CHAPTER 1
MTF IN OPTICAL SYSTEMS

Linear-systems theory provides a powerful set of tools with which we can analyze optical and electro-optical systems. The spatial impulse response of the system is Fourier transformed to yield the spatial-frequency optical transfer function. Simply expressing the notion of image quality in the frequency domain does not by itself generate any new information. However, the conceptual change in viewpoint – instead of a spot size, we now consider a frequency response – facilitates additional insight into the behavior of an imaging system, particularly in the common situation where several subsystems are combined. We can multiply the individual transfer function of each subsystem to give the overall transfer function. This procedure is easier than the repeated convolutions that would be required for a spatial-domain analysis, and allows immediate visualization of the performance limitations of the aggregate system in terms of the performance of each of the subsystems. We can see where the limitations of performance arise and which crucial components must be improved to yield better overall image quality.

In Chapter 1, we develop this concept and apply it to classical optical systems, that is, imaging systems alone without detectors or electronics. We will first define terms and then discuss image-quality issues.

1.1 Impulse response

The impulse response $h(x,y)$ is the smallest image detail that an optical system can form. It is the blur spot in the image plane when a point source is the object of an imaging system. The finite width of the impulse response is a result of the combination of diffraction and aberration effects. We interpret $h(x,y)$ as an irradiance (W/cm^2) distribution as a function of position. Modeling the imaging process as a convolution operation (denoted by *), we express the image irradiance distribution $g(x,y)$ as the ideal image $f(x,y)$ convolved with the impulse response $h(x,y)$:

$$g(x,y) = f(x,y) * h(x,y) \quad . \tag{1.1}$$

The ideal image $f(x,y)$ is the irradiance distribution that would exist in the image plane (taking into account the system magnification) if the system had perfect image quality, in other words, a delta-function impulse response. The ideal image is thus a magnified version of the input-object irradiance, with all detail preserved. For conceptual discussions, we typically assume that the imaging system has unit magnification, so that we can directly take $f(x,y)$ as the object irradiance distribution, albeit as a function of image-plane coordinates. We can see from Eq. (1.1) that if $h(x,y) = \delta(x,y)$, the image is a perfect replica of the object. It is within this context that $h(x,y)$ is also known as the point-spread function (PSF). A perfect optical system is capable of forming a point image of a point object. However, because of the blurring effects of diffraction and aberrations, a real imaging system has an impulse response that is not a point. For any real system $h(x,y)$ has finite spatial extent. The narrower the PSF, the less blurring occurs in the image-forming process. A more compact impulse response indicates better image quality.

As Fig. 1.1 illustrates, we represent mathematically a point object as a delta function at location (x',y') in object-plane coordinates

$$f(x_{obj},y_{obj}) = \delta(x' - x_{obj}, y' - y_{obj}) . \tag{1.2}$$

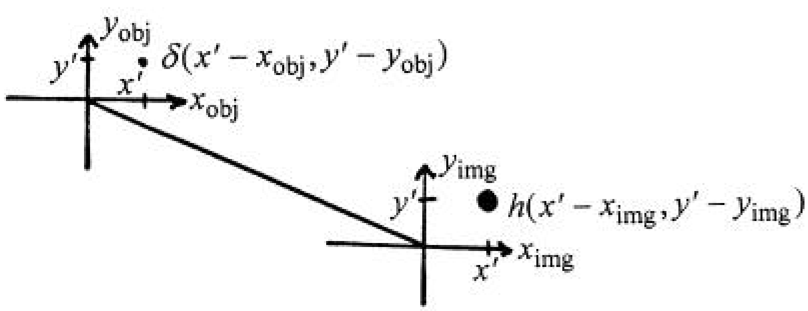

Figure 1.1 A delta function in the object is mapped to a blur function, the impulse response, in the image plane.

