Information Transfer from Incoherently Radiating Objects

Carl W. Helstrom*

Department of Applied Physics and Information Science

University of California, San Diego

La Jolla, California 92037

NGL-05-009-079

ABSTRACT

The mutual information between the radiance of an incoherently radiating object plane and the field at the aperture of an observing system is calculated on a classical basis. A formula of the Shannon type is obtained when the time-bandwidth product of the object light is large, the average radiance from the object plane is small enough to permit the threshold approximation, and the radiance of the object plane is modeled as a spatial gaussian process. It includes the dependence on the bandwidth of the object light and the effective temperature of the background, assumed spatially and temporally white. A set of sufficient statistics for the aperture field, based on Fourier sampling of the object plane, is introduced, and its bearing on the resolution of fine details of the object is brought out.

The information content of optical images has mostly been calculated on the assumption that both the desired image and the corrupting noise can be described as independent and additive gaussian random processes.[1-10] The result is a formula of the type introduced by Shannon for the capacity of a band-limited channel with additive gaussian noise.[11] It has been generally recognized that the assumption of additivity and gaussian statistics is open to objection when applied to an incoherently radiating object and to an image treated as an illuminance or as the density of developed photographic film.[8] Only treatments based on photon statistics,[12-14] and those limited to coherent light, have avoided this assumption.

Under a classical theory of light the information in an image can be no greater than that present in the values of the electromagnetic field at the aperture of the image-forming optical instrument during the interval of observation. This aperture field has two additive components, the light from the object plane and the random background light. The field created by an incoherently radiating object plane cannot, however, be precisely controlled by a transmitter of information. Only through the radiance of the object plane can information be imparted to the incoherent light emitted. The object field itself is a spatio-temporal gaussian random process whose mutual coherence function depends on the radiance of the object plane. Although the background light can be treated as gaussian noise, it does not add directly to the information-bearing quantities for the object light, which are the values of its mutual coherence function and not the object field itself. The model of an additive gaussian channel cannot, therefore, be simply applied. Since only the object radiance can be controlled, the mutual information of significance is not that between the field at the object plane and the field $\psi_a(\underset{\sim}{r}, t)$ at the aperture of the observing system. Rather, it is that between the radiance $B(\underset{\sim}{u})$ of the object plane and the field $\psi_a(\underset{\sim}{r}, t)$; we denote it by $I(B; \psi_a)$.

In this paper we shall show how to calculate the mutual information $I(B; \psi_a)$. At the end a formula of the Shannon type will be obtained, but the advantage of this treatment is that the approximations necessary are clearly brought out. There are three principal assumptions required. (1) The time T during which the field is observed is much greater than the reciprocal of the bandwidth W of the object light; $WT \gg 1$. (2) The average total energy $\bar{E}$ received from the object plane during the observation must be much less than NMWT, where N is the spectral density of the background light and M is the number of spatial degrees of freedom in the aperture field. (3) The object radiance $B(\underset{\sim}{u})$ is treated as a spatial gaussian random process, an assumption requiring low contrast in the object plane in order to avoid negative values of the radiance.

The second assumption, that $\bar{E}/NMWT \ll 1$, underlies the application of the threshold approximation to determining the detectability of incoherently radiating objects,[15] as well as to calculating their information transfer. The number M is given by[16]

$$M = A\, A_o/(\lambda R)^2,$$

where A is the area of the aperture, A_o the area of the object plane, λ the dominant wavelength of the object light, assumed quasimonochromatic, and R the distance of the object. In terms of the average radiance $\bar{B}$ of the object plane,

$$\bar{E} = \bar{B}\, A\, A_o\, T/(4\pi R^2),$$

and the critical ratio

$$\bar{E}/NMWT = (\bar{B}/NW)\, (\lambda^2/4\pi) \ll 1$$

is independent of the observation time, the sizes of object plane and aperture, and the distance between them. The ratio $\overline{B}/W$ is the average radiance spectral density of the object light (watts/cm^2 Hz), and $N = K\mathcal{T}$, where K is Boltzmann's constant and $\mathcal{T}$ is the effective absolute temperature of the background light.

The mutual information $I(B; \psi_a)$ refers to a class or ensemble of patterns $B(\underset{\sim}{u})$ of radiance in the object plane, not to a single pattern or scene. Although at the end of our analysis this ensemble will be taken as made up of gaussian spatial random processes, it is convenient to think of it during the analysis as a discrete set of patterns, each of which represents a symbol of some alphabet into which messages have been coded. The optical instrument is then treated as a device that must decide, on the basis of its aperture field $\psi_a(\underset{\sim}{r}, t)$ during an observation interval (0, T), which of the ensemble of patterns $B(\underset{\sim}{u})$ is actually present. It bases this decision on a set of likelihood ratios,[17] each corresponding to one of the possible radiance patterns $B(\underset{\sim}{u})$.

Under assumptions (1) and (2), these likelihood ratios can be expressed in terms of a set of sufficient statistics that depend quadratically on the aperture field and contain all the information in it relevant to deciding, in the presence of white gaussian background light, which among the ensemble of radiance patterns $B(\underset{\sim}{u})$ currently exists over the object plane. These statistics are, in this approximation, independent gaussian random variables. The mutual information $I(B; \psi_a)$ can be expressed in terms of them, and when assumption (3) holds, the Shannon-type formula for $I(B; \psi_a)$ is easily obtained.

In this way, the transmission of information from object plane to observing instrument can be regarded as taking place via a large number--roughly WT--of independent, additive gaussian channels. The channels correspond to what has been called *Fourier sampling* of the object plane,[18] and channels associated with the finer details in the object possessthe smaller effective signal-to-noise

ratios. For details smaller then the conventional resolution limit $\lambda R/D$--
R = distance of object, D = diameter of aperture--, the signal-to-noise ratios are vanishingly small, and information associated with them is lost. This aspect of the analysis bears out previous conclusions about the resolvability of fine details in the object plane.[18]

I. The Object Plane as a Transmitter of Information

As mutual information is most directly interpretable in terms of communication, it is useful to consider the object plane 0 as part of a communication system transmitting messages to a distant receiver, as shown in Fig. 1. Because the object plane radiates incoherently, the transmitter can control only its radiance, not the actual values of the electromagnetic field that it radiates. In sending messages, the transmitter selects radiance patterns $B(\underset{\sim}{u})$ from a predetermined set, each pattern corresponding to a symbol of some alphabet into which messages have been coded. The patterns are changed every T seconds as new message symbols appear.

The receiver is an optical instrument admitting the light from 0 through an aperture A. It knows the patterns $B(\underset{\sim}{u})$ and the symbols for which they stand. It must be able to decide which patterns $B(\underset{\sim}{u})$ have been exposed at the object plane during some sequence of intervals, each of duration T. It issues a stream of symbols identifying the patterns it believes were transmitted, and a subsequent decoder interprets them in terms of messages that might have been sent, occasionally making errors because the receiver decided incorrectly about some of the patterns. The receiver bases its decisions on the values of the field $\psi_a(\underset{\sim}{r}, t)$ at its aperture, processing it--as by lenses, stops, and a photosensitive surface--in such a way as to make the decisions with minimum probability of error.[17] The mutual information of significance is that between the ensemble of radiance patterns $B(\underset{\sim}{u})$ and the field values $\psi_a(\underset{\sim}{r}, t)$ at the aperture. In calculating it we shall assume for simplicity a classical and scalar theory of the electromagnetic field.

