OPTICA ACTA, 1970, VOL. 17, NO. 1, 59–76

On image evaluation and enhancement†

LEO LEVI

Physics Department, The City College of the City University of New York, New York ‡

(*Received* 16 *December* 1968)

Abstract. Many techniques have been applied in image enhancement. Some of these are briefly reviewed here. Note that many of them may be classed under the general heading of 'linear filtering'.

To evaluate any image enhancement procedure, *quantitative* image assessment criteria are required. The criteria applicable to systems meant to convey information are reviewed here. It is noted that large areas of application can be covered by two, relatively simple, criteria: signal to noise ratio and mean squared deviation.

Image optimization is investigated from the point of view of least squared deviation with both linear filters and scale change considered. The latter is treated in some detail. The term 'image restoration' is proposed to cover efforts to minimize the squared deviation of the actual image from its ideal version.

1. Introduction

Whenever an image is used to convey information about an object, the object radiance pattern must be transformed into an image *irradiance* pattern, yielding the desired information§. To optimize this process it is desirable to operate on the object radiance in such a way as to maximize the information conveyed. Many techniques have been developed to accomplish this. The oldest and best-known of these is applied before the image is formed and is called lens design. More recently various techniques have been developed to operate on the image after it has been formed. These are generally referred to as *image enhancement*. Many of them can be classified as linear filtering techniques and it is these that concern us here. Both linear and non-linear filtering techniques may employ scanning to convert the primary image into an electric signal which may then be operated on by electric filters [1–3].

Linear optical techniques operating in the image plane are usually some version of the 'unsharp masking' technique in which a blurred replica of the image is subtracted from the sharp image. Among these are:

(1) Photographic techniques in which a blurred negative is printed in register with a sharp positive print [4, 5].

(2) 'Flurododge', in which a phosphor is excited by means of a sharp ultraviolet image and partially quenched by a blurred infra-red image [5, 6] (a similar photographic technique employing the Herschel effect to erase part of the exposure has also been used [7]).

† This work was supported, in part, by the Office of Naval Research.

‡ Present address: 435 Ft. Washington Ave., New York, N.Y.

§ We use the word 'information' here in its colloquial, qualitative sense. Later, when we use it in its quantitative, communication–theoretical sense, it will be put in quotation marks.

(3) A 'velocity modulation dodging' technique [8]. Here photographic printing takes place by means of a large scanning spot which reads the average transmittance over a sizeable primary image area; the velocity of this spot is controlled by the average transmittance thus obtained.

Because they operate in the image plane all these techniques must be dynamic or tailored to the specific image being enhanced. This difficulty may be avoided by operating in the aperture of the imaging system, where the spatial frequency spectrum can be operated on directly. Early efforts in this direction used a reduction of iris transmittance toward the edge of the aperture; these were called 'apodization' [9].

More recently extremely powerful optical filtering techniques have been developed [10–12], These are quite analogous to electronic filters except that they are spatial rather than temporal. Their spatial character does, of course, result in some important additional differences: spatial filters are generally two-dimensional and therefore somewhat more complicated than the one-dimensional time series of electric signals. At the same time, certain limitations inherent in the 'arrow of time' are thus eliminated : filters which are 'physically unrealizable' electrically, may be readily realizable in optics. Although such processing can be done to some extent with incoherent illumination, greater control is attained only with coherent illumination [11]. The advent of the laser which can provide the required levels of coherent light has given considerable impetus to these techniques and practical methods for generating such filters have been developed [11, 12]. Perhaps the most flexible of these is computer generation of filters which has been described in detail [13].

Non-linear techniques of image enhancement, too, have been used. Perhaps the best-known of these is photographic printing on high-contrast paper, which represents a 'super-linear' transformation over a limited signal amplitude range, with a 'sub-linear' transformation for signal amplitudes below and above this range. These sub-linear operations are known as 'thresholding' and 'clipping'. In 'digitizing', portions of the original signals, which may occur on a continuous scale, are grouped into a common output level [14]. The 'logetronic' technique [5, 15] is also non-linear. (This technique is similar to the velocity modulated dodging described above, under (3), except that the scanning spot *intensity* is controlled by the integrated transmittance—and this is done in a closed loop.) We make, however, no attempt to cover these here.

With such powerful techniques at our disposal the main problem becomes that of deciding what the transform should be, a decision which will depend to some extent on the particular application. Especially when the undesired background has a form significantly different from the desired signal, spatial filtering techniques become obvious. Such techniques have been used, for instance, to eliminate 'beam tracks' in bubble chamber photographs, while leaving the desired 'event tracks' intact [16]. Inversely, when one shape is to be selected from a large assemblage of shapes, all the others may be considered 'background' and a filter enhancing the desired shape might be optimum [17].

Such filters must be designed for each occasion as it arises. There remains, however, a large class of situations where the whole luminance pattern in the original object is of interest but has not been fully preserved in the available image. It is this class of situations with which we are primarily concerned here—for which we seek the optimum transformation. The first step in this direction is *image evaluation* to which the next section is devoted. An expression for the optimum linear transformation, including scale change, is developed in § 3 and pure scale change is treated in § 4.

2. Image evaluation

Considerable material has been published concerning image evaluation†. Since the variety of results and criteria may be confusing to the casual reader, we give here a brief summary of the various results and their areas of applicability.

The treatment is based on linear filtering theory. This is generally applicable in optical systems where, at most, the detector will introduce non-linearities. But, even here, many of the common detectors such as photo-multipliers and photographic emulsions (developed to $\gamma = 1$) closely approximate a linear response.

Concerning noise, we assume that it is strictly additive. This assumption is less generally valid. In photographic emulsions the granularity varies with density and in phototubes the shot noise is directly proportional to the square root of the signal. Allowing for such signal–noise interactions would require individual treatments for each of a number of cases and, furthermore, the resulting analyses would be far more complicated; this more general situation is therefore not treated here. We note, however, that for small signals, i.e. for signals near the detection threshold, noise usually is additive and that this does represent an interesting class of situations.

To permit a smoother presentation we first introduce symbols for the various quantities we must discuss. To treat two-dimensional distribution in a compact form, we use the vector notation both in the spatial and in the spatial frequency domains:

$$\mathbf{x} = x, y; \quad d^2\mathbf{x} = dx\,dy,$$
$$\boldsymbol{\nu} = \nu_x, \nu_y; \quad d^2\boldsymbol{\nu} = d\nu_x\,d\nu_y.$$

It is also convenient to introduce a 'pseudo-object' with radiance $s_0(x)$; this is actually the irradiance in an ideal image and is defined in terms of the actual object radiance $\hat{s}_0(\mathbf{X})$:

$$\hat{s}_0(\mathbf{x}) = k\hat{s}_0(\mathbf{x}/m_0),$$

where $\mathbf{X}$ and $\mathbf{x}$ represent the coordinates in the object and image planes, respectively, m_0 is the magnification in the imaging process, and k is a constant relating image irradiance to object radiance‡. The image irradiance is denoted by $s_i(\mathbf{x})$ and the irradiance noise in the image by $n_i(\mathbf{x})$.