Assuming that the system has unit magnification, the ideal image is a delta function located at (x', y') in image-plane coordinates

$$g(x_{obj},y_{obj}) = \delta(x' - x_{img}, y' - y_{img}) . \tag{1.3}$$

In a real imaging system, instead of a delta function at the ideal image point, the impulse response is centered at $x' = x_{img}$ and $y' = y_{img}$ in the image plane $g(x_{img},y_{img}) = h(x' - x_{img}, y' - y_{img})$, in response to the delta-function object of Eq. (1.2). We represent a continuous function $f(x_{obj},y_{obj})$ of object coordinates, by breaking the continuous object into a set of point sources at specific locations, each with a strength proportional to the object brightness at that particular location. Any given point source has a weighting factor $f(x', y')$, which we find using the sifting property of the delta function:

$$f(x', y') = \iint d(x' - x_{\text{obj}}, y' - y_{\text{obj}})\, f(x_{\text{obj}}, y_{\text{obj}})\, dx_{\text{obj}}\, dy_{\text{obj}} \quad . \tag{1.4}$$

The image of each discrete point source will be the impulse response of Eq. (1.1) at the conjugate image-plane location, weighted by corresponding object brightness. The image irradiance function $g(x_{\text{img}}, y_{\text{img}})$ becomes the summation of weighted impulse responses. This summation can be written as a convolution of the ideal image function $f(x_{\text{img}}, y_{\text{img}})$ with the impulse response

$$g(x_{\text{img}}, y_{\text{img}}) = \iint h(x' - x_{\text{img}}, y' - y_{\text{img}})\, f(x_{\text{img}}, y_{\text{img}})\, dx'\, dy' \,, \tag{1.5}$$

which is equivalent to Eq. (1.1). Figure 1.2 illustrates the imaging process using two methods: the clockwise loop demonstrates the weighted superposition of the impulse responses and the counterclockwise loop demonstrates a convolution with the impulse response. Both methods are equivalent.

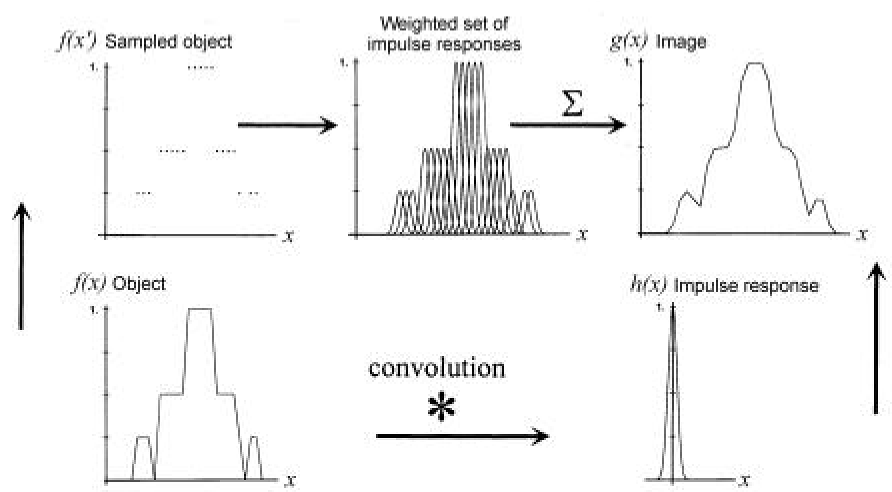

Figure 1.2 Image formation can be modeled as a convolutional process. The clockwise loop is a weighted superposition of impulse responses and the counterclockwise loop is a convolution with the impulse response.

Representing image formation as a convolutional process assumes linearity and shift invariance (LSI). To model imaging as a convolutional process, we must have a unique impulse response that is valid for any position or brightness of the point-source object. Linearity is necessary for us to be able to superimpose the individual impulse responses in the image plane into the final image. Linearity requirements are typically accurately satisfied for the irradiance distribution itself (the so-called aerial image). However, certain detectors such as photographic film, detector arrays (especially in the IR), and xerographic media are particularly nonlinear in their impulse response. In these cases, the impulse response is a function of the input irradiance level. We can only perform LSI analysis for a restricted range of input irradiances. Another linearity consideration is that coherent optical systems (optical processors) are linear in electric field (V/cm), while incoherent systems (imaging systems) are linear in irradiance (W/cm^2). We will deal exclusively with incoherent imaging systems. Note that partially coherent systems are not linear in either electric field or irradiance and their analysis—as a convolutional system—is more complicated, requiring definition of the mutual coherence function.[1]

Shift invariance is the other requirement for a convolutional analysis. According to the laws of shift invariance, a single impulse response can be defined that is not a function of image-plane position. Shift invariance assumes that the functional form of $h(x,y)$ does not change over the image plane. This shift invariance allows us to write the impulse response as $h(x' - x_{\text{img}}, y' - y_{\text{img}})$, a function of distance from the ideal image point, rather than as a function of image-plane position in general. Aberrations violate the assumption of shift invariance because typically the impulse response is a function of field angle. To preserve a convolutional analysis in this case, we segment the image plane into isoplanatic regions over which the functional form of the impulse response does not change appreciably.