The field $\psi_{so}(\underset{\sim}{u}, t)$ immediately in front of the object plane 0 is a circular-complex gaussian spatio-temporal stochastic process,[19] quasimonochromatic and having a central frequency $\Omega/2\pi = c/\lambda$; c is the velocity of light and λ the dominant wavelength. The mean values of the field are zero; its statistical distributions are determined entirely by its mutual coherence function[15,20]

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

$$\pi k^{-2}\ B(\underset{\sim}{u}_1)\ \delta(\underset{\sim}{u}_1 - \underset{\sim}{u}_2)\ \chi(t_1 - t_2)\ \exp[-\ i\Omega(t_1 - t_2)], \qquad (1.1)$$

where $B(\underset{\sim}{u})$ is the radiance of the object plane and $\chi(\tau)$ is the temporal coherence function, normalized so that $\chi(0) = 1$. All patterns have the same color; the temporal coherence function $\chi(\tau)$ is invariable. The two-dimensional delta-function $\delta(\underset{\sim}{u}_1 - \underset{\sim}{u}_2)$ indicates the absence of correlation between field values at points separated by distances much smaller than the extent of significant details in the patterns $B(\underset{\sim}{u})$, in accordance with the incoherent nature of the emitted light.

The bandwidth W of the object light is conveniently defined by[15]

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

$$= |\chi(0)|^2 \Bigg/ \int_{-\infty}^{\infty} |\chi(\tau)|^2\ d\tau, \qquad (1.2)$$

where

$$X(\omega) = \int_{-\infty}^{\infty} \chi(\tau)\ e^{-i\omega\tau}\ d\tau \qquad (1.3)$$

is its spectral density, with angular frequencies ω referred to $\Omega = 2\pi c/\lambda$. This bandwidth W for natural sources of light is so great that any effective means of altering the radiance $B(\underset{\sim}{u})$ requires a time much longer than 1/W, and we can assume that each pattern is exposed for a time $T >> 1/W$. The basic observation interval for each transmitted symbol can be taken as $(0, T)$, and the mutual information will be referred to this interval of T seconds' duration.

II. The Representation of the Aperture Field

The field $\psi_a(\underset{\sim}{r}, t)$ at the aperture A consists of two independent parts, the field $\psi_s(\underset{\sim}{r}, t)$ of the light from the object plane, and the background field $\psi_n(\underset{\sim}{r}, t)$:

$$\psi_a(\underset{\sim}{r}, t) = \psi_s(\underset{\sim}{r}, t) + \psi_n(\underset{\sim}{r}, t). \tag{2.1}$$

Both are circular-complex gaussian random processes.[19] The mutual coherence function of the aperture field has two corresponding terms,

$$\frac{1}{2} \underset{\sim}{E}[\psi_a(\underset{\sim}{r}_1, t_1)\, \psi_a^*(\underset{\sim}{r}_2, t_2)] = \varphi_s(\underset{\sim}{r}_1, \underset{\sim}{r}_2)\, \chi(t_1 - t_2) \exp[-i\Omega(t_1 - t_2)] + N\, \delta(\underset{\sim}{r}_1 - \underset{\sim}{r}_2)\, \delta(t_1 - t_2), \tag{2.2}$$

the former referring to the object light, or signal, the latter to the background light, or noise.

The spatial coherence function $\varphi_s(\underset{\sim}{r}_1, \underset{\sim}{r}_2)$ of the light from the object plane depends on the concurrent radiance distribution $B(\underset{\sim}{u})$,

$$\varphi_s(\underset{\sim}{r}_1, \underset{\sim}{r}_2) = \pi k^{-2} \int_0 S(\underset{\sim}{r}_1, \underset{\sim}{u})\, S^*(\underset{\sim}{r}_2, \underset{\sim}{u})\, B(\underset{\sim}{u})\, d^2\underset{\sim}{u}, \tag{2.3}$$

where $S(\underset{\sim}{r}, \underset{\sim}{u})$ is the point-spread function for light propagating from object to aperture.[15] We suppose these separated by such a long distance R that the rays from the object can be treated as paraxial, and the point-spread function is to good approximation

$$S(\underset{\sim}{r}, \underset{\sim}{u}) = (ik/2\pi R) \exp[ikR + ik|\underset{\sim}{r} - \underset{\sim}{u}|^2/2R]. \tag{2.4}$$

Then the spatial coherence function of the object light is

$$\varphi_s(\underset{\sim}{r}_1, \underset{\sim}{r}_2) = (4\pi R^2)^{-1} \exp[ik(\underset{\sim}{r}_1{}^2 - \underset{\sim}{r}_2{}^2)/2R]\, \beta(\underset{\sim}{r}_1 - \underset{\sim}{r}_2), \tag{2.5}$$

where

$$\beta(\underset{\sim}{r}) = \int_0 B(\underset{\sim}{u}) \exp(- ik \underset{\sim}{r} \cdot \underset{\sim}{u}/R) \, d^2\underset{\sim}{u} \tag{2.6}$$

is the Fourier transform of the object radiance $B(\underset{\sim}{u})$. The total energy E received from the object during a typical observation interval (0, T) is

$$E = T \int_A \varphi_s(\underset{\sim}{r}, \underset{\sim}{r}) \, d^2\underset{\sim}{r} = \beta(0) \, AT/4\pi R^2, \tag{2.7}$$

where

$$\beta(0) = \int_0 B(\underset{\sim}{u}) \, d^2\underset{\sim}{u} \tag{2.8}$$

is the integrated radiance of the object plane.

The second term in Eq. (2.2) is the mutual coherence function of the background light, which has a distribution in frequency and angle much broader than that of the object light and is taken as spatially and temporally white. Its spectral density N is in our present units equal to $K\mathcal{T}$.

In order to deal with the probability density functions of the aperture field $\psi_a(\underset{\sim}{r}, t)$, we must express it in terms of a countable set of random variables. To this end we expand the field in a series of functions orthonormal over the aperture A and over the observation interval (0, T). Since all the patterns have the same temporal coherence function $\chi(\tau)$, it is convenient to write $\psi_a(\underset{\sim}{r}, t)$ in terms of the eigenfunctions $\gamma_m(t)$ of the integral equation

$$g_m \, \gamma_m(t) = T^{-1} \int_0^T \chi(t - s) \, \gamma_m(s) \, ds, \tag{2.9}$$

and we write the aperture field in the form

$$\psi_a(\underset{\sim}{r}, t) = \sum_p \sum_m \psi_{pm} \, f_p(\underset{\sim}{r}) \, \gamma_m(t), \tag{2.10}$$

where $f_p(\underset{\sim}{r})$ form an arbitrary complete set of functions orthonormal over the aperture A. The field coefficients ψ_{pm} will then be statistically independent for different values of m.[21]

Because we assume that WT >>1, the eigenvalues g_m are given approximately by[16]

$$g_m = T^{-1} X(2\pi m/T), \; m = \ldots -1, 0, 1, 2, \ldots \tag{2.11}$$

in terms of the spectral density $X(\omega)$ of the object light. They sum to 1,

$$\sum_m g_m = 1 , \tag{2.12}$$

and by Eqs. (1.2) and (2.9) and the orthonormality of the eigenfunctions $\gamma_m(t)$,

$$\sum_m g_m^2 = 1/WT. \tag{2.13}$$

There are roughly WT significant temporal modes in the expansion of $\psi_a(\underset{\sim}{r}, t)$ as in Eq. (2.10).

The coefficients ψ_{pm} for each temporal mode labeled by m are conveniently listed as a column vector $\underset{\sim}{\psi}_m$, whose hermitian conjugate row vector $\underset{\sim}{\psi}_m^+$ lists their complex conjugates ψ_{pm}^*. The ψ_{pm} are also circular-complex gaussian random variables[19]; in terms of the aperture field they are given by

$$\psi_{pm} = \int_A f_p^*(\underset{\sim}{r}) \; \gamma_m^*(t) \; \psi_a(\underset{\sim}{r}, t) \; d^2\underset{\sim}{r} \; dt. \tag{2.14}$$

By using Eqs. (2.2) and (2.9) it can be shown that their covariances are

$$\frac{1}{2} \underset{\sim}{E}(\psi_{pm} \; \psi_{qn}^*) = (T \; g_m \; \varphi_{pq} + N \; \delta_{pq}) \; \delta_{nm}, \tag{2.15}$$

where δ_{nm} is the Kronecker delta symbol and

$$\varphi_{pq} = \int_A \int_A f_p^*(\underset{\sim}{r}_1)\, \varphi_s(\underset{\sim}{r}_1, \underset{\sim}{r}_2)\, f_q(\underset{\sim}{r}_2)\, d^2\underset{\sim}{r}_1 d^2\underset{\sim}{r}_2 \tag{2.16}$$

is an element of a covariance matrix $\underset{\sim}{\varphi}$ depending on the radiance $B(\underset{\sim}{u})$ of the object plane through Eqs. (2.5) and (2.6).

