Resolvability of Objects from the
Standpoint of Statistical Parameter Estimation

Carl W. Helstrom*

Department of Applied Physics and Information Science

University of California, San Diego

La Jolla, California, 92037

NGR- 05-009-079

Abstract

The values of the radiance at points of an incoherently radiating object are considered as parameters of the statistical description of the field at the aperture of an observing optical instrument. By means of the Cramér-Rao inequality a lower bound is set to the mean-square errors of unbiased estimates of the radiance values. The errors are shown to increase rapidly when the object is sampled at points separated by less than a conventional resolution interval.

The ability of an optical instrument to resolve details of an object can be evaluated in various ways. All of them lead to the general conclusion that if the instrument takes in light of wavelength λ from the object through an aperture of diameter a, it will blur details having an angular subtense smaller than λ/a--a conclusion that might be drawn from the undular nature of light simply by dimensional analysis.

A common explanation of how the optical instrument obliterates fine details views it as a linear spatial filter transforming the light field at the object plane into the light field at the image plane. The aperture limits the spatial bandpass of this filter, and the loss of high spatial frequencies prevents the reconstruction of features of inversely proportional dimensions in the object. This viewpoint has led to the proposal that the finite size of the object and the resultant analyticity of its spatial Fourier transform should permit reconstruction of the entire object by mathematical operations on those spatial frequencies that do pass the aperture.[1] The inevitable presence of random background light and the unavoidable introduction of random noise in recording the image light subject any such procedure to deleterious errors.[2] The influence of noise on other linear estimation schemes has been analyzed by Rushforth and Harris.[3]

This approach to resolution through linear filtering is most suitable for coherent object fields. Ordinary objects, however, radiate or reflect incoherent light. Except in microscopy, where coherent illumination may be utilized for the sake of certain phase effects, the light emanating from most objects possesses a very low degree of spatial coherence. With such incoherently illuminated or radiating objects it is not the field of the light that is of interest, for

that field is best described as a random process having zero mean value and a most erratic spatio-temporal variation. Rather it is the mean-square value of the field, averaged over many cycles of the dominant temporal frequency Ω, that characterizes the object in the most informative way. Specifically, the mutual coherence function of the light field $\psi_0(u, t)$ immediately in front of the object has the form

$$\frac{1}{2} \underset{\sim}{E} \psi_0(\underset{\sim}{u}_1, t_1) \psi_0^*(\underset{\sim}{u}_2, t_2) =$$

$$= CB(\underset{\sim}{u}_1) \, \delta(\underset{\sim}{u}_1 - \underset{\sim}{u}_2) \, \chi(t_1 - t_2) \, \exp[-i\Omega(t_1 - t_2)],$$

where $B(\underset{\sim}{u})$ is the radiance of the object at point $\underset{\sim}{u}$, $\chi(\tau)$ is the temporal auto-covariance of the field, C is a suitable constant, and $\underset{\sim}{E}$ denotes the statistical expectation. The presence of the two-dimensional delta-function $\delta(\underset{\sim}{u}_1 - \underset{\sim}{u}_2)$ signifies that the coherence length of the light is much smaller than the extent of any details of interest. It is the radiance function $B(\underset{\sim}{u})$ that describes the object for us.

As the light propagates toward the aperture of our optical instrument, its coherence function changes in a predictable fashion,[4] but its field remains a stochastic process, to which is usually added another random field, referred to as the background. This background field has a distribution in frequency and angle that is generally much broader than that of the object light. The values of the radiance $B(\underset{\sim}{u})$ at various points of the object, which are the quantities we really want to know, are related to the net field at the aperture not in a deterministic fashion, but only in a statistical sense. They are *parameters* of the joint probability density functions of the aperture field. The function of the optical instrument is to estimate them by some operation on that field, and the estimates will be subject to error because of the stochastic nature of the light from the background and from the object itself.

In a previous paper we pointed out how the resolvability of details in the object plane might be treated as a problem in decision theory.[5] The optical instrument is required to decide whether two close objects of a specific kind are present, or only one. The probability of its deciding correctly, as a function of the separation of the test objects, measures the resolvability of details having the same size and form.

A subsequent paper introduced a modal expansion of the aperture field arising naturally in an analysis of the detectability of incoherently radiating objects.[6] The strengths of the several modes are directly related to the radiance of the object plane at points separated by a conventional resolution interval $\lambda R/a$, R being the distance of the object. The minimum mean-square errors in unbiased estimates of these radiance values were derived from a quantum-statistical description of the field.

Here we shall develop the statistical theory of resolvability further by viewing the function of an optical instrument as one of estimating the radiance of the object plane. The radiance $B(\underset{\sim}{u})$ is sampled in a suitable manner, and the samples are regarded as parameters of the joint pdf's of the aperture field. By means of the Cramér-Rao inequality lower bounds are set to the mean-square errors of unbiased estimates of those parameters. We shall demonstrate how these minimum mean-square errors soar when the radiance is sampled at points closer together than the conventional resolution interval $\lambda R/a$.

4

I. Sampling the Object Plane

It is the radiance $B(\underset{\sim}{u})$ of the object plane that is to be estimated as a function of position $\underset{\sim}{u} = (u_x, u_y)$. Since it is impossible to estimate $B(\underset{\sim}{u})$ at all points of the plane, the plane must be sampled, and there are several ways by which this can be done.

The most definitive methods employ a set of functions $F_m(\underset{\sim}{u})$ that are orthogonal over some part O of the object plane, or over the whole of it,

$$\int_0 F_m(\underset{\sim}{u}) \; F_n^*(\underset{\sim}{u}) \; d^2\underset{\sim}{u} = C_{\underset{\sim}{m}} \; \delta_{\underset{\sim}{mn}}, \tag{1.1}$$

where $C_{\underset{\sim}{m}}$ is a suitable normalization constant. The functions are distinguished by a two-vector index $\underset{\sim}{m} = (m_x, m_y)$. In terms of them the object radiance is written

$$B(\underset{\sim}{u}) = \sum_{\underset{\sim}{m}} B_{\underset{\sim}{m}} \; F_{\underset{\sim}{m}}(\underset{\sim}{u}), \tag{1.2}$$

and in general only a finite number of coefficients, or "samples", $B_{\underset{\sim}{m}}$ will be estimated. The sampling functions $F_m(\underset{\sim}{u})$ are taken dimensionless so that the samples $B_{\underset{\sim}{m}}$ have the dimensions of radiance.