1.2 Spatial frequency

We can also consider the imaging process from a frequency-domain (modulation-transfer-function) viewpoint, as an alternative to the spatial-domain (impulse-response) viewpoint. An object- or image-plane irradiance distribution is composed of "spatial frequencies" in the same way that a time-domain electrical signal is composed of various frequencies: by means of a Fourier analysis. As seen in Fig. 1.3, a given profile across an irradiance distribution (object or image) is composed of constituent spatial frequencies. By taking a one-dimensional profile across a two-dimensional irradiance distribution, we obtain an irradiance-*vs*-position waveform, which can be Fourier decomposed in exactly the same manner as if the waveform was in the more familiar form of volts *vs* time. A Fourier decomposition answers the question of what frequencies are contained in the waveform in terms of spatial frequencies with units of cycles (cy) per unit distance, analogous to temporal frequencies in cy/s for a time-domain waveform. Typically for optical systems, the spatial frequency is in cy/mm.

An example of one basis function for the one-dimensional waveform of Fig. 1.3 is shown in Fig. 1.4. The spatial period X (crest-to-crest repetition distance) of the waveform can be inverted to find the *x*-domain spatial frequency denoted by $\xi \equiv 1/X$.

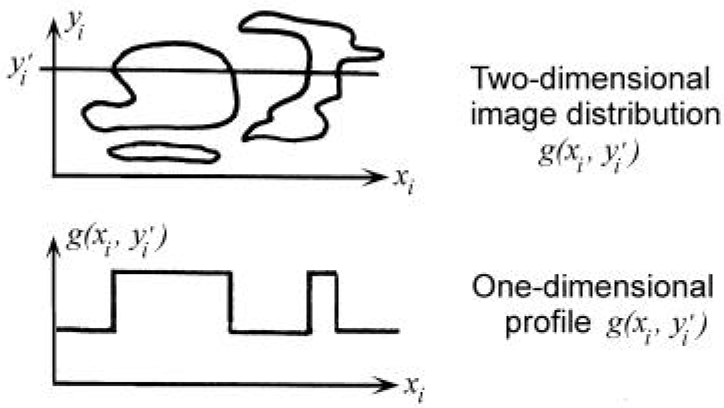

Figure 1.3 Definition of a spatial-domain irradiance waveform.

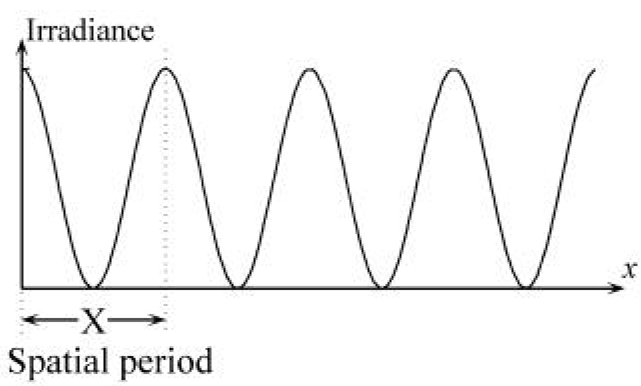

Figure 1.4 One -dimensional spatial frequency.

Fourier analysis of optical systems is more general than that of time-domain systems because objects and images are inherently two-dimensional, and thus the basis set of component sinusoids is also two-dimensional. Figure 1.5 illustrates a two-dimensional sinusoid of irradiance. The sinusoid has a spatial period along both the *x* and *y* directions, *X* and *Y* respectively. If we invert these spatial periods we find the two spatial frequency components that describe this waveform: $\xi = 1/X$ and $\eta = 1/Y$. Two pieces of information are required for specification of the two-dimensional spatial frequency. An alternate representation is possible using polar coordinates, the minimum crest-to-crest distance, and the orientation of the minimum crest-to-crest distance with respect to the *x* and *y* axes.