The coefficients ψ_{pm} contain all the information present in the aperture field $\psi_a(\underset{\sim}{r}, t)$ during the basic interval $(0, T)$. The joint probability density function (p. d. f.) of n of these coefficients, conditional on the actual radiance pattern concurrently exposed at the aperture plane, has the complex gaussian form[19]

$$p(\psi|B) = (2\pi)^{-n} \prod_m \det(T\, g_m\, \underset{\sim}{\varphi} + N\, \underset{\sim}{I})^{-1}$$

$$\times \exp[-\frac{1}{2}\, \underset{\sim}{\psi}_m^{+} (T\, g_m\, \underset{\sim}{\varphi} + N\, \underset{\sim}{I})^{-1}\, \underset{\sim}{\psi}_m], \tag{2.17}$$

where $\underset{\sim}{I}$ is the identity matrix and det stands for the determinant. Later it will be possible to increase n beyond all bounds.

III. The Likelihood Ratios and Their Decomposition

If the receiver is to decide with minimum probability of error which pattern $B(\underset{\sim}{u})$ is being exposed at the object plane, it must determine the posterior probability of each pattern, given the actual field $\psi_a(\underset{\sim}{r}, t)$ observed during the interval (0, T), and it must select the pattern and its associated message symbol having the greatest posterior probability.[17] The posterior probability of a given radiance pattern $B(\underset{\sim}{u})$ is proportional to the conditional p. d. f. $p(\underset{\sim}{\psi}|B)$ and to the prior probability of the pattern. Equivalently, it is proportional to the likelihood ratio

$$\Lambda(\underset{\sim}{\psi}|B) = p(\underset{\sim}{\psi}|B)/p_0(\underset{\sim}{\psi}) \tag{3.1}$$

obtained by dividing $p(\underset{\sim}{\psi}|B)$ by the joint p. d. f. $p_0(\underset{\sim}{\psi})$ of the coefficients ψ_{pm} when no light arrives from the object plane and background light alone is present at the aperture. This p. d. f. is obtained by setting $\underset{\sim}{\varphi} = \underset{\sim}{0}$ in Eq. (2.17),

$$p_0(\underset{\sim}{\psi}) = (2\pi)^{-n} \prod_m (\det N \underset{\sim}{I})^{-1} \exp[-\frac{1}{2} \underset{\sim}{\psi}_m{}^+ N^{-1} \underset{\sim}{I} \underset{\sim}{\psi}_m]. \tag{3.2}$$

The likelihood ratio can then be written as

$$\Lambda(\underset{\sim}{\psi}|B) = e^Z \tag{3.3}$$

where the log-likelihood ratio Z is given by

$$Z = \frac{1}{2}\sum_m \underset{\sim}{\psi}_m{}^+[N^{-1} \underset{\sim}{I} - (T g_m \underset{\sim}{\varphi} + N \underset{\sim}{I})^{-1}] \underset{\sim}{\psi}_m$$

$$- \sum_m \ln \det(\underset{\sim}{I} + N^{-1} T g_m \underset{\sim}{\varphi}) =$$

$$\frac{1}{2} N^{-2} \sum_m \underset{\sim}{\psi}_m^{+} T g_m \underset{\sim}{\varphi} (\underset{\sim}{I} + N^{-1} T g_m \underset{\sim}{\varphi})^{-1} \underset{\sim}{\psi}_m$$

$$- \sum_m \text{Tr } \ell n(\underset{\sim}{I} + N^{-1} T g_m \underset{\sim}{\varphi}); \tag{3.4}$$

we have used the rule[22] that for any positive definite matrix $\underset{\sim}{M}$, $\ell n \det \underset{\sim}{M} = \text{Tr}(\ell n \underset{\sim}{M})$, Tr standing for the trace of a matrix. At this point the total number n of coefficients ψ_{pm} included can be made infinite, so that all the information in the aperture field is taken into account. The log-likelihood ratio Z remains finite in this passage to the limit $n \rightarrow \infty$.

The receiver can determine by operations on its aperture field $\psi_a(\underset{\sim}{r}, t)$ the statistic Z associated with each pattern $B(\underset{\sim}{u})$ in the ensemble, and it can use the Z's to calculate the posterior probability of each pattern. The Z's, therefore, contain all the information in the aperture field relevant to the receptive function of the optical instrument as part of the communication system. We now wish to express each Z as a weighted sum of independent random variables that do not depend on the patterns $B(\underset{\sim}{u})$. This will require the assumption that the average energy $\bar{E}$ received from each pattern during the interval (0, T) is much less than NMWT.

Each radiance pattern $B(\underset{\sim}{u})$ is expanded in a common series of orthonormal functions $\mathcal{B}_j(\underset{\sim}{u})$,

$$B(\underset{\sim}{u}) = \sum_j b_j \mathcal{B}_j(\underset{\sim}{u}). \tag{3.5}$$

The patterns are now distinguished by their sets $\{b_j\}$ of expansion coefficients. Through Eqs. (2.5), (2.6), and (2.16) there is a corresponding decomposition of the covariance matrices $\underset{\sim}{\varphi}$,

$$\underset{\sim}{\varphi} = \sum_j b_j \underset{\sim}{\varphi}_j, \tag{3.6}$$

and we require the matrices $\underset{\sim}{\varphi}_j$ to be orthogonal in the sense that

$$T^2 \operatorname{Tr} \underset{\sim}{\varphi}_k \underset{\sim}{\varphi}_\ell = \mu_k^2 \delta_{k\ell}, \tag{3.7}$$

where the μ_k are certain positive constants. From Eqs. (2.5), (2.6), and (2.16) it is not hard to show that

$$\operatorname{Tr} \underset{\sim}{\varphi}_k \underset{\sim}{\varphi}_\ell = A^2(4\pi R^2)^{-2} \int_0 \int_0 \mathcal{B}_k(\underset{\sim}{u}_1) \, |\mathcal{J}(\underset{\sim}{u}_1 - \underset{\sim}{u}_2)|^2 \, \mathcal{B}_\ell(\underset{\sim}{u}_2) \, d^2\underset{\sim}{u}_1 d^2\underset{\sim}{u}_2, \tag{3.8}$$

where[22]

$$\mathcal{J}(\underset{\sim}{u}) = A^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I_A(\underset{\sim}{r}) \exp(ik\underset{\sim}{u} \cdot \underset{\sim}{r}/R) \, d^2\underset{\sim}{r} \tag{3.9}$$

is the Fourier transform of the indicator function of the aperture,

$$I_A(\underset{\sim}{r}) = 1, \qquad \underset{\sim}{r} \in A,$$

$$= 0, \qquad \underset{\sim}{r} \notin A.$$

In order for Eq. (3.7) to hold, the functions $\mathcal{B}_j(\underset{\sim}{u})$ must be eigenfunctions of the integral equation

$$p_j^2 \mathcal{B}_j(\underset{\sim}{u}) = \int_0 |\mathcal{J}(\underset{\sim}{u} - \underset{\sim}{v})|^2 \mathcal{B}_j(\underset{\sim}{v}) \, d^2\underset{\sim}{v}. \tag{3.10}$$

This same equation arose in analyzing the resolution of details in the object plane.[18] The constants μ_j in Eq. (3.7) are related to the eigenvalues p_j^2 through

$$\mu_j = p_j AT/4\pi R^2. \tag{3.11}$$

In terms of the matrices $\underset{\sim}{\varphi}_j$ we define the random variables

$$z_j = \frac{1}{2} N^{-2} \sum_m T\, g_m\, \underset{\sim}{\psi}_m^{+}\, \underset{\sim}{\varphi}_j\, \underset{\sim}{\psi}_m - c_j, \tag{3.12}$$

where

$$c_j = N^{-1}\, T \sum_m g_m \,\mathrm{Tr}\, \underset{\sim}{\varphi}_j \tag{3.13}$$

is a constant making the expected value of z_j vanish when no object light is present.