Since optical systems are usually used with more than one object, any evaluation should include consideration of all of these. We do this by using average values, such as $\overline{s_i^2(\mathbf{x})}$, where the bar indicates the average over the whole

† The relationship between the various quantitative quality criteria and the subjective evaluation of image quality has been investigated extensively [18] but is irrelevant to the present investigation.

‡ In an optical system where the object and image spaces have the same refractive index and the relative aperture is not very large:

$$k = \tau A_E / b^2 = \pi\tau / 4F^2,$$

where τ is the transmittance of the optical media (taking account of both reflection and absorption losses), A_E is the area of the exit pupil, b is the distance from this pupil to the image, and F is the effective F/number in the image space.

'ensemble' of objects. (In practice it may not be practical—or even possible—to include all objects; it is then usually possible, however, to employ a large enough representative sample.)

We must refer to the point spread function, $P(\mathbf{x})$, which represents the irradiance in the image of a point object, normalized so that

$$\int P(\mathbf{x})\, d^2\mathbf{x} = 1. \tag{1}$$

The two-dimensional optical transfer function (otf) is:

$$T(\boldsymbol{\nu}) = \int P(\mathbf{x}) \exp(i2\pi\boldsymbol{\nu} \cdot \mathbf{x})\, d^2\mathbf{x}. \tag{2}$$

Note that we have defined $T(\boldsymbol{\nu})$ to be double-sided with $-\infty < \boldsymbol{\nu} < \infty$.

The signal amplitude spectra are denoted by capital letters:

$$S(\boldsymbol{\nu}) = \lim_{X, Y \to \infty} \frac{1}{4XY} \int_{-X}^{X} \int_{-Y}^{Y} s(\mathbf{x}) \exp(i2\pi\boldsymbol{\nu} \cdot \mathbf{x})\, d^2\mathbf{x}, \tag{3}$$

$$N(\boldsymbol{\nu}) = \lim_{X, Y \to \infty} \frac{1}{4XY} \int_{-X}^{X} \int_{-Y}^{Y} n(\mathbf{x}) \exp(i2\pi\boldsymbol{\nu} \cdot \mathbf{x})\, d^2\mathbf{x} \tag{4}$$

and the mean Wiener, or '*power*', spectra by subscripted W: (5)

$$W_{\text{o}}(\boldsymbol{\nu}) = \overline{|S(\boldsymbol{\nu})|^2}, \quad W_{\text{n}}(\boldsymbol{\nu}) = \overline{|N(\boldsymbol{\nu})|^2}. \tag{6}$$

When there is no danger of ambiguity, we occasionally omit the arguments x and ν and the limits of the integrals.

We now present a review of the communication–theoretical criteria which have been proposed and then assign to each its place in the general scheme of image evaluation.

(1) '*Information*' *capacity*

'Information' capacity in its communication–theoretical sense has a specific meaning which is briefly summarized in Appendix 1. This has a very simple form when the following common conditions are met: the noise is normally distributed, the signal 'power' is limited and the noise is stationary (i.e. has fixed statistics) over a sufficiently large area (cf. [19, p. 151]). The 'information' capacity can then be written:

$$H = 2A_{\text{i}} \int_{-\infty}^{\infty} \log\left[1 + \frac{W_{\text{o}}}{W_{\text{n}}}\right] d^2\boldsymbol{\nu}, \tag{7}$$

where A_{i} is the image area.

The capacity for carrying information has been used as a quality criterion—especially for detectors [20, 21].

(2) *Linfoot's figures of merit*

Linfoot proposed three quality criteria for optical imaging systems [22, 23]:

(*a*) *Correlation quality* measures the correlation between image and object:

$$Q_1 = \frac{\int \overline{s_{\text{o}} s_{\text{i}}\, d^2\mathbf{x}}}{\int \overline{s_{\text{o}}^{\,2}\, d^2\mathbf{x}}} = \frac{\int W_{\text{o}} T\, d^2\boldsymbol{\nu}}{\int W_{\text{o}}\, d^2\boldsymbol{\nu}}. \tag{8}$$

(*b*) *Relative structural content* measures the image 'power' relative to the object 'power':

$$Q_2 = \frac{\int \overline{s_i^2 d^2\mathbf{x}}}{\int \overline{s_o^2 d^2\mathbf{x}}} = \frac{\int W_o T^2 d^2\boldsymbol{\nu}}{\int W_o d^2\boldsymbol{\nu}}. \quad (9)$$

(*c*) *Fidelity* measures the closeness of correspondence of image with object:

$$Q_3 = 1 - \frac{\int \overline{(s_i s_o)^2 d^2\mathbf{x}}}{\int \overline{s_o^2 d^2\mathbf{x}}} = 1 - \frac{\int W_o(1-T)^2 d\boldsymbol{\nu}}{\int W_o d^2\boldsymbol{\nu}}. \quad (10)$$

Note that the correlation quality is the mean between the other two quantities:

$$Q_1 = \tfrac{1}{2}(Q_2 + Q_3). \quad (11)$$

In equations (8), (9) and (10) we have made use of Parseval's theorem which implies that the integral of a function squared equals the integral of its Fourier transform squared [19, p. 158].

(3) *Signal to noise ratio*

The signal to noise ratio:

$$R = \frac{\int \overline{s_i^2 d^2\mathbf{x}}}{\int \overline{n_i^2 d^2\mathbf{x}}} = \frac{\int W_o T^2 d^2\boldsymbol{\nu}}{\int W_n d^2\boldsymbol{\nu}}. \quad (12)$$

too, has been used as a quality indicator of an image [24–27].

When the detection process involves measuring individual area elements, the noise density may be represented by $\sigma^2/\Delta^2\mathbf{x}$, where σ is the standard deviation of the random irradiance fluctuations measured over the area element $\Delta^2\mathbf{x}$.

(4) *Mean squared deviation*

Another criterion is the square of the deviation of the output, signal plus noise, from the idealized output signal, integrated over the total image:

$$E = \int \overline{[s_o - (s_i + n_i)]^2} \, d^2\mathbf{x} = \int \overline{[S_o - (S_o T + N_i)]^2} \, d^2\boldsymbol{\nu}.$$

Assuming the noise to be independent of the signal, this becomes:

$$E = \int [W_o(1-T)^2 + W_n] \, d^2\boldsymbol{\nu} \quad (13)$$

This quantity, too, is very useful in image evaluation [28, 29].