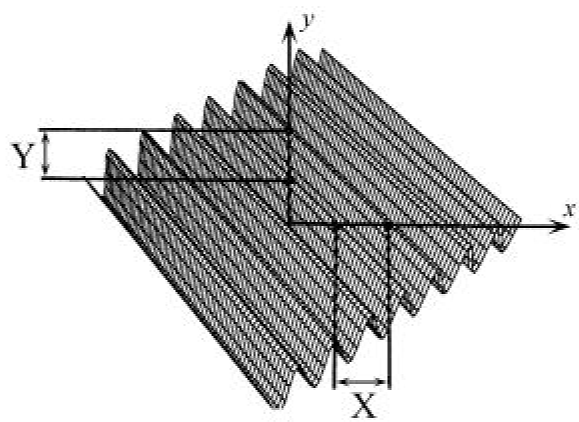

Figure 1.5 Two-dimensional spatial frequency.

Angular spatial frequency is typically encountered in the specification of imaging systems designed to observe a target at a long distance. If the target is far enough away to be in focus for all distances of interest, then it is convenient to specify system performance in angular units, that is, without having to specify a particular range distance. Angular spatial frequency is most often specified in cy/mrad. It can initially be a troublesome concept because both cycles and milliradians are dimensionless quantities but, with reference to Fig. 1.6, we find that the angular spatial frequency ξ_{ang} is simply the range *R* multiplied by the target spatial frequency ξ. For a periodic target of spatial period *X*, we define an angular period $\theta \equiv X/R$, an angle over which the object waveform repeats itself. The angular period is in radians if *X* and *R* are in the same units. Inverting this angular period gives angular spatial frequency $\xi_{ang} = R/X$. Given the resolution of optical systems, often *X* is in meters and *R* is in kilometers, for which the ratio *R*/*X* is then in cy/mrad.

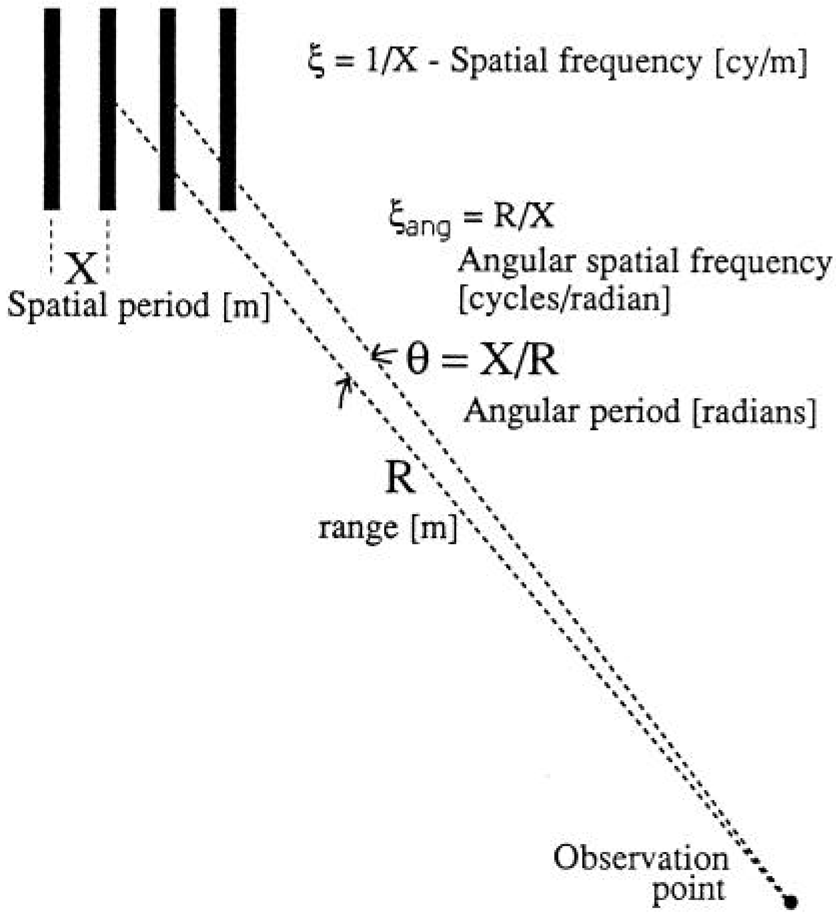

Figure 1.6 Angular spatial frequency.