Under the assumption that E/NMWT<< 1, the log-likelihood ratio Z in Eq. (3.4) can be written approximately as

$$Z \cong \frac{1}{2} N^{-2} \sum_m \underset{\sim}{\psi}_m^{+}\, T\, g_m\, \underset{\sim}{\varphi}\, \underset{\sim}{\psi}_m + C_B'$$

$$= \sum_j b_j\, z_j + C_B, \tag{3.14}$$

where C_B and C_B' are known constants depending on the radiance pattern $B(\underset{\sim}{u})$, but not on the aperture field $\psi_a(\underset{\sim}{r}, t)$. This is the threshold approximation.[15] Given the data z_j, therefore, the receiver can determine, through the Z's defined by Eq. (3.14), the posterior probability of each pattern and thence make its decisions in the optimum fashion. The z_j's constitute in this approximation a set of sufficient statistics.

When as assumed here WT >> 1, furthermore, the z_j's are statistically independent gaussian random variables. In the presence of a particular radiance pattern $B(\underset{\sim}{u})$, the expected value of z_j is, by Eqs. (2.13), (2.15), (3.6), (3.7), and (3.11),

$$\underset{\sim}{E}(z_j|B) = N^{-2} \sum_m T^2\, g_m^{\,2}\, \mathrm{Tr}\, \underset{\sim}{\varphi}_j\, \underset{\sim}{\varphi}$$

$$= b_j \, N^{-2} \, T^2 \, (WT)^{-1} \, \mathrm{Tr} \, \underset{\sim}{\varphi}_j{}^2$$

$$= b_j \, \mu_j{}^2/N^2WT \quad = \quad b_j \, d_j{}^2, \tag{3.15}$$

where

$$d_j{}^2 = \mu_j{}^2/N^2WT, \tag{3.16}$$

and their covariances are

$$\{z_k, z_\ell\} = \underset{\sim}{E}(z_k \, z_\ell | B) - \underset{\sim}{E}(z_k | B) \, \underset{\sim}{E}(z_\ell | B)$$

$$= N^{-2} \sum_m T^2 \, g_m^2 \, \mathrm{Tr} \, \underset{\sim}{\varphi}_k (\underset{\sim}{I} + N^{-1} \, T \, g_m \, \underset{\sim}{\varphi}) \, \underset{\sim}{\varphi}_\ell (\underset{\sim}{I} + N^{-1} \, T \, g_m \, \underset{\sim}{\varphi})$$

$$\cong N^{-2} \, (WT)^{-1} \, T^2 \, \mathrm{Tr} \, \underset{\sim}{\varphi}_k \, \underset{\sim}{\varphi}_\ell$$

$$= (\mu_k{}^2/N^2WT) \, \delta_{k\ell} = d_k{}^2 \, \delta_{k\ell}, \tag{3.17}$$

independently of $B(\underset{\sim}{u})$, when $E \ll NMWT$. The demonstration that our approximations are valid when $WT \gg 1$ and $E/NMWT \ll 1$ is given in Appendix A. The z_j's are approximately gaussian as a consequence of the central limit theorem of statistics; Eq. (3.12) defines them as the sum of a large number--roughly WT--of independent random variables. Being gaussian and uncorrelated, they are, in this approximation, statistically independent.

IV. The Mutual Information

Since the patterns $B(\underset{\sim}{u})$ correspond to specific sets of coefficients b_j, the communication system can be regarded as one that transmits, every T seconds, a set of numbers b_j and receives the set of gaussian random variables z_j. It is equivalent to a set of parallel channels with additive gaussian noise; the signal-to-noise ratio in the j-th channel is

$$[\underset{\sim}{E}(z_j|B)]^2/\text{Var } z_j = b_j^2 \, d_j^2, \qquad (4.1)$$

where by Eqs. (3.11) and (3.16) d_j^2 is proportional to the eigenvalue p_j^2 of the integral equation, Eq. (3.10).

When the object plane O is rectangular and its length b_x and its width b_y are much greater than the width of the kernel $|\mathcal{G}(\underset{\sim}{u})|^2$ of Eq. (3.10), the eigenvalues p_j^2 are approximately given by the spatial Fourier transform of $|\mathcal{G}(\underset{\sim}{u})|^2$ evaluated at the points of a rectangular lattice whose cells measure $(2\pi/b_x) \times (2\pi/b_y)$. Thus[18]

$$p_{\underset{\sim}{j}}^2 = (\lambda R/A)^2 \, I_A^{(2)} \, (j_x \gamma_x, j_y \gamma_y),$$

$$\gamma_x = \lambda R/b_x, \quad \gamma_y = \lambda R/b_y, \qquad (4.2)$$

where $\underset{\sim}{j} = (j_x, j_y)$ is a pair of integers replacing the previous index j, and[22]

$$I_A^{(2)}(\underset{\sim}{r}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I_A(\underset{\sim}{r}') \, I_A(\underset{\sim}{r}' - \underset{\sim}{r}) \, d^2\underset{\sim}{r}' \qquad (4.3)$$

is the self-convolution of the indicator function of the aperture. The signal-to-noise ratio in the channel $\underset{\sim}{j}$ is

$$b_{\underset{\sim}{j}}^2 \, d_{\underset{\sim}{j}}^2 = (N^2 WT)^{-1} \left(\frac{b_{\underset{\sim}{j}} AT}{4\pi R^2}\right)^2 (\lambda R/A)^2 \, I_A^{(2)}(j_x \gamma_x, j_y \gamma_y). \qquad (4.4)$$

For a rectangular aperture $a_x \times a_y$

$$I_A^{(2)}(x, y) = (a_x - |x|)(a_y - |y|), \quad |x| < a_x, \ |y| < a_y;$$

$$I_A^{(2)}(x, y) = 0, \quad |x| \geq a_x \text{ or } |y| \geq a_y. \tag{4.5}$$

The channels for which $|j_x| \gamma_x > a_x$ or $|j_y| \gamma_y > a_y$ have, therefore, negligible signal-to-noise ratio. (Eq. (4.4) is not exact, but we can be sure that the true eigenvalues, though not zero in this region, are exceedingly small.) These channels, as discussed previously,[18] correspond to details in the object whose sizes are smaller than $(\delta_x, \delta_y) = (\lambda R / a_x, \lambda R / a_y)$, that is, smaller than a conventional resolution interval in the object plane 0. Information in such fine details is lost in the noise.

The mutual information, per T-second interval, between the object plane and the aperture plane is in this approximation given by[23]

$$I(B; \psi_a) = \iint q(\{b_j\}) \, p(\{z_j\}|\{b_j\}) \times \ln [p(\{z_j\}|\{b_j\})/p(\{z_j\})] \prod_j (dz_j \, db_j), \tag{4.6}$$

where

$$p(\{z_j\}) = \int q(\{b_j\}) \, p(\{z_j\}|\{b_j\}) \prod_j db_j \tag{4.7}$$

and $p(\{z_j\}|\{b_j\})$ is the joint conditional p. d. f. of the z_j's, given the coefficients b_j specifying the pattern $B(\underset{\sim}{u})$; $p(\{z_j\}|\{b_j\})$ has the gaussian form with conditional means and covariances given by Eqs. (3.15) and (3.17). Here $q(\{b_j\})$ is the joint p. d. f. of the coefficients b_j for the ensemble of patterns $B(\underset{\sim}{u})$ in terms of which messages are transmitted. By using the gaussian form of $p(\{z_j\}|\{b_j\})$ we can write the information as

$$I(B;\ \psi_a) = -\frac{1}{2}\sum_j \ell n\ (2\pi e\ d_j^{\ 2}) - \int p(\{z_j\})\ \ell n\ p(\{z_j\}) \prod_j dz_j; \tag{4.8}$$

a finite number n_b of coefficients b_j are included here, and finally the limit $n_b \to \infty$ is taken.

