PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

# An Information Theory of Visual Communication

Friedrich O. Huck, Carl L. Fales and Zia-Ur Rahman

*Phil. Trans. R. Soc. Lond. A* 1996 **354**, 2193-2248
doi: 10.1098/rsta.1996.0098

# An information theory of visual communication

By Friedrich O. Huck[1], Carl L. Fales[1] and Zia-ur Rahman[2]

[1]*NASA Langley Research Center, Hampton, VA 23681, USA*

[2]*Science and Technology Corporation, Hampton, VA 23666, USA*

## Contents



The fundamental problem of visual communication is that of producing the best possible picture at the lowest data rate. We address this problem by extending information theory to the assessment of the visual communication channel as a whole, from image gathering to display. The extension unites two disciplines, the electro-optical design of image gathering and display devices and the digital processing for image coding and restoration. The mathematical development leads to several intuitively attractive figures of merit for assessing the visual communication channel as a function of the critical limiting factors that constrain its performance. Multiresolution decomposition is included in the mathematical development to optimally

*Phil. Trans. R. Soc. Lond.* A (1996) **354**, 2193–2248
*Printed in Great Britain*

combine the economical encoding of the transmitted signal with image gathering and restoration.

Quantitative and qualitative assessments demonstrate that a visual communication channel ordinarily can be expected to produce the best possible picture at the lowest data rate only if the image-gathering device produces the maximum-realizable information rate and the image-restoration algorithm properly accounts for the critical limiting factors that constrain the visual communication. These assessments encompass (*a*) the electro-optical design of the image-gathering device in terms of the trade-off between blurring and aliasing in the presence of photodetector and quantization noises, (*b*) the compression of data transmission by redundancy reduction, (*c*) the robustness of the image restoration to uncertainties in the statistical properties of the captured radiance field, and (*d*) the enhancement of particular features or, more generally, of the visual quality of the observed image. The 'best visual quality' in this context normally implies a compromise among maximum-realizable fidelity, sharpness, and clarity which depends on the characteristics of the scene and the purpose of the visual communication (e.g. diagnosis versus entertainment).

## 1. Introduction

The fundamental problem of communication, as Shannon (1948) and Shannon & Weaver (1964) stated it, is that of reproducing at one point either exactly or approximately a message selected at another point. In the classical model of communication (figure 1), the information source selects a desired message out of a set of possible messages. The transmitter changes this message into the signal that is actually sent over the communication channel to the receiver. The receiver changes this signal back into a message, and hands this message to the destination. Ordinarily, the signal is perturbed by noise during transmission or at one of the terminals. Consequently, the received signal is not necessarily the same as that sent out by the transmitter.

In visual communication, in the guise of telephotography or television, for example, it seems reasonable to start with the spatially varying radiance field that is either reflected or emitted by a scene. Then the source of the message is that particular patch of the radiance field that resides within the field of view of the image-gathering device, and the destination of the message is the observed image that the image-display device reconstructs from the received signal. Consequently, the transmitter becomes the image-gathering device that transforms the captured radiance field (i.e. the selected message) into the signal that is transmitted, and the receiver becomes the image-display device that transforms the received signal into an image (i.e. the observed message).

Visual communication is now increasingly carried out with digital image processing. As figure 2 depicts, image gathering is combined with digital encoding to compress data transmission, and image display is combined with digital restoration to enhance image quality. Consequently, visual communication can be described in terms of the following three major processes: the image-gathering process that transforms the captured radiance field into an economically encoded signal, the data-transmission process that conveys this signal to the image-restoration system, and the image-restoration process that transforms the received signal into an enhanced image. This paper deals only with the input and output transformations. It does not




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

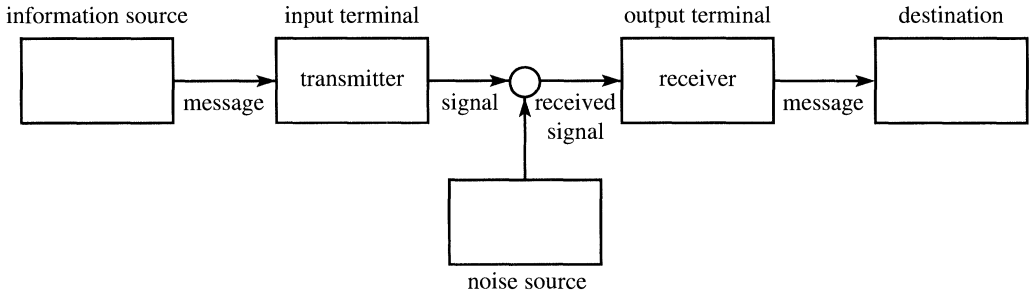



Figure 1. Classical model of communication channel (Shannon 1948; Shannon & Weaver 1964).

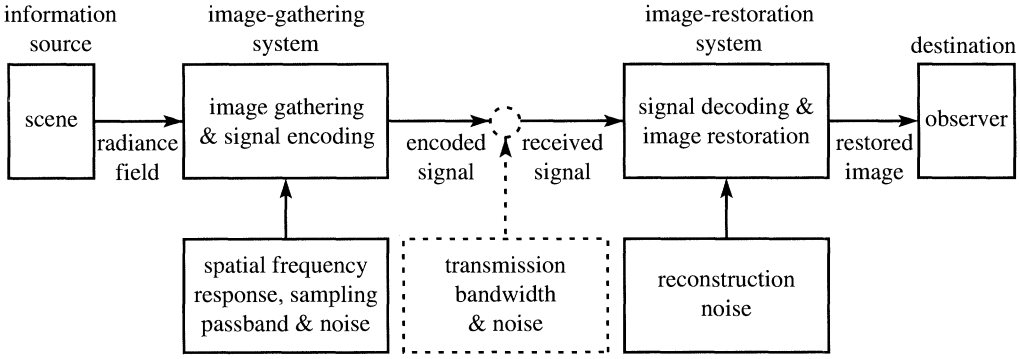



Figure 2. Model of visual communication channel together with the critical limiting factors that constrain its performance.

include the effects of perturbations in the data-transmission process, which is the traditional subject of communication theory.

In most applications, which range from telescope to microscope, the resolution of the image-gathering device is coarser than the finest detail in the scene, so that the spatial frequency spectrum of the radiance field extends beyond the sampling passband of the image-gathering device. Normally, therefore, the task of visual communication becomes that of reproducing an image of the scene for just that portion of the radiance-field spectrum that is contained within the sampling passband. Moreover, the image-gathering device inevitably perturbs this within-passband component when it transforms the captured radiance field into a digital signal. These factors inexorably bar the encoder from direct access to the scene (i.e. the original source). This precludes the application of information theory directly to the scene for the analysis of data compression and rate distortion. The effects of image gathering must be included for this analysis to correlate with actual performance.

Image gathering and reconstruction are performed by electro-optical devices. As figure 3 suggests, these devices constrain the performance of the visual communication channel by the following critical limiting factors: (*a*) the spatial frequency response of the optical (objective lens and photodetector) apertures, (*b*) the sampling passband of the photodetection mechanism, and (*c*) the photodetector, quantization, and reconstruction noises. For optical apertures, unlike electronic filters, the spatial frequency response (to the incoherent radiance field) cannot approximate an ideal bandpass response, i.e. a response that is unity within a finite bandwidth and zero outside. Instead, the spatial frequency response of optical apertures de-

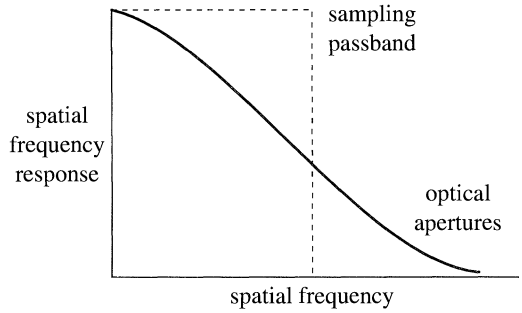



Figure 3. Critical constraints. The image-gathering process is inevitably constrained by the trade-off between blurring and aliasing in the presence of noise.

creases smoothly with increasing frequency and may even exhibit some oscillation at its final decay. Therefore, the electro-optical design of the image-gathering device is inevitably confronted by a trade-off between the blurring that is due to the spatial frequency response of the optical apertures and the aliasing that is due to the insufficient sampling by the photodetection mechanism. Aliasing can be substantially decreased only at the cost of increasing blurring and vice versa. Similarly, the design of the image-display device is confronted by a trade-off between blurring and raster effects. However, if the image reproduction is carried out with the aid of digital processing, then interpolation can be used to overcome this constraint. The critical constraint that remains then is the reconstruction noise (e.g. film granularity).

The critical limiting factors that constrain image gathering and reconstruction have long been of concern in the design of telephotography and television systems. As early as 1934, Mertz & Gray (1934) studied the trade-offs between blurring and aliasing in image gathering and between blurring and raster effects in image reconstruction. Subsequently, among many others who have studied these problems, Schade (1951, 1952, 1953, 1955) and Schreiber (1993) presented particularly comprehensive evaluations of electronic imaging systems. However, the focus of these studies remained on image gathering and reconstruction without the aid of digital processing. Meanwhile, most researchers involved in image coding and restoration have consistently neglected to adequately account for these input and output constraints. In the prevailing digital image processing literature (Andrews & Hunt 1977; Pratt 1978; Huang 1979; Rosenfeld & Kak 1982; Gonzalez & Wintz 1987; Jain 1989), the constraints of image gathering are still incompletely modelled solely by blurring plus noise, while the constraints of image reconstruction are entirely ignored.

So far, then, the major obstacle to a definitive analysis of visual communication has been that the electro-optical design of image gathering and display devices and the digital processing for image coding and restoration have remained independent disciplines that follow distinctly separate traditions. Yet without an adequate theory that embraces these two disciplines, one cannot effectively optimize visual communication. As Gabor put it succinctly, 'experiments unguided by theory do not appear very promising' (Gabor 1946).

The aim of this paper, therefore, is to extend information theory to the assessment of the visual communication channel as a whole by rigorously uniting electro-optical design with digital image processing. This channel can be regarded to be of high quality only if the information rate from the scene to the observer approaches the maximum possible and the required data rate approaches the minimum possible.

Clearly, the goal to produce 'the best possible picture at the lowest data rate' can be attained only by jointly optimizing image gathering, coding, and restoration.

The mathematical development is based on the two classical works that are the foundation of modern communication theory. Shannon (1948) and Shannon & Weaver (1964) introduced the concept of the rate of transmission of information in a noisy channel, and Wiener (1949) introduced the concept of the minimum mean-square error restoration of signals corrupted by noise. By including the constraints of image gathering and display into these familiar concepts, we can quantitatively assess visual communication in terms of the following six figures of merit:

1. The rate of transmission of information, or information rate, $\mathcal{H}$ that the image-gathering system produces for the radiance field that resides within its field of view.
2. The theoretical minimum data rate $\mathcal{E}$, i.e. the entropy of completely decorrelated data, which is associated with the information rate $\mathcal{H}$.
3. The information efficiency $\mathcal{H}/\mathcal{E}$ of the completely decorrelated data.
4. The maximum-realizable fidelity $\mathcal{F}$ of the digital image that can be restored from the received information, unconstrained by the image-display medium.
5. The information rate $\mathcal{H}_o$ of the observed image that the image-restoration system produces from the received information on an image-display medium.
6. The maximum-realizable fidelity $\mathcal{F}_o$ of the observed image, with information rate $\mathcal{H}_o$, that can be restored in continuous form on an image-display medium.

The first four criteria account for the perturbations that occur when the image-gathering system transforms the continuous radiance field into the digitally encoded signal that is transmitted. The last two criteria include the perturbations that occur when the image-display device transforms the digitally restored image into the continuous observed image. The latter help to (*a*) correlate quantitative assessments with perceptual performance and (*b*) match the design of the image gathering and restoration systems with each other. For example, the constraints on the complexity and cost of image restoration are obviously more severe in commercial television than in medical diagnosis and military or planetary reconnaissance.