The functions $F_m(\underset{\sim}{u})$ might conveniently be the indicator functions of contiguous rectangles $\Delta_x \times \Delta_y$ in the plane,

$$F_m(\underset{\sim}{u}) = 1, \quad \left\{ \begin{array}{l} (m_x - \frac{1}{2})\Delta_x < u_x < (m_x + \frac{1}{2})\Delta_x, \\[2mm] (m_y - \frac{1}{2})\Delta_y < u_y < (m_y + \frac{1}{2})\Delta_y \end{array} \right. \tag{1.3}$$

$$F_m(\underset{\sim}{u}) = 0, \quad \underset{\sim}{u} \text{ elsewhere.}$$

$B_{\underset{\sim}{m}}$ is then the average radiance over the rectangle centered at $(m_x \Delta_x, m_y \Delta_y)$.

Alternatively we might use the sampling functions

$$F_{\underset{\sim}{m}}(\underset{\sim}{u}) = \text{sinc}(u_x \Delta_x^{-1} - m_x) \, \text{sinc}(u_y \Delta_y^{-1} - m_y), \tag{1.4}$$

with sinc $z = (\sin \pi z)/\pi z$. If the radiance $B(\underset{\sim}{u})$ has a spatial Fourier transform lying entirely within a rectangle $\Delta_x^{-1} \times \Delta_y^{-1}$ in the spatial-frequency plane, the coefficients in Eq. (1.2) are samples of $B(\underset{\sim}{u})$ at the lattice points,

$$B_{\underset{\sim}{m}} = B(m_x \, \Delta_x, \, m_y \, \Delta_y), \tag{1.5}$$

by the two-dimensional version of the Whittaker-Shannon sampling theorem.[7]

Because the functions $F_{\underset{\sim}{m}}(\underset{\sim}{u})$ in Eqs. (1.3) and (1.4) are centered at points of a lattice, we refer to these forms of sampling as lattice sampling. For both, $C_{\underset{\sim}{m}} = \Delta_x \Delta_y = A_\Delta$. The region O is the entire object plane, although in practice only a finite number of samples $B_{\underset{\sim}{m}}$ will be estimated. The smaller Δ_x and Δ_y, the finer the details in the object plane that can be described by Eq. (1.2).

A third representation of the radiance of the object plane can be obtained from a Fourier series, with

$$F_{\underset{\sim}{m}}(\underset{\sim}{u}) = \exp[2\pi i(m_x u_x b_x^{-1} + m_y u_y b_y^{-1})],$$

$$-\frac{1}{2} b_x < u_x < \frac{1}{2} b_x, \quad -\frac{1}{2} b_y < u_y < \frac{1}{2} b_y, \tag{1.6}$$

The region O is now a rectangle $b_x \times b_y$ with area $A_o = b_x b_y$, and $C_{\underset{\sim}{m}} = A_o$. The greater the number of terms retained in Eq. (1.2), the finer the detail it can describe. We call this "Fourier sampling".

II. The Aperture Field

The object plane radiates incoherently, creating at the aperture of the optical system a field $\psi_s(\underset{\sim}{r}, t)$--assumed for simplicity to be a scalar--that is a circular-complex gaussian spatio-temporal random process.[8] The probability density functions of this process are completely determined by the mutual coherence function of the object field,

$$\varphi_s(\underset{\sim}{r}_1, t_1; \underset{\sim}{r}_2, t_2) = E[\psi_s(\underset{\sim}{r}_1, t_1) \psi_s^*(\underset{\sim}{r}_2, t_2)]. \qquad (2.1)$$

Also present is background light whose field $\psi_n(\underset{\sim}{r}, t)$ has the same statistical character, but is spatially and temporally white with spectral density N. On the basis of the total aperture field

$$\psi_+(\underset{\sim}{r}, t) = \psi_s(\underset{\sim}{r}, t) + \psi_n(\underset{\sim}{r}, t) \qquad (2.2)$$

observed during a finite interval $(0, T)$, the samples $B_{\underset{\sim}{m}}$ of the object radiance are to be estimated.

For convenience of discussion we assume that the object light is quasi-monochromatic and spectrally pure, so that its mutual coherence function can be factored into spatial and temporal parts,[4]

$$\varphi_s(\underset{\sim}{r}_1, t_1; \underset{\sim}{r}_2, t_2) = \varphi_s(\underset{\sim}{r}_1, \underset{\sim}{r}_2) \chi(t_1 - t_2)$$

$$\times \exp[-i\Omega(t_1 - t_2)], \qquad (2.3)$$

where $\Omega = 2\pi c/\lambda$ is the central angular frequency of the object light and λ is its wavelength.

The temporal autocovariance function $\chi(\tau)$ is normalized so that $\chi(0) = 1$. Its Fourier transform

$$X(\omega) = \int_{-\infty}^{\infty} \chi(\tau) e^{i\omega\tau} d\tau, \qquad (2.4)$$

7

which is positive and real, represents the spectral density of the object light, with angular frequencies ω referred to Ω as origin. The bandwidth W of the object light is conveniently defined by

$$W = \left[\int_{-\infty}^{\infty} X(\omega) \ d\omega/2\pi \right]^2 \bigg/ \int_{-\infty}^{\infty} [X(\omega)]^2 \ d\omega/2\pi$$

$$= |\chi(0)|^2 \bigg/ \int_{-\infty}^{\infty} |\chi(\tau)|^2 \ d\tau. \tag{2.5}$$

In optics the product WT is normally much greater than 1; indeed, it may be 10^5 or more.

The spatial autocovariance function $\varphi_s(r_1, r_2)$ is so normalized that $2\Omega^2 c \ \varphi_s(r, r)$ is the illuminance at point r of the aperture. The total energy E_s received from the object during the observation interval (0, T) is

$$E_s = 2\Omega^2 c \ T \int_A \varphi_s(r, r) \ d^2r, \tag{2.6}$$

where $\int_A$ indicates an integration over the aperture.