(5) *Transcorrelation*

The correlation between the ideal and the actual image, too, has been used as a quality criterion. Specifically, the *transcorrelation* has been defined [30]†:

$$C = \frac{\dfrac{1}{A}\displaystyle\int \overline{s_i' s_o}\, d^2\mathbf{x} - \dfrac{1}{A^2}\int \overline{s_i'}\, d^2\mathbf{x} \int \overline{s_o}\, d^2\mathbf{x}}{\left\{\left[\dfrac{1}{A}\displaystyle\int \overline{s_o^2}\, d^2\mathbf{x} - \left(\dfrac{1}{A}\int \overline{s_o}\, d^2\mathbf{x}\right)^2\right]\left[\dfrac{1}{A}\displaystyle\int \overline{s_i'^2}\, d^2\mathbf{x} - \left(\dfrac{1}{A}\int s_i\, d^2\mathbf{x}\right)^2\right]\right\}^{1/2}}, \tag{14}$$

where s_i' is primed to indicate that it includes the noise in the image $s_i' = s_i + n_i$.

On writing the signal as the sum of its mean (μ) and deviation (δ):

$$s_o = \mu_o + \delta_o, \quad s_i' = \mu_i + \delta_i,$$

with

$$\int \delta_o\, d^2\mathbf{x} = \int \delta_i\, d^2\mathbf{x} = 0,$$

we readily find that equation (14) can be written:

$$C = \int \overline{\delta_i \delta_o}\, d^2\mathbf{x} \Big/ \sqrt{\left(\int \overline{\delta_i^2}\, d^2\mathbf{x} \int \overline{\delta_o^2}\, d^2\mathbf{x}\right)}. \tag{14a}$$

From Parseval's theorem this can be written in terms of the spectra:

$$C = \int \overline{S_o S_i'^*}\, d^2\boldsymbol{\nu} \Big/ \sqrt{\left(\int \overline{|S_o|^2}\, d^2\boldsymbol{\nu} \int \overline{|S_i'|^2}\, d^2\boldsymbol{\nu}\right)}. \tag{15}$$

Noting that $S_i' = TS_o + N_i$, and again assuming the noise independent of the signal, this yields:

$$C = \int W_o T\, d^2\boldsymbol{\nu} \Big/ \sqrt{\left(\int W_o\, d^2\boldsymbol{\nu} \int (W_o T^2 + W_n)\, d^2\boldsymbol{\nu}\right)}. \tag{15a}$$

Having presented the 'candidates', let us now evaluate them in conjunction with their task—'communication'. The general communication process has as its purpose the transmission of information. It can be broken up into three fundamental elements: (1) generation and coding; (2) transmission; and (3) reception, including decoding of the received signal. Once information has been generated, it can be lost, but not increased, by the other components of the system. (This is, of course, true by definition; but it is also true if we use 'information' in the specific sense in which it is defined in Appendix 1. If it were not, we would have to scrap that definition!) Thus, the best we can hope to do in phase (2) is to maintain the information constant or, if physical restrictions prevent this, to maximize the transmitted information. At the receiving end, too, information may be lost; but there we have two distinct tasks, detection and decoding, which we must separate conceptually even though they may actually be combined in a single physical process. Part of the confusion in connection with 'image enhancement' seems to stem from this combination of tasks: the detector must maximize *information* detected—the decoder must maximize *fidelity*. In order to evaluate systems in respect to these tasks we must first have quantitative definitions of them. The definition of 'information' given in Appendix 1 seems

† Because we do not consider temporal variations here I have dropped, in equation (14), the time dependence from the original expression.

to be satisfactory, having passed all the relevant tests. (We must, of course, be careful to distinguish between 'significance' and 'information'; the diary of a little girl may carry as much 'information' as an encyclopaedia.)

An appropriate criterion for 'fidelity' is less obvious. The decoding process depends on the range of input signal possibilities. On the one hand, if only a small number of binary decisions are necessary, high fidelity may be possible simply on the basis of these decisions. In that case, a weighted area integral of the total output will provide a number on which a decision concerning the input signal can be based, a different weighting function being required for each choice (cf. [31]). (The resolution of spectral lines or double stars falls into this category also.) Here the output may still be optical but 'fidelity' may no longer be defined in the general image sense—only the parameter of interest must enter the judgment of fidelity.

However, when the possible input signals are largely unrestricted, or continuously distributed, an attempt must be made to approximate them in the decoding—at least within some practical transformation. In such situations, the least squared deviation is a good candidate for 'fidelity'. It yields the best estimate of the signal in the following sense: assuming that nothing is known about signal and noise, except their mean Wiener spectra, and the fact that they have a Gaussian distribution, then the output filtered for least squared deviation represents the signal more likely responsible for the received output than any other signal. (The *a posteriori* input signal probability density is a maximum at the signal represented by the filtered output.) This criterion has been criticized [29] for giving equal weight to one large deviation and to many small ones, but this criticism does not seem to be obviously weighty.

Thus we may conclude that in transmission and in pure detection criterion (1) is appropriate and in decoding of continuous-range signals criterion (4) is appropriate. It is proposed that an effort to minimize the squared deviation be called 'image restoration' in contrast to the more general term 'image enhancement' covering all improvement techniques.

To understand the significance of the other criteria, we note that at the threshold of detectability (small signal to noise ratio) information capcity, as given by equation (7), approaches the signal to noise ratio, so that, in practice, the latter, criterion (3), may be used as a criterion of detectability [31].

When the system noise spectrum is uniform or totally unknown, Linfoot's 'relative structural content' (criterion 2 (b)) becomes a measure of signal to noise ratio. When noise is negligible, his 'fidelity' (criterion 2 (c)) becomes equivalent to the mean squared deviation (E) criterion. More specifically, his 'fidelity' is:

$$Q_3 = 1 - \epsilon,$$

where ϵ is the integrated squared deviation relative to the total input signal 'power'.

The specific significance of transcorrelation is less obvious. It does become unity in an ideal system and it is clearly a measure of how closely the system approaches the ideal. But it is not clear which aspect of the system is optimized when we maximize transcorrelation rather than, say, squared deviation. Note, however, that Linfoot's 'correlation quality' (criterion 2 (a)), which is the noise-free equivalent of transcorrelation, is simply the arithmetic mean between 'fidelity' and 'relative structural content'; this indicates for transcorrelation a role intermediate to optimizing information content and fidelity. It might, therefore, be appropriate for systems which must be used in both types of application.

3. Image enhancement

3.1. *General considerations*

Once criteria for evaluating a given image are available, optimization is the next step. As pointed out in the introduction, linear filtering is both convenient and a very powerful technique in such optimization [31 *a*]. It has been treated thoroughly in the theory of electric signals and analogous, two-dimensional techniques have been analysed extensively with many of the results from electric signal theory directly applicable [10, 11].