This paper builds on several earlier papers (Huck *et al.* 1985, 1988, 1993, 1994; McCormick *et al.* 1989; Fales & Huck 1991) that deal with the application of information theory to the assessment of visual communication. However, it does not merely assimilate results from these papers, but, instead, it presents a new, more complete and cohesive development. This includes (*a*) the derivation of information rate and entropy directly from conditional probability, (*b*) the examination of insufficient sampling and quantization in terms of information theory, (*c*) the effects of the image-display constraints on the transformation of the digitally restored image to the continuous observed image, and (*d*) the integration of wavelet decomposition with image gathering and restoration.

Section 2 presents the mathematical model of image gathering and reproduction. For the image reproduction, it clearly distinguishes between *reconstruction*, which is carried out without digital processing, and *restoration*, which is carried out with digital processing. Whereas the image *reconstruction* is only concerned with producing a continuous representation of the discrete *output* of the image-gathering device, the image *restoration* is concerned with producing a representation of the *input* to this device. The restoration is implemented with a Wiener filter that uses an interpolation lattice to suppress the blurring and raster effects of the image-reconstruction process and, thus, produces images with the absolutely maximum-realizable fidelity $\mathcal{F}_o$ for any particular scene and image-gathering system. This interpolation also sim-





PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

plifies the task of correlating quantitative assessments with perceptual performance, for then the visual quality is distorted by fewer and, hence, more tractable perturbations.

Section 3, the central part of this paper, presents the mathematical development that applies information theory to the assessment of visual communication. The development starts with conditional probability to formulate $\mathcal{H}$, $\mathcal{E}$, and $\mathcal{H}_o$. It then ties the information rate to the Wiener filter presented in § 2 to express $\mathcal{F}$ and $\mathcal{F}_o$ as functions of the spectral distribution of $\mathcal{H}$. Together, these relationships form the desired mathematical basis for assessing the performance of the visual communication channel as a function of its critical constraints.

Section 4 includes multiresolution decomposition into the mathematical development. This technique facilitates signal analysis at a hierarchy of scales by decomposing the acquired signal into components that are localized in space and spatial frequency. It has become for this reason an attractive tool for economically encoding signals in a wide range of applications, notably in compressing data transmission bandwidth in visual communication channels. Of the numerous decomposition schemes that have been developed recently, we use the familiar wavelet transform as a specific example for integrating signal analysis and synthesis with image gathering and restoration. This integration is required to maximize the information rate and efficiency of the transmitted wavelet coefficients and the fidelity and visual quality of the restored image.

Finally, § 5 characterizes the upper bounds of performance of visual communication channels in terms of the above figures of merit and correlates these quantitative assessments with experimental results. These results are produced by combining computer simulations of the image-gathering process with the digital-processing algorithms for image coding and restoration. These simulations enable us to generate random targets with easily prescribed statistical properties and to characterize a wide range of electro-optical designs in terms of spatial frequency response, sampling passband, and signal-to-noise ratio.

Four appendices are added for details that, although important in the mathematical development, would obscure the main thread of the presentation. Appendix A relates the mathematical model of the image-gathering device to its electro-optical design parameters. Appendices B and C examine the treatment of insufficient sampling and quantization, respectively, in terms of information theory. And Appendix D extends the formulations of $\mathcal{H}_o$ and $\mathcal{F}_o$ to image restorations for which the blurring and raster effects of the image-reconstruction process cannot be entirely suppressed with fine digital interpolation. This constraint, for example, becomes an important factor in the design of (digital) high-definition television (HDTV), in which a buffer will be added to the receiver to decode and process the received signal.

## 2. Image gathering and reproduction

This section presents (*a*) the mathematical model of image gathering and reconstruction, (*b*) the constraints that these input and output transformations impose on visual communication, (*c*) the assumptions that are made to permit the use of linear system analysis, (*d*) the Wiener filter that restores the image with the minimum mean-square restoration error (MSRE), and (*e*) an interactive enhancement filter that, when combined with the Wiener filter, can be used to adjust the visual quality of the




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

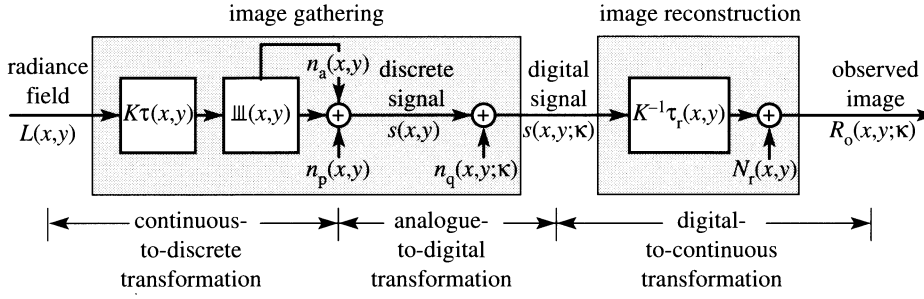



Figure 4. Model of image gathering and reconstruction.

restored image. This material is largely taken from our companion paper (Fales *et al.* 1996), which deals with digital image restoration in far greater detail. However, we include two factors that that paper did not account for, namely, the quantization noise due to the analogue-to-digital transformation and the reconstruction noise due to the image-display medium.

### (*a*) *Image gathering*

Figure 4 depicts a model of image gathering and reconstruction. The image-gathering process transforms the continuous radiance field $L(x,y)$ into the digital signal $s(x,y;\kappa)$, as defined by

$$s(x,y;\kappa) = [KL(x,y) * \tau(x,y)]\,\underline{\amalg}(x,y) + n_p(x,y) + n_q(x,y;\kappa), \tag{2.1 a}$$

where $K$ is the steady-state gain of the linear radiance-to-signal transformation (Appendix A), $\tau(x,y)$ is the spatial response of the image-gathering device, and $n_p(x,y)$ and $n_q(x,y;\kappa)$ are the additive, discrete photodetector and quantization noises, respectively. The symbol $*$ denotes convolution, and the function

$$\underline{\amalg} \equiv \underline{\amalg}(x,y) = \sum_{m,n} \delta(x-m, y-n)$$

denotes the square sampling lattice with unit intervals. The analogue-to-digital transformation is done for $\kappa$ levels with $\eta$-bit quantization, where $\eta = \log \kappa$ and log denotes logarithm to base 2. The Fourier transform of equation (2.1 *a*) is

$$\tilde{s}(\upsilon,\omega;\kappa) = [K\hat{L}(\upsilon,\omega)\hat{\tau}(\upsilon,\omega)] * \underline{\hat{\amalg}} + \tilde{n}_p(\upsilon,\omega) + \tilde{n}_q(\upsilon,\omega;\kappa), \tag{2.1 b}$$

where $\hat{L}(\upsilon,\omega)$ is the continuous radiance-field transform, $\hat{\tau}(\upsilon,\omega)$ is the spatial frequency response (SFR) of the image-gathering device, $\tilde{n}_p(\upsilon,\omega)$ and $\tilde{n}_q(\upsilon,\omega;\kappa)$ are the discrete noise transforms, and $(\upsilon,\omega)$ are the spatial frequencies with units of cycles per sample. The tilde '~' is used instead of the caret '^' to emphasize that the transformed function is periodic in the spatial frequency domain. The function $\underline{\hat{\amalg}}$ is the Fourier transform of the sampling lattice, as given by

$$\begin{aligned}\underline{\hat{\amalg}} \equiv \underline{\hat{\amalg}}(\upsilon,\omega) &= \sum_{m,n} \delta(\upsilon-m, \omega-n) \\ &= \delta(\upsilon,\omega) + \underline{\hat{\amalg}}_s(\upsilon,\omega),\end{aligned}$$

where $\delta(\upsilon,\omega)$ is the Dirac delta function and $\underline{\hat{\amalg}}_s \equiv \underline{\hat{\amalg}}_s(\upsilon,\omega)$ accounts for the sampling sidebands. The associated sampling passband

$$\hat{B} = [(\upsilon,\omega); |\upsilon| \leqslant \tfrac{1}{2}, |\omega| \leqslant \tfrac{1}{2}]$$

has unit area, i.e. $|\hat{B}| = 1$. Thus equation (2.1 *b*) also can be written in the form

$$\tilde{s}(\upsilon,\omega;\kappa) = K\hat{L}(\upsilon,\omega)\hat{\tau}(\upsilon,\omega) + \hat{n}_{\mathrm{a}}(\upsilon,\omega) + \tilde{n}_{\mathrm{p}}(\upsilon,\omega) + \tilde{n}_{\mathrm{q}}(\upsilon,\omega;\kappa), \tag{2.1 c}$$

where $\hat{L}(\upsilon,\omega)\hat{\tau}(\upsilon,\omega)$ is the blurred component of the digital signal and

$$\hat{n}_{\mathrm{a}}(\upsilon,\omega) = [K\hat{L}(\upsilon,\omega)\hat{\tau}(\upsilon,\omega)] * \underline{\hat{|||}}_s$$

are the aliased signal components that insufficient sampling folds into the sampling passband.

Mertz & Gray (1934) observed over 60 years ago that 'the complete process of transmission (in telephotography and television) may be divided into two parts, (*a*) the reproduction of the original picture with a blurring similar to that caused in general by an optical system of only finite perfection, and (*b*) the superposition on it of an extraneous pattern not present in the original, but which is a function of both the original and the scanning system.' Since then, Schade (1951, 1952, 1953, 1955) and Schreiber (1993), among many others, have studied this problem in considerable detail for electronic imaging systems. However, these studies have not included digital processing for image restoration.

The prevailing digital image processing literature (Andrews & Hunt 1977; Pratt 1978; Huang 1979; Rosenfeld & Kak 1982; Gonzalez & Wintz 1987; Jain 1989) largely ignores this 'extraneous pattern', or aliasing, and deals only with the blurring and the noise. This neglect arises directly from the fundamental difference between the model of image gathering given by equations (2.1) and the model found in the general literature. In this literature, the continuous variables $(x, y)$ become discrete variables and the continuous convolution $*$ becomes the discrete convolution, so that (without distinguishing between photodetector and quantization noises), equation (2.1 *a*) becomes $s(x,y) = L(x,y) * \tau(x,y) + n(x,y)$. Clearly, this simplistic transition from continuous to discrete formulation misses the problem of insufficient sampling that confronts the physical continuous-to-discrete transformation in the image-gathering process, as Appendix B further delineates.

To assess visual communication in terms of information theory, the image-gathering process is constrained to be linear and isoplanatic (spatially invariant), and the radiance-field and noise amplitudes are constrained to be Gaussian, wide-sense stationary, and statistically independent. In addition to these constraints, we make the following assumptions:

1. The radiance field $L(x,y)$ is effectively confined to some isoplanatism patch $A$ centred at $x = y = 0$. Because it is the variation of $L(x,y)$ from the mean radiance level of patch $A$ (with area $|A|$) that is of interest, we can let the mean of $L(x,y)$ be zero. For a sufficiently large area $|A|$, the power spectral density (PSD) $\hat{\Phi}_L(\upsilon,\omega)$ of $L(x,y)$ can be approximated by

$$\hat{\Phi}_L(\upsilon,\omega) = \sigma_L^2\,\hat{\Phi}'_L(\upsilon,\omega) = \frac{1}{|A|}\overline{|\hat{L}(\upsilon,\omega)|^2},$$

where $\overline{|(\cdot)|^2}$ denotes the expected value, or average, of $|(\cdot)|^2$ over an ensemble of radiance fields,

$$\sigma_L^2 = \int\!\!\int_{-\infty}^{\infty} \hat{\Phi}_L(\upsilon,\omega)\,\mathrm{d}\upsilon\mathrm{d}\omega$$

is the variance, and $\hat{\Phi}'_L(\upsilon,\omega)$ is the normalized PSD.