The object plane, we assume, is so far away that the light rays from the part of it being estimated are paraxial. The spatial autocovariance $\varphi_s(r_1, r_2)$ can be expressed in terms of the radiance $B(u)$ through the Fresnel-Kirchhoff approximation,[9]

$$\varphi_s(r_1, r_2) = (8\pi R^2 \Omega^2 c)^{-1} \int_0 B(u) \ \mathcal{E}(r_1, u) \mathcal{E}^*(r_2, u) \ d^2u,$$

$$\mathcal{E}(r, u) = \exp\left(\frac{ik}{2R}|r - u|^2\right), \tag{2.7}$$

where $k = \Omega/c = 2\pi/\lambda$ and R is the distance between object and aperture planes. Through Eq. (1.2) the spatial autocovariance function depends on the set $B = \{B_m\}$ of radiance samples,

8

$$\varphi_s(\underset{\sim}{r}_1,\ \underset{\sim}{r}_2;\ \underset{\sim}{B}) = (8\pi R^2\Omega^2 c)^{-1}\sum_{\underset{\sim}{m}} B_{\underset{\sim}{m}} \int_0^{\cdot} F_{\underset{\sim}{m}}(\underset{\sim}{u})\ \mathscr{E}(\underset{\sim}{r}_1,\ \underset{\sim}{u})\ \mathscr{E}^*(\underset{\sim}{r}_2,\ \underset{\sim}{u})\ d^2\underset{\sim}{u}.$$

$$(2.8)$$

In this way the samples $\underset{\sim}{B} = \{B_{\underset{\sim}{m}}\}$ are parameters of the joint probability density functions of values of the aperture field $\psi_+(\underset{\sim}{r},\ t)$ at various points $\underset{\sim}{r}$ and times t.

The foregoing description is based on classical physics and requires $N = K\mathcal{T} \gg \hbar\Omega$, where K = Boltzmann's constant, $\mathcal{T}$ is the effective absolute temperature of the background, and $\hbar$ = Planck's constant $h/2\pi$. When $K\mathcal{T} \ll \hbar\Omega$, the observations are said to be quantum-limited, a condition requiring an easy modification of our results.

III. Errors in Estimates of Radiance Samples

The samples $\underset{\sim}{B} = \{B_m\}$ specifying the radiance distribution of the object plane are to be estimated from observation of the field $\psi_+(\underset{\sim}{r}, t)$ at the aperture of an optical system. The system is to be designed to make the estimates as accurately as possible. How well it can be expected to perform can be assessed by the mean-square errors ε_m in the estimates $\hat{B}_m$ of the samples B_m,

$$\varepsilon_m = E(\hat{B}_m - B_m)^2. \tag{3.1}$$

By restricting ourselves to unbiased estimates,

$$E(\hat{B}_m) = B_m, \tag{3.2}$$

we can set a lower bound to ε_m by means of the Cramér-Rao inequality,[10,11]

$$\varepsilon_m = E(\hat{B}_m - B_m)^2 \geq L_{mm} \tag{3.3}$$

where the matrix $\underset{\sim}{L} = ||L_{mn}|| = \underset{\sim}{H}^{-1}$ is inverse to the matrix $\underset{\sim}{H}$, whose (mn)-element is

$$H_{mn} = - E\left[\frac{\partial^2}{\partial B_m \partial B_n} \ell n\, p(\underset{\sim}{\psi};\, \underset{\sim}{B})\right]. \tag{3.4}$$

Here $p(\underset{\sim}{\psi};\, \underset{\sim}{B})$ is the joint probability density function of samples $\underset{\sim}{\psi} = \{\psi_+(\underset{\sim}{r}_p,\, t_q)\}$ of the aperture field at points $\underset{\sim}{r}_p \in A$ and at times $t_q \in (0,\, T)$. After forming the expectation E, the right-hand side of Eq. (3.4) must be taken to the limit of an infinitely dense sampling of A and $(0,\, T)$.

The off-diagonal elements of the matrix $\underset{\sim}{L}$ are related to the covariances of unbiased estimates $\hat{B}_m$ at different points. Specifically, if $\underset{\sim}{X}$ is an arbitrary column vector of real elements and $\underset{\sim}{\tilde{X}}$ is its transposed row vector,

$$\underset{\sim}{\tilde{X}}\, \underset{\sim}{V}\, \underset{\sim}{X} \geq \underset{\sim}{\tilde{X}}\, \underset{\sim}{L}\, \underset{\sim}{X}, \tag{3.5}$$

where the elements of the matrix $\underset{\sim}{V} = ||V_{mn}||$ are the covariances

$$V_{mn} = E(\hat{B}_m - B_m)(\hat{B}_n - B_n). \tag{3.6}$$

Eq. (3.3) is a special case of Eq. (3.5).

The significance of the multidimensional Cramér-Rao inequality is best understood in terms of the concentration ellipsoid of the errors, a quadric surface whose equation is $\tilde{Y} \, V^{-1} Y = m + 2$ in an m-dimensional space with coordinates $\tilde{Y} = (y_1, y_2, \ldots, y_m)$, where m is the number of radiance samples.[12] Equivalent to Eq. (3.5) is the inequality

$$\tilde{Y} \, V^{-1} Y \le \tilde{Y} \, H \, Y, \tag{3.7}$$

which asserts that the concentration ellipsoid lies outside the ellipsoid whose equation is

$$\tilde{Y} \, H \, Y = m + 2. \tag{3.8}$$

When the errors are uncorrelated, the axes of the concentration ellipsoid are proportional to their r.m.s. values.