We can proceed no farther without knowing the joint p. d. f. $q(\{b_j\})$, which is needed in Eq. (4.7) for calculating $p(\{z_j\})$. In general, this joint p. d. f. will be difficult to determine, involving, for instance, extensive measurements of a large number of typical scenes as represented by radiance distributions $B(\underset{\sim}{u})$ on the object plane. The constraint that $B(\underset{\sim}{u})$ must be non-negative will restrict the class of possible joint p. d. f.'s of the coefficients b_j in the expansion of Eq. (3.5).

The simplest assumption at this point is that the radiance $B(\underset{\sim}{u})$ is a realization of a homogeneous spatial gaussian random process with a covariance function $\sigma_B(\underset{\sim}{u}_1 - \underset{\sim}{u}_2)$ whose width is of the order of the size of typical details in the scene. The mean of the gaussian distribution is taken large enough so that $B(\underset{\sim}{u})$ becomes negative only with negligible probability. The average contrast of the resulting scenes will be of the order of

$$[\sigma_B(0)]^{1/2}/\underset{\sim}{E}(B)$$

and necessarily small.

When this gaussian assumption is made, the calculation of mutual information is standard[24] and leads to the result

$$I(B;\ \psi_a) = \frac{1}{2}\ \mathrm{Tr}\ \ell n(\underset{\sim}{I} + \underset{\sim}{\sigma}\ \underset{\sim}{D}_2), \tag{4.9}$$

where $\underset{\sim}{D}_2$ is a matrix whose diagonal elements are $d_k^{\ 2}$ and where the matrix $\underset{\sim}{\sigma}$ has elements

$$\sigma_{k\ell} = \int \mathcal{B}_k(\underset{\sim}{u}_1)\ \sigma_B(\underset{\sim}{u}_1 - \underset{\sim}{u}_2)\ \mathcal{B}_\ell(\underset{\sim}{u}_2)\ d^2\underset{\sim}{u}_1 d^2\underset{\sim}{u}_2. \tag{4.10}$$

As shown in Appendix B, the mutual information can be written as

$$I(B;\, \psi_a) = (A_o/2\lambda^2R^2) \int_A \ln[1 + \alpha(\lambda R/A)^2\, \Sigma_B(\underset{\sim}{r})\, I_A^{(2)}(\underset{\sim}{r})]\, d^2\underset{\sim}{r}, \tag{4.11}$$

where A_o of the area of the object,

$$\alpha = (N^2WT)^{-1}\, (AT/4\pi R^2)^2, \tag{4.12}$$

and

$$\Sigma_B(\underset{\sim}{r}) = \int_0 \sigma_B(\underset{\sim}{u})\, \exp(-\, ik\underset{\sim}{u} \cdot \underset{\sim}{r}/R)\, d^2\underset{\sim}{u} \tag{4.13}$$

is the Fourier transform of the covariance function $\sigma_B(\underset{\sim}{u})$ of the object radiance. This result has a form that has frequently appeared in the literature; only the constants involved are somewhat different. In particular, the bandwidth W of the object light and the time T of observation appear explicitly.

It has been assumed in going from Eq. (4.9) to Eq. (4.11) that the object plane is much larger than the conventional resolution element $\lambda R/D$, where D is the diameter of the aperture. The mutual information is then proportional, as might be expected, to the area of the object plane. If Δ is the diameter of typical details in the scene, that is, if Δ is the width of the covariance function $\sigma_B(\underset{\sim}{u})$, the width of its transform $\Sigma_B(\underset{\sim}{r})$ is of the order of $\lambda R/\Delta$; and when $\lambda R/\Delta >> D$, or $\Delta << \lambda R/D$, the value of the integral in Eq. (4.11) depends chiefly on the diameter of the aperture. That is, details of size Δ much less than $\lambda R/D$ do not contribute to the mutual information, and as we have seen,[18] their amplitudes cannot be estimated by the receiver.

For a gaussian covariance

$$\sigma_B(\underset{\sim}{u}) = s_B{}^2\, \exp(-\, \underset{\sim}{u}^2/2\Delta^2) \tag{4.14}$$

and a circular aperture of radius a, for which

$$I_A^{(2)}(\underset{\sim}{r}) = (2A/\pi)\ (\theta - \sin\theta\cos\theta),$$
$$\theta = \cos^{-1}\ (|\underset{\sim}{r}|/2a), \tag{4.15}$$

the mutual information is given by

$$I(B;\ \psi_a) = M\ \pi \int_0^1 \sin(\pi x)\ \ell n\{1 + \eta\ \exp[-\ 4\pi^2 y^2 (1 + \cos\pi x)]$$
$$\times\ (x - \pi^{-1}\sin\pi x)\}\,dx,$$
$$y = \Delta/\delta, \quad \delta = \lambda R/a,$$
$$\eta = (2y^2/\pi^2)\ D_\delta^{\ 2},$$
$$D_\delta^{\ 2} = E_\delta^{\ 2}/N^2 WT,$$
$$E_\delta = (\pi\ s_B\ \delta^2)\ AT/4\pi R^2,$$
$$M = A\ A_o/(\lambda R)^2. \tag{4.16}$$

Here M is the number of spatial degrees of freedom in the aperture field.[16]

In Fig. 2 we have plotted $I(B;\ \psi_a)/M$, the mutual information per spatial degree of freedom per T-second interval, versus $D_\delta^{\ 2}$ for various values of $y = \Delta/\delta$. Here $D_\delta^{\ 2}$ is the signal-to-noise ratio associated with a received energy E_δ, which is the total energy that would be received from a circle of radius $\delta = \lambda R/a$ on the object plane were the radiance uniformly equal to s_B, the standard deviation of the radiance patterns $B(\underset{\sim}{u})$. At each value of $D_\delta^{\ 2}$ there is a value of the ratio Δ/δ that yields maximum mutual information; this optimum ratio decreases as the quantity $D_\delta^{\ 2}$ increases. The greater the variance σ_B of the radiance $B(\underset{\sim}{u})$, the more the fine details contribute to the information transferred.

Appendix A. The Sufficient Statistics

When in the log-likelihood ratio Z of Eq. (3.4) the terms $(\underset{\sim}{I} + N^{-1} T g_m \underset{\sim}{\varphi})^{-1}$ and $\ell n(\underset{\sim}{I} + N^{-1} T g_m \underset{\sim}{\varphi})$ are expanded in power series, we obtain

$$
\begin{aligned}
Z &= \frac{1}{2} N^{-2} \sum_m \underset{\sim}{\psi}_m^{+} T g_m \underset{\sim}{\varphi}(\underset{\sim}{I} - N^{-1} T g_m \underset{\sim}{\varphi} + \ldots) \underset{\sim}{\psi}_m \\
&\qquad - \sum_m \mathrm{Tr}[N^{-1} T g_m \underset{\sim}{\varphi} - N^{-2} T^2 g_m^{\,2} \underset{\sim}{\varphi}^2 + \ldots] \\
&= \sum_j b_j z_j - \frac{1}{2} N^{-3} \sum_m \underset{\sim}{\psi}_m^{+} T^2 g_m^{\,2} \underset{\sim}{\varphi}^2 \underset{\sim}{\psi}_m + \ldots \\
&\qquad + N^{-2} \sum_m \mathrm{Tr}\, T^2 g_m^{\,2} \underset{\sim}{\varphi}^2 - \ldots
\end{aligned}
\tag{A1}
$$

where z_j is defined in Eq. (3.12) and b_j in Eq. (3.5). The terms omitted contain higher powers of $\underset{\sim}{\varphi}$ than the second. We wish to show that the terms in $\underset{\sim}{\varphi}^2$ can be neglected when $\bar{E} \ll NMWT$, where $\bar{E}$ is the average energy received from the object during the observation interval (0, T). Calling those terms Z', we find for their expected value, from Eqs. (2.9), (2.15),

$$
\begin{aligned}
\underset{\sim}{E}(Z'|B) &= N^{-3} \sum_m T^3 g_m^{\,3}\, \mathrm{Tr}\, \underset{\sim}{\varphi}^3 = \\
&\quad N^{-3} \int_0^T \int_0^T \int_0^T \chi(t_1 - t_2)\, \chi(t_2 - t_3)\, \chi(t_3 - t_1)\; dt_1\, dt_2\, dt_3 \\
&\quad \times \int_A \int_A \int_A \varphi_s(\underset{\sim}{r}_1, \underset{\sim}{r}_2)\, \varphi_s(\underset{\sim}{r}_2, \underset{\sim}{r}_3)\, \varphi_s(\underset{\sim}{r}_3, \underset{\sim}{r}_1)\; d^2\underset{\sim}{r}_1 d^2\underset{\sim}{r}_2 d^2\underset{\sim}{r}_3 .
\end{aligned}
\tag{A2}
$$