2. Similarly, the PSD $\tilde{\Phi}_{\mathrm{p}}(\upsilon,\omega)$ of the photodetector noise $n_{\mathrm{p}}(x,y)$ can be approximated by

$$\tilde{\Phi}_{\mathrm{p}}(\upsilon,\omega)=\frac{1}{|A|}\overline{|\tilde{n}_{\mathrm{p}}(\upsilon,\omega)|^2},$$

where

$$\sigma_{\mathrm{p}}^2=\iint\limits_{\tilde{B}}\tilde{\Phi}_{\mathrm{p}}(\upsilon,\omega)\,\mathrm{d}\upsilon\mathrm{d}\omega$$

is the variance. For white noise, $\tilde{\Phi}_{\mathrm{p}}(\upsilon,\omega)=\sigma_{\mathrm{p}}^2$.

3. The PSD $\tilde{\Phi}_s(\upsilon,\omega)$ of the discrete signal $s(x,y)$ with analogue amplitude variations (i.e. before quantization) is defined by

$$\tilde{\Phi}_s(\upsilon,\omega)=\frac{1}{|A|}\,\overline{|\tilde{s}(\upsilon,\omega)|^2}.$$

Consequently, according to equation (2.1 *b*) (without the quantization noise),

$$\tilde{\Phi}_s(\upsilon,\omega)=[K^2\hat{\Phi}_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2]*\underline{\hat{|\,|}}+\tilde{\Phi}_{\mathrm{p}}(\upsilon,\omega).\tag{2.2}$$

The corresponding signal variance $\sigma_s^2$ is

$$\sigma_s^2=\iint\limits_{\hat{B}}\tilde{\Phi}_s(\upsilon,\omega)\,\mathrm{d}\upsilon\mathrm{d}\omega.\tag{2.3 a}$$

Ordinarily, the contribution of the photodetector noise $n_{\mathrm{p}}(x,y)$ to $\sigma_s^2$ is small and can be ignored, so that

$$\sigma_s^2\approx K^2\sigma_L^2\int\limits_{-\infty}^{\infty}\hat{\Phi}'_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2\,\mathrm{d}\upsilon\mathrm{d}\omega.\tag{2.3 b}$$

4. For the analogue-to-digital transformation, the quantization is uniform and the signal is equally likely to occur anywhere in the quantization interval $-c\sigma_s/\kappa$ to $c\sigma_s/\kappa$, where $c$ is a constant that adjusts the quantization intervals and $\sigma_s^2$ is the variance of the signal before quantization. Consequently, the acquired signal is quantized over the dynamic range of $-c\sigma_s$ to $c\sigma_s$, and signal values outside this range are assigned to either 0 or $\kappa-1$. For $c=\sqrt{3}$, the dynamic range encompasses 92% of the signal. (The motivation for this particular adjustment is given in § 3 *c*.) These assumptions imply that the quantization noise $n_{\mathrm{q}}(x,y,;\kappa)$ has the uniform probability density

$$p(n_{\mathrm{q}})=\begin{cases}\kappa/2c\sigma_s, & -c\sigma_s/\kappa\leqslant n_{\mathrm{q}}\leqslant c\sigma_s/\kappa,\\ 0, & \text{elsewhere},\end{cases}$$

with a mean equal to zero and a variance given by

$$\sigma_{\mathrm{q}}^2=\int_{-c\sigma_s/\kappa}^{c\sigma_s/\kappa}n_{\mathrm{q}}^2p(n_{\mathrm{q}})\,\mathrm{d}n_{\mathrm{q}}=\frac{1}{3}\left(\frac{c\sigma_s}{\kappa}\right)^2=\left(\frac{\sigma_s}{\kappa}\right)^2.\tag{2.4 a}$$

Moreover, the quantization error of any one sample is uncorrelated with that of any other sample (although the present formulation allows for autocorrelation), and the error signal is uncorrelated with the discrete signal $s(x,y)$ or, equivalently, with

$L(x,y)$ and $n_{\mathrm{p}}(x,y)$. The quantization error signal is then a wide-sense stationary process with a PSD $\tilde{\Phi}_{\mathrm{q}}(\upsilon,\omega;\kappa)$ equal to its variance $\sigma_{\mathrm{q}}^2$; that is,

$$\tilde{\Phi}_{\mathrm{q}}(\upsilon,\omega;\kappa) = \tilde{\Phi}_{\mathrm{q}}(\kappa) = \sigma_{\mathrm{q}}^2 = (\sigma_s/\kappa)^2. \tag{2.4\,b}$$

Finally, the quantization noise is treated Gaussian, which immediately implies the independence on $n_{\mathrm{q}}(x,y;\kappa)$ from $L(x,y)$ and $n_{\mathrm{p}}(x,y)$ and imposes an approximation of the uniform probability density by a Gaussian shape.

### (*b*) *Image reconstruction*

The image-reconstruction process transforms the digital signal $s(x,y;\kappa)$ into the observed continuous image $R_{\mathrm{o}}(x,y;\kappa)$, as defined by

$$R_{\mathrm{o}}(x,y;\kappa) = K^{-1}s(x,y;\kappa) * \tau_{\mathrm{r}}(x,y) + N_{\mathrm{r}}(x,y), \tag{2.5\,a}$$

where $\tau_{\mathrm{r}}(x,y)$ is the spatial response of the reconstruction process and $N_{\mathrm{r}}(x,y)$ is the continuous reconstruction noise. The normalization by the constant $K$ accounts for the steady-state gain of the linear signal-to-image transformation in the reconstruction process. In practice, of course, this gain differs from that of the radiance-to-signal transformation introduced in equations (2.1). However, it is not necessary to account for this difference here. By using the same gain constant for both image gathering and reconstruction, the average reflectance of the image is equated to that of the scene. The Fourier transform of equation (2.5 *a*) is

$$\hat{R}_{\mathrm{o}}(\upsilon,\omega;\kappa) = K^{-1}\tilde{s}(\upsilon,\omega;\kappa)\hat{\tau}_{\mathrm{r}}(\upsilon,\omega) + \hat{N}_{\mathrm{r}}(\upsilon,\omega), \tag{2.5\,b}$$

where $\hat{\tau}_{\mathrm{r}}(\upsilon,\omega)$ is the SFR of the reconstruction process and $\hat{N}_{\mathrm{r}}(\upsilon,\omega)$ is the noise transform.

The reconstructed image can be explicitly expressed in terms of three major components:

$$\hat{R}_{\mathrm{o}}(\upsilon,\omega;\kappa) = \hat{R}_{\mathrm{b}}(\upsilon,\omega) + \hat{R}_{\mathrm{a}}(\upsilon,\omega) + \hat{R}_{\mathrm{n}}(\upsilon,\omega;\kappa), \tag{2.5\,c}$$

where

$$\hat{R}_{\mathrm{b}}(\upsilon,\omega) = \hat{L}(\upsilon,\omega)\hat{\tau}(\upsilon,\omega)\hat{\tau}_{\mathrm{r}}(\upsilon,\omega)$$

represents the blurred component,

$$\hat{R}_{\mathrm{a}}(\upsilon,\omega) = [\hat{L}(\upsilon,\omega)\hat{\tau}(\upsilon,\omega) * \underline{|\hat{|}|}_s(\upsilon,\omega)]\hat{\tau}_{\mathrm{r}}(\upsilon,\omega)$$

represents the aliased components, and

$$\hat{R}_{\mathrm{n}}(\upsilon,\omega;\kappa) = K^{-1}[\tilde{n}_{\mathrm{p}}(\upsilon,\omega) + \tilde{n}_{\mathrm{q}}(\upsilon,\omega;\kappa)]\hat{\tau}_{\mathrm{r}}(\upsilon,\omega) + \hat{N}_{\mathrm{r}}(\upsilon,\omega)$$

represents the photodetector, quantization, and reconstruction noises.

The reconstruction noise $N_{\mathrm{r}}(x,y)$, as treated here, is implicitly a function of the reconstruction response $\tau_{\mathrm{r}}(x,y)$ as well as of the image-display medium. Consider, for example, film granularity. This granularity may range from fine to coarse (typically for slow to fast films, respectively). It lies dormant in the unexposed emulsion of the film until this film is exposed and developed. The exposure is made with the spot of light that the image-display device scans across the film. If this spot of light is large relative to the granularity of the film, then the spatial details of the signal that this spot records on the film are correspondingly large compared with the granularity. The reverse, of course, is true also. Therefore, the noise $N_{\mathrm{r}}(x,y)$ is that of the reconstruction process rather than just of the film. Similarly, the reconstruction

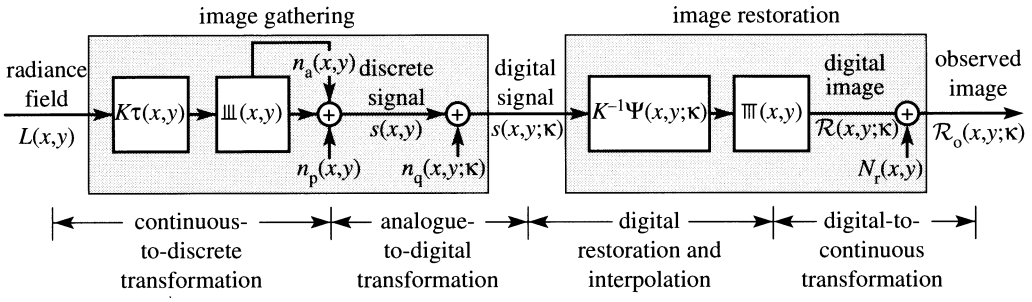



Figure 5. Model of image gathering and restoration.

response $\tau_r(x,y)$ is not just that of the spot of light, but it includes the response of the film (or image-display medium in general).

### (*c*) *Image restoration*

Figure 5 depicts a model of image gathering and restoration. We assume that the linear, isoplanatic operator $\Psi(x,y;\kappa)$ restores the image on an interpolation lattice, depicted by the symbol $\overline{|||}(x,y)$, which is much denser than the sampling lattice $|||(x,y)$ of the image-gathering process. If this interpolation is sufficiently dense, then the operator $\Psi(x,y;\kappa)$ effectively suppresses the blurring and raster effects of the image-display device. For this condition the image-restoration process can be modelled as transforming the digital signal $s(x,y;\kappa)$ into the digital image $\mathcal{R}(x,y;\kappa)$, as given by the Fourier transform pair

$$\mathcal{R}(x,y;\kappa) = K^{-1}s(x,y;\kappa) * \Psi(x,y;\kappa), \tag{2.6\,a}$$

$$\hat{\mathcal{R}}(\upsilon,\omega;\kappa) = K^{-1}\tilde{s}(\upsilon,\omega;\kappa)\,\hat{\Psi}(\upsilon,\omega;\kappa). \tag{2.6\,b}$$

In further transforming this digital image into the observed image $\mathcal{R}_o(x,y;\kappa)$, the image-display medium adds the reconstruction noise $N_r(x,y)$ so that

$$\mathcal{R}_o(x,y;\kappa) = \mathcal{R}(x,y;\kappa) + K^{-1}N_r(x,y) \tag{2.7\,a}$$

$$\hat{\mathcal{R}}_o(\upsilon,\omega;\kappa) = \hat{\mathcal{R}}(\upsilon,\omega;\kappa) + K^{-1}\hat{N}_r(\upsilon,\omega). \tag{2.7\,b}$$

The restoration SFR $\hat{\Psi}(\upsilon,\omega;\kappa)$ now replaces the reconstruction SFR $\hat{\tau}_r(\upsilon,\omega)$ in equation (2.5 *b*). This is mathematically justifiable because the interpolation lattice separates its sideband components of the digital image far enough apart from each other so that, in effect, $\tilde{\Psi}(\upsilon,\omega;\kappa)\hat{\tau}_r(\upsilon,\omega) \to \tilde{\Psi}(\upsilon,\omega;\kappa)$, i.e. the SFR $\hat{\tau}_r(\upsilon,\omega)$ passes the fundamental component of $\tilde{\Psi}(\upsilon,\omega;\kappa)$ without perturbing it while entirely rejecting the other components. The expression for the restored image $\mathcal{R}_o(x,y;\kappa)$ in terms of its three major components has exactly the same form as equation (2.5 *c*) for the reconstructed image $R_o(\upsilon,\omega;\kappa)$.