At large signal-to-noise ratio, the maximum-likelihood estimates of the parameters B_m have approximately a joint gaussian distribution, and the level surfaces of this distribution are ellipsoids parallel to the concentration ellipsoid. The Cramér-Rao inequality in this limit becomes asymptotically an equality.[10]

When as here the density functions $p(\psi; B)$ have the circular gaussian form, the elements of the matrix H are given by

$$H_{mn} = \frac{\partial}{\partial B_m^{(1)}} \frac{\partial}{\partial B_n^{(2)}} H(B^{(1)}, B^{(2)}) \Big|_{B^{(1)} = B^{(2)} = B}, \tag{3.9}$$

where the ambiguity function $H(B^{(1)}, B^{(2)})$ is

$$H(B^{(1)}, B^{(2)}) = (E_s/N)^2 (WT)^{-1} \int_A \int_A \varphi_s(r_1, r_2; B^{(1)}) \, \varphi_s(r_2, r_1; B^{(2)}) \, d^2r_1 d^2r_2$$

$$\times \; \left| \int_A \varphi_s(\underset{\sim}{r}, \; \underset{\sim}{r}; \; B) \; d^2\underset{\sim}{r} \right|^{-2}. \tag{3.10}$$

After the differentiations in Eq. (3.9), $\underset{\sim}{B}^{(1)}$ and $\underset{\sim}{B}^{(2)}$ are set equal to the true set $\underset{\sim}{B}$ of radiance samples.[13,14]

If we now substitute from Eqs. (2.7) and (2.8) into Eq. (3.10) and differentiate as in Eq. (3.9), we obtain

$$H_{\underset{\sim\sim}{mn}} = (E_s/N)^2 \; (WT)^{-1} \; B_T^{-2} \; J_{\underset{\sim\sim}{mn}} \tag{3.11}$$

where

$$B_T = \int_0 B(\underset{\sim}{u}) \; d^2\underset{\sim}{u}$$

is the total radiant power of the object plane and

$$J_{\underset{\sim\sim}{mn}} = \int_0 \!\! \int_0 F_m(\underset{\sim}{u}_1) \; F_n(\underset{\sim}{u}_2) \; |\mathscr{J}(\underset{\sim}{u}_1 - \underset{\sim}{u}_2)|^2 \; d^2\underset{\sim}{u}_1 d^2\underset{\sim}{u}_2 \tag{3.12}$$

with

$$\mathscr{J}(\underset{\sim}{u}) = A^{-1} \int_A I_A(\underset{\sim}{r}) \; \exp(-ik\underset{\sim}{r}\cdot\underset{\sim}{u}/R) \; d^2\underset{\sim}{r} \tag{3.13}$$

proportional to the Fourier transform of the indicator function $I_A(\underset{\sim}{r})$ of the aperture.[15] The matrix element $J_{\underset{\sim\sim}{mn}}$ can also be written

$$J_{\underset{\sim\sim}{mn}} = A^{-2} \int_A \!\! \int_A K_m(\underset{\sim}{r}_1 - \underset{\sim}{r}_2) \; K_n^*(\underset{\sim}{r}_1 - \underset{\sim}{r}_2) \; d^2\underset{\sim}{r}_1 d^2\underset{\sim}{r}_2 \tag{3.14}$$

where

$$K_m(\underset{\sim}{r}) = \int_0 F_m(\underset{\sim}{u}) \; \exp(ik\underset{\sim}{u}\cdot\underset{\sim}{r}/R) \; d^2\underset{\sim}{u} \tag{3.15}$$

is the Fourier transform of the sampling function $F_m(\underset{\sim}{u})$. In order to evaluate the minimum mean-square errors as in Eq. (3.3) it is necessary to invert the matrix $\underset{\sim}{J} = ||J_{\underset{\sim\sim}{mn}}||$.

Under quantum-limited conditions, $N = K\mathcal{T} \ll \hbar\Omega$, the factor $(E_s/N)^2 \, (WT)^{-1}$ in Eq. (3.11) must be replaced by $N_s M f_1(\mathcal{D})$, with[14]

$$f_1(\mathcal{D}) = \mathcal{D}W \int_{-\infty}^{\infty} [X(\omega)]^2 \, [1 + \mathcal{D}WX(\omega)]^{-1} \, d\omega/2\pi, \qquad (3.16)$$

where $X(\omega)$ is given by Eq. (2.4), $\mathcal{D} = N_s/\mathfrak{N}MWT$, $N_s = E_s/\hbar\Omega$ is the average total number of photons received from the object during the interval $(0, T)$, M is the number of spatial degrees of freedom in the object light at the aperture, and

$$\mathfrak{N} = [\exp(\hbar\Omega/K\mathcal{T}) - 1]^{-1}. \qquad (3.17)$$

The number M is given by

$$M = \left[\int_A \varphi_s(\underset{\sim}{r}, \underset{\sim}{r}; \underset{\sim}{B}) \, d^2\underset{\sim}{r} \right]^2$$

$$\times \left[\int_A\int_A |\varphi_s(\underset{\sim}{r}_1, \underset{\sim}{r}_2; \underset{\sim}{B})|^2 \, d^2\underset{\sim}{r}_1 d^2\underset{\sim}{r}_2 \right]^{-1} \qquad (3.18)$$

and is roughly equal to $AA_o/\lambda^2 R^2$, where A is the area of the aperture and A_o is the area of the part of the object plane whose radiance is being estimated. When $\mathcal{D} \gg 1$, an extreme quantum-limited condition, $f_1(\mathcal{D}) \doteq 1$ and the factor $(E_s/N)^2 \, (WT)^{-1}$ in Eq. (3.11) is replaced by $N_s M$.

In deriving the ambiguity function in Eq. (3.10), the threshold approximation was made.[5] The error thus introduced can be shown to be of the order of $E_s/NMWT$ under a background limitation and of the order of N_s/MWT under a quantum limitation,[14] provided $M \gg 1$. Under ordinary conditions $WT \gg 1$.

13

IV. Lattice Sampling

Under lattice sampling the functions $F_m(\underset{\sim}{u})$ are obtained by translation of the central function $F_0(\underset{\sim}{u})$,

$$F_m(\underset{\sim}{u}) = F_0(u_x + m_x \Delta_x, u_y + m_y \Delta_y). \tag{4.1}$$

