition using the following wavelets in the MATLAB environment:
- HAAR
 - db2
 - db4
 - db8
- 2.25.** Use the denoised operation on the brain image used in Exercise 2.24 with db6 (level 3) wavelet in the MATLAB wavelet graphical user interface (GUI). Use three sets of different soft thresholds at each level with the fixed thresholding method to obtain three denoised images. Compare these images and explain the variations in the denoised images with respect to the threshold values. (You may add Gaussian noise in the original image and then repeat the denoised operations to observe performance variations better.)

2.13. REFERENCES

1. J.C. Russ, *The Image Processing Handbook*, 2nd Edition, CRC Press, Boca Raton, FL, 1995.
2. B. Jahne, *Practical Handbook on Image Processing for Scientific Applications*, CRC Press, Boca Raton, FL, 1997.
3. R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, 2nd Edition, Prentice Hall, Englewood Cliffs, NJ, 2002.
4. A.P. Dhawan and A. Sicsu, "Segmentation of images of skin-lesions using color and surface pigmentation," *Comput. Med. Imaging Graph.*, Vol. 16, pp. 163–177, 1992.
5. L.K. Arata, A.P. Dhawan, A.V. Levy, J. Broderick, and M. Gaskil, "Three-dimensional anatomical model based segmentation of MR brain images through principal axes registration," *IEEE Trans. Biomed. Eng.*, Vol. 42, pp. 1069–1078, 1995.
6. A.P. Dhawan, L.K. Arata, A.V. Levy, and J. Mantil, "Iterative principal axes registration method for analysis of MR-PET brain images," *IEEE Trans. Biomed. Eng.*, Vol. 42, pp. 1079–1087, 1995.
7. H.C. Andrews and B.R. Hunt, *Digital Image Restoration*, Prentice Hall, Englewood Cliffs, NJ, 1977.
8. H. Barrett and W. Swindell, *Radiological Imaging: The Theory of Image Formation, Detection and Processing, Volumes 1–2*, Academic Press, New York, 1981.
9. R.N. Bracewell, *Two-Dimensional Imaging*, Prentice Hall, Englewood Cliffs, NJ, 1995.
10. Z.H. Cho, J.P. Jones, and M. Singh, *Fundamentals of Medical Imaging*, John Wiley & Sons, New York, 1993.
11. M.P. Ekstrom, *Digital Image Processing Techniques*, Academic Press, Inc, New York, 1984.
12. P. Stucki, *Advances in Digital Image Processing*, Plenum Press, New York, 1979.
13. A. Rosenfield and A.C. Kak, *Digital Picture Processing*, 2nd Edition, Volumes 1 and 2, Academic Press, Englewood Cliffs, NJ, 1982.
14. J. Radon, "Über die Bestimmung von Funktionen durch ihre Integralwerte langs gewisser Mannigfaltigkeiten," *Ber. Verb. Saechs. AKAD. Wiss., Leipzig, Math. Phys.*, KI 69, pp. 262–277, 1917.
15. G.N. Hounsfield, "A method and apparatus for examination of a body by radiation such as X or gamma radiation," Patent 1283915, The Patent Office, London, 1972.
16. A.M. Cormack, "Representation of a function by its line integrals with some radiological applications," *J. Appl. Phys.*, Vol. 34, pp. 2722–2727, 1963.
17. G.T. Herman, *Image Reconstruction from Projections*, Academic Press, New York, 1980.
18. A. Papoulis and S.U. Pillai, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York, 2001.

19. S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 11, No. 7, pp. 674–693, 1989.
20. I. Daubechies, *Ten Lectures on Wavelets*, Society for Applied Mathematics, Philadelphia, PA, 1992.
21. S. Mallat, "Wavelets for a vision," *Proc. IEEE*, Vol. 84, No. 4, pp. 604–614, 1996.
22. M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*, Prentice Hall, Englewood Cliffs, 1995.
23. A. Raheja and A.P. Dhawan, "Wavelet based multiresolution expectation maximization reconstruction algorithm for emission tomography," *Comput. Med. Imaging Graph.*, Vol. 24, No. 6, pp. 359–376, 2000.
24. J.B. Weaver, X. Yansun, D.M. Healy Jr., and L.D. Cromwell, "Filtering noise from images with wavelet transforms," *Magn. Reson. Med.*, Vol. 21, No. 2, pp. 288–295, 1991.
25. S. Patwardhan, A.P. Dhawan, and P. Relue, "Classification of melanoma using tree-structured wavelet transform," *Comput. Methods Programs Biomed.*, Vol. 72, No. 3, pp. 223–239, 2003.
26. A. Laine, "Wavelets in spatial processing of biomedical images," *Annu. Rev. Biomed. Eng.*, Vol. 2, pp. 511–550, 2000.
27. M. Unser, A. Aldroubi, and A. Laine, "IEEE transactions on medical imaging," Special Issue on Wavelets in Medical Imaging, 2003.