1.3 Transfer function

Equation (1.1) describes the loss of detail inherent in the imaging process as the convolution of the ideal image function with the impulse response. The convolution theorem[2] states that a convolution in the spatial domain is a multiplication in the frequency domain. Taking the Fourier transform (denoted $\mathcal{F}$) of both sides of Eq. (1.1) yields

$$\mathcal{F}[g(x,y)] = \mathcal{F}[f(x,y) * h(x,y)] \tag{1.6}$$

and

$$G(\xi,\eta) = F(\xi,\eta) \times H(\xi,\eta) \quad , \tag{1.7}$$

where uppercase functions denote the Fourier transforms of the corresponding lowercase functions: F denotes the object spectrum, G denotes the image spectrum, and H denotes the spectrum of the impulse response. $H(\xi,\eta)$ is the transfer function, in that it relates the object and image spectra multiplicatively. The Fourier transform changes the irradiance waveform from a spatial-position function to the spatial-frequency domain, but generates no new information. The appeal of the frequency-domain viewpoint is that the multiplication of Eq. (1.7) is easier to perform and visualize than the convolution of Eq. (1.1). This convenience is most apparent in the analysis of imaging systems consisting of several subsystems, each with its own impulse response. As Eq. (1.8) demonstrates, each subsystem has its own transfer function as the Fourier transform of its impulse response.

The final result of all the subsystems operating on the input object distribution is a multiplication of their respective transfer functions. Figure 1.7 illustrates that we can analyze a combination of several subsystems by the multiplication of transfer functions of Eq. (1.9) rather than the convolution of impulse responses of Eq. (1.8):

$$f(x,y) * h_1(x,y) * h_2(x,y) * \ldots * h_n(x,y) = g(x,y) \qquad (1.8)$$

and

$$F(\xi,\eta) \times H_1(\xi,\eta) \times H_2(\xi,\eta) \times \ldots \times H_n(\xi,\eta) = G(\xi,\eta) \ . \qquad (1.9)$$

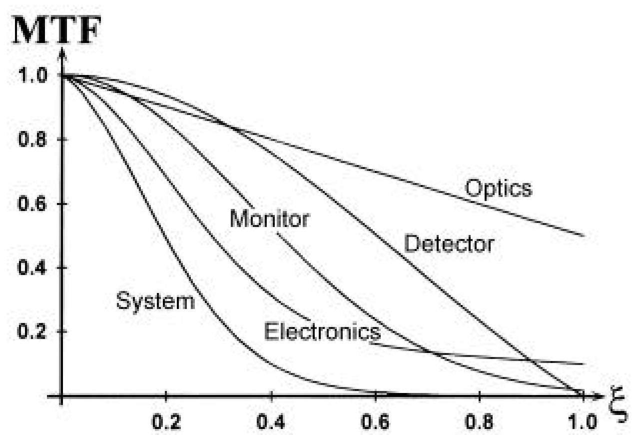

Figure 1.7 The aggregate transfer function of several subsystems is a multiplication of their transfer functions.

For the classical optical systems under discussion in this first chapter, we ignore the effects of noise and we typically assume that $H(\xi,\eta)$ has been normalized to have unit value at zero spatial frequency (a uniform image-irradiance distribution). This normalization yields a relative transmittance for the various frequencies and ignores attenuation factors that are independent of spatial frequency, such as Fresnel reflections or material absorption. Although this normalization is common, when we use it, we lose information about the absolute signal levels. For some cases we may want to keep the signal-level information, particularly when electronics noise is a significant factor.

