1

Introduction

1.1 Signals, Operators, and Imaging Systems

As a simple definition, we may consider an imaging system to map the distribution of the input "object" to a "similar" distribution at the output "image" (where the meaning of "similar" is to be determined). Often the input and output amplitudes are represented in different units. For example, the input is often electromagnetic radiation with units of, say, watts per unit area, while the output may be a transparent negative emulsion measured in dimensionless units of "density" or "transmittance". In other words, the system often changes the form of the energy; it is a "transducer". The goal of this book is to mathematically describe the properties of imaging systems, and it is often convenient to use the model of a system as a chain of links.

1.1.1 The Imaging Chain

An imaging system is often modeled as a "chain" of links that transfer information in the form of energy from an input (the *object*) to the output (the *image*) of the system. Many schemes of links in an imaging system are plausible, depending on details, but one eight-link model is appropriate in many imaging systems:

1. The source of energy (usually in the form of electromagnetic radiation).
2. The object to be imaged, which interacts with the energy from the source by reflection, refraction, absorption, scattering, and/or other mechanism.
3. Propagation of the energy to the imaging system.
4. Energy collection (often using an optical system composed of lenses and/or mirrors).
5. Sensing or detection by a transducer (converts incident energy to a measurable form, e.g., photons to electrons).
6. Image processing, including data compression (if any).
7. Storage and/or transmission (if any).
8. Display.

The source and object often are one and the same, e.g., radiating objects such as stars. Sometimes it is useful to model the imaging system with a second stage of energy propagation after collection, as when evaluating optical imaging systems composed of a single thin lens.

Fourier Methods in Imaging Roger L. Easton, Jr.

In this book, we consider a simplified picture of the imaging chain that combines the source, object, and energy propagation into one entity that we will call the "input function" (or just the "input"); it usually is specified by a single-valued physical quantity f that varies over continuous coordinates in space, and perhaps in time t and color, specified by wavelength λ (or equivalently by temporal frequency ν):

$$\textit{Input to imaging system} = f[x, y, z, t, \lambda] \textit{ or } f[x, y, z, t, \nu] \tag{1.1a}$$

Though these quantities have explicit spectral and temporal coordinates in addition to the spatial dimensions, the signals considered in this book most often will be functions of one or two spatial coordinates only and are specified by $f[x]$ or $f[x, y]$. In other words, the input and output signals will be constant in time and wavelength. The number of coordinates necessary to specify the function is the *dimensionality*, and the set of all possible such coordinates defines its *domain*. We will use the shorthand notation of "1-D" and "2-D" for one- and two-dimensional signals, respectively; 2-D signals such as $f[x, y]$ are of greatest interest in imaging applications, but the study of 1-D signals will be considered in depth in this book as well. This is because 1-D systems often are easy to visualize and the results may be directly transferable to problems with higher dimensionality.

The output of the imaging system also will usually be specified by a single-valued physical quantity that will be denoted by g. Though the domain of the image g may be identical to that of the object f, it is more common to require different coordinates, and these will be denoted, when necessary, by primed coordinates. In many cases, the output will be a function of two spatial coordinates with no dependence on wavelength:

$$\textit{Output of imaging system} = g[x, y, z, t, \lambda] \rightarrow g[x, y] \tag{1.1b}$$

In imaging applications, the numerical value assigned to the dependent variable of the input or output signals in Equation (1.1) represents a measurable physical quantity. Based on the familiarity of optical imaging, a common descriptive name for f is the "brightness" of the scene, though this terminology is not used in some subdisciplines of optics, such as radiometry. Regardless of the nomenclature, the appropriate quantity (i.e., the units of f) depends on the particular imaging system. In optical imaging, the relevant quantity is the *irradiance* at each coordinate: the time average of the square of the magnitude of the electric field. In X-ray or gamma-ray imaging, the measured quantity is the number of quanta incident on the sensor as a function of spatial location. In acoustic imaging (sonar or ultrasound), the acoustic power radiated, transmitted, or reflected by an object is the quantity of record.

Now consider some important examples in a bit more detail. For example, optical images may be generated by light that is *coherent* or *incoherent*. At this point, we can think of coherent light as composed of a single wavelength λ and incoherent light as composed of many wavelengths, such as *natural light*. The relevant quantity in optical imaging with coherent illumination is the complex-valued amplitude of the electric field (including both its magnitude and the phase). The appropriate measured quantity in incoherent (natural) light is the time average of the square of the magnitude of the complex-valued amplitude of the electromagnetic field; this is the *irradiance* and may be denoted by $\langle |f|^2 \rangle$. In still other applications, the physical quantity represented by f may have a very different form. Though the signal and the system may have different forms, most of the principles discussed in this book will be applicable to some extent in all imaging situations.

The description of the imaging system requires a mathematical model of its action upon the input function f to generate the output g. This action will be denoted by an operator represented as an upper-case script character, such as $\mathcal{O}\{f[x, y, \ldots]\}$. The operator symbolizes the mathematical rule that assigns a particular output amplitude g to every location in its domain. In many cases, it will be possible to describe the action of the system as the combination of a specific function associated with the imaging system (the "system function") and a particular mathematical operation (e.g., multiplication or integration).