Thus, the filter maximizing the signal to noise ratio at a point $\mathbf{x}_1$ is simply the 'matched' filter [32, 33]:

$$T_M(\boldsymbol{\nu}) = kS_o^*(\boldsymbol{\nu}) \exp(i2\pi\boldsymbol{\nu}\cdot\mathbf{x}_1)/W(\boldsymbol{\nu}). \tag{16}$$

Note that this filter maximizes the signal to noise ratio at only one point, making it useful for the detection of a signal at a known location. It does not, in general, maximize the overall signal to noise ratio†. The filter minimizing the squared deviation is [25, 26]:

$$T_f = W_o T^*/(W_o T^2 + W_n). \tag{17}$$

Derivations of these results are readily available, but they are given again, in a simplified form, in Appendix 2 for the insight that these derivations give.

In the absence of noise ($W_n = 0$), the least squared deviation filter, equation (17) becomes:

$$T_f = T^{-1},$$

simply the reciprocal of the original imaging otf. This filter has been suggested to compensate for the degradation due to the original optical system [34], *ad hoc* improvements on it have been implemented [35, 36], and a relatively simple manner of constructing it has been described [37].

On the other hand, if the signal level is low compared to that of the noise and the signal spectrum is uniform over the range of significant otf (T), equation (17) approaches the matched filter (equation (16)) with $\mathbf{x}_1 = 0$ and the otf taking the place of S_o.

Note that formulae (16) and (17) apply to a given object spectrum. We would like to apply them to ensemble averages of such spectra to obtain a filter which will be optimum for that ensemble. This is a tempting step to take; but I cannot justify it. We are thus restricting ourselves for the present, to finding the optimum transformation of the 'average' object.

In addition to linear filtering, we treat here a linear scale change which is very simply implemented in optics, where it is called 'magnification'. This important

† Specification of the filter maximizing the overall signal to noise ratio would involve finding the frequency at which the signal to noise ratio attains its maximum value and then constructing a very narrow band-pass filter centred at that frequency. This was pointed out to the author by Professor A. Laemmel.

degree of freedom does not seem to have been treated thoroughly—at least not in communication theory terms (perhaps, because it is not readily applicable to electric signals).

In contrast to earlier treatments, we also consider here the effect due to the final detector—after the transformations have been completed. This detector will generally introduce both spectral and noise effects—it will act both as a filter and as a noise generation. In the following treatment we restrict the final detector only in that we assume linearity. We are, however, primarily interested in visual detection and eventually make the attempt to apply our results to this important detector. This attempt is based on the assumption that the human visual system responds linearly—an assumption which is open to serious challenge. It does seem, however, justified as a first step in a treatment which will, eventually, have to take account of the non-linearity in the visual system.

3.2. *Least squared error transformation*

In this section we find the filter characteristic and the magnification which will minimize the squared deviation of the detected signal (s_d) from the original object (s_o). In addition to the symbols defined in § 2 we introduce the following. The optical transfer functions (otf) of the original imaging system, the filter, and the detector are denoted by $T_o(\nu), T_f(\nu)$ and $T_d(\nu)$, respectively. For convenience we also introduce the total processing otf:

$$T_t(\nu, m) = T_f(\nu) T_d(\nu/m), \tag{18}$$

where m is the magnification relating image diameter at the detector to that at the original image plane.

The noise introduced by the detector, together with the unavoidable noise introduced by the filter, is denoted by $n_d(\mathbf{x}')$, its amplitude spectrum by $N_d(\nu')$ and its integrated mean Wiener spectrum by:

$$P_{dn} = \int \overline{|N_d(\nu)|^2}\, d^2\nu. \tag{19}$$

The optimizing magnification (m) is treated as part of the filtering operation. The reference scale is fixed in the plane of the original image ($\mathbf{x}$ plane or ν plane) and the filter transfer function is given referred to that plane. This requires, then, a scale change only in the detector spectra, which must be referred to the original image plane (see T_d in equation (18)). The general imaging, filtering, and detection process is illustrated in the block diagram, figure 1, which also summarizes nomenclature.

$$\text{Object} \xrightarrow[\text{[}1,\, m_o\text{] pseudo-object [}T_o,\, 1\text{]}]{\text{OPTICAL SYSTEM}} \text{image} \xrightarrow[\text{[}T_f,\, m\text{]}]{\text{FILTER}} \text{filtered image} \xrightarrow[\text{[}T_d, 1\text{]}]{\text{DETECTOR}} \text{detected image}$$

Figure 1. The brackets represent the transformations occurring at each stage. In each bracket, the first symbol represents the mtf and the second symbol the magnification.

It is assumed here that edge effects may be neglected. Consideration of these is not too difficult (cf. [29]) but very awkward and might obscure the basic considerations presented here.

Tracing the signal amplitude spectrum through the various stages we have:

$$\left.\begin{array}{lll} (1) & \text{object:} & S_o'(m_o\nu), \\ (2) & \text{`pseudo-object':} & S_o(\nu), \\ (3) & \text{original image:} & S_i'(\nu) = T_o(\nu)S_o(\nu) + N_i(\nu), \\ (4) & \text{filtered image:} & T_f(\nu)S_i'(\nu), \\ (5) & \text{detected image:} & S_d(\nu/m) = T_d(\nu/m)T_f(\nu)[T_o(\nu)S_o(\nu) \\ & & \qquad + N_i(\nu)] + N_d(\nu/m), \end{array}\right\} \tag{20}$$

where all the ν refer to the spatial frequency at the original image plane.

From equation (2) we can write immediately the expression for the integrated squared error:

$$\begin{aligned} E &= \int \overline{|S_o - S_d|^2}\, d^2\nu \\ &= \int \overline{|S_o(\nu) - T_t(\nu, m)[T_o(\nu)S_o(\nu) + N_i(\nu)] - N_d(\nu/m)|^2}\, d^2\nu \\ &= \int^* \overline{|S_o(\nu) - T(\nu, m)[S_o(\nu) + \hat{N}_i(\nu)] - N_d(\nu/m)|^2}\, d^2\nu, \end{aligned} \tag{21}$$

where we have written:

$$T(\nu, m) = T_t(\nu, m)T_o(\nu) \tag{22}$$

and

$$\hat{N}_i(\nu) = N_i(\nu)/T_o(\nu) \tag{23}$$

and the integral extends only over the regions in which T_o exists.

Assuming, again, that N_i, N_d are independent of S_o and each other, this becomes:

$$E = \int \overline{|S_o(\nu) - T(\nu, m)[S_o(\nu) + N_i(\nu)]|^2}\, d^2\nu + mP_{nd}, \tag{24}$$

where the integral of $N_d(\nu/m)$ was evaluated by means of a simple transformation and reference to equation (19).