The mean-square error $e^2$ between the incident radiance field $L(x,y)$ and any image $R(x,y;\kappa)$ is defined by

$$e^2 = \frac{1}{|A|}\int\!\!\int_{-\infty}^{\infty} \overline{|\hat{L}(\upsilon,\omega) - \hat{R}(\upsilon,\omega;\kappa)|^2}\,\mathrm{d}\upsilon\mathrm{d}\omega = \int\!\!\int_{-\infty}^{\infty} \hat{e}^2(\upsilon,\omega;\kappa)\,\mathrm{d}\upsilon\mathrm{d}\omega, \tag{2.8}$$

where $\hat{e}^2(\upsilon,\omega;\kappa)$ is the mean-square error PSD. The Wiener filter $\hat{\Psi}(\upsilon,\omega;\kappa)$, which

minimizes the mean-square error $e^2$, is (Fales *et al.* 1996)

$$\hat{\Psi}(\upsilon,\omega;\kappa) = \frac{K^2\,\hat{\Phi}_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2}{\tilde{\Phi}_s(\upsilon,\omega;\kappa)\hat{\tau}(\upsilon,\omega)} = \frac{K^2\,\hat{\Phi}_L(\upsilon,\omega)\hat{\tau}^*(\upsilon,\omega)}{\tilde{\Phi}_s(\upsilon,\omega;\kappa)}, \tag{2.9 a}$$

$$= \frac{\hat{\Phi}_L(\upsilon,\omega)\hat{\tau}^*(\upsilon,\omega)}{\hat{\Phi}_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2 * \underline{|\hat{|}|} + K^{-2}\,\tilde{\Phi}_{\mathrm{p}}(\upsilon,\omega) + K^{-2}\,\tilde{\Phi}_{\mathrm{q}}(\upsilon,\omega;\kappa)}. \tag{2.9 b}$$

These expressions for $\hat{\Psi}(\upsilon,\omega;\kappa)$ are independent of the reconstruction noise $\hat{N}_{\mathrm{r}}(\upsilon,\omega)$. The digital image $\mathcal{R}(x,y;\kappa)$ differs from the observed image $\mathcal{R}_{\mathrm{o}}(x,y;\kappa)$ that is recorded on the image-display medium only by the noise that this medium adds. For photodetector noise that is white so that the PSD $\tilde{\Phi}_{\mathrm{p}}(\upsilon,\omega)$ is equal to its variance $\sigma_{\mathrm{p}}^2$, and for the assumptions made about the analogue-to-digital transformation so that the PSD $\tilde{\Phi}_{\mathrm{q}}(\upsilon,\omega;\kappa)$ is equal to its variance $\sigma_{\mathrm{q}}^2$, equation (2.9 *b*) can be expressed as a function of the SNRs $(K\sigma_L/\sigma_{\mathrm{p}})$ and $(K\sigma_L/\sigma_s)\kappa$ as

$$\hat{\Psi}(\upsilon,\omega;\kappa) = \frac{\hat{\Phi}'_L(\upsilon,\omega)\hat{\tau}^*(\upsilon,\omega)}{\hat{\Phi}'_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2 * \underline{|\hat{|}|} + (K\sigma_L/\sigma_{\mathrm{p}})^{-2} + (K\sigma_L/\sigma_s)^{-2}\kappa^{-2}}, \tag{2.9 c}$$

where $\hat{\Phi}'_L(\upsilon,\omega) = \sigma_L^{-2}\,\hat{\Phi}_L(\upsilon,\omega)$.

The minimum mean-square restoration error (MSRE) $\epsilon^2$ of $\mathcal{R}(x,y;\kappa)$ is

$$\epsilon^2 = \iint_{-\infty}^{\infty} \hat{\epsilon}^2(\upsilon,\omega;\kappa)\,\mathrm{d}\upsilon\mathrm{d}\omega,$$

where $\hat{\epsilon}(\upsilon,\omega;\kappa)$ is the corresponding minimum MSRE PSD given by

$$\hat{\epsilon}^2(\upsilon,\omega;\kappa) = \hat{\Phi}_L(\upsilon,\omega)[1 - \hat{\Gamma}(\upsilon,\omega;\kappa)] \tag{2.10}$$

and

$$\hat{\Gamma}(\upsilon,\omega;\kappa) = \hat{\tau}(\upsilon,\omega)\,\hat{\Psi}(\upsilon,\omega;\kappa) \tag{2.11}$$

is the throughput SFR. When the aliasing, photodetector, and quantization noises approach zero, then the Wiener filter $\hat{\Psi}(\upsilon,\omega;\kappa)$ approaches the inverse filter $\hat{\tau}^{-1}(\upsilon,\omega)$ and the SFR $\hat{\Gamma}(\upsilon,\omega;\kappa)$ approaches the ideal bandpass response. The minimum MSRE $\epsilon_{\mathrm{o}}^2$ of $\mathcal{R}_{\mathrm{o}}(x,y;\kappa)$ is constrained by the reconstruction noise $N_{\mathrm{r}}(x,y)$, so that

$$\epsilon_{\mathrm{o}}^2 = \iint_{-\infty}^{\infty} \hat{\epsilon}_{\mathrm{o}}^2(\upsilon,\omega;\kappa)\,\mathrm{d}\upsilon\mathrm{d}\omega,$$

where $\hat{\epsilon}_{\mathrm{o}}^2(\upsilon,\omega;\kappa)$ is the corresponding minimum MSRE PSD given by

$$\hat{\epsilon}_{\mathrm{o}}^2(\upsilon,\omega;\kappa) = \hat{\epsilon}^2(\upsilon,\omega;\kappa) + \hat{\Phi}_{\mathrm{r}}(\upsilon,\omega) \tag{2.12}$$

and $\hat{\Phi}_{\mathrm{r}}(\upsilon,\omega)$ is the PSD of $N_{\mathrm{r}}(x,y)$.

The phrase 'dense interpolation lattice' is not intended to suggest that the spot of light which records the restored image on the image-display medium should actually be reduced in size relative to the one used for the reconstruction in equations (2.5). Such a reduction would increase the effective reconstruction noise $N_{\mathrm{r}}(x,y)$. Instead, the restored image is recorded on a larger frame. For example, if the interpolation lattice is 4 times 'denser' than the sampling lattice (which, in practice, is sufficient to effectively suppress the blurring and raster effects of the reconstruction process (Fales




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

*et al.* 1996)), then the recorded image grows 4 times larger. If this interpolation is not practical, then it becomes necessary to account for the blurring and raster effects of the reconstruction process, as Appendix D formulates. In addition, Fales *et al.* (1996) present Wiener filters that minimize the MSRE for coarse digital interpolation and for small-kernel operators. We avoid these filters here because the associated degradations, even when small, distract from the correlation of quantitative assessments with the perceptual image quality.

### (*d*) *Image enhancement*

The minimum MSRE representation that the Wiener filter produces often includes some perceptual defects caused by (*a*) the aliasing artifacts that are due to the insufficient sampling in the image-gathering process, (*b*) the photodetector and quantization noises that are amplified and coloured by the high-frequency enhancement that is due to the Wiener filter, and (*c*) the ringing near sharp edges (Gibbs phenomenon) that is due to the steep roll-off in the throughput SFR $\hat{\Gamma}(\upsilon, \omega; \kappa)$. It is not possible to have maximum sharpness without either aliasing or ringing. Some overshoot near sharp edges is not only acceptable but also improves sharpness. However, excessive overshoot with ringing is clearly undesirable (McCormick *et al.* 1989; Schreiber 1993; Fales *et al.* 1996).

Hence, it is often desirable to combine the Wiener filter with some algorithm for interactively controlling the perceptual image quality. This control can be provided by the Wiener–Gaussian enhancement (WIGE) filter (McCormick *et al.* 1989; Fales *et al.* 1996), which is the product of the Wiener filter $\hat{\Psi}(\upsilon, \omega; \kappa)$ and an enhancement filter $\hat{\tau}_{\mathrm{e}}(\upsilon, \omega)$, so that

$$\hat{\Psi}_{\mathrm{e}}(\upsilon, \omega; \kappa) = \hat{\Psi}(\upsilon, \omega; \kappa)\hat{\tau}_{\mathrm{e}}(\upsilon, \omega), \tag{2.13}$$

where

$$\hat{\tau}_{\mathrm{e}}(\upsilon, \omega) = \exp[-2(\pi\sigma_i\rho)^2] + \zeta(2\pi\sigma_{\mathrm{e}}\rho/\sigma_{\mathrm{o}})^2 \exp[-2(\pi\sigma_{\mathrm{e}}\rho)^2], \tag{2.14}$$

$\rho^2 = \upsilon^2 + \omega^2$, and $\sigma_{\mathrm{o}} = (2/\pi\mathrm{e})^{1/2} = 0.48$. The first term on the right side of equation (2.14) is a low-pass Gaussian filter with standard deviation $\sigma_i$. This parameter controls the degree of smoothing that may be desired to suppress visually annoying defects. The second term is the Fourier transform of the familiar Laplacian-of-Gaussian ($\nabla^2 G$) operator commonly used to enhance the radiance-field transitions for subsequent edge detection, as given by (Marr & Hildreth 1980; Marr 1982)

$$\nabla^2 G(x, y) = \frac{1}{\pi\sigma_{\mathrm{e}}^2}\left(1 - \frac{r^2}{2\sigma_{\mathrm{e}}^2}\right) \exp\left(-\frac{r^2}{2\sigma_{\mathrm{e}}^2}\right),$$

where $r^2 = x^2 + y^2$ and $\sigma_{\mathrm{e}}$ is the standard deviation of the Gaussian function. This parameter controls the resolution of the edge enhancement. Finally, the parameter $\zeta$ controls the degree of edge enhancement. The ratio $\sigma_{\mathrm{e}}^2/\sigma_{\mathrm{o}}^2$ is included so that the enhancement with $\zeta = 1$ is directly proportional to the change of intensity at an edge transition (Alter-Gartenberg *et al.* 1992). The Gaussian function is chosen for its desirable filter characteristics (Gabor 1946). It is simultaneously optimally localized in both the spatial and spatial-frequency domains.

The minimum MSRE $\hat{\epsilon}_{\mathrm{e}}^2$ of the enhanced digital image is

$$\hat{\epsilon}_{\mathrm{e}}^2 = \int_{-\infty}^{\infty}\int \hat{\epsilon}_{\mathrm{e}}^2(\upsilon, \omega; \kappa)\, \mathrm{d}\upsilon\mathrm{d}\omega,$$

where

$$\hat{\epsilon}_{\mathrm{e}}^2(\upsilon,\omega;\kappa) = \hat{\Phi}_L(\upsilon,\omega)[1-\hat{\tau}_{\mathrm{e}}(\upsilon,\omega)\hat{I}(\upsilon,\omega;\kappa)]. \tag{2.15}$$

Similarly, the minimum MSRE $\hat{\epsilon}_{\mathrm{oe}}^2$ of the enhanced observed image is

$$\hat{\epsilon}_{\mathrm{oe}}^2 = \int_{-\infty}^{\infty}\int \hat{\epsilon}_{\mathrm{o}}^2(\upsilon,\omega;\kappa)\,\mathrm{d}\upsilon\mathrm{d}\omega,$$

where

$$\hat{\epsilon}_{\mathrm{oe}}^2(\upsilon,\omega) = \hat{\epsilon}_{\mathrm{e}}^2(\upsilon,\omega;\kappa) + \hat{\Phi}_{\mathrm{r}}(\upsilon,\omega). \tag{2.16}$$

## 3. Figures of merit

This section develops the six figures of merit outlined in the Introduction. The development is meaningful only for the image restoration and enhancement presented in §§ 2 *c*, *d* for which there evolves a close relationship between information rate and maximum-realizable image quality. This relationship does not extend to the image reconstruction presented in § 2 *b*.

The mathematical development departs from the traditional application of information theory to the assessment of communication channels by including the effects of insufficient sampling and analogue-to-digital transformation. These effects are carefully examined in Appendices B and C, respectively. The results of this examination show that, for the assumptions in § 2 *a*, the aliased signal components and the quantization errors can both be treated as independent, additive noise. This treatment, of course, greatly facilitates the quantitative assessment of visual communication.