It is obvious that the output image generally is affected both by the mathematical form of the specific input object and by the characteristics of the system. The functional expression for a common type of simple image would be:

$$\mathcal{O}\{f[x, y, z, t, \lambda]\} = g[x', y'] \tag{1.2}$$

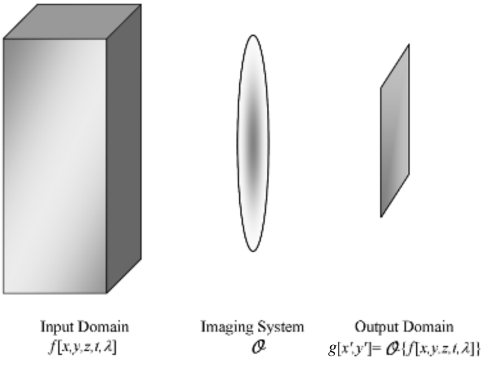

Figure 1.1 Schematic of an imaging system that acts on a time-varying input with three spatial dimensions and color, $f[x, y, z, t, \lambda]$ to produce a 2-D monochrome (gray-scale) image $g[x', y']$.

The schematic of the imaging process is shown in Figure 1.1. The spatial domain of the output image is often different from that of the input object, hence the use of primed characters in Equation (1.2). In realistic situations, the amplitude g also is affected by other parameters, such as the time, the exposure time Δt, and the spectral response of the sensor. In Equation (1.2), the effects of these additional parameters could be considered to be implicit in the system operator $\mathcal{O}$.

1.2 The Three Imaging Tasks

In many imaging applications, input objects and output images are functions of spatial dimensions only. Examples of mathematical relations for 1-D and 2-D systems are:

$$\mathcal{O}\{f[x]\} = g[x'] \tag{1.3a}$$

$$\mathcal{O}\{f[x, y]\} = g[x', y'] \tag{1.3b}$$

Simply put, the imaging chain relates three "entities": the input object, the action of the imaging system, and the output image. These three entities are denoted by the symbols f, $\mathcal{O}\{\ \}$, and g, respectively. A simple description of an imaging "task" is the process of specifying one of the three entities from knowledge of the other two. Three cases are evident:

1. The *forward* or *direct* problem: to find the mathematical expression for the image $g[x', \ldots]$ given complete knowledge of the input object $f[x, \ldots]$ and the system $\mathcal{O}$.
2. The *inverse problem*: to evaluate the input $f[x, \ldots]$ from the measured image $g[x', \ldots]$ and the system $\mathcal{O}$.
3. The *system analysis* problem: to determine the action of the operator $\mathcal{O}$ from the input $f[x, \ldots]$ and the image $g[x', \ldots]$ (the solution is often very similar in form to that of the inverse problem).

The solution of the direct task is often rather easy, while the others may be difficult or even mathematically impossible. Other and more complicated variants of these imaging problems are common, including the cases where knowledge of the entities (f, g, and/or $\mathcal{O}$) may be incomplete or contaminated by random noise. Some of the variants of the imaging problem will be considered in later chapters.

The additional complexity of the more general imaging system model is perhaps evident just from observation of the form of the general 1-D imaging relation in Equation (1.3a); the operator $\mathcal{O}$ must be a function of x and x' because it relates the object to the image. The most general form of $\mathcal{O}$

in Equation (1.3a) may modify the "brightness" f and/or the "location" x of all or part of the input signal by rearrangement, amplification, attenuation, or removal in an arbitrary fashion. For example, the image amplitude g at a specific location could be derived from the input amplitude at the corresponding location, from that at a different location, or from amplitudes at multiple locations in the input $f[x, \ldots]$. The functional form of the relationship between $f[x, \ldots]$ and $g[x', \ldots]$ may be linear or nonlinear, deterministic or random.

Though it is desirable to mathematically represent the action of system operators so that they are both concise and generally applicable, these two characteristics usually are mutually exclusive. In other words, a general system operator appropriate for the imaging task likely is impossible to specify in a concise mathematical notation.

Perhaps these examples give the readers some flavor of the difficulties to be attacked when specifying the action of an imaging system that is more general than the usual simple cases.

1.3 Examples of Optical Imaging

At this point we introduce a few simplified examples of optical and medical imaging systems to illustrate the imaging "tasks" and the mathematical concepts introduced in this book.