The matrices $\underset{\sim}{H}$ and $\underset{\sim}{J}$ then have the Toeplitz form; that is, their elements depend only on the differences of their indices, $J_{\underset{\sim}{mn}} = J_{\underset{\sim}{m-n}}$, where

$$J_{\underset{\sim}{p}} = \iint_0 \int_0 F_0(\underset{\sim}{v}_1) F_0^*(\underset{\sim}{v}_2) \mid \mathscr{G}(\underset{\sim}{v}_1 - \underset{\sim}{v}_2 - \Delta_{\underset{\sim}{p}}) \mid^2 d^2\underset{\sim}{v}_1 d^2\underset{\sim}{v}_2$$

$$= A^{-2} \iint_A \int_A \mid K_0(\underset{\sim}{r}_1 - \underset{\sim}{r}_2) \mid^2 \exp[-ik\Delta_{\underset{\sim}{p}} \cdot (\underset{\sim}{r}_1 - \underset{\sim}{r}_2)/R] \, d^2\underset{\sim}{r}_1 d^2\underset{\sim}{r}_2,$$

$$\Delta_{\underset{\sim}{p}} = (p_x \Delta_x, p_y \Delta_y). \tag{4.2}$$

In practice the matrix $\underset{\sim}{J} = \mid\mid J_{\underset{\sim}{m-n}}\mid\mid$ will be finite, and there will be no simple analytical form for its inverse. Under certain conditions, however, an approximate formula for the elements of the inverse matrix $\underset{\sim}{G} = \underset{\sim}{J}^{-1}$ can be obtained by assuming that $\underset{\sim}{J}$ is infinite in extent. We denote this infinite extension of $\underset{\sim}{J}$ by $\underset{\sim}{J}^{\infty}$. Its inverse $\underset{\sim}{G}^{\infty}$ also has the Toeplitz form, and the elements of $\underset{\sim}{G}^{\infty}$ are solutions of the array of simultaneous equations

$$\sum_q G_{\underset{\sim}{p-q}}^{\infty} \, J_{\underset{\sim}{q-m}}^{\infty} = \delta_{\underset{\sim}{pm}}. \tag{4.3}$$

These equations can be solved by a Fourier transformation, provided the inverse of $\underset{\sim}{J}^{\infty}$ exists. The diagonal elements $G_{\underset{\sim}{m}}^{\infty}$ of $\underset{\sim}{G}^{\infty}$ will be good approximations to the diagonal elements of $\underset{\sim}{G}$ when $\underset{\sim}{G}$ has many elements, that is, when the diameter of the part O of the object plane being estimated is many times greater

than the sampling intervals Δ_x and Δ_y. When Δ_x and Δ_y are too small, as we shall see, the approximation breaks down because $\underline{J}^{\infty-1}$ no longer exists.

We define the discrete Fourier transforms

$$g(\underset{\sim}{\omega}) = \sum_{\underset{\sim}{p}} G_{\underset{\sim}{p}}^{\infty} \exp(i\underset{\sim}{p} \cdot \underset{\sim}{\omega}), \tag{4.4}$$

$$j(\underset{\sim}{\omega}) = \sum_{\underset{\sim}{p}} J_{\underset{\sim}{p}}^{\infty} \exp(i\underset{\sim}{p} \cdot \underset{\sim}{\omega}), \tag{4.5}$$

which have periods 2π in ω_x and ω_y. Then, as in the convolution theorem, Eq. (4.3) is equivalent to

$$g(\underset{\sim}{\omega}) \, j(\underset{\sim}{\omega}) = 1, \tag{4.6}$$

and the diagonal elements of G^{∞} are

$$G_0^{\infty} = \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} [j(\underset{\sim}{\omega})]^{-1} \, d^2\underset{\sim}{\omega}/(2\pi)^2. \tag{4.7}$$

The lower bound on the mean-square error $\varepsilon_{\underset{\sim}{m}}$ of an unbiased estimate of $B_{\underset{\sim}{m}}$ is then approximately

$$\varepsilon_{\underset{\sim}{m}} \geq (N/E_s)^2 \, WT \, B_T^2 \, G_0^{\infty}. \tag{4.8}$$

The factor G_0^{∞} in the right-hand side of this inequality is actually larger than the factor $G_{\underset{\sim}{mm}} = (J^{-1})_{\underset{\sim}{mm}}$ that should appear there, and Eq. (4.8) therefore does not constitute a true lower bound.

Substituting the second part of Eq. (4.2) into Eq. (4.5), we obtain

$$j(\underset{\sim}{\omega}) = A_\gamma \, A^{-2} \int_A \int_A \sum_{\underset{\sim}{m}} |K_0(\underset{\sim}{r}_1 - \underset{\sim}{r}_2)|^2$$

$$\times \, \delta(x_1 - x_2 + \gamma_x(m_x - \omega_x/2\pi)) \, \delta(y_1 - y_2 + \gamma_y(m_y - \omega_y/2\pi)) \, dx_1 dy_1 dx_2 dy_2,$$

$$\gamma_x = \lambda R/\Delta_x, \quad \gamma_y = \lambda R/\Delta_y, \quad A_\gamma = \gamma_x \, \gamma_y, \tag{4.9}$$

15

where m_x and m_y are integers, $\underset{\sim}{m} = (m_x, m_y)$, and $\underset{\sim}{r}_j = (x_j, y_j)$, $j = 1, 2$. Taking the aperture A to be rectangular with sides a_x and a_y, we find after integrating over it

$$j(\omega_x, \omega_y) = A_\gamma A^{-2} \sum_{\underset{\sim}{m}} [a_x - \gamma_x |m_x - \omega_x/2\pi|]$$

$$\times [a_y - \gamma_y |m_y - \omega_y/2\pi|] |K_0(\gamma_x(m_x - \omega_x/2\pi), \gamma_y(m_y - \omega_y/2\pi))|^2 \quad (4.10)$$

a term of the sum being set equal to zero whenever either of its square-bracketed factors is negative.

The inverse $[j(\underset{\sim}{\omega})]^{-1}$ ceases to exist when

$$\gamma_x = \lambda R/\Delta_x > 2a_x$$

or

$$\gamma_y = \lambda R/\Delta_y > 2a_y,$$

that is, when $\Delta_x < \lambda R/2a_x = \frac{1}{2} \delta_x$ or $\Delta_y \leqslant \lambda R/2a_y = \frac{1}{2} \delta_y$, for then $j(\underset{\sim}{\omega})$ vanishes over a finite area of the rectangle $-\pi \leqslant (\omega_x, \omega_y) < \pi$. The infinite form of $\underset{\sim}{J}$ no longer has an inverse, and the diagonal elements of $\underset{\sim}{G}$ cannot be approximated by Eq. (4.7).