### (*a*) *Information rate* $\mathcal{H}$

The first two transformations in figure 5 convert the continuous radiance field $L(x,y)$ into the digital signal $s(x,y;\kappa)$, as defined by equations (2.1). According to Shannon (1948) and Shannon & Weaver (1964), the rate of transmission of information $\mathcal{H}$ through these transformations or, equivalently, the mutual information between the input radiance field $L(x,y)$ within the sampling passband and the output digital signal $s(x,y;\kappa)$ is defined by

$$\mathcal{H} = \mathcal{E}[s(x,y;\kappa)] - \mathcal{E}[s(x,y;\kappa)|L(x,y)] \tag{3.1 a}$$

$$= \mathcal{E}[\tilde{s}(\upsilon,\omega;\kappa)] - \mathcal{E}[\tilde{s}(\upsilon,\omega;\kappa)|\hat{L}(\upsilon,\omega)], \tag{3.1 b}$$

where the first term, $\mathcal{E}[\cdot]$, in each equation represents the entropy of the received signal in spatial and spatial frequency coordinates, respectively, and the second term, $\mathcal{E}[\cdot|\cdot]$, represents the conditional entropy of the received signal given the radiance field. The conditional entropy is the uncertainty of the digital signal $s(x,y;k)$ when the radiance field $L(x,y)$ is known. Thus, the information rate $\mathcal{H}$ given by this defining equation measures the amount of information received less the part of this which is due to noise. Because the sources of noise can be treated as independent and additive, we can, according to Shannon's theorem 16, simplify the above expressions for $\mathcal{H}$ to

$$\mathcal{H} = \mathcal{E}[s(x,y;\kappa)] - \mathcal{E}[n(x,y;\kappa)] \tag{3.2 a}$$

$$= \mathcal{E}[\tilde{s}(\upsilon,\omega;\kappa)] - \mathcal{E}[\hat{n}(\upsilon,\omega;\kappa)], \tag{3.2 b}$$




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

where, from equation (2.1 *c*),

$$\tilde{s}(\upsilon,\omega;\kappa) = K\hat{L}(\upsilon,\omega)\hat{\tau}(\upsilon,\omega) + \hat{n}(\upsilon,\omega;\kappa)$$

and

$$\hat{n}(\upsilon,\omega;\kappa) = \hat{n}_{\mathrm{a}}(\upsilon,\omega) + \tilde{n}_{\mathrm{p}}(\upsilon,\omega) + \tilde{n}_{\mathrm{q}}(\upsilon,\omega;\kappa).$$

A full interpretation of equation (3.2 *b*) in terms of the probability densities

$$p_s[\tilde{s}(\upsilon,\omega;\kappa)] \quad \text{and} \quad p_n[\hat{n}(\upsilon,\omega;\kappa)]$$

of the signal and the noise, respectively, is given by the expression

$$\mathcal{H} = -\frac{1}{2}\iint_{\hat{B}}\left\{\iint p_s[\tilde{s}(\upsilon,\omega;\kappa)]\log p_s[\tilde{s}(\upsilon,\omega;\kappa)]\,\mathrm{d}\tilde{s}(\upsilon,\omega;\kappa) - \iint p_n[\hat{n}(\upsilon,\omega;\kappa)]\log p_n[\hat{n}(\upsilon,\omega;\kappa)]\,\mathrm{d}\hat{n}(\upsilon,\omega;\kappa)\right\}\mathrm{d}\upsilon\mathrm{d}\omega, \tag{3.2 c}$$

where $\mathrm{d}\tilde{s}(\upsilon,\omega;\kappa)$ and $\mathrm{d}\hat{n}(\upsilon,\omega;\kappa)$ denote real differential area elements comprising the real and imaginary parts of $\tilde{s}(\upsilon,\omega;\kappa)$ and $\hat{n}(\upsilon,\omega;\kappa)$, respectively. The Gaussian probability density $p_s[\tilde{s}(\upsilon,\omega;\kappa)]$ of the signal $\tilde{s}(\upsilon,\omega;\kappa)$ with the PSD $\tilde{\Phi}_s(\upsilon,\omega;\kappa)$ is

$$p_s[\tilde{s}(\upsilon,\omega;\kappa)] = \frac{1}{\pi\tilde{\Phi}_s(\upsilon,\omega;\kappa)}\exp\left\{-\frac{|\tilde{s}(\upsilon,\omega;\kappa)|^2}{\tilde{\Phi}_s(\upsilon,\omega;\kappa)}\right\}.$$

Similarly, the Gaussian probability density $p_n[\hat{n}(\upsilon,\omega;\kappa)]$ of the noise $\hat{n}(\upsilon,\omega;\kappa)$ with the PSD $\hat{\Phi}_{\mathrm{n}}(\upsilon,\omega;\kappa)$ is

$$p_n[\hat{n}(\upsilon,\omega;\kappa)] = \frac{1}{\pi\hat{\Phi}_{\mathrm{n}}(\upsilon,\omega;\kappa)}\exp\left\{-\frac{|\hat{n}(\upsilon,\omega;\kappa)|^2}{\hat{\Phi}_{\mathrm{n}}(\upsilon,\omega;\kappa)}\right\}.$$

The PSD $\tilde{\Phi}_s(\upsilon,\omega;\kappa)$ is the sum of equations (2.2) and (2.4 *b*), as given by

$$\tilde{\Phi}_s(\upsilon,\omega;\kappa) = [K^2\hat{\Phi}_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2] * \underline{\hat{|||}} + \tilde{\Phi}_{\mathrm{p}}(\upsilon,\omega) + \tilde{\Phi}_{\mathrm{q}}(\upsilon,\omega;\kappa).$$

Recalling from § 2 *a* that

$$\underline{\hat{|||}} \equiv \underline{\hat{|||}}(\upsilon,\omega) = \delta(\upsilon,\omega) + \underline{\hat{|||}}_s(\upsilon,\omega),$$

the PSD $\hat{\Phi}_{\mathrm{n}}(\upsilon,\omega;\kappa)$ is given by

$$\hat{\Phi}_{\mathrm{n}}(\upsilon,\omega;\kappa) = [K^2\hat{\Phi}_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2] * \underline{\hat{|||}}_s + \tilde{\Phi}_{\mathrm{p}}(\upsilon,\omega) + \tilde{\Phi}_{\mathrm{q}}(\upsilon,\omega;\kappa).$$

Substituting the above into equation (3.2 *c*) yields the following results:

$$\mathcal{H} = \frac{1}{2}\iint_{\hat{B}}\log\tilde{\Phi}_s(\upsilon,\omega;\kappa)\,\mathrm{d}\upsilon\mathrm{d}\omega - \frac{1}{2}\iint_{\hat{B}}\log\hat{\Phi}_{\mathrm{n}}(\upsilon,\omega;\kappa)\,\mathrm{d}\upsilon\mathrm{d}\omega \tag{3.2 d}$$

$$= \frac{1}{2}\iint_{\hat{B}}\log\frac{\tilde{\Phi}_s(\upsilon,\omega;\kappa)}{\hat{\Phi}_{\mathrm{n}}(\upsilon,\omega;\kappa)}\,\mathrm{d}\upsilon\mathrm{d}\omega = \frac{1}{2}\iint_{\hat{B}}\log\left[1+\frac{K^2\hat{\Phi}_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2}{\tilde{\Phi}_{\mathrm{n}}(\upsilon,\omega;\kappa)}\right]\mathrm{d}\upsilon\mathrm{d}\omega \tag{3.2 e}$$

$$= \frac{1}{2}\iint_{\hat{B}}\log\left[1+\frac{\hat{\Phi}_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2}{\hat{\Phi}_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2 * \underline{\hat{|||}}_s + K^{-2}\tilde{\Phi}_{\mathrm{p}}(\upsilon,\omega) + K^{-2}\tilde{\Phi}_{\mathrm{q}}(\upsilon,\omega;\kappa)}\right]\mathrm{d}\upsilon\mathrm{d}\omega. \tag{3.2 f}$$

For the assumptions that $\tilde{\Phi}_{\mathrm{p}}(\upsilon,\omega)=\sigma_{\mathrm{p}}^2$ and $\tilde{\Phi}_{\mathrm{q}}(\upsilon,\omega;\kappa)=\sigma_{\mathrm{q}}^2$,

$$\mathcal{H}=\frac{1}{2}\iint_{\hat{B}}\log\left[1+\frac{\hat{\Phi}_L'(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2}{\tilde{\Phi}_L'(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2 * \underline{\underline{|\hat{|}|}}_s+(K\sigma_L/\sigma_{\mathrm{p}})^{-2}+(K\sigma_L/\sigma_s)^{-2}\kappa^{-2}}\right]\mathrm{d}\upsilon\mathrm{d}\omega. \tag{3.2 g}$$

The above formulations of $\mathcal{H}$ can be tied to the Wiener filter $\hat{\Psi}(\upsilon,\omega;\kappa)$ given by equations (2.9), the minimum-realizable MSRE PSD $\hat{\epsilon}^2(\upsilon,\omega;\kappa)$ given by equation (2.10), and the throughput SFR $\hat{T}(\upsilon,\omega;\kappa)$ given by equation (2.11) as follows. Comparing the PSD $\tilde{\Phi}_s(\upsilon,\omega;\kappa)$ of the digital signal with the PSD $\hat{\Phi}_{\mathrm{n}}(\upsilon,\omega;\kappa)$ of the noise that this signal contains, it can be seen that

$$\hat{\Phi}_{\mathrm{n}}(\upsilon,\omega;\kappa)=\tilde{\Phi}_s(\upsilon,\omega;\kappa)-K^2\hat{\Phi}_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2.$$

Substituting this expression for $\hat{\Phi}_{\mathrm{n}}(\upsilon,\omega;\kappa)$ into the first form of equation (3.2 e) yields

$$\mathcal{H}=-\frac{1}{2}\iint_{\hat{B}}\log\left[1-\frac{K^2\hat{\Phi}_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2}{\tilde{\Phi}_s(\upsilon,\omega;\kappa)}\right]\mathrm{d}\upsilon\mathrm{d}\omega.$$

Next, solving equation (2.9 a) for $\tilde{\Phi}_s(\upsilon,\omega;\kappa)$ and substituting the result into the above equation yields

$$\mathcal{H}=-\frac{1}{2}\iint_{\hat{B}}\log[1-\hat{\tau}(\upsilon,\omega)\hat{\Psi}(\upsilon,\omega;\kappa)]\,\mathrm{d}\upsilon\mathrm{d}\omega.$$

Finally, comparing this result to equation (2.10) for $\hat{\epsilon}^2(\upsilon,\omega;\kappa)$ and equation (2.11) for $\hat{T}(\upsilon,\omega;\kappa)$ leads to the following intuitively attractive expressions:

$$\mathcal{H}=-\frac{1}{2}\iint_{\hat{B}}\log[1-\hat{T}(\upsilon,\omega;\kappa)]\,\mathrm{d}\upsilon\mathrm{d}\omega \tag{3.2 h}$$

$$=\frac{1}{2}\iint_{\hat{B}}\log\frac{\hat{\Phi}_L(\upsilon,\omega)}{\hat{\epsilon}^2(\upsilon,\omega;\kappa)}\,\mathrm{d}\upsilon\mathrm{d}\omega \tag{3.2 i}$$

$$=\frac{1}{2}\iint_{\hat{B}}\log\hat{\Phi}_L(\upsilon,\omega)\,\mathrm{d}\upsilon\mathrm{d}\omega-\frac{1}{2}\iint_{\hat{B}}\log\hat{\epsilon}^2(\upsilon,\omega;\kappa)\,\mathrm{d}\upsilon\mathrm{d}\omega. \tag{3.2 j}$$

The first integral in equation (3.2 j), which is a function of the radiance-field PSD $\hat{\Phi}_L(\upsilon,\omega)$, can be regarded either as the entropy of the radiance field within the sampling passband $\hat{B}$ or, equivalently, as the information rate that the image-gathering system would produce if it were constrained only by its sampling passband. The second integral in equation (3.2 j), which is a function of the minimum MSRE PSD $\hat{\epsilon}^2(\upsilon,\omega;\kappa)$, accounts for the loss of information that the perturbations of the image-gathering process cause within the sampling passband.