When $WT \gg 1$, the integral involving the temporal coherence function is approximately equal to

$$T \int_{-\infty}^{\infty} [X(\omega)]^3 \, d\omega/2\pi,$$

where $X(\omega)$ is the spectral density defined in Eq. (1.3). By using Eqs. (2.5), (2.6), and (3.9) we can write the triple spatial integral as

$$\begin{aligned}
\mathrm{Tr}\, \underset{\sim}{\varphi}^3 &= (4\pi R^2)^{-3} \int_A \int_A \int_A \beta(\underset{\sim}{r}_1 - \underset{\sim}{r}_2)\, \beta(\underset{\sim}{r}_2 - \underset{\sim}{r}_3)\, \beta(\underset{\sim}{r}_3 - \underset{\sim}{r}_1)\, d^2\underset{\sim}{r}_1 d^2\underset{\sim}{r}_2 d^2\underset{\sim}{r}_3 \\
&= (4\pi R^2)^{-3} \int_0 \int_0 \int_0 \int_A \int_A \int_A B(\underset{\sim}{u}_1)\, B(\underset{\sim}{u}_2)\, B(\underset{\sim}{u}_3) \\
&\quad \times \exp\{-ik[(\underset{\sim}{r}_1 - \underset{\sim}{r}_2) \cdot \underset{\sim}{u}_1 + (\underset{\sim}{r}_2 - \underset{\sim}{r}_3) \cdot \underset{\sim}{u}_2 + (\underset{\sim}{r}_3 - \underset{\sim}{r}_1) \cdot \underset{\sim}{u}_3]/R\} \\
&\quad \times d^2\underset{\sim}{u}_1 d^2\underset{\sim}{u}_2 d^2\underset{\sim}{u}_3 d^2\underset{\sim}{r}_1 d^2\underset{\sim}{r}_2 d^2\underset{\sim}{r}_3 \\
&= (4\pi R^2)^{-3} A^3 \int_0 \int_0 \int_0 B(\underset{\sim}{u}_1)\, B(\underset{\sim}{u}_2)\, B(\underset{\sim}{u}_3) \\
&\quad \times \mathcal{J}(\underset{\sim}{u}_1 - \underset{\sim}{u}_2)\, \mathcal{J}(\underset{\sim}{u}_2 - \underset{\sim}{u}_3)\, \mathcal{J}(\underset{\sim}{u}_3 - \underset{\sim}{u}_1)\, d^2\underset{\sim}{u}_1 d^2\underset{\sim}{u}_2 d^2\underset{\sim}{u}_3. \qquad \text{(A3)}
\end{aligned}$$

The principal contribution to this integral comes from the mean radiance $\overline{B}$, which is constant over an area much broader than the width of the kernel $\mathcal{J}(\underset{\sim}{u})$. The functions $\mathcal{J}(\underset{\sim}{u}_2 - \underset{\sim}{u}_3)$ and $\mathcal{J}(\underset{\sim}{u}_3 - \underset{\sim}{u}_1)$ can therefore be replaced by $(\lambda R)^2 A^{-1} \delta(\underset{\sim}{u}_2 - \underset{\sim}{u}_3)$ and $(\lambda R)^2 A^{-1} \delta(\underset{\sim}{u}_3 - \underset{\sim}{u}_1)$ without much error, and Eq. (A3) is approximately

$$\begin{aligned}
\mathrm{Tr}\, \underset{\sim}{\varphi}^3 &= (4\pi R^2)^{-3} (\lambda R)^4 A \int_0 [B(\underset{\sim}{u})]^3 \, d^2\underset{\sim}{u} \\
&\cong (4\pi R^2)^{-3} (\lambda R)^4 A A_0 \overline{B}^3. \qquad \text{(A4)}
\end{aligned}$$

For the term of first order in $\underset{\sim}{\varphi}$, similarly

$$\underset{\sim}{E}(Z|B) = N^{-2} \sum_m T^2 g_m{}^2 \,\mathrm{Tr}\, \underset{\sim}{\varphi}^2 = N^{-2}(T/W)\, \mathrm{Tr}\, \underset{\sim}{\varphi}^2 \qquad \text{(A5)}$$

from Eq. (2.13), where by the same method as was used to obtain Eq. (A4)

$$\mathrm{Tr}\ \underset{\sim}{\varphi}^2 = (4\pi R^2)^{-2}\ (\lambda R)^2\ A \int_0 [B(\underset{\sim}{u})]^2\ d^2\underset{\sim}{u}$$

$$\cong (4\pi R^2)^{-2}\ (\lambda R)^2\ A\ A_o\ \overline{B}^2. \qquad (A6)$$

The ratio of the expected values of Z' and Z is, therefore, by Eqs. (A2), (A4), (A5), and (A6),

$$\underset{\sim}{E}(Z'|B)/\underset{\sim}{E}(Z|B) = N^{-1}(\lambda R)^2\ \overline{B}(4\pi R^2)^{-1}\ W \int_{-\infty}^{\infty} [X(\omega)]^3\ d\omega/2\pi. \qquad (A7)$$

For a unimodal spectral density $X(\omega)$ such as a rectangular or a Lorentz spectrum, the integral over ω in Eq. (A7) will be of the order of W^{-2}, and we find

$$\underset{\sim}{E}(Z'|B)/\underset{\sim}{E}(Z|B) = \overline{E}/NMWT = (\overline{B}/NW)\ (\lambda^2/4\pi), \qquad (A8)$$

where

$$\overline{E} = \overline{B}\ A_o\ A\ T/4\pi R^2 \qquad (A9)$$

is the average energy received from the object plane and M is the number of spatial degrees of freedom in the aperture field, given by Eq. (4.16).[16] Thus the terms with an extra factor $N^{-1}\ T\ g_m\ \underset{\sim}{\varphi}$ in the log-likelihood ratio, Eq. (A1), contribute on the average a fraction ($\overline{E}/NMWT$) of the main terms, and when $\overline{E} << NMWT$ we can neglect them. It can be shown that the error in the variance of Z so made is also of relative order $\overline{E}/NMWT$. The approximation made in Eq. (3.17) also requires neglecting terms of the form $N^{-1}\ T\ g_m\ \underset{\sim}{\varphi}$ relative to those proportional to the identity matrix $\underset{\sim}{I}$ and is valid whenever $\overline{E}/NMWT << 1$.