1.3.1 Ray Optics

Solution of a particular imaging task demands that the available "entities" f, g, and $\mathcal{O}$ be represented or modeled as mathematical expressions, which are then manipulated to derive an expression for the unknown entity. To illustrate the concept, consider the particularly simple, yet still very useful, mathematical model of optical imaging from introductory optics. A point source of energy emits geometrical "rays" of light that propagate in straight lines to infinity in all directions. The imaging "system" is an optical "element" that interacts with any ray it intercepts. The interaction mechanism is a physical process (usually refraction or reflection) that "diverts" the ray from its original direction. In this example, the optical element is a single thin lens located at a distance z_1 from the source. If the diameter of the lens is infinite, then all rays that move at all from left to right will intercept the lens and be diverted. Such a system may be described by the single parameter, the "focal length", which determines the "power" of the system to redirect rays. We will denote the focal length by the bold-faced roman "**f**" to distinguish it from the italic character "f" that will be used to represent the input amplitude. In the example of Figure 1.2, **f** is positive and the system redirects the light rays that emerge from the same object point so that they converge to an image point located at some distance z_2 from the lens. The mathematical descriptions relevant to the input "object", the imaging system, and output image are respectively the distance z_1 from the object to the lens, the focal length **f**, and the distance z_2 from the lens to the image point. The relationship of these three distances is the mathematical model of the imaging system, which is most commonly presented in the form:

$$\frac{1}{z_1} + \frac{1}{z_2} = \frac{1}{\mathbf{f}} \tag{1.4a}$$

This simple model of the optical system is "perfect" in the sense that all light rays emerging from a single source point of infinitesimal size at the object are assumed to converge to a single infinitesimal area about the location: the point "image". The relation of Equation (1.4a) may be rearranged into forms for each of the three tasks by placing the known parameters on one side of the equation and the unknown value on the other. The relations for the three tasks are trivial to derive from Equation (1.4a):

1. Direct task: given the object distance z_1 and the parameter **f** of the imaging system, find the output image distance z_2. The mathematical expression to be solved is a simple rearrangement of Equation (1.4a) with the two known quantities on the left-hand side and the unknown on the

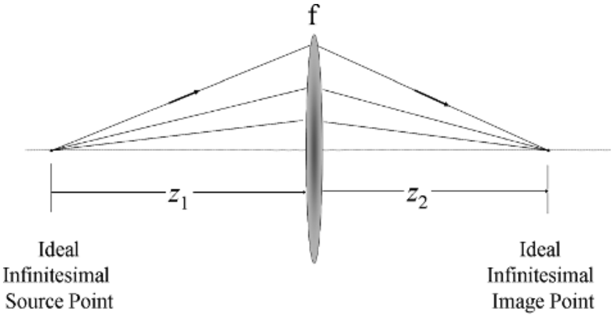

Figure 1.2 Optical imaging system in the ray optics model with no aberrations. All rays from a single source point converge to a single *image* point.

right:

$$\left(\frac{1}{\mathbf{f}} - \frac{1}{z_1}\right)^{-1} = z_2 \tag{1.4b}$$

2. Inverse task: given the output image point located at z_2 and the description of the system in the form of **f**, find the input object point via:

$$\left(\frac{1}{\mathbf{f}} - \frac{1}{z_2}\right)^{-1} = z_1 \tag{1.4c}$$

3. Analysis task: given the input object and output image points, determine the specification of the imaging system:

$$\left(\frac{1}{z_1} + \frac{1}{z_2}\right)^{-1} = \mathbf{f} \tag{1.4d}$$

In this particularly simple model, the mathematical relations to be solved in the direct and inverse tasks differ only in which parameter appears on the right-hand side, which means that the direct and inverse imaging tasks are equally "difficult" (or rather, equally trivial) to solve.

For a "fixed system" with focal length **f**, the imaging equations in Equation (1.4) "pair up" object and image planes located at distances z_1 and z_2. In other words, the imaging system constructs a "mapping" of source planes to image planes. However, we rarely seem to think about the action of the lens on objects with depth, or even on planar objects located at some distance z_1 that does not satisfy the imaging equation for a fixed focal length **f** and image distance z_2. In other words, we rarely even consider "out-of-focus" images in this simple model, though it is easy to imagine situations where we might want to calculate the appearance of such an image. It is possible to evaluate the "quality" of the image created by a nonideal imaging system using this simple ray optics model, but there are significant benefits from the more sophisticated model of light as a wave that is introduced next.

1.3.2 Wave Optics

Models of light that are more sophisticated than simple "rays" require a different mathematical description of how the "brightness" (the amplitude) generated by the object propagates through space. The resulting system operator $\mathcal{O}$ is significantly more complicated, and thus so are the corresponding equations that relate the input object and output image. A more complete model of optical imaging considers the physical observation that light "rays" are subject to optical "diffraction" that makes the energy deviate from straight-line propagation. A simple extension of the example already considered produces a finite-sized "patch" of light instead of an infinitesimal point, as suggested by Figure 1.3.

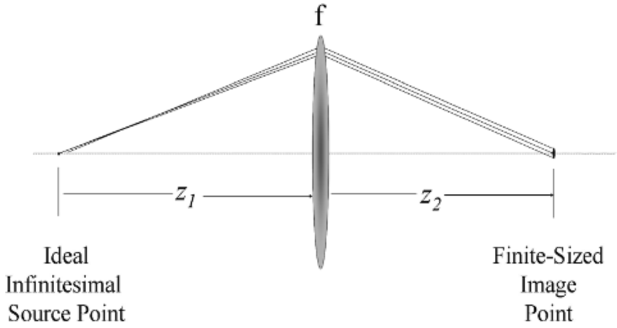

Figure 1.3 Ray model of optical imaging that includes *diffraction*, so that rays from a single source point "spread" while propagating to and from the lens. The image is not a single image "point", but rather a "blurry" spot.