The theoretical upper bound of $\mathcal{H}$ is Shannon's channel capacity $\mathcal{C}$ for a bandwidth-limited system with an average power limitation, as given by

$$\mathcal{C}=\tfrac{1}{2}|\hat{B}|\log[1+(K\sigma_L/\sigma_{\mathrm{n}})^2], \tag{3.3}$$

where $\sigma_{\mathrm{n}}^2$ accounts for the photodetector and quantization noises, i.e. $\sigma_{\mathrm{n}}^2=\sigma_{\mathrm{p}}^2+\sigma_s^2\kappa^{-2}$.

Equations (3.2) for $\mathcal{H}$ reduce to equation (3.3) for $\mathcal{C}$ only when

$$\hat{\Phi}_L(v,\omega) = \begin{cases} \sigma_L^2, & (v,\omega) \in \hat{B} \\ 0, & \text{elsewhere,} \end{cases} \qquad \hat{\tau}(v,\omega) = \begin{cases} 1, & (v,\omega) \in \hat{B} \\ 0, & \text{elsewhere.} \end{cases} \tag{3.4}$$

Because these two conditions cannot occur in practice, the best that can be done, as § 5 illustrates, is to maximize $\mathcal{H}$ by appropriately matching both $\hat{B}$ to $\hat{\Phi}_L(v,\omega)$ and $\hat{\tau}(v,\omega)$ to $\hat{B}$. Thus, while $\mathcal{C}$ still sets the *theoretical* upper bound on $\mathcal{H}$, the maximum-realizable information rate $\mathcal{H}_m$ sets a *practical* upper bound as a function of the critical limiting factors that constrain the image-gathering process.

### (*b*) *Theoretical minimum data rate* $\mathcal{E}$

The theoretical minimum data rate $\mathcal{E}$ that is required to sustain the information rate $\mathcal{H}$ may be represented, as Appendix C delineates, as the information rate through the analogue-to-digital transformation (figure 5). In short, the appendix first shows that for a general model of this transformation (i.e. one that is not constrained by the assumptions in § 2 *a* about the quantization process), the mutual information between the discrete input signal (with analogue amplitude) and the digital output signal is precisely equal to the entropy of the latter. Based on this result, we use the entropy symbol $\mathcal{E}$ to denote the mutual information between the discrete signal $s(x,y)$ and the digital signal $s(x,y;\kappa)$ for the approximate model of the quantization process in § 2 *a*. Accordingly, the entropy $\mathcal{E}$ is defined as

$$\mathcal{E} = \mathcal{E}[s(x,y;\kappa)] - \mathcal{E}[s(x,y;\kappa)|s(x,y)],$$

where the conditional entropy $\mathcal{E}[s(x,y;\kappa)|s(x,y)]$ is the uncertainty of $s(x,y;\kappa)$ when $s(x,y)$ is known. For the assumption that the quantization noise $n_{\mathrm{q}}(x,y;\kappa)$ is additive and independent, the conditional entropy becomes, according to Shannon's theorem 16, the entropy $\mathcal{E}[n_{\mathrm{q}}(x,y;\kappa)]$ of the quantization noise, so that

$$\mathcal{E} = \mathcal{E}[s(x,y;\kappa)] - \mathcal{E}[n_{\mathrm{q}}(x,y;\kappa)] \tag{3.5 a}$$

$$= \mathcal{E}[\tilde{s}(v,\omega;\kappa)] - \mathcal{E}[\tilde{n}_{\mathrm{q}}(v,\omega;\kappa)]. \tag{3.5 b}$$

A full interpretation of equation (3.5 *b*) in terms of the probability densities

$$p_s[\tilde{s}(v,\omega;\kappa)] \quad \text{and} \quad p_{\mathrm{q}}[\tilde{n}_{\mathrm{q}}(v,\omega;\kappa)]$$

of the digital signal and the quantization noise, respectively, follows the same steps that lead from equation (3.2 *b*) to equation (3.2 *c*) for $\mathcal{H}$, except that $\tilde{n}_{\mathrm{q}}(v,\omega;\kappa)$ now replaces $\hat{n}(v,\omega;\kappa)$. When this is done, then entropy $\mathcal{E}$ given by equation (3.5 *b*) is found to be

$$\mathcal{E} = \frac{1}{2}\iint_{\hat{B}} \log \tilde{\Phi}_s(v,\omega;\kappa)\,\mathrm{d}v\mathrm{d}\omega - \frac{1}{2}\iint_{\hat{B}} \log \tilde{\Phi}_{\mathrm{q}}(v,\omega;\kappa)\,\mathrm{d}v\mathrm{d}\omega \tag{3.5 c}$$

$$= \frac{1}{2}\iint_{\hat{B}} \log \frac{\tilde{\Phi}_s(v,\omega;\kappa)}{\tilde{\Phi}_{\mathrm{q}}(v,\omega;\kappa)}\,\mathrm{d}v\mathrm{d}\omega = \frac{1}{2}\iint_{\hat{B}} \log \frac{\tilde{\Phi}_s(v,\omega) + \tilde{\Phi}_{\mathrm{q}}(v,\omega;\kappa)}{\tilde{\Phi}_{\mathrm{q}}(v,\omega;\kappa)}\,\mathrm{d}v\mathrm{d}\omega \tag{3.5 d}$$

$$= \frac{1}{2}\iint_{\hat{B}} \log\left[1 + \frac{\tilde{\Phi}_s(v,\omega)}{\tilde{\Phi}_{\mathrm{q}}(v,\omega;\kappa)}\right]\mathrm{d}v\mathrm{d}\omega. \tag{3.5 e}$$

Substituting the PSD $\tilde{\Phi}_s(v,\omega)$ given by equation (2.2) yields

$$\mathcal{E} = \frac{1}{2}\iint_{\hat{B}} \log\left[1 + \frac{\hat{\Phi}_L(v,\omega)|\hat{\tau}(v,\omega)|^2 * \hat{\underline{\underline{|||}}} + K^{-2}\tilde{\Phi}_{\mathrm{p}}(v,\omega)}{K^{-2}\tilde{\Phi}_{\mathrm{q}}(v,\omega;\kappa)}\right] \mathrm{d}v\mathrm{d}\omega. \tag{3.5 f}$$

For the assumptions that $\tilde{\Phi}_{\mathrm{p}}(v,\omega) = \sigma_{\mathrm{p}}^2$ and $\tilde{\Phi}_{\mathrm{q}}(v,\omega;\kappa) = \sigma_{\mathrm{q}}^2$,

$$\mathcal{E} = \frac{1}{2}\iint_{\hat{B}} \log\left[1 + \frac{\hat{\Phi}'_L(v,\omega)|\hat{\tau}(v,\omega)|^2 * \hat{\underline{\underline{|||}}} + (K\sigma_L/\sigma_{\mathrm{p}})^{-2}}{(K\sigma_L/\sigma_s)^{-2}\kappa^{-2}}\right] \mathrm{d}v\mathrm{d}\omega. \tag{3.5 g}$$

These expressions for $\mathcal{E}$ represent the entropy of completely decorrelated data. They set the theoretical lower bound on the data rate that is associated with $\mathcal{H}$.

### (c) *Information efficiency* $\mathcal{H}/\mathcal{E}$

It is intuitively attractive to define the information efficiency of completely decorrelated data by the ratio $\mathcal{H}/\mathcal{E}$. Thus, whereas $\mathcal{H}$ ties the performance of the visual communication channel to the *quality* with which images can be restored, $\mathcal{H}/\mathcal{E}$ ties its performance to the *quantity* of data that must be transmitted and stored.

The conditions for which $\mathcal{H}/\mathcal{E}$ reaches its upper bound are easily found. Both equation (3.2 *f*) for $\mathcal{H}$ and equation (3.5 *f*) for $\mathcal{E}$ reduce to the same expression

$$\mathcal{E}_{\mathrm{q}} = \mathcal{H}_{\mathrm{q}} = \frac{1}{2}\iint_{\hat{B}} \log\left[1 + \frac{\hat{\Phi}_L(v,\omega)|\hat{\tau}(v,\omega)|^2}{K^{-2}\tilde{\Phi}_{\mathrm{q}}(v,\omega;\kappa)}\right] \mathrm{d}v\mathrm{d}\omega$$

when the aliasing and photodetector noises become very small relative to the quantization noise for $\mathcal{H}$ and to the within-passband signal component for $\mathcal{E}$. These conditions, therefore, are the ones for which $\mathcal{H}/\mathcal{E}$ reaches its maximum possible ratio $\mathcal{H}_{\mathrm{q}}/\mathcal{E}_{\mathrm{q}} = 1$. This result suggests that the electro-optical design of the image-gathering device that maximizes $\mathcal{H}$ also maximizes $\mathcal{H}/\mathcal{E}$. However, there always remains a trade-off between $\mathcal{H}$ and $\mathcal{H}/\mathcal{E}$ in the selection of the number of quantization levels: $\mathcal{H}$ favours fine quantization and $\mathcal{H}/\mathcal{E}$ favours coarse quantization.

Furthermore, if two additional conditions are imposed on the above equation for $\mathcal{E}_{\mathrm{q}} = \mathcal{H}_{\mathrm{q}}$, namely, the ones for which $\mathcal{H}$ reaches $\mathcal{C}$ given by equation (3.3), then

$$\mathcal{E}_{\mathrm{q}} = \mathcal{C}_{\mathrm{q}} = \tfrac{1}{2}|\hat{B}|\log[1 + (K\sigma_L/\sigma_{\mathrm{q}})^2].$$

Finally, for $\sigma_{\mathrm{q}}^2 = \sigma_s^2\kappa^{-2}$ from equation (2.4 *a*), where $\sigma_s^2 = K^2\sigma_L^2$,

$$\mathcal{E}_{\mathrm{q}} = \mathcal{C}_{\mathrm{q}} = \tfrac{1}{2}|\hat{B}|\log(1 + 3\kappa^2/c^2) \approx |\hat{B}|\log(\sqrt{3}\kappa/c).$$

Comparing this result with the maximum possible entropy $\mathcal{E}_m = |\hat{B}|\eta = |\hat{B}|\log\kappa$, it can be seen that $\mathcal{C}_{\mathrm{q}} = \mathcal{E}_{\mathrm{q}} = \mathcal{E}_m$ when $c = \sqrt{3}$. This constant is implicitly included in equation (2.9 *c*) for $\hat{\Psi}(v,\omega)$, equation (3.2 *g*) for $\mathcal{H}$, and equation (3.5 *g*) for $\mathcal{E}$.

### (d) *Maximum-realizable fidelity* $\mathcal{F}$

The fidelity $\boldsymbol{F}$ is a measure of the similarity between the radiance field $L(x,y)$ and the image $R(x,y;\kappa)$, as defined by (Fellgett & Linfoot 1955; Linfoot 1955)

$$\boldsymbol{F} = 1 - \frac{\iint_A \overline{|L(x,y) - R(x,y;\kappa)|^2}\,\mathrm{d}x\mathrm{d}y}{\iint_A \overline{|L(x,y)|^2}\,\mathrm{d}x\mathrm{d}y} \tag{3.6 a}$$




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

$$= 1 - \frac{|A|^{-1} \iint \overline{|\hat{L}(v,\omega) - \hat{R}(v,\omega;\kappa)|^2}\, \mathrm{d}v\mathrm{d}\omega}{|A|^{-1} \iint \overline{|\hat{L}(v,\omega)|^2}\, \mathrm{d}v\mathrm{d}\omega} \tag{3.6 b}$$

$$= 1 - \sigma_L^{-2} \int\!\!\int_{-\infty}^{\infty} \hat{e}^2(v,\omega;\kappa)\, \mathrm{d}v\mathrm{d}\omega. \tag{3.6 c}$$

Substituting the minimum MSRE PSD $\hat{\epsilon}^2(v,\omega;\kappa)$ given by equation (2.10) into equation (3.6 *c*) for $\hat{e}^2(v,\omega;\kappa)$ yields the maximum-realizable fidelity

$$\mathcal{F} = \sigma_L^{-2} \int\!\!\int_{-\infty}^{\infty} \hat{\Phi}_L(v,\omega) \hat{\Gamma}(v,\omega;\kappa)\, \mathrm{d}v\mathrm{d}\omega \tag{3.7 a}$$

with which the Wiener filter $\hat{\Psi}(v,\omega;\kappa)$ restores the image $\hat{\mathcal{R}}(v,\omega;\kappa)$ in digital form.