16

Sinc-function Sampling

The analysis is simplest for sinc-function sampling as in Eq. (1.4). From Eq. (3.15),

$$K_0(\underset{\sim}{r}) = A_\Delta = \Delta_x \Delta_y, \quad -\frac{1}{2}\gamma_x < x < \frac{1}{2}\gamma_x, \quad -\frac{1}{2}\gamma_y < y < \frac{1}{2}\gamma_y, \qquad (4.11)$$

$$K_0(\underset{\sim}{r}) = 0 \text{ elsewhere.}$$

Putting this into Eq. (4.10) we find

$$j(\underset{\sim}{\omega}) = A_\delta A_\Delta [1 - \delta_x|\omega_x|/2\pi\Delta_x] [1 - \delta_y|\omega_y|/2\pi\Delta_y],$$

$$A_\delta = \delta_x \delta_y, \qquad (4.12)$$

the square-bracketed expressions again vanishing when their contents are negative. Substituting this into Eq. (4.7), we obtain

$$G_0^\infty = 4A_\delta^{-2} \ell n(1 - \frac{\delta_x}{2\Delta_x}) \ell n(1 - \frac{\delta_y}{2\Delta_y}),$$

$$\Delta_x > \frac{1}{2}\delta_x, \quad \Delta_y > \frac{1}{2}\delta_y. \qquad (4.13)$$

The relative mean-square error in an unbiased estimate of the radiance $B_{\underset{\sim}{m}}$ at point $\Delta_{\underset{\sim}{m}} = (m_x \Delta_x, m_y \Delta_y)$ is now, by Eq. (4.8),

$$\epsilon_{\underset{\sim}{m}}/B_{\underset{\sim}{m}}^2 \geq (N/E_\Delta)^2 WT(A_\Delta^2/A_\delta^2) \ell n(1 - \frac{\delta_x}{2\Delta_x}) \ell n(1 - \frac{\delta_y}{2\Delta_y})$$

$$= (N/E_\Delta)^2 WT (A_\Delta/A_\delta) G_0^{'\infty}, \qquad (4.14)$$

with

$$G_0^{'\infty} = 4(A_\Delta/A_\delta) \ell n(1 - \frac{\delta_x}{2\Delta_x}) \ell n(1 - \frac{\delta_y}{2\Delta_y}) \qquad (4.15)$$

In Eq. (4.14) E_Δ is the total energy that would be received from the area $A_\Delta = \Delta_x \Delta_y$ of the object if it radiated uniformly. The bound is approximately valid only when $\Delta_x > \frac{1}{2}\delta_x$ and $\Delta_y > \frac{1}{2}\delta_y$, where

$$\delta_x = \lambda R/a_x, \quad \delta_y = \lambda R/a_y \qquad (4.16)$$

are the conventional resolution elements in the x- and y-directions on the object plane.

When the sampling intervals Δ_x and Δ_y are much greater than the resolution elements δ_x and δ_y, respectively, the relative mean-square error in Eq. (4.14) becomes approximately

$$\epsilon_{\underset{\sim}{m}}/B_{\underset{\sim}{m}}^2 \gtrsim (N/E_s)^2 \text{ (MWT) } M_\Delta \tag{4.17}$$

where M_Δ is the number of sampling rectangles into which the object is divided, $E_s = M_\Delta E_\Delta$ is the total energy received from the object, and

$$M = A_o/A_\delta = M_\Delta A_\Delta/A_\delta \tag{4.18}$$

is the number of spatial degrees of freedom in the object field. Since the total radiant power B_T is proportional to the sum

$$\sum_{\underset{\sim}{m}} B_{\underset{\sim}{m}},$$

which contains M_Δ terms, we conclude that the relative mean-square error in an estimate of B_T is bounded by

$$\underset{\sim}{E}(\hat{B}_T - B_T)^2/B_T^2 \gtrsim (N/E_s)^2 \text{ MWT}, \tag{4.19}$$

in agreement with a previous result.[17]

With sinc-function sampling the matrix elements $J_{\underset{\sim}{p}}$ themselves can be evaluated in closed form from Eqs. (4.2) and (4.11). They are

$$J_{\underset{\sim}{p}} = A_\delta A_\Delta J_{\underset{\sim}{p}}' \tag{4.20}$$

with

$$J_{\underset{\sim}{p}}' = (A_\Delta/A_\delta) \text{ sinc}^2(p_x \Delta_x/\delta_x) \text{ sinc}^2(p_y \Delta_y/\delta_y),$$

$$\Delta_x \leq \tfrac{1}{2}\delta_x, \ \Delta_y \leq \tfrac{1}{2}\delta_y, \tag{4.21}$$

and

$$J_{\underset{\sim}{p}}' = (2\pi^2)^{-2} (A_\delta/A_\Delta) [1 - (-1)^{p_x}] [1 - (-1)^{p_y}]/p_x^2 p_y^2,$$

$$J_{\underset{\sim}{0}}' = \left(1 - \frac{\delta_x}{4\Delta_x}\right)\left(1 - \frac{\delta_y}{4\Delta_y}\right), \quad \Delta_x > \tfrac{1}{2}\delta_x, \ \Delta_y > \tfrac{1}{2}\delta_y. \tag{4.22}$$

One-Dimensional Object

The accuracy of our approximation to $\underset{\sim}{G}^{\infty}$ can be most easily assessed numerically for estimates of the radiance of a one-dimensional object. We consider the matrix $\underset{\sim}{J'} = ||J'_{m-n}||$ whose elements are the one-dimensional versions of those in Eqs. (4.21), (4.22), which were derived for sinc-function sampling. Here, after subscripts x and y are dropped,

$$J_p' = (2\pi^2)^{-1} (\delta/\Delta) [1 - (-1)^p]/p^2$$

$$J_0' = (1 - \delta/4\Delta), \quad \Delta > \tfrac{1}{2}\delta, \qquad (4.23)$$

$$J_p' = (\Delta/\delta) \, \text{sinc}^2(p\Delta/\delta), \quad \Delta < \tfrac{1}{2}\delta. \qquad (4.24)$$