With this normalization, $H(\xi,\eta)$ is referred to as the optical transfer function (OTF). Unless the impulse response function $h(x,y)$ satisfies certain symmetry conditions, its Fourier transform $H(\xi,\eta)$ is in general a complex function, having both a magnitude and a phase portion, referred to as the modulation transfer function (MTF) and the phase transfer function (PTF) respectively:

$$\text{OTF} \equiv H(\xi,\eta) = \left| H(\xi,\eta) \right| \exp[-j\theta(\xi,\eta)] \tag{1.10}$$

and

$$\text{MTF} \equiv \left| H(\xi,\eta) \right| \qquad \text{PTF} \equiv \theta(\xi,\eta)\ . \tag{1.11}$$

1.3.1 Modulation transfer function

The modulation transfer function is the magnitude response of the optical system to sinusoids of different spatial frequencies. When we analyze an optical system in the frequency domain, we consider the imaging of sinewave inputs (Fig. 1.8) rather than point objects.

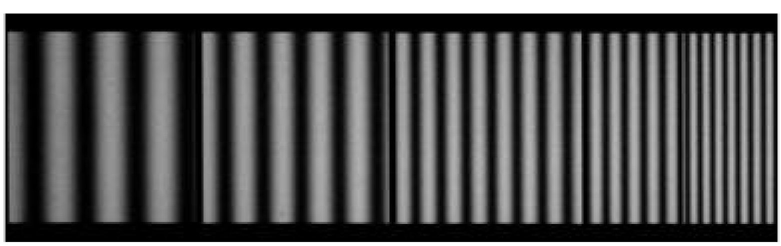

Figure 1.8 Sinewave target of various spatial frequencies.

A linear shift-invariant optical system images a sinusoid as another sinusoid. The limited spatial resolution of the optical system results in a decrease in the modulation depth M of the image relative to what is was in the object distribution (Fig. 1.9). Modulation depth is defined as the amplitude of the irradiance variation divided by the bias level:

$$M = \frac{A_{\max} - A_{\min}}{A_{\max} + A_{\min}} = \frac{2\times ac\text{ component}}{2\times dc\text{ component}} = \frac{ac}{dc} \quad . \tag{1.12}$$

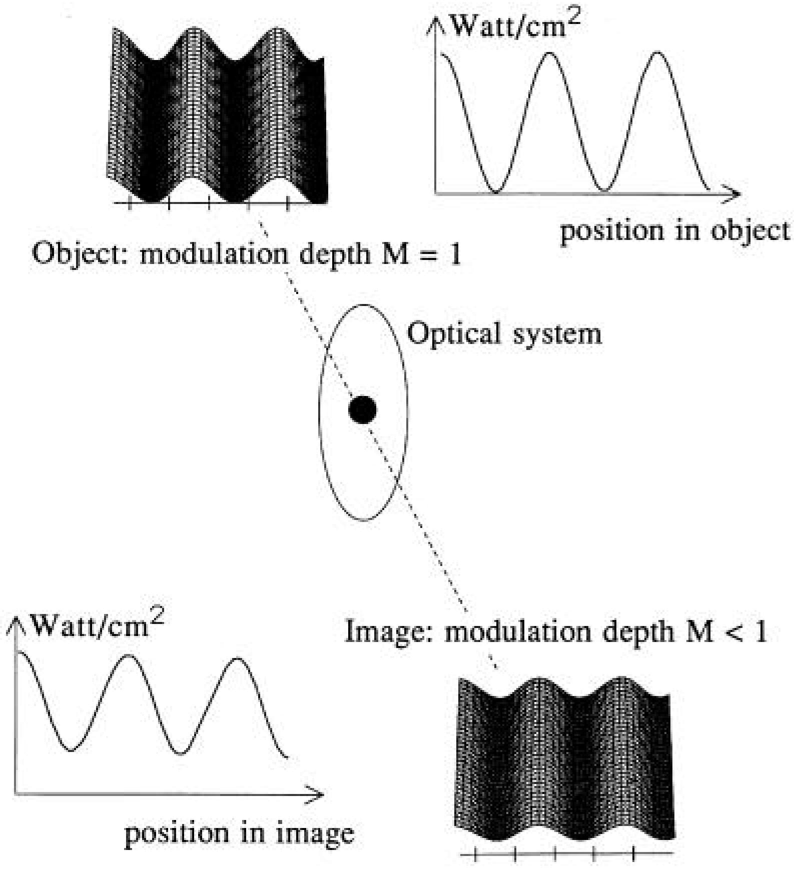

Figure 1.9 Modulation depth decreases going from object to image.