It is more convenient to describe the light as a "wave" instead of as a "ray" in models of imaging systems that include diffraction. Each infinitesimal source point in this model emits spherical *wavefronts* of electromagnetic radiation that propagate outward from the source at the velocity of light. The radiation at all points on a particular such wave surface was emitted at the same instant of time. In other words, the vectors perpendicular to the local wave surface are the *rays*. One benefit of the wave model is due to the fact that a mathematical function exists that describes the spherical wave everywhere in space, thus making possible a "large-scale" or "global" picture of the radiation.

An optical element of the system (typically a lens or mirror) intercepts a portion of each spherical wave and acts to change its radius of curvature, perhaps even "reversing" the curvature so that successive propagating wave surfaces then *converge* toward an "image" of the point before diverging again (Figure 1.4). This model suggests another interpretation of the action of the optical system as an attempt to "replicate" the infinitesimal energy distribution of the point source. The fidelity of the reproduction is determined by the size of the image; a more faithful image exhibits a more compact distribution of energy. In real life, the size of the image produced by a "flawless" optical system decreases and the fidelity improves if the system intercepts are a larger area of the outgoing wave. In other words, the size of the image of the point object decreases as the area of the optic increases (assuming no defects in the optical system known as "aberrations"). We note at this point that it is physically impossible to replicate the infinitesimal area of the original source; even a "perfect" image has a finite area. The difference is ascribed to the phenomenon of optical "diffraction" and the finite-sized image of an infinitesimal point source is a "diffraction spot".

Images produced by multiple sources may be calculated in this model if the output of the system is the sum of the individual outputs from the point-source inputs; this is the first introduction to the concept of a *linear* system. In such a case, the image of two equally bright and closely spaced point sources in the wave optics model is the sum of two finite-sized "diffraction patterns" that may overlap. As the distance between the sources is decreased, the ability of the observer to distinguish the individual sources from the image will become more difficult. This leads to the concept of "spatial resolution" of an optical imaging system, as shown in Figure 1.4. The objects are pairs of point sources separated by different angles. The first pair of images is clearly distinguishable, while the overlap of the two images in the second example makes it more difficult to discern the true nature of the original object.

The wave model of light also allows estimation of the "pattern of energy" generated by the spherical wave at locations other than the image point, i.e., the distribution of light at an "out-of-focus" image may be calculated (Figure 1.5). This provides a means to establish the "appearance" or "quality" of the image at locations that are "in" or "out of" focus.

Optical imaging models that include diffraction are considered in Chapters 21–23.

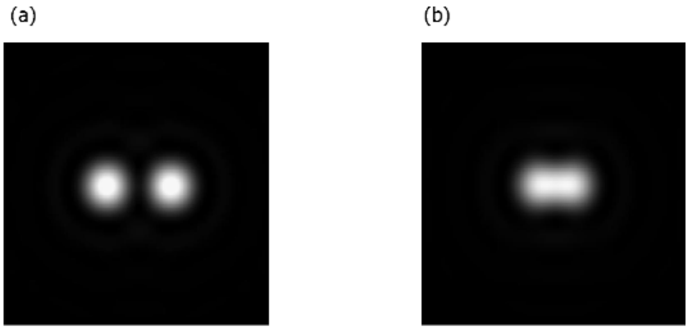

Figure 1.4 The effect of diffraction on the ability to resolve objects and fine structure in the image. The original objects are pairs of point objects (e.g., double stars) at two different separations viewed through an aberration-free system. In (a), the object clearly consists of two disjoint stars, but the result is less convincing in (b).

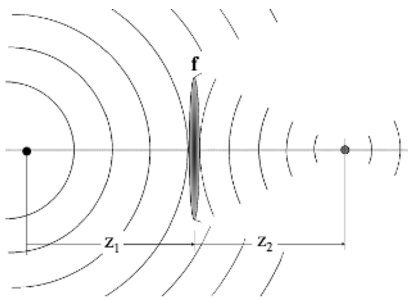

Figure 1.5 Schematic of the wave model of light propagation. Several spherical waves are shown that were emitted by the infinitesimal point source at times separated by equal intervals. The lens intercepts a portion of each wave and changes the curvatures of the wavefronts to make them converge. The size of the resulting image point is finite, not infinitesimal. The waves continue to propagate and expand away from the image.