Eleven- and fifteen-rowed matrices $||J'_{m-n}||$ were inverted by a digital computer. The central--and largest--element of the inverse matrix $\underset{\sim}{G'}$ is plotted in Fig. 1 as a function of the ratio δ/Δ; these are the curves marked "11" and "15". The curve marked "∞" displays the diagonal elements $G_0'^{\infty}$ of the infinite matrix $\underset{\sim}{J'}^{\infty}$; they are given by the one-dimensional form of Eq. (4.15),

$$G_0'^{\infty} = 2(\Delta/\delta)|\ln(1 - \delta/2\Delta)|, \quad \delta/\Delta < 2. \qquad (4.25)$$

The graph illustrates the extremely rapid increase in the mean-square error when an attempt is made to estimate the radiance at points closer than $\tfrac{1}{2}\delta = \lambda R/2a$, where a is the width of the aperture. The central element of $\underset{\sim}{J'}^{-1}$ soon reaches astronomical values, rising the faster, the larger the number of points at which the radiance samples are unknown. For $\Delta < \tfrac{1}{2}\delta$, on the other hand, the minimum mean-square error of the estimate is insensitive to the number of points involved and is close to the value calculated by assuming the number to be infinite.

The diagonal elements of $\underset{\sim}{J'}^{-1}$ decrease from the center to the corners of the matrix. The reason for this is that the assumption of a finite number of

19

sample points requires the radiance to be known precisely at points outside the sampled area. The estimates of B_m for points near the edge are influenced by a smaller number of unknown radiance values and can therefore be made more accurately.

At large signal-to-noise ratio the elements G'_{mn} of the inverse matrix $\underset{\sim}{J}'^{-1}$ are nearly proportional to the covariances of the errors in the estimates $\hat{B}_m$ and $\hat{B}_n$. For $\Delta > \frac{1}{2}\delta$ these are approximately equal to the elements G'^{∞}_{m-n} of $\underset{\sim}{G}'^{\infty} = \underset{\sim}{J}'^{\infty-1}$, where from Eq. (4.4)

$$G'^{\infty}_p = \int_{-\pi}^{\pi} [j'(\omega)]^{-1} \exp(-ip\omega) \, d\omega/2\pi. \tag{4.26}$$

In Table 1 we have listed the asymptotic correlation coefficients $G'^{\infty}_p/G'^{\infty}_o$ for several values of δ/Δ. When Δ is much larger than δ, the estimates are approximately uncorrelated. As Δ increases, the correlation spreads over more and more adjacent elements.

Table 1

Correlation Coefficient of Radiance Estimates

Number of Intervals, p

δ/Δ	0	1	2	3	4	5	6	7
0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.5	1.0	-0.0582	0.00417	-0.00676	0.00107	-0.00245	0.00048	-0.00126
1.0	1.0	-0.139	0.0237	-0.0194	0.00658	-0.00720	0.00301	-0.00373
1.5	1.0	-0.269	0.0882	-0.0576	0.0307	-0.0242	0.0152	-0.0131
1.9	1.0	-0.505	0.304	-0.230	0.174	-0.142	0.115	-0.0983
3.0	1.0	-0.936	0.764	-0.538	0.321	-0.155	0.0556	-0.00117

The last line of Table 1 lists the correlation coefficients G'_{p0}/G'_{00} obtained from the central row of the inverse $\underset{\sim}{G}' = \underset{\sim}{J}'^{-1}$ of the 15×15 matrix $\underset{\sim}{J}'$, whose elements were calculated from Eq. (4.24) for $\delta/\Delta = 3$. Adjacent estimates are much more strongly correlated than for $\delta/\Delta < 2$. The small value for $p = 7$ reflects the fact that the two sample points 7 units on each side of center lie at the edge of the observed area, where the estimates are influenced by somewhat fewer unknown radiance values than for points near the center.

Indicator-Function Sampling

When the indicator functions of Eq. (1.3) are used for sampling in one or two dimensions, the matrix elements J_p' are integrals that cannot be expressed in closed form. The Fourier transform of the central indicator function is

$$K_0(\underset{\sim}{r}) = A_\Delta \; \text{sinc}(x/\gamma_x) \; \text{sinc}(y/\gamma_y), \tag{4.27}$$

with $\gamma_x = \lambda R/\Delta_x$, $\gamma_y = \lambda R/\Delta_y$. For a one-dimensional object the approximate values of the central diagonal elements of $\underset{\sim}{J}'^{-1}$ are

$$G_0'^\infty = (\pi^2/\delta) \int_{-\frac{1}{2}}^{\frac{1}{2}} \text{csc}^2(\pi u) \; \{ \sum_{m=-\infty}^{\infty} (u - m)^{-2} \; [1 - |u - m| \; \delta/\Delta]\}^{-1} \; du,$$
$$\Delta < 2\delta, \tag{4.28}$$

corresponding to Eq. (4.25). This has been plotted as a dashed curve in Fig. 1; numerical integration was required.

V. Fourier Sampling

The minimum mean-square error of an unbiased estimate of any of the coefficients B_m in the expansion of the object radiance, Eq. (1.2), is given by Eqs. (3.3) and (3.9),

$$\varepsilon_m \geq (N/E_s)^2 \ WT \ B_T^2 \ G_{mm}, \tag{5.1}$$

where G_{mm} is a diagonal element of the matrix $G = J^{-1}$. The elements of J are in turn given by Eq. (3.10). The matrix inversion would be simple if J were a diagonal matrix, and it is natural to look for the kind of sampling for which it is. The matrix J will be diagonal if the orthogonal functions $F_m(u)$ are eigenfunctions of the integral equation

$$\lambda_m \ F_m(u) = \int_0 | \mathcal{G}(u - v)|^2 \ F_m(v) \ d^2v, \tag{5.2}$$

whose kernel $| \mathcal{G}(u - v)|^2$ is defined by Eq. (3.11).

The object whose radiance is to be estimated is assumed to occupy a finite region O of the object plane. Its diameter will in general be much larger than the width of the kernel $| \mathcal{G}(u)|^2$, which is of the order of $\lambda RA^{-\frac{1}{2}}$, A being the area of the aperture. As discussed previously,[18] the eigenfunctions are then, for a rectangular object, sinusoids as in Eq. (1.6); and the eigenvalues λ_m, which will be the diagonal elements of J, are the values of the Fourier transform of the kernel $| \mathcal{G}(u)|^2$ evaluated at points separated in the x- and y-directions by $2\pi/b_x$ and $2\pi/b_y$, respectively, where b_x and b_y are the length and breadth of the object. Thus we are led to Fourier sampling.