Appendix B. The Mutual Information

When the coefficients b_j in Eq. (3.14) have gaussian distributions with means $\bar{b}_j$ and covariances

$$\{b_k, b_\ell\} = \sigma_{k\ell}, \tag{B1}$$

the statistics z_j are also gaussian with means

$$\bar{z}_j = \bar{b}_j \, d_j^2,$$

according to Eq. (3.15), and covariances

$$\begin{aligned} \{z_k, z_\ell\} &= d_k^2 \, \delta_{k\ell} + d_k^2 \, d_\ell^2 \, \{b_k, b_\ell\} \\ &= d_k^2 \, \delta_{k\ell} + \sigma_{k\ell} \, d_k^2 \, d_\ell^2. \end{aligned} \tag{B2}$$

The entropy of the distribution of the z_j's is then

$$H(Z) = \frac{1}{2} \sum_k \ln(2\pi e) + \frac{1}{2} \ln \det \left| d_k^2 \, \delta_{k\ell} + \sigma_{k\ell} \, d_k^2 \, d_\ell^2 \right|. \tag{B3}$$

This is the second term in Eq. (4.8), which when combined with the first term yields for the mutual information

$$\begin{aligned} I(B; \psi_a) &= \frac{1}{2} \ln \det \left| \delta_{k\ell} + \sigma_{k\ell} \, d_\ell^2 \right| \\ &= \frac{1}{2} \operatorname{Tr} \ln(\underset{\sim}{I} + \underset{\sim}{\sigma} \, \underset{\sim}{D}_2) \end{aligned} \tag{B4}$$

as in Eq. (4.9), where

$$(\underset{\sim}{D}_2)_{k\ell} = d_k^2 \, \delta_{k\ell}. \tag{B5}$$

Define the matrix

$$\underset{\sim}{H}(z) = \underset{\sim}{\sigma} \, \underset{\sim}{D}_2 (\underset{\sim}{I} + z \, \underset{\sim}{\sigma} \, \underset{\sim}{D}_2)^{-1}; \tag{B6}$$

it is the solution of the matrix equation[25]

$$\underset{\sim}{H}(z) + z\, \underset{\sim}{\sigma}\, \underset{\sim}{D}_2\, \underset{\sim}{H}(z) = \underset{\sim}{\sigma}\, \underset{\sim}{D}_2 . \tag{B7}$$

In terms of it the mutual information is

$$I(B; \psi_a) = \frac{1}{2} \int_0^1 \mathrm{Tr}\, \underset{\sim}{H}(z)\, dz. \tag{B8}$$

By means of the function

$$H(\underset{\sim}{u}, \underset{\sim}{v}; z) = \sum_{k,\ell} \mathcal{B}_k(\underset{\sim}{u})\, H_{k\ell}(z)\, \mathcal{B}_\ell(\underset{\sim}{u}) \tag{B9}$$

and the orthonormality of the functions $\mathcal{B}_j(\underset{\sim}{u})$ we can write Eq. (B7) as

$$\begin{aligned} &H(\underset{\sim}{u}, \underset{\sim}{v}; z) \\ &\quad + \alpha z \int_0 \int_0 \sigma_B(\underset{\sim}{u} - \underset{\sim}{w})\, |\mathcal{G}(\underset{\sim}{w} - \underset{\sim}{x})|^2\, H(\underset{\sim}{x}, \underset{\sim}{v}; z)\, d^2\underset{\sim}{w}\, d^2\underset{\sim}{x} \\ &\quad = \alpha \int_0 \sigma_B(\underset{\sim}{u} - \underset{\sim}{w})\, |\mathcal{G}(\underset{\sim}{w} - \underset{\sim}{v})|^2\, d^2\underset{\sim}{w}, \end{aligned} \tag{B10}$$

where as in Eq. (4.12),

$$\alpha = A^2\, T^2/N^2 WT(4\pi R^2)^2 . \tag{B11}$$

Here we have used the expansion

$$|\mathcal{G}(\underset{\sim}{u} - \underset{\sim}{v})|^2 = \sum_j p_j{}^2\, \mathcal{B}_j(\underset{\sim}{u})\, \mathcal{B}_j(\underset{\sim}{v}), \tag{B12}$$

which follows from the integral equation, Eq. (3.10), and the orthonormality of the functions $\mathcal{B}_j(\underset{\sim}{u})$ over the object plane 0.

If we now assume that the object plane 0 is much greater than the region over which the kernel $|\mathcal{G}(\underset{\sim}{u})|^2$ is significant--this region is of the order of $\delta_x \times \delta_y$--, Eq. (B10) can be approximately solved by Fourier transforms, and

$H(\underset{\sim}{u}, \underset{\sim}{v}; z) = H'(\underset{\sim}{u} - \underset{\sim}{v}; z)$ is a function only of $\underset{\sim}{u} - \underset{\sim}{v}$. We introduce its Fourier transform

$$h(\underset{\sim}{r}; z) = \int H'(\underset{\sim}{u}; z) \exp(- i k \underset{\sim}{u} \cdot \underset{\sim}{r}/R) d^2\underset{\sim}{u}, \tag{B13}$$

which with Eqs. (3.9) and (4.13) permits the Fourier transform of Eq. (B10) to be written

$$h(\underset{\sim}{r}; z) [1 + \alpha z(\lambda R/A)^2 \Sigma_B(\underset{\sim}{r}) I_A^{(2)} (\underset{\sim}{r})]$$

$$= \alpha(\lambda R/A)^2 \Sigma_B(\underset{\sim}{r}) I_A^{(2)}(\underset{\sim}{r}) , \tag{B14}$$

where $I_A^{(2)}(\underset{\sim}{r})$ is defined in Eq. (4.3), $\Sigma_B(\underset{\sim}{r})$ in Eq. (4.13).

With $H(\underset{\sim}{u}, \underset{\sim}{v}; z)$ a function only of $\underset{\sim}{u} - \underset{\sim}{v}$, the trace in Eq. (B8) is

$$\mathrm{Tr}\ \underset{\sim}{H}(z) = \int_0 H(\underset{\sim}{u}, \underset{\sim}{u}; z) d^2\underset{\sim}{u} = A_o H'(\underset{\sim}{0}; z) =$$

$$A_o(\lambda R)^{-2} \int h(\underset{\sim}{r}; z) d^2\underset{\sim}{r} =$$

$$(\alpha/A^2) \int \Sigma_B(\underset{\sim}{r}) I_A^{(2)}(\underset{\sim}{r}) [1 + \alpha z(\lambda R/A)^2 \Sigma_B(\underset{\sim}{r}) I_A^{(2)}(\underset{\sim}{r})]^{-1} d^2\underset{\sim}{r} \tag{B15}$$

by Eq. (B14) and the inverse transform to Eq. (B13). This can now be substituted into Eq. (B8) and integrated over z to yield Eq. (4.11).

Footnote

* This research was carried out under Grant NGL 05-009-079 from the National Aeronautics and Space Administration.

References

1. A. Blanc-Lapierre, M. Perrot, G. Péri, Optica Acta 2, 1 (1955).
2. G. Toraldo di Francia, Optica Acta 2, 5 (1955).
3. E. H. Linfoot, J. Opt. Soc. Am. 45, 808 (1955).
4. P. B. Fellgett and E. H. Linfoot, Phil. Trans. Roy. Soc. (London), 247, no. A. 931, 369 (1955).
5. E. H. Linfoot, Optica Acta 5, 1 (1958).
6. E. H. Linfoot, Physica 24, 476 (1958).
7. R. C. Jones, J. Opt. Soc. Am. 50, 1166 (1960).
8. R. C. Jones, J. Opt. Soc. Am. 51, 1159 (1961).
9. R. C. Jones, J. Opt. Soc. Am. 52, 1193 (1962).
10. B. R. Frieden, J. Opt. Soc. Am. 58, 1105 (1968).
11. C. Shannon, Bell System Tech J. 27, 623 (1948).
12. D. Gabor, Progress in Optics, E. Wolf, ed. (North-Holland Publishing Co., Amsterdam) 1, 109 (1961).
13. R. C. Jones, J. Opt. Soc. Am. 52, 493 (1962).
14. E. Hisdal, J. Opt. Soc. Am. 55, 1446 (1965); 57, 35 (1967); 58, 977, 985 (1968); 59, 921 (1969).
15. C. W. Helstrom, J. Opt. Soc. Am. 59, 164 (1969).
16. C. W. Helstrom, J. Opt. Soc. Am. 60, 521 (1970).