1.3.3 System Evaluation of Hubble Space Telescope

One of the more impressive achievements of imaging science was the successful diagnosis of the optical system in the Hubble Space Telescope (HST) from images taken while in orbit. The HST was deployed from the Space Shuttle in 1990 after more than a decade of design and construction. Readers are probably aware that the primary mirror was improperly figured due to faulty testing and the quality of resulting images was much poorer than specified. Before the design and construction of corrective optics, the action of the optical system had to be characterized "at a distance" (to put it very mildly!). In the terms of our imaging-system description, the system operator $\mathcal{O}\{\ \}$ was only partly known when HST was deployed. By combining information from output images $g[x, y]$ of known object functions $f[x, y]$ with educated guesses of the cause of the optical faults, the action of the existing system was determined to sufficient accuracy to enable engineers to design the optical compensator dubbed COSTAR (the *Corrective Optics Space Telescope Axial Replacement*). COSTAR was retrofitted into HST during a Shuttle servicing mission in 1993. The results of this team effort to fix the problem were nothing short of spectacular; the quality of the corrected images met the original specifications. Two

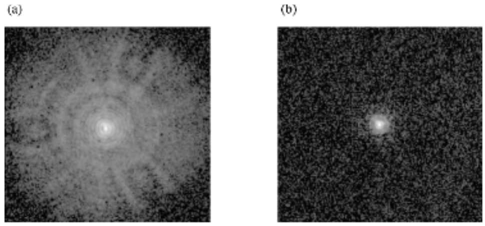

Figure 1.6 Images of a star from HST before and after adding the COSTAR optical corrector: (a) image before correction that exhibits "blur" due to the spherical aberration of the primary mirror; (b) image obtained after correction with COSTAR, where the light is concentrated in a much smaller spot.

images that compare the quality before and after the retrofit are shown in Figure 1.6, and other examples of the results before and after correction may be accessed at the HST website, hubble.stsci.edu.

1.3.4 Imaging by Ground-Based Telescopes

Now consider a practical example of an imaging system – the imaging of stellar objects through the atmosphere by ground-based telescopes. Stars seen from Earth are excellent approximations of point sources, so that their wavefronts are effectively spherical waves with $z_1 \cong \infty$. In other words, the wavefronts are effectively planar before encountering the atmosphere. At visible wavelengths and at standard temperature and pressure, the atmosphere is a transmissive medium that refracts light according to Snell's well-known law with index of refraction $n \cong 1.0002$, which varies with air density and thus with temperature. Local temporal variations in the temperature (and thus the density) of air produce local temporal variations in the refractive index, and thus change the angle of refraction. Since the air temperature varies locally and with depth in the atmosphere, the plane waves emitted by star(s) are refracted in a random pattern. The deformations translate and "defocus" the images over time intervals of the order of hundredths of seconds. The resulting image is a jumble of energy since light from a star is recorded at different locations in the image plane while light from nearby parts of the sky may be recorded at the same detector site. The clusters of recorded energy are commonly called *speckles*. The resolution of the recorded images is limited by these atmospheric aberrations, often to a very large degree. Fortunately, mathematical tools have been developed to utilize these speckle images to recover useful information out to the diffraction limit of the telescope. An early imaging tool for this purpose that was developed by Labeyrie in 1972 processed multiple short-exposure images taken through the atmosphere to provide high-resolution information that would normally be averaged to invisibility. The technique was modified by Knox and Thompson in 1974. A simulated example of Labeyrie's algorithm for "stellar speckle interferometry" is shown in Figure 1.7.

1.4 Imaging Tasks in Medical Imaging

Medical diagnostic imaging is another classic example of the inverse problem where the need is to determine the unknown "input" from the measured output and knowledge of the system. One essential difference between medical and astronomical imaging is due to the fact that the electromagnetic radiation used in the former is very much more "energetic"; the radiation is envisioned to be in the form of "photons" with very large energies that therefore propagate in straight lines in the manner of perfect geometrical rays. We now consider simplified versions of a few imaging systems with very different properties that are used in medical diagnostic applications.

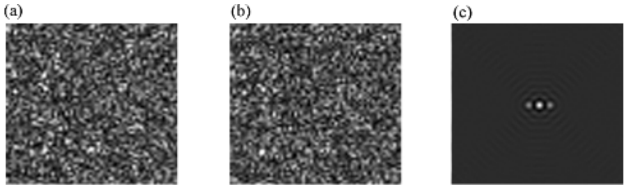

Figure 1.7 Simulation of Labeyrie's algorithm for stellar speckle interferometry: (a) image obtained from a single star through a turbulent atmosphere, showing "speckles" due to local variations in atmospheric refraction; (b) simulated image of a double star taken through the same atmosphere, showing "paired" speckles; (c) processed image from several instances of (b), which provides evidence of the existence and separation of the double star.

1.4.1 Gamma-Ray Imaging

In this first example, the individual photons, generated by nuclear decay in radioactive materials, are called *gamma rays*. Gamma-ray imaging is the basis for diagnostic nuclear medicine, where chemicals selectively absorbed by specific organs are tagged with radioactive atoms. For example, it is possible to attach radioactive technetium to iodine atoms that are injected into the body. The iodine is then selectively absorbed by the thyroid gland. It is medically useful to image gamma rays emitted by radioactive decay of the technetium to visualize the pathology of the thyroid gland. The object function may be written as $f[x, y, z, \lambda, t, \theta, \phi]$, where f is the number of photons emitted from the location in unit time in the spherical direction defined by the azimuth angle θ and latitude ϕ.