The Fourier transform of the kernel $| \mathcal{G}(u)|^2$ is, by the convolution theorem, the self-convolution of the indicator function $I_A(r)$ of the aperture. For a

23

rectangular aperture $a_x \times a_y$, this is

$$\Phi(\underset{\sim}{\omega}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\mathcal{G}(\underset{\sim}{u})|^2 \exp(i\underset{\sim}{\omega} \cdot \underset{\sim}{u}) \, d^2\underset{\sim}{u} =$$

$$(\lambda R/A)^2 \int_A I_A(\underset{\sim}{r}) \, I_A(k^{-1} R\underset{\sim}{\omega} - \underset{\sim}{r}) \, d^2\underset{\sim}{r} =$$

$$(\lambda R/A)^2 \, [a_x - |k^{-1} R\omega_x|] \, [a_y - |k^{-1} R\omega_y|],$$

$$|\omega_x| < ka_x/R, \quad |\omega_y| < ka_y/R, \tag{5.3}$$

and $\Phi(\underset{\sim}{\omega}) = 0$ for $\underset{\sim}{\omega}$ outside the rectangle $(ka_x/R) \times (ka_y/R)$. Hence the eigen-values of Eq. (5.2) are approximately

$$\lambda_{\underset{\sim}{m}} = (\lambda R)^2 \, [1 - |m_x| \, \delta_x/b_x] \, [1 - |m_y| \, \delta_y/b_y]/A,$$

$$\delta_x = \lambda R/a_x \ll b_x, \quad \delta_y = \lambda R/a_y \ll b_y, \tag{5.4}$$

with $[x] = 0$ for $x < 0$.

The minimum relative mean-square error of an unbiased estimate of the coefficient

$$B_{\underset{\sim}{m}} = (b_x \, b_y)^{-1} \int_0 B(x, y) \exp 2\pi i(m_x \, xb_x^{-1} + m_y \, yb_y^{-1}) \, dxdy \tag{5.5}$$

is

$$\epsilon_{\underset{\sim}{m}}/B_0^2 \gtrsim (N/E_s)^2 \, MWT[1 - |m_x| \, \delta_x/b_x]^{-1} \, [1 - |m_y| \, \delta_y/b_y]^{-1}, \tag{5.6}$$

where $M = AA_0/(\lambda R)^2 = A_0/A_\delta$ is the number of spatial degrees of freedom in the object O. $(A_0 = b_x \, b_y = $ the area of the object.) We have referred the mean-square errors to the value of the central coefficient $B_0 = B_T/A_0$. For $\underset{\sim}{m} = \underset{\sim}{0}$ Eq. (5.6) agrees with Eq. (4.19).

As $|m_x|$ and $|m_y|$ increase, so does the minimum relative mean-square error $\epsilon_{\underset{\sim}{m}}/B_0^2$. It appears to become infinite for $|m_x| = b_x/\delta_x$ or $|m_y| = b_y/\delta_y$, but the exact eigenvalues $\lambda_{\underset{\sim}{m}}$ do not go to zero at that point, although they become

extremely small for $|m_x| > b_x/\delta_x$, $|m_y| > b_y/\delta_y$. The complex exponential sampling function $F_m(\underset{\sim}{u})$ in Eq. (1.6) will bear a significantly large coefficient $B_{\underset{\sim}{m}}$ when the object contains many details whose widths in the x- and y-directions are of the order of b_x/m_x and b_y/m_y, respectively. This coefficient will be subject to a very large error when the details have widths smaller than $\delta_x = \lambda R/a_x$ and $\delta_y = \lambda R/a_y$.

Footnotes

*   This research was carried out under NASA Grant No. 05-009-079.

1.   H. Wolter, Progress in Optics, E. Wolf, Ed. (North-Holland Publ. Co.,
      Amsterdam) 1, 157 (1961).

2.   Y. T. Lo, J. Appl. Phys. 32, 2052 (1961).

3.   C. K. Rushforth and R. W. Harris, J. Opt. Soc. Am. 58, 539 (1968).

4.   L. Mandel and E. Wolf, Rev. Mod. Phys. 37, 231 (1965).

5.   C. W. Helstrom, J. Opt. Soc. Am. 59, 164 (1969).

6.   C. W. Helstrom, "The Modal Decomposition of Aperture Fields in the Detec-
      tion and Estimation of Incoherent Objects", submitted to J. Opt. Soc. Am.

7.   D. Gabor, Progress in Optics, E. Wolf, Ed. (North-Holland Publ. Co.,
      Amsterdam) 1, 109 (1961).

8.   C. W. Helstrom, Statistical Theory of Signal Detection (Pergamon Press, Ltd.,
      Oxford, 1968) 2nd Ed. See pp. 69-72.

9.   Ref. 5, Eqs. (1.8), (1.9).

10.   H. Cramér, Mathematical Methods of Statistics (Princeton University Press,
      1946), pp. 473 ff.

11.   C. R. Rao, Bull. Calcutta Math. Soc. 37, 81 (1945).

12.   L. Schmetterer, Mathematische Statistik (Springer-Verlag, New York, 1966), p. 63.

13.   Ref. 5, Appendix B.

14.   C. W. Helstrom, "Estimation of Object Parameters by a Quantum-Limited
      Optical System", submitted to J. Opt. Soc. Am.

15.   Ref. 1, Eq. (A4).

16.   C. W. Helstrom, J. Opt. Soc. Am. 59, 924 (1969). See Eq. (5.10)

17. Ref. 5, Eq. (6.3), in which $\mathcal{J} = M^{-\frac{1}{2}}$.

18. Ref. 6, Section I. See especially Eqs. (1.15), (1.21).

Figure Caption

Fig. 1. Factor G_0' in lower bound to mean-square error of an unbiased estimate of the sample value of the radiance function of the object plane. Solid curves: sinc-function sampling; dashed curve: indicator-function sampling. Curves are labeled with the number of sample points. $\delta = \lambda R/a$, Δ = sampling interval, a = width of aperture, λ = wavelength, R = distance to object plane.