Clearly, the last term adds an amount to the error which is independent of T; hence the filter which minimizes the integral in equation (24) also minimizes E. Comparison of this integral with equation (A 4) in Appendix 2 shows that it is in the exact form for which the least squared error filter was derived. This filter, thus, has the form:

$$T = T_o T_{to} = \frac{W_o}{W_o + |\hat{N}_d|^2} = \frac{|T_o|^2 W_o}{|T_o|^2 W_o + W_n} \tag{25}$$

and the optimum total processing otf:

$$T_{to} = \frac{T_o^* W_o}{|T_o|^2 W_o + W_n}, \tag{26}$$

where the asterisk indicates the complex conjugate.

Reference to equation (18) shows that the optimum filter otf alone is:

$$T_{\mathrm{fo}}(\boldsymbol{\nu}) = \frac{T_{\mathrm{o}}^*(\boldsymbol{\nu})W_{\mathrm{o}}(\boldsymbol{\nu})/T_{\mathrm{d}}(\boldsymbol{\nu}/m)}{|T_{\mathrm{o}}(\boldsymbol{\nu})|^2 W_{\mathrm{o}}(\boldsymbol{\nu}) + W_{\mathrm{n}}(\boldsymbol{\nu})}, \quad T_{\mathrm{d}}(\boldsymbol{\nu}/m) \neq 0$$
$$= 0, \qquad T_{\mathrm{d}}(\boldsymbol{\nu}/m) = 0. \tag{27}$$

The resulting minimum integrated squared error is readily obtained from the final result of Appendix 2:

$$E_{\min} = \int \frac{W_{\mathrm{o}}|\hat{N}_{\mathrm{i}}|^2}{W_{\mathrm{o}} + |\hat{N}_{\mathrm{i}}|^2}\, d^2\boldsymbol{\nu} + mP_{\mathrm{dn}}$$
$$= \int \frac{W_{\mathrm{o}}W_{\mathrm{n}}}{|T_{\mathrm{o}}|^2 W_{\mathrm{o}} + W_{\mathrm{n}}}\, d^2\boldsymbol{\nu} + mP_{\mathrm{dn}}. \tag{28}$$

The expression for the minimum squared error has two terms. The first of these is independent of m and the second is directly proportional to it. Hence m should be as small as possible for best signal estimation. Since it implies a small image on the detector this conclusion may seem strange on first sight—on two counts:

(1) In view of the tendency of otf's to decrease with increasing spatial frequency, we expect the detector to degrade a small-scale image more than a large-scale one.

(2) The same signal image, when spread over a larger area, will permit averaging the random fluctuations over a larger area and should, therefore, yield superior imaging in terms of signal to noise ratio and, therefore, also fidelity.

Both of these arguments are irrelevant, however. In the system model under consideration, the degradation due to $T_{\mathrm{d}}(\boldsymbol{\nu}')$ is fully compensated for in the filtering otf (cf. equations (18), (27)); thus the quality of the signal at the detector is independent of the scale. It is then also clear that the reduced detector area, implied by the smaller m, will cause less noise to be detected. The possibility of averaging over a larger area is irrelevant here, since this is part of the decoding process and the decoding has been completed earlier in the system.

Although m does not enter the first term of equation (28) explicitly, it does affect it indirectly because of its effect on the filter. Especially the limits on this filter must be considered carefully.

By way of illustration we given here a method for finding the limit on m when the otf's are one-dimensional (or radially symmetric) and have cut-off frequencies (ν_{c}) beyond which they vanish. If the detector cut-off frequency is less than that of the filter ($m\nu_{\mathrm{cd}} < \nu_{\mathrm{co}}$), equation (27) may no longer be optimum. Thus we obtain the following limitation on m:

$$m \geqslant \nu_{\mathrm{co}}/\nu_{\mathrm{cd}}, \tag{29}$$

where ν_{co} denotes the cut-off frequency of $T_{\mathrm{o}}S_{\mathrm{o}}$.

In practice, it may be desirable to impose an upper limit, L, on the absolute value of T_{fo}, i.e.

$$T_{\mathrm{d}}(\nu/m) \geqslant \frac{1}{L}\, T_{\mathrm{to}}(\nu) \quad \text{for all } \nu < \nu_{\mathrm{co}}. \tag{30}$$

To find the limitation on m implied by this condition we must evaluate:

$$m = \frac{\nu_t(LT)}{\nu_d(T)}, \quad \nu_t < \nu_{co} \tag{31}$$

for all $T < 1/L$. Here $\nu_i(T_1)$ denotes the value of ν for which $T_i(\nu) = T_1$. The maximum value of m occurring within the indicated ranges represents the lower limit on m (see figure 2).

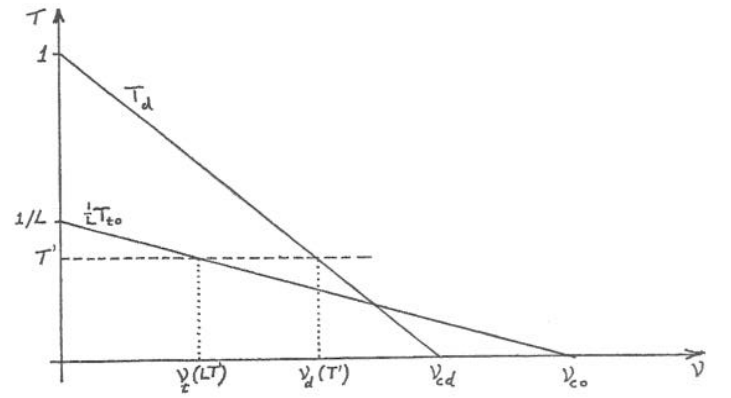

Figure 2. Superimpose plots of $T_d(\nu)$ and $(1/L)T_{to}(\nu)$ and draw a horizontal line corresponding to any $T < 1/L$. The intersections of this line with the two plots yield, respectively, $\nu_d(T)$ and $\nu_t(LT)$ and, hence, $m(T) = \nu_t(LT)/\nu_d(T)$.

4. Pure scale change

In this section we treat the image 'enhancement'—or better optimization—possible by means of a simple scale change, without any other processing. In other words, given a 'noisy' picture and a detector with a given otf, at what magnification should the image be formed on the detector to optimize reception ?

Except for the pioneering work of Selwyn [38] very little work seems to have been done on this problem. Selwyn attempted to find the visual magnification which optimizes resolution of a photographically recorded sine wave pattern. Even though this work was done before the communication theoretic techniques had been developed for optics, his method was remarkably sophisticated even by today's standards. We present here the more general treatment which has now become feasible.