The kinetic energy of the gamma rays emitted by the technetium is approximately 0.14 MeV, which translates to an energy $E \cong 2.24 \times 10^{-7}$ ergs per photon. Imaging these gamma-ray photons presents a problem that does not exist in the imaging of visible light because the corresponding wavelength of light for the gamma ray is $\lambda_\gamma = hc/E \cong 9 \times 10^{-12}$ m, which is very much shorter than visible wavelengths with $\lambda \cong 5 \times 10^{-7}$ m. These gamma-ray photons are sufficiently energetic to pass through optical lenses and mirrors without effect (unless incident at very shallow ("grazing") angles of incidence, where they may be deviated by a small change in angle). Since the usual mechanisms for redirecting light do not work for gamma-ray photons, a different physical interaction must be used between energy and matter, such as absorption by dense materials. The "pinhole gamma camera" shown in Figure 1.8 is an example of such a system; it consists of an absorbing "plate" (typically made of lead) with a small hole of diameter d. The plate "selects" those photons from the source that travel along very specific paths through the hole to the sensor.

As a first example, consider a planar object that emits f photons with wavelength λ_γ per unit time from the location $[x, y]$ and that is located at a distance z_1 from the absorbing plate. The photons emitted by the object that pass through the hole continue to propagate to a planar photosensitive detector located at a distance z_2 from the absorber. The image is created by counting the photons with wavelength λ_γ over some exposure time to produce an "image" $g[x, y]$; the system equation has the form:

$$\mathcal{O}\{f[x, y, \lambda, t, \theta, \phi]\} = g[x, y] \tag{1.5}$$

The diameter of the pinhole and the distances z_1 and z_2 are parameters of the system operator $\mathcal{O}$. The goal of this modality of diagnostic imaging is to derive the spatial form of the planar object function f from the planar image g to solve the inverse problem.

The "quality" of the image is determined by whether the measured number of photons at each position of the detector is proportional to the number of photons emitted by the corresponding position of the object. Two parameters directly affect this measurement: the diameter d of the pinhole and the

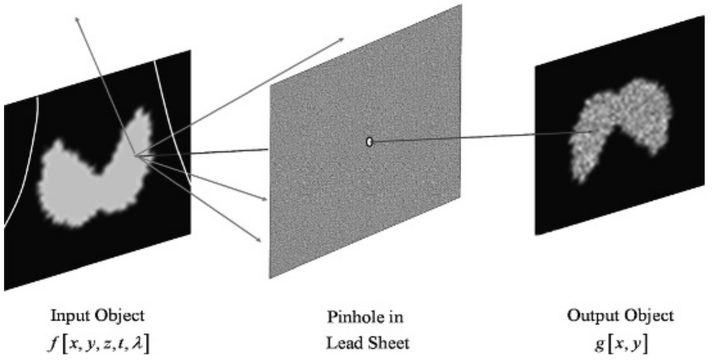

Figure 1.8 Schematic of gamma-ray pinhole imaging. The planar object $f[x, y, z, t, \lambda]$ emits energetic gamma-ray photons in all directions. An occasional emitted photon travels along a path that passes through the pinhole and is absorbed by the sensor, thus forming the image $g[x, y]$.

number of photons counted at each position in the detector, which is determined by the number of photons emitted by the object.

The action of the pinhole in the imaging system is to constrain the path of photons that reach the detector. If d is very small, the position of emission within the planar object may be determined very accurately and we might expect the resulting image to be a "faithful" replica of the original object. As d is increased, photons emitted from the same location in the object can expose different locations on the sensor and photons from different locations on the object expose the same location on the sensor. This ambiguity in location of sources degrades the image; it becomes "blurry", which means that the "spatial resolution" is somehow proportional to d^{-1}. It also is easy to see that the number of photons that reach the detector also is determined by the area of the pinhole; the number is small if $d \gtrapprox 0$ and larger if $d \gg 0$. For reasons not discussed here, the measured number of photons becomes more certain if more photons are counted. The resulting reduced variation in the number of counted photons means that the actual number of photons emitted by different points on the object may be more accurately estimated. In other words, the image becomes less "noisy" by counting more photons because the "brightness resolution" is improved.

The discussions of the last few paragraphs show that image quality in the pinhole gamma camera is determined by two countervailing requirements due to the hole diameter d; a smaller hole improves the spatial resolution but increases the statistical noise in the image. Simulations of images that illustrate these principles are shown in Figure 1.9.

Note also that the image is "inverted" by the gamma camera and that the relative sizes of the distances z_1 and z_2 determine the "magnification" of the image. If $z_1 > z_2$, then a particular area of the object is imaged onto a smaller area of the detector and the image is "minified" (smaller than the original). Obviously the image is "magnified" if $z_2 > z_1$.