17. C. W. Helstrom, Statistical Theory of Signal Detection, 2nd ed. (Pergamon Press, Oxford, 1968), pp. 193ff.

18. C. W. Helstrom, J. Opt. Soc. Am. 60, (1970).

19. Ref. 17, pp. 69ff.

20. E stands for expected value.

21. Ref. 17, pp. 124ff.

22. Ref. 15, p. 174

23. J. B. Thomas, An Introduction to Statistical Communication Theory (J. Wiley & Sons, Inc. New York, 1969), p. 560.

24. Ref. 23, pp. 566f.

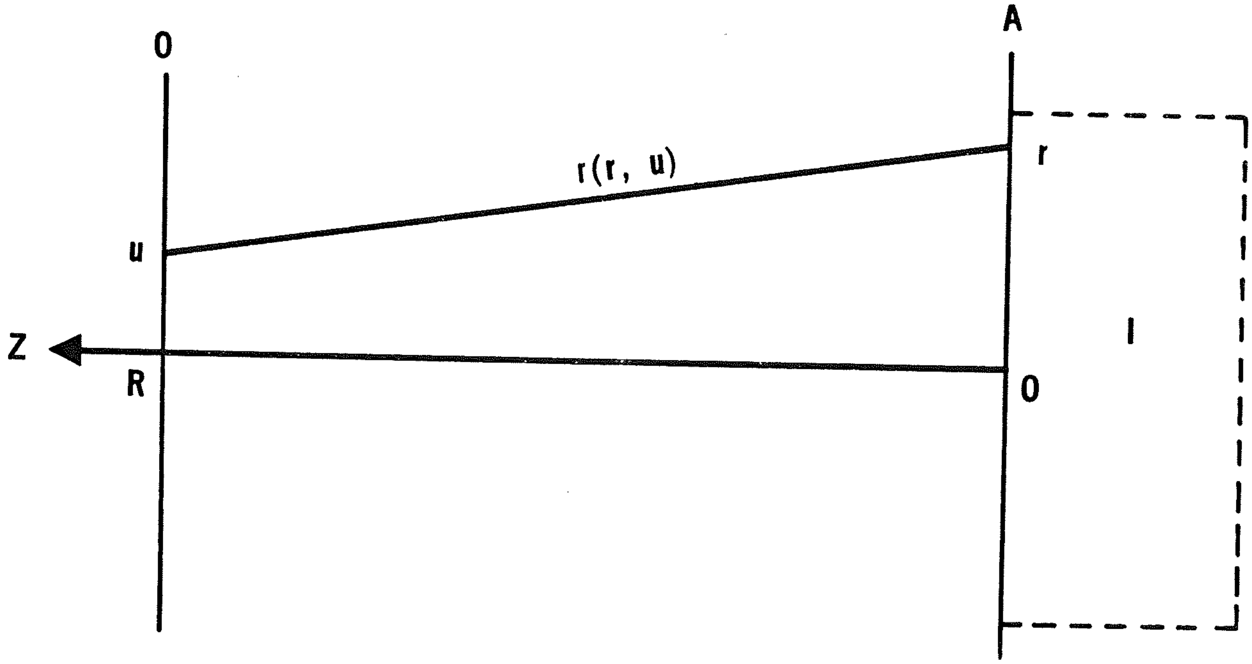
25. Ref. 17, pp. 394ff.

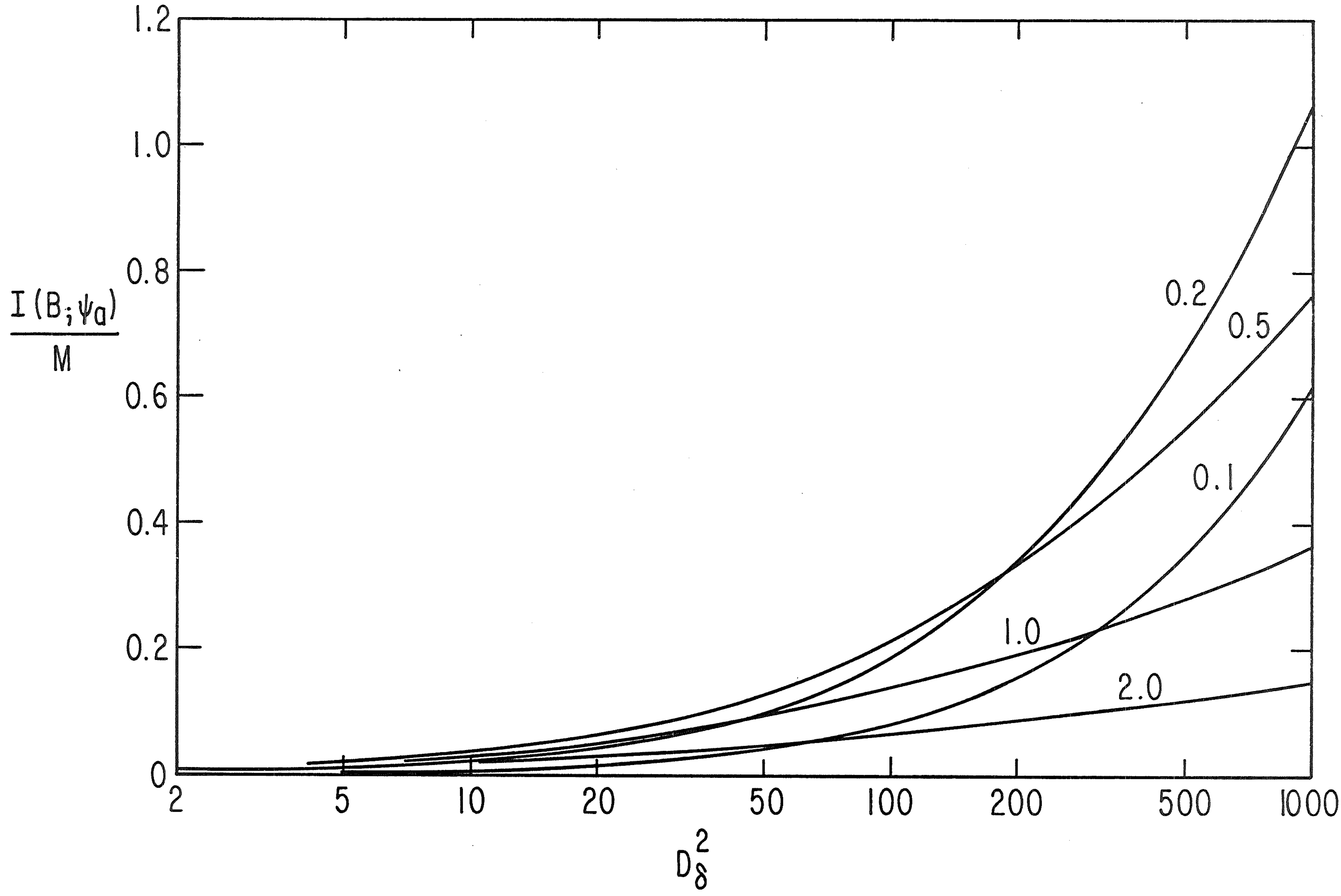
26. Ref. 17, pp. 219ff, 401f.

Figure Captions

Fig. 1. Geometrical configuration of object plane O and aperture plane A. Light from the object and the background falls on plane A from the left. I is an optical instrument processing the field $\psi_a(\underset{\sim}{r}, t)$ on plane A and functioning as a receiver.

Fig. 2. Mutual information $I(B; \psi_a)/M$ per spatial degree of freedom versus signal-to-noise ratio D_δ^2. Curves are labeled by the ratio Δ/δ, Δ measuring the sizes of object details, δ the resolution interval on the object plane.


O
A
r(r, u)
r
u
Z
I
R
O


1.2
1.0
0.8
0.6
0.4
0.2
0
I (B; ψa) / M
0.2
0.5
0.1
1.0
2.0
2
5
10
20
50
100
200
500
1000
D_δ^2