The Wiener filter $\hat{\Psi}(v,\omega;\kappa)$ given by equation (2.9 *a*) and $\mathcal{F}$ given by equation (3.7 *a*) can be expressed in terms of the spectral distribution $\hat{\mathcal{H}}(v,\omega;\kappa)$ of $\mathcal{H}$ as follows (Frieden 1970). Solving the integrand of equation (3.2 *h*) given by

$$\hat{\mathcal{H}}(v,\omega;\kappa) = -\log[1 - \hat{\Gamma}(v,\omega;\kappa)]$$

for the throughput SFR $\hat{\Gamma}(v,\omega;\kappa)$ and equating the result to equation (2.11) yields

$$\hat{\Gamma}(v,\omega;\kappa) = \hat{\tau}(v,\omega)\, \hat{\Psi}(v,\omega;\kappa) = 1 - 2^{-\hat{\mathcal{H}}(v,\omega;\kappa)}.$$

Hence,

$$\hat{\Psi}(v,\omega;\kappa) = \frac{1}{\hat{\tau}(v,\omega)}[1 - 2^{-\hat{\mathcal{H}}(v,\omega;\kappa)}]$$

and

$$\mathcal{F} = \sigma_L^{-2} \int\!\!\int_{-\infty}^{\infty} \hat{\Phi}_L(v,\omega)[1 - 2^{-\hat{\mathcal{H}}(v,\omega;\kappa)}]\, \mathrm{d}v\mathrm{d}\omega. \tag{3.7 b}$$

Consequently, for the conditions of very high information rates $\mathcal{H}$, for which the aliasing, photodetector, and quantization noises approach zero, $\hat{\Psi}(v,\omega;\kappa)$ approaches the inverse filter $\hat{\tau}^{-1}(v,\omega)$, and $\mathcal{F}$ approaches its absolute maximum value for a particular scene.

In many applications, the goal may be to restore some enhanced representation $\hat{L}(v,\omega)\hat{\tau}_e(v,\omega)$. If $\hat{\tau}_e(v,\omega)$ is a linear filter, such as the enhancement filter given by equation (2.14), then the relationships given by equations (3.7) for $\mathcal{F}$ can readily be extended to similar relationships for $\mathcal{F}_e$ that the filter $\hat{\Psi}(v,\omega;\kappa)\hat{\tau}_e(v,\omega)$ produces, as given by

$$\mathcal{F}_e = \frac{\sigma_L^2}{\sigma_{Le}^2} \int\!\!\int_{-\infty}^{\infty} \hat{\Phi}'_L(v,\omega)|\hat{\tau}_e(v,\omega)|^2 \hat{\Gamma}(v,\omega;\kappa)\, \mathrm{d}v\mathrm{d}\omega \tag{3.8 a}$$

$$= \frac{\sigma_L^2}{\sigma_{Le}^2} \int\!\!\int_{-\infty}^{\infty} \hat{\Phi}'_L(v,\omega)|\hat{\tau}_e(v,\omega)|^2 [1 - 2^{-\hat{\mathcal{H}}(v,\omega;\kappa)}]\, \mathrm{d}v\mathrm{d}\omega, \tag{3.8 b}$$




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

where

$$\sigma_{Le}^2 = \sigma_L^2 \iint_{-\infty}^{\infty} \hat{\Phi}'_L(v,\omega)|\hat{\tau}_e(v,\omega)|^2 \,\mathrm{d}v\mathrm{d}\omega.$$

### (e) *Information rate* $\mathcal{H}_o$

The reconstruction noise $N_r(x,y)$ that the third transformation in figure 5 adds when the digitally restored image is recorded on an image-display medium can be directly included with the other perturbations that occur in the visual communication channel. When this is done by repeating the steps that lead from equation (3.2 *b*) to equation (3.2 *f*) for $\mathcal{H}$, except that $\hat{n}(v,\omega;\kappa)$ now includes $\hat{N}_r(v,\omega)$, then the information rate $\mathcal{H}_o$ of the observed image is found to be

$$\mathcal{H}_o = \frac{1}{2}\iint_{\hat{B}} \log\left[1 + \frac{\hat{\Phi}_L(v,\omega)|\hat{\tau}(v,\omega)|^2}{\hat{\Phi}_L(v,\omega)|\hat{\tau}(v,\omega)|^2 * \underline{|\hat{|}|}_s + K^{-2}[\tilde{\Phi}_p(v,\omega) + \tilde{\Phi}_q(v,\omega;\kappa)] + \hat{\Phi}'_r(v,\omega;\kappa)}\right] \mathrm{d}v\mathrm{d}\omega, \tag{3.9 a}$$

where

$$\hat{\Phi}'_r(v,\omega;\kappa) = \hat{\Phi}_r(v,\omega)|\,\hat{\Psi}(v,\omega;\kappa)|^{-2}.$$

For the assumptions that

$$\tilde{\Phi}_p(v,\omega) = \sigma_p^2, \qquad \tilde{\Phi}_q(v,\omega;\kappa) = \sigma_q^2, \qquad \hat{\Phi}'_r(v,\omega;\kappa) = \sigma_r^2,$$

$$\mathcal{H}_o = \frac{1}{2}\iint_{\hat{B}} \log\left[1 + \frac{\hat{\Phi}'_L(v,\omega)|\hat{\tau}(v,\omega)|^2}{\hat{\Phi}'_L(v,\omega)|\hat{\tau}(v,\omega)|^2 * \underline{|\hat{|}|}_s + (K\sigma_L/\sigma_p)^{-2} + (K\sigma_L/\sigma_s)^{-2}\kappa^{-2} + (\sigma_L/\sigma_r)^{-2}}\right] \mathrm{d}v\mathrm{d}\omega. \tag{3.9 b}$$

These expressions for $\mathcal{H}_o$ are based on the assumption that $\hat{\Psi}(v,\omega;\kappa)$ restores the observed image $\mathcal{R}_o(x,y;\kappa)$ with an interpolation lattice that effectively suppresses the blurring and raster effects of the reconstruction process. If this dense digital interpolation is not practical, then it becomes necessary to account for the reconstruction response and raster, as formulated in Appendix D.

The effect of the reconstruction noise may be assessed in terms of the number of distinguishable grey levels $\iota$. The last result in Appendix C suggests that the number of levels $\iota$ and the SNR $\sigma_s/K\sigma_r$ can be related by the expression

$$\iota = (1 + \sigma_s^2/K^2\sigma_r^2)^{1/2} \approx \sigma_s/K\sigma_r.$$

Substituting $\sigma_r = \sigma_s/\iota$ into equation (3.9 *c*) yields

$$\mathcal{H}_o = \frac{1}{2}\iint_{\hat{B}} \log\left[1 + \frac{\hat{\Phi}'_L(v,\omega)|\hat{\tau}(v,\omega)|^2}{\hat{\Phi}'_L(v,\omega)|\hat{\tau}(v,\omega)|^2 * \underline{|\hat{|}|}_s + (K\sigma_L/\sigma_p)^{-2} + (K\sigma_L/\sigma_s)^{-2}(\kappa^{-2} + \iota^{-2})}\right] \mathrm{d}v\mathrm{d}\omega. \tag{3.9 c}$$

This expression shows that the number of quantization levels $\kappa$ and the number of distinguishable grey levels $\iota$ are equivalent in terms of their significance to $\mathcal{H}_o$. This equivalence suggests that the information rate to the observer cannot be significantly improved by increasing only $\kappa$ or $\iota$, one without regard to the other.

### (*f*) *Maximum-realizable fidelity* $\mathcal{F}_o$

The maximum-realizable fidelity $\mathcal{F}_o$ of the observed image that is restored with dense interpolation differs from the maximum-realizable fidelity $\mathcal{F}$ of the digital image only by the addition of the reconstruction noise $N_r(x, y)$. Substituting the minimum MSRE PSD $\hat{\epsilon}_o^2(\upsilon, \omega; \kappa)$ given by equation (2.12) for $\hat{e}^2(\upsilon, \omega; \kappa)$ into equation (3.6*c*) yields the maximum-realizable fidelity

$$\mathcal{F}_o = \sigma_L^{-2} \int_{-\infty}^{\infty}\int [\hat{\Phi}_L(\upsilon, \omega)\hat{\Gamma}(\upsilon, \omega; \kappa) - \hat{\Phi}_r(\upsilon, \omega)]\, \mathrm{d}\upsilon \mathrm{d}\omega \tag{3.10a}$$

$$= \sigma_L^{-2} \int_{-\infty}^{\infty}\int \{\hat{\Phi}_L(\upsilon, \omega)[1 - 2^{-\hat{\mathcal{H}}(\upsilon,\omega;\kappa)}] - \hat{\Phi}_r(\upsilon, \omega)\}\, \mathrm{d}\upsilon \mathrm{d}\omega. \tag{3.10b}$$

The spectral information density $\hat{\mathcal{H}}(\upsilon, \omega; \kappa)$ in equation (3.10*b*) is the integrand of $\mathcal{H}$ given by equation (3.2*f*). For the assumption that $\hat{\Phi}_r(\upsilon, \omega) = \sigma_r^2$,

$$\mathcal{F}_o = \sigma_L^{-2} \int_{-\infty}^{\infty}\int \hat{\Phi}_L(\upsilon, \omega)\hat{\Gamma}(\upsilon, \omega; \kappa)\, \mathrm{d}\upsilon \mathrm{d}\omega - (\sigma_L/\sigma_r)^{-2}. \tag{3.10c}$$

Finally, if the effect of the reconstruction noise is approximated by the number of distinguishable grey levels $\iota$, then the term $(\sigma_L/\sigma_r)^{-2}$ in equation (3.10*c*) becomes $(K\sigma_L/\sigma_s)^{-2}\iota^{-2}$.

## 4. Multiresolution decomposition

Multiresolution processing has been implemented with many different architectures (tree structures) and filters (operators) for signal decomposition (analysis) and reconstruction (synthesis) (Gabor 1946; Burt & Adelson 1983; Daubechies 1988, 1990; Mallat 1989*a*, *b*; Ebrahimi & Kunt 1991; Woods 1991; Vetterli & Herley 1992; Akansu & Haddad 1992). This section starts with a single-level decomposition and then extends this decomposition to a particular multi-level realization, the wavelet transform. As above in § 2, we clearly distinguish between image reconstruction and restoration. For image reconstruction, we follow the traditional treatment of signal synthesis and image reconstruction as independent processes; whereas for image restoration, we integrate signal synthesis with the Wiener filter to again produce images with the maximum-realizable fidelity. The information-theoretic assessment applies, of course, only to the integrated synthesis and restoration.