At this point, we can practice being imaging scientists by using these observations to redesign the pinhole camera to compensate for its limitations. We might first want to decrease the statistical noise in the image by increasing the number of counted photons. One means to do this is to increase the dose of radioactivity to the patient so that more photons are emitted by the thyroid. Since this strategy creates additional problems for the patient, we seek other means. First, recognize that gamma-ray photons should be emitted from the technetium in any direction with equal likelihood. This allows us to record more of the emitted photons by adding more pinholes and more detectors. The resulting system creates the same "on-axis" image as before, but also additional "off-axis" images that contain similar information. The differences in geometry of the 3-D object and the sensors produce distortions in the images, but these may be corrected and the images combined to synthesize a single image formed from

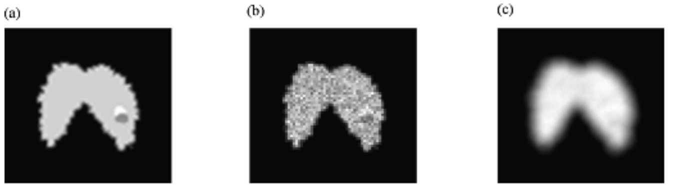

Figure 1.9 Simulation of the effect of pinhole size on spatial resolution. (a) Original object $f[x, y]$ is a simulated thyroid with a white "hot" spot (that emits more gamma rays) and an adjacent dark gray "cold" spot (the object has been rotated by 180° so that its orientation matches the images); (b) simulated image obtained using a small pinhole to produce good spatial resolution but with statistical noise due to the small number of counted photons; (c) simulated image with larger pinhole that reduces the statistical noise but "blurs" the image due to overlapping of different source points at the same location on the sensor.

a larger number of recorded photons, and thus with less statistical variation. For example, consider a system constructed from pinholes that are positioned sufficiently close together to produce overlapping images, meaning that photons from different locations on the object may expose the same location on the detector. Obviously, it is necessary to "unmix" the overlapping images to produce an image with improved quality. An example is shown in Figure 1.10. The technique for "unmixing" the overlapped images will be considered in the discussion of image filtering later in this book. This concept may be further extended by drilling more and more "pinholes" in the lead absorber. The pinholes may even merge together to form regions of "open space". In the example shown in Figure 1.11, 50% of the lead has been removed, thus transmitting many more photons to the detector. This pattern of detected photons is processed by a mathematical algorithm based on the pattern of pinhole apertures to "reconstruct" an approximation of the original object with less statistical variation. The process of collection and reconstruction is called *coded aperture imaging*.

1.4.2 Radiography

The next example of a medical imaging system is conventional radiography, which is what most people mean when they say "My dentist took an X-ray today". The X-radiation consists of photons with energies of the same order as gamma rays, but that are emitted from a source distinct from the object that may be characterized by its linear dimension d (area $\propto d^2$). The photons are selectively removed from the beam by structures within the 3-D object that absorb and/or scatter the incident radiation. The mathematical description of the 3-D object is a function f that measures its "ability" to remove X-ray photons from the beam, which may be called the X-ray *attenuation coefficient*. Also note that the dependence on exposure time t is ignored in the usual case of an object that does not change over the time scale of the image measurement. Just as was the case for gamma-ray imaging, the "spatial resolution" of the image is determined by the certainty that a recorded photon traveled a specific path from the source through the object, and thus the number reflects a physical feature of the object. The limiting factor in radiography is the diameter d of the X-ray source, which means that the uncertainty in the path traveled by a photon from a small source is small, while that uncertainty in the path is larger if d is larger.

In our simple example, we may consider that the photons that pass through the object are imaged using the same kind of sensor as in gamma-ray imaging. The form of the imaging equation is:

$$\mathcal{O}\{f[x, y, z]\} = g[x, y] \tag{1.6}$$

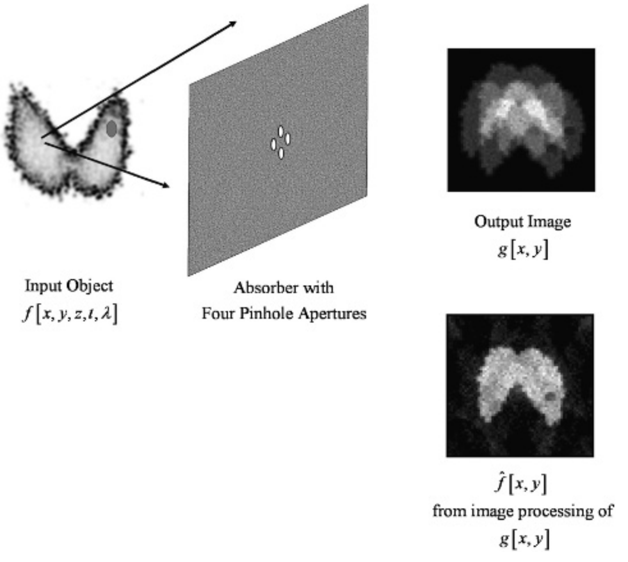

Figure 1.10 Simulation of output image $g[x, y]$ obtained of the object $f[x, y]$ through four pinholes. The overlapping gamma-ray images are then digitally processed in a computer to produce the estimate $\hat{f}[x, y]$ that more closely resembles $f[x, y]$.