We return to equation (24), where we must set:

$$T_f \equiv 1$$

and thus obtain for the integrated squared error:

$$E = \int \{W_o(\mathbf{v}) + T_d^2(\mathbf{v}/m)[T_o^2(\mathbf{v})/W_o(\mathbf{v}) + W_n(\mathbf{v})] - 2T_d(\mathbf{v}/m)T_o(\mathbf{v})W_o(\mathbf{v})\} d^2\mathbf{v} + mP_{dn}. \tag{32}$$

Since we have eliminated the possibility of filtering, m remains in the integral and we must differentiate E with respect to m to find its minimum. Denoting $(d/d\nu)T(\mathbf{v})$ by $T'(\mathbf{v})$, we find, on differentiating under the integral sign:

$$\frac{dE}{dm} = -\frac{2}{m^2}\int \{T_d(\mathbf{v}/m)[T_o^2(\mathbf{v})W_o(\mathbf{v}) + W_n(\mathbf{v})] - T_o(\mathbf{v})W_o(\mathbf{v})\}T_d'(\mathbf{v}/m)\mathbf{v}\, d^2\mathbf{v} + P_{dn}. \tag{33}$$

This equation must be equated to zero and solved for m in order to find the optimum magnification.

To illustrate the use of this equation we apply it to a one-dimensional purely sinusoidal object:

$$W_o(\boldsymbol{\nu}) = W\delta(\nu_x - \nu_o)\delta(\nu_y), \tag{34}$$

where W is a constant, and δ is the Dirac delta 'function'.

Writing $T_o(\nu_o) = \tau$, we find then

$$\frac{dE}{dm} = -\frac{2}{m^2}\left\{\tau W[\tau T_d(\nu_o/m) - 1]T_d'(\nu_o/m)\nu_o + \int T_d(\boldsymbol{\nu}/m)W_n(\boldsymbol{\nu})\right\} T_d'(\boldsymbol{\nu}/m)\boldsymbol{\nu}\, d\boldsymbol{\nu} + P_{dn}. \tag{35}$$

If we can assume that $W_n(\boldsymbol{\nu}) = W_n$ is constant over the frequency range over which $T_d(\boldsymbol{\nu}/m)$ is significant we can remove it from the integral and integrate the remainder by parts:

$$\int T_d(\boldsymbol{\nu}/m)T_d'(\boldsymbol{\nu}/m)\boldsymbol{\nu}\, d\boldsymbol{\nu} = -\frac{m^2}{2}\int T_d^2(\boldsymbol{\nu})\, d\boldsymbol{\nu} = -\frac{m^2}{2}P_{dT}. \tag{36}$$

The last integral represents the 'equivalent passband' [39] of the detector, which we denote by P_{dT}. We can thus write the condition on m for minimizing the integrated squared error:

$$\frac{dE}{dm} = \frac{2\tau W\nu_o}{m^2}[1 - \tau T_d(\nu_o/m)]T_d'(\nu_o/m) + P_{dT}W_n + P_{dn} = 0. \tag{37}$$

Note that the optimum value of m depends on the total noise 'power' and on τ (i.e. T_o). Under these same conditions equation (32) for the integrated squared error becomes:

$$E = W[1 - \tau T_d(\nu_o/m)]^2 + m[W_n P_{dT} + P_{dn}]. \tag{38}$$

Instead of minimizing the integrated squared error we may wish to maximize the signal to noise ratio:

$$R = \int T_d^2(\boldsymbol{\nu}/m)T_o^2(\boldsymbol{\nu})W_o(\boldsymbol{\nu})\, d\boldsymbol{\nu} \Big/ \int [T^2 d(\boldsymbol{\nu}/m)W_n(\boldsymbol{\nu}) + \overline{|N_d(\boldsymbol{\nu}/m)|^2}]\, d^2\boldsymbol{\nu}. \tag{39}$$

Again assuming a sinusoial object (equation (34)) and a constant image noise spectrum, W_n, this becomes:

$$R_s = \tau^2 W T^2{}_d(\nu_o/m)/[W_n P_{dT} + P_{dn}]. \tag{40}$$

Maximizing this with respect to m is equivalent to maximizing:

$$r = \frac{1}{m}T_d^2(\nu_o/m).$$

On setting

$$\frac{dr}{dm} = 0,$$

we find:

$$\frac{T'd(\nu_o/m)}{T_d(\nu_o/m)} = -\frac{m}{2\nu_o}.$$

On denoting the spatial frequency on the detector by:

$$\nu_o' = \nu_o/m, \tag{41}$$

the condition becomes simply:

$$2\nu_o' \frac{T_d'(\nu_o')}{T_d(\nu_o')} = -1 \tag{42}$$

This is identical to the result obtained by Selwyn [38, equation (17)] on the basis of a rather complicated detection mechanism.

To make the illustration even more concrete, we assume now an otf of the form:

$$T_d(\nu') = \exp(-|\nu'|/c), \tag{43}$$

which is an approximation valid for a number of photographic emulsions [40] as well as for the human visual system at the higher spatial frequencies [41]. When this is substituted into equation (42), we find:

$$\nu_o' = \tfrac{1}{2}c. \tag{44}$$

From recently published data of the human visual system [42], the value of c is equal to about 32 cycles/mm. Thus the optimum spatial frequency on the retina is about 16 cycles/mm.

This seems in satisfactory agreement with Selwyn's experimental finding of 18·4 cycles/mm†—a finding which was only a rough estimate. (His theoretical work called for an optimum at 57·4 cycles/mm.)

The value of 16 cycles/mm is valid for a situation in which the noise is effectively 'white'. It is not related to the fact that the human visual system has a peak sensitivity at about 35 lines/mm. This fact implies that our approximation (43) is no longer valid at this low frequency and explains the upward shift in the observed optimum frequency.

5. Conclusions

The form of the optimum filter depends not only on the spectral characteristics of the class of objects considered, but also on the 'decoding' to be employed. When only a very limited number of objects are to be differentiated, optimum detection techniques, based on decision theory [31] can be tailored to the particular objects involved. If, on the other hand, the objects are largely unrestricted, least squared deviation techniques become advisable.

The filter minimizing the integrated squared deviation for a quite general optical imaging system can be stated in relatively simple form.

When the human visual system is used as the detector, the analysis predicts correctly the results found when a sinusoidal object is to be detected.

Acknowledgments

The helpful discussions and valuable constructive criticism of Professor Arthur Laemmel are here gratefully acknowledged, as is the encouragement received from Dr. Tolhurst, chief of the Physiological Psychology Branch of ONR.

† Selwyn's result was 1·25 cycles/mm at a normal viewing distance of 250 mm. Assuming an effective ocular focal length of 17 mm, this corresponds to 18·4 cycles/mm on the retina.

Appendix 1

Quantity of information

Since it is now well established, we give here only a very brief introduction to the concept of 'information'.

If reception of a signal changes the knowledge of the receiver about some subject, the signal is said to have conveyed information. More specifically, consider a subject having attribute A. Now let P_i represent the probability that the receiver assigned to this fact initially and P_f the probability he assigns to it after receiving the signal. Then

$$H = \log (P_f / P_i) \tag{A 1}$$

is a measure of the information conveyed by the signal by virtue of this change.

The following considerations explain the use of the logarithmic function in equation (A 1).

When quantities of information are received about a number of independent attributes or objects, we expect the total quantity received to equal the sum of the individual quantities. At the same time, the probabilities must be multiplied together to yield the probability of the combination. By using the logarithm, this multiplication is converted into the desired addition.

Note also that if the received signal is misleading, i.e. it is such that the probability assigned to A decreases as a result of the signal, then $P_f < P_i$ and the information becomes negative. This again corresponds to the intuitive notion that the receiver is less well informed as a result of receiving the signal.

The definition (A 1) is very similar in form to the physical concept of entropy; the physical connection between 'information' and entropy has been investigated extensively [43, 44].

Appendix 2

The matched and least squared deviation filters

Denote the irradiance in the 'pseudo-object' and in the final image by $s_o(x)$, $s_f(x)$ and their amplitude spectra by $S_o(\nu)$, $S_f(\nu)$, respectively, the noise irradiance and its amplitude spectrum by $n(x)$ and $N(\nu)$, the mean Wiener spectra corresponding to s_o and n by $W_o(\nu)$ and the filter function by $T(\nu) = T_F(\nu) \exp [it(\nu)]$, where T_F and t are both real. Note that

$$W_o = \overline{S_o^2}, \quad W_n = \overline{N^2}. \tag{A 1}$$

We now seek the filter function maximizing the signal to noise ratio at point x_1. To do this we write the signal level as the Fourier transform of the signal spectrum and, assuming the noise to be stationary, the noise-level-squared proportional to the integrated noise spectrum:

$$R = \frac{s_f(x_1)^2}{n(x_1)^2} = \frac{\left| \int S_o T \exp (i2\pi\nu \, . \, x_1) \, d\nu \right|^2}{k \int W_n T_F^2 \, d\nu}. \tag{A 2}$$

We do this for the case where $W_n(\nu) = W_n$ is constant over the range of T_F. Then the filter function maximizing R will be identical with the one maximizing:

$$\left| \int S_o T \exp (i2\pi\nu \, . \, x_1) \, d\nu \right|^2 \Big/ \left[\int W_o \, d\nu \int T_F^2 \, d\nu \right]. \tag{A 2a}$$

Now, combining the triangle and Schwarz inequalities:

$$\left|\int f(x)g(x)\,dx\right|^2 \leq \left[\int |f(x)g(x)|\,dx\right]^2 \leq \int |f(x)|^2\,dx \int |g(x)|^2\,dx,$$

with the left-hand member attaining its maximum value—namely equality with the right-hand member—when $f(x)=cg^*(x)$. Applying this result to equations (A 2), we find that R is maximum when

$$T=S_o{}^*\exp(-i2\pi\nu.x) \tag{A 3}$$

and that this is the filter function maximizing the signal to noise ratio at x_1.

The filter minimizing the squared deviation can be derived readily even for a general noise spectrum. Note that

$$S_f=(S_o+N)T. \tag{A 4}$$

Therefore the integrated mean squared deviation may be written:

$$E=\int \overline{(s_o-s_F)^2}\,d^2\mathbf{x}=\int \overline{(S_o-S_F)^2}\,d^2\boldsymbol{\nu}$$

$$=\int \overline{|S_o(1-T)-NT|^2}\,d^2\boldsymbol{\nu}. \tag{A 5$'$}$$

If signal and noise are totally uncorrelated:

$$E=\int (\overline{|S_o|^2}\,|1-T|^2+\overline{|N|^2}\,|T|^2)\,d^2\boldsymbol{\nu}$$

$$=\int [W_o(1+T_F{}^2-2T_F\cos t)+W_nT_F{}^2]\,d^2\boldsymbol{\nu}.$$

For minimum E, $\cos t$ must equal unity and

$$E=\int [(W_o+W_n)T_F{}^2-2W_oT_F+W_o]\,d^2\boldsymbol{\nu}.$$

Completing the square of the terms involving T_F:

$$E=\int\left[\left(\sqrt{(W_o+W_n)}T_F-\frac{W_o}{\sqrt{(W_o+W_n)}}\right)^2+\left(W_o-\frac{W_o{}^2}{W_o+W_n}\right)\right]d^2\boldsymbol{\nu}$$

$$=\int\left[\left(\sqrt{(W_o+W_n)}T_F-\frac{W_o}{\sqrt{(W_o+W_n)}}\right)^2+\frac{W_oW_n}{W_o+W_n}\right]d^2\boldsymbol{\nu}. \tag{A 5}$$

The term in parentheses—the only part of the integrand depending on T—cannot be negative, therefore E will be a minimum when that term vanishes, yielding for the desired filter function:

$$T_F=\frac{W_o}{W_o+W_n}. \tag{A 6}$$

When this condition is met, the mean squared deviation will have its minimum value as given by the last term of the integral of equation (A 5):

$$E_{\min}=\int \frac{W_oW_n}{W_o+W_n}\,d^2\boldsymbol{\nu}. \tag{A 7}$$

If the signal has passed a filter T_o before the noise is added, equation (A 4) will become:

$$S_F = (S_o T_o + N)T.$$

We may write this in the original form:

$$S_F = (S_o + \hat{N})T_T, \quad T_o \neq 0,$$

by putting $T_T = T_o T$ and $\hat{N} = N/T_o$, $T_o \neq 0$. With this notation the previous analysis applies and equation (A 6) becomes:

$$T_{TF} = \frac{W_o}{W_o + (W_n/T_o T_o^*)} = T_F T_o$$

and

$$T_F = \frac{W_o}{W_o T_o + W_n/T_o^*} = \frac{W_o T_o^*}{W_o|T_o|^2 + W_n}, \quad T_o \neq 0,$$

$$T_F = 0, \qquad T_o \neq 0.$$

Similarly, equation (7) becomes:

$$E_{\min} = \int^* \frac{W_i W_n}{W_i + W_n} d^2\nu,$$

where we have written W_i for the Wiener spectrum of the intermediate image:

$$W_i = W_o|T_o|^2$$

and the integral extends only over the regions where $T_o \neq 0$.

Beaucoup de techniques ont été appliquées pour le renforcement des images. On passe rapidement en revue certaines d'entre elles. Il convient de remarquer que beaucoup d'entre elles peuvent être classées sous la rubrique ' filtrage linéaire '.

Afin d'évaluer un procédé de restitution de l'image, il faut avoir des critères quantitatifs pour la qualité de celle-ci. On passe en revue les critères applicables aux systèmes destinés à transmettre l'information. On remarque que beaucoup de domaines d'application peuvent être couverts par deux critères relativement simples: rapport signal sur bruit et écart quadratique moyen.