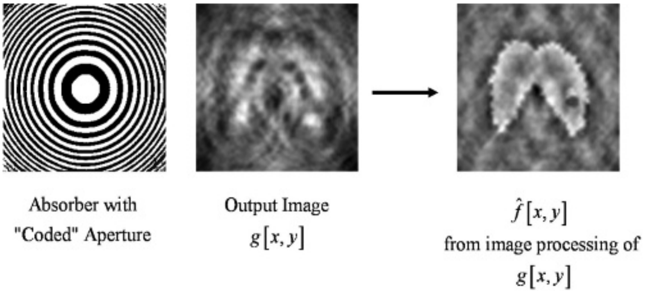

Figure 1.11 Simulated imaging through a continuum of pinholes that forms the "coded aperture": (a) aperture; (b) image of photons generated by the simulated thyroid through this aperture; (c) output image after processing based on knowledge to merge the continuum of raw images.

where the dependence on wavelength λ and time t has been deleted for clarity. Again, the parameters of the imaging system, including the source diameter d and the distances z_1 and z_2, are implicit in the system operator $\mathcal{O}\{\ \}$. The task is to determine the 3-D input function $f[x, y, z]$ from the 2-D output $g[x, y]$. The signal recorded by the sensor at a particular location is due to the integrated attenuation of the X-ray beam along the associated path through the object. This means that information about the third spatial dimension (the position in "depth") is lost by the recording process. For example, a dark spot on an X-ray image indicates only that the total attenuation of X-rays along that particular path is large, but it does not, by itself, tell us how the X-ray absorption was distributed along the path; it may have been concentrated in single region, in more than one location, or uniformly distributed along the path. This may be conveniently demonstrated on a 2-D function with "width" along the x-axis and

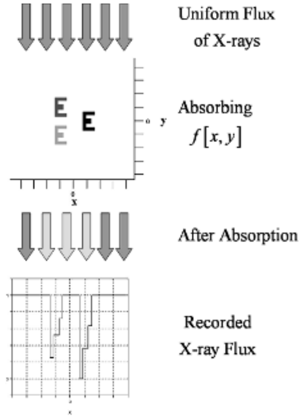

Figure 1.12 Simulation of imaging by a radiographic system. The object is a 2-D function $f[x, y]$ that describes the X-ray attenuation of a body. In this example, "white" represents a region that is completely transparent to X-rays and black represents a perfect absorber. The 1-D image $g[x]$ at the bottom is the line integral of the X-ray absorption by the object. The "depth" information of the structure is lost.

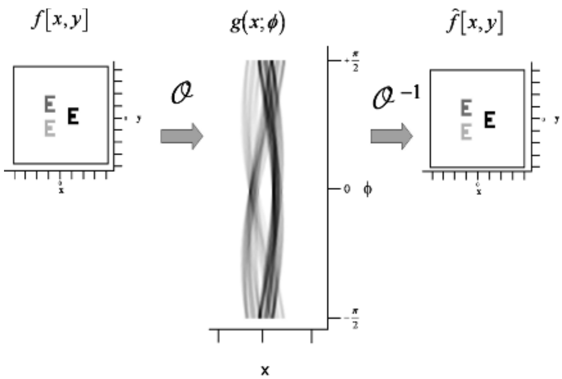

Figure 1.13 Simulation of X-ray CT system. The X-ray transmission of the object is measured at each of a set of angles ϕ to compute the 2-D image $g(x;\phi)$. An estimate $\widehat{f}[x, y]$ of the original object is computed from $g(x;\phi)$.

"depth" along z, i.e., $f[x, z]$. The X-ray absorption integrated along the z-direction yields a 1-D image $g[x]$, as shown in Figure 1.12.

This fundamental limitation of the radiographic system gives us a second opportunity to act as imaging scientists by modifying the system to recover the information about the absorber distribution in "depth" within the object. The machine that first accomplished this goal, the CT scanner, won a Nobel prize for its inventors.

1.4.3 Computed Tomographic Radiography

The modification to the radiographic system to retrieve depth information is somewhat similar to that used in gamma-ray imaging in the sense that more than one image of the same object is collected and processed. Because the source is distinct from the object in radiography, the images are not obtained simultaneously, but rather the source location must be changed between images. Consider the situation if two radiographs of the same object are made so that the X-rays pass through the object at two angles, say ϕ_1 and ϕ_2, relative to some coordinate system. In other words, we measure the X-ray attenuation of the object at two different locations about the object, as shown in Figure 1.13. The system operator now has the form:

$$\mathcal{O}\{f[x, y, z]\} = g[x, y, \phi_n] \tag{1.7}$$

In short, we have constructed a system that generates a third *angular* dimension of image data, though in this case the third dimension is sampled at only two locations and differs from the desired spatial coordinate z. Again, we need to solve the inverse problem by evaluating $f[x, y, z]$. As the next step, we can "fill in" the spaces between the azimuthal samples by gathering data at more angles ϕ. Finally, it is necessary to determine the appropriate mathematical operation that will "reconstruct" the X-ray attenuation at each location in the original object $f[x, y, z]$ from the measurements $g(x, y, \phi)$. The fact that we now have three dimensions of data should indicate that this gives us a "fighting chance" to solve the problem, but the mathematical derivation requires tools yet to be developed. We will derive one means to solve the inverse problem for this system in Chapter 12.