On examine l'optimisation de l'image du point de vue du moindre carré de l'écart lorsqu'on considère à la fois des filtres linéaires et un changement d'échelle. Ce dernier est traité assez en détail. On propose le terme ' restauration de l'image ' pour désigner les efforts pour minimiser l'écart quadratique de l'image par rapport à sa version idéale.

Es sind verschiedene Verfahren zur Bildverbesserung beschrieben worden. Einige davon werden hier kurz zusammengefasst. Viele davon fallen unter das Stichwort ' lineare Filterung '.

Zur Beurteilung der Verfahren zur Bildverbesserung bedarf es eines quantitativen Bewertungskriteriums. Über die Kriterien, die auf Systeme zur Informationsübertragung angewendet werden können, wird ein Überblick gegeben. Man stellt fest, dass ein weiter Anwendungsbereich von nur zwei, verhältnismässig einfachen Kriterien gedeckt wird: dem Verhältnis vom Signal zum Rauschen und die Abweichung vom quadratischen Mittel.

Mittels der Methode der kleinsten quadratischen Abweichung wird die Bildoptimierung behandelt für die zwei Fälle der linearen Filterung und der Massstabsänderung. Die letztere wird ausführlich untevsucht. Der Ausdruck ' Bildverbesserung ' sollte dem Bestreben, die quadratische Abweichung des tatsächlichen Bildes von der Vorlage klein zu machen, vorbehalten bleiben.

REFERENCES

[1] GOLDMARK, P. C., and HOLLYWOOD, J. M., 1951, *Proc. Inst. Radio Engrs*, **39**, 1314.
[2] KOVASZNAY, L. S. G., and JOSEPH, H. M., 1955, *Proc. Inst. Radio Engrs*, **43**, 560.
[3] GRAHAM, R. E., 1962, *I.R.E. Trans. Inf. Theory*, **8**, 129.
[4] SPIEGLER, G., and JURIS, K., 1931, *Photogr. Korr.*, **67**, 4.
[5] LOHMANN, A., 1959, *Optica Acta*, **6**, 319.
[6] CLARK, A. B., 1962, *Photogramm. Engng*, **28**, 393.
[7] KELLY, D. H., 1961, *J. opt. Soc. Am.*, **51**, 1095.
[8] PARGAS, P., 1965, *Photogr. sci. Engng*, **9**, 219.
[9] JACQUINOT, P., and ROIZEN-DOSSIER, B., 1964, *Progress in Optics*, Vol. 3, edited by E. Wolf (Amsterdam: North-Holland), p. 29.
[10] FRANÇON, M., 1963, *Modern Applications of Physical Optics* (New York: Wiley).
[11] VANDER LUGT, A., 1968, *Optica Acta*, **15**, 1.
[12] TSUJUCHI, J., 1963, *Progress in Optics*, Vol. 2, edited by E. Wolf (Amsterdam: North-Holland), p. 133.
[13] LOHMANN, A. W., and PARIS, D. P., 1967, *Appl. Optics*, **6**, 1739.
[14] RICHARDS, G. P., and BISIGNANI, W. T., 1967, *Proc. Inst. elect. electron. Engrs*, **55**, 1707.
[15] CRAIG, D. R., 1954, *Photogr. Engng*, **5**, 219.
[16] FALCONER, D. G., 1966, *Appl. Optics*, **5**, 1365.
[17] VANDER LUGT, A., ROTZ, F. B., and KLOSTER, A., Jr., 1965, *Optical and Electro-optical Processing*, edited by J. T. Tippett, D. A. Berkowitz, L. C. Clapp, C. J. Koester and A. Vanderburgh, Jr. (Cambridge, Massachusetts: M.I.T.), p. 125.
[18] MEES, C. E. K., and JAMES, T. H., 1966, *The Theory of the Photographic Process*, third edition (New York: Macmillan), p. 538.
[19] LEVI, L., 1968, *Appl. Optics* (New York : Wiley).
[20] FELLGETT, F. B., and LINFOOT, E. H., 1955, *Phil. Trans. R. Soc.* A, **247**, 369.
[21] JONES, R. C., 1962, *J. opt. Soc. Am.*, **52**, 1193.
[22] LINFOOT, E. H., 1956, *J. opt. Soc. Am.*, **46**, 740.
[23] LINFOOT, E. H., 1958, *Optica Acta*, **5**, 1.
[24] HARRIS, J. L., 1964, *J. opt. Soc. Am.*, **54**, 606.
[25] RINDFLEISCH, T., and WILLINGHAM, D., 1966, *Adv. Electronics Electron. Phys.* A, **22**, 341.
[26] ROETLING, P. G., TRABKA, E. A., and KINZLEY, R. E., 1968, *J. opt. Soc. Am.*, **58**, 342.
[27] KINZLEY, R. E., MAZUROWSKI, M. J., and HOLLADAY, T. M., 1968, *Appl. Optics*, **8**, 1577.
[28] HELSTROM, C. W., 1967, *J. opt. Soc. Am.*, **57**, 297.
[29] SLEPIAN, D., 1967, *J. opt. Soc. Am.*, **57**, 918.
[30] MACMILLAN, R. S., and YOUNG, G. O., 1968, *J. opt. Soc. Am.*, **58**, 346.
[31] HELSTROM, C. W., 1964, *I.E.E.E. Trans. Inf. Theory*, **10**, 275.
[31 a] FRIEDEN, B. R., 1968, *J. opt. Soc. Am.*, **58**, 1272.
[32] TURIN, G. L., 1960, *I.R.E. Trans. Inf. Theory*, **6**, 311.
[33] KOZMA, A., and KELLY, D. L., 1965, *Appl. Optics*, **4**, 387.
[34] ELIAS, P., GREY, D. S., and ROBINSON, D. Z., 1952, *J. opt. Soc. Am.*, **42**, 127.
[35] HARRIS, J. L., Sr., 1966, *J. opt. Soc. Am.*, **56**, 569.
[36] MCGLAMERY, B. L., 1967, *J. opt. Soc. Am.*, **57**, 293.
[37] STROKE, G. W., 1968, *Photogr. Korr.*, **104**, 82.
[38] SELWYN, E. W. H., 1948, *Photogr. J.* B, **88**, 46.
[39] SCHADE, O. H., 1954, *Circ. U.S. natn. Bur. Stand.*, No. 526, p. 231.
[40] LEWIS, N. W., 1962, *J. opt. Soc. Am.*, **52**, 948.
[41] VAN NES, F. L., and BOUMAN, M. A., 1967, *J. opt. Soc. Am.*, **57**, 401.
[42] CAMPBELL, F. W., 1968, *Proc. Inst. elect. electron. Engrs*, **56**, 1009.
[43] BRILLOUIN, L., 1962, *Science and Information Theory*, second edition (New York: Academic Press).
[44] GABOR, D., 1961, *Progress in Optics*, Vol. 1, edited by E. Wolf (Amsterdam: North-Holland), p. 111.