### (*a*) *Single-level transform*

(i) *Image gathering and signal analysis*

Figure 6*a* depicts a model of image gathering with a single level of multiresolution signal analysis. The discrete Fourier transform (DFT) allows the analysis and requantization to be performed in the spatial frequency domain. The quantization noise

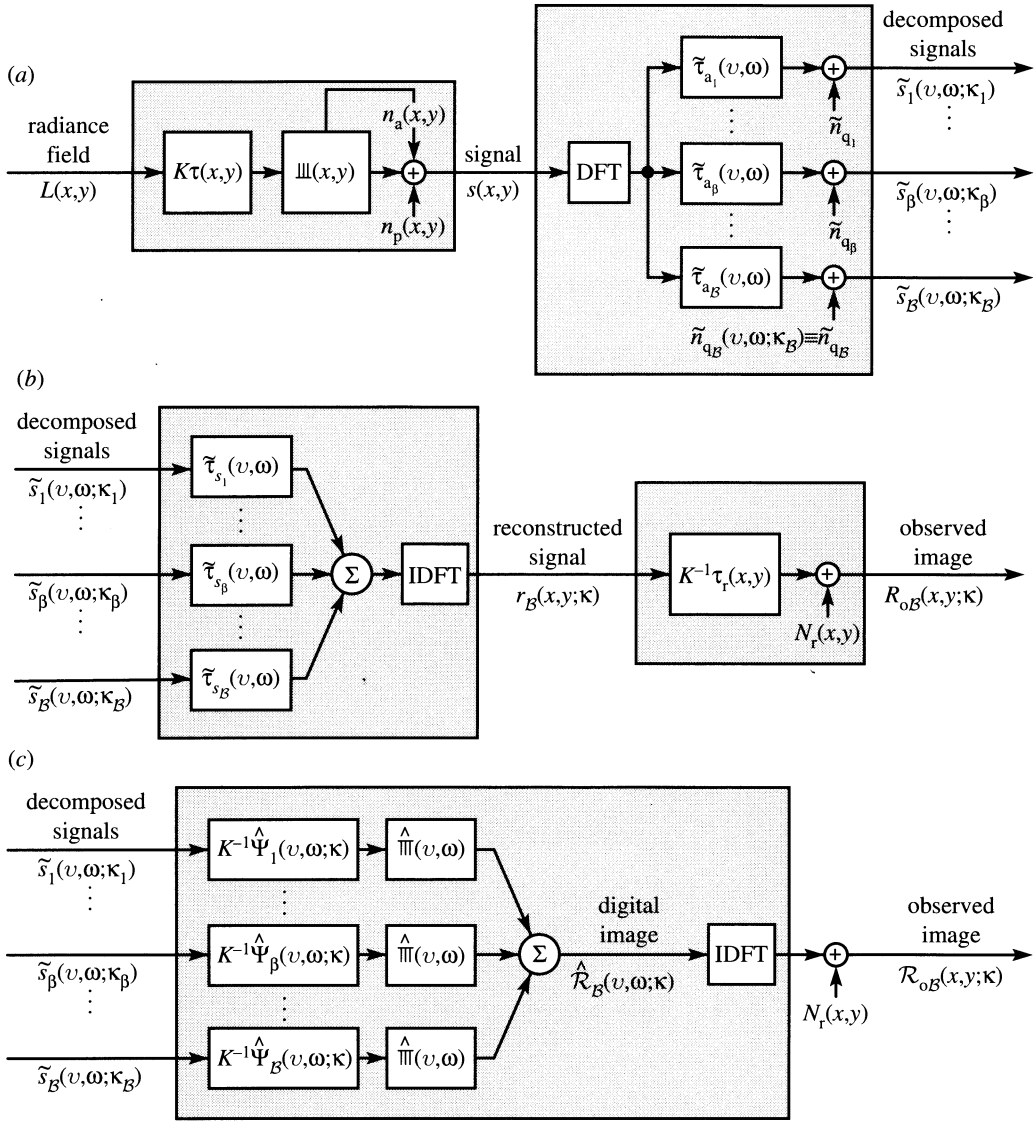



Figure 6. Model of image gathering and reproduction with single-level multiresolution decomposition. (*a*) Image gathering and signal analysis, (*b*) signal synthesis and image reconstruction, (*c*) integrated signal synthesis and image restoration. The requantization replaces the quantization in the analogue-to-digital transformation for the trade-off between information rate and data rate. The processing and requantization are performed in the spatial frequency domain.

$n_q(x,y;\kappa)$ of the analogue-to-digital transformation (figures 4 and 5) is replaced by the requantization noises $n_{q\beta}(x,y;\kappa_\beta)$ because it is the number of requantization levels that now critically constrains the trade-off between the information rate $\mathcal{H}$ and associated theoretical minimum data rate $\mathcal{E}$. With this simplification, the image gathering and signal analysis together transform the continuous radiance field $L(x,y)$ into the decomposed digital signals $s_\beta(x,y;\kappa_\beta)$, as defined by

$$s_\beta(x,y;\kappa_\beta) = [KL(x,y) * \tau(x,y) * \tau_{a\beta}(x,y)]\,\underline{|||}_\beta + [n_p(x,y) * \tau_{a\beta}(x,y)]\,\underline{|||}_\beta + n_{q\beta}(x,y;\kappa_\beta), \quad (4.1\,a)$$




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

where $\beta = 1, 2, \ldots, \mathcal{B}$, $\tau_{a\beta}(x, y)$ are the spatial responses of the analysis filters, and the functions

$$\underline{\underline{|||}}_\beta \equiv \underline{\underline{|||}}_\beta(x, y) = X_\beta Y_\beta \sum_{m,n} \delta(\upsilon - X_\beta m, \omega - Y_\beta n)$$

are the associated decimation lattices with sampling intervals $(X_\beta, Y_\beta)$. Note that the intervals $X_\beta = Y_\beta \geqslant 1$ for the image-gathering process with unit sampling intervals, and that equation (4.1 *a*) reduces to equation (2.1 *a*) when the decomposition is removed by letting $\tau_{a\beta}(x, y) = \delta(x, y)$ and $X_\beta = Y_\beta = 1$. The Fourier transform of equation (4.1 *a*) is

$$\begin{aligned}\tilde{s}_\beta(\upsilon, \omega; \kappa_\beta) = {} & [K\hat{L}(\upsilon, \omega)\hat{\tau}(\upsilon, \omega)\tilde{\tau}_{a\beta}(\upsilon, \omega)] * \underline{\underline{\hat{|||}}}_\beta \\ & + [\tilde{n}_p(\upsilon, \omega)\tilde{\tau}_{a\beta}(\upsilon, \omega)] * \underline{\underline{\hat{|||}}}'_\beta + \tilde{n}_{q\beta}(\upsilon, \omega; \kappa_\beta), \end{aligned} \tag{4.1 b}$$

where $\tilde{\tau}_{a\beta}(\upsilon, \omega)$ are the SFRs of the analysis filters, $\tilde{n}_{q\beta}(\upsilon, \omega; \kappa_\beta)$ are the discrete transforms of the requantization noises, and

$$\underline{\underline{\hat{|||}}}_\beta \equiv \underline{\underline{\hat{|||}}}_\beta(\upsilon, \omega) = \sum_{m,n} \delta\left(\upsilon - \frac{m}{X_\beta}, \omega - \frac{n}{Y_\beta}\right)$$

$$\underline{\underline{\hat{|||}}}'_\beta \equiv \underline{\underline{\hat{|||}}}'_\beta(\upsilon, \omega) = \sum_{m=0}^{X_\beta - 1} \sum_{n=0}^{Y_\beta - 1} \delta\left(\upsilon - \frac{m}{X_\beta}, \omega - \frac{n}{Y_\beta}\right)$$

account for the decimation. The associated sampling passband for each decomposed digital signal is

$$\hat{B}_\beta = \left[(\upsilon, \omega); |\upsilon| \leqslant \frac{1}{2X_\beta}, |\omega| \leqslant \frac{1}{2Y_\beta}\right]$$

with area $|\hat{B}_\beta| = (X_\beta Y_\beta)^{-1}$.

The PSDs $\tilde{\Phi}_{s\beta}(\upsilon, \omega)$ of the decomposed signals before requantization are

$$\tilde{\Phi}_{s\beta}(\upsilon, \omega) = \tilde{\Phi}_s(\upsilon, \omega)|\tilde{\tau}_{a\beta}(\upsilon, \omega)|^2, \tag{4.2}$$

where $\tilde{\Phi}_s(\upsilon, \omega)$ is given by equation (2.2). The corresponding variances $\sigma^2_{s\beta}$ are

$$\begin{aligned}\sigma^2_{s\beta} = {} & \iint_{\hat{B}_\beta} \tilde{\Phi}_{s\beta}(\upsilon, \omega)\, \mathrm{d}\upsilon \mathrm{d}\omega \\ \approx {} & K^2\sigma_L^2 \int\!\!\int_{-\infty}^{\infty} \hat{\Phi}'_L(\upsilon, \omega)|\hat{\tau}(\upsilon, \omega)|^2|\tilde{\tau}_{a\beta}(\upsilon, \omega)|^2\, \mathrm{d}\upsilon \mathrm{d}\omega. \end{aligned} \tag{4.3}$$

Moreover, consistent with the assumptions about analogue-to-digital transformation in § 2 *a*, the PSDs $\tilde{\Phi}_{q\beta}(\upsilon, \omega; \kappa_\beta)$ of the requantization noises are

$$\tilde{\Phi}_{q\beta}(\upsilon, \omega; \kappa_\beta) = X_\beta Y_\beta \left[\frac{\sigma_{s\beta}(\upsilon, \omega)}{\kappa_\beta(\upsilon, \omega)}\right]^2, \tag{4.4}$$

where $\sigma_{s\beta}(\upsilon, \omega)$ are the standard deviations of the discrete spatial frequency components $(\upsilon, \omega)$ in band $\beta$, and $\kappa_\beta(\upsilon, \omega)$ are the number of requantization levels assigned to these components.

(ii) *Signal synthesis and image reconstruction*

The signal synthesis depicted in figure 6*b* reassembles the $\mathcal{B}$ decomposed signals into the reconstructed signal $\tilde{r}_{\mathcal{B}}(\upsilon,\omega;\kappa)$, as defined by

$$\tilde{r}_{\mathcal{B}}(\upsilon,\omega;\kappa)=\sum_{\beta=1}^{\mathcal{B}}\tilde{s}_{\beta}(\upsilon,\omega;\kappa_{\beta})\tilde{\tau}_{s\beta}(\upsilon,\omega), \tag{4.5}$$

where $\tilde{\tau}_{s\beta}(\upsilon,\omega)$ are the SFRs of the synthesis filters. The conditions for perfect reconstruction, for which $\tilde{r}_{\mathcal{B}}(\upsilon,\omega;\kappa)=\tilde{s}(\upsilon,\omega)$, are as follows: the requantization noises $\tilde{n}_{q\beta}(\upsilon,\omega;\kappa_{\beta})$ are negligible; the synthesis filters are given by

$$\tilde{\tau}_{s\beta}(\upsilon,\omega)=\sum_{\alpha=1}^{\mathcal{B}}\tilde{\tau}^{*}_{a\alpha}(\upsilon,\omega)[\tilde{\mathcal{T}}_{s}^{-1}(\upsilon,\omega)]_{\alpha\beta}, \tag{4.6}$$

where

$$[\tilde{\mathcal{T}}_{s}(\upsilon,\omega)]_{\beta\alpha}=[\tilde{\tau}^{*}_{a\beta}(\upsilon,\omega)\tilde{\tau}_{a\alpha}(\upsilon,\omega)]*\underline{\hat{|||}}'_{\gamma}$$

and $\gamma=\max(\alpha,\beta)$; and $[\tilde{\mathcal{T}}_{s}^{-1}(\upsilon,\omega)]_{\alpha\beta}$ exists. The inverse discrete Fourier transform (IDFT) and the image-display device transform the signal $\tilde{r}_{\mathcal{B}}(\upsilon,\omega;\kappa)$ into the observed image $R_{o\mathcal{B}}(x,y;\kappa)$, as given by

$$R_{o\mathcal{B}}(x,y;\kappa)=K^{-1}r_{\mathcal{B}}(x,y;\kappa)*\tau_{r}(x,y)+N_{r}(x,y). \tag{4.7}$$

(iii) *Integrated signal synthesis and image restoration*

The integrated signal synthesis and image restoration depicted in figure 6*c* accounts for both the image gathering and signal analysis to minimize the mean-square error $e^{2}_{\mathcal{B}}$ between the radiance field $\hat{L}(\upsilon,\omega)$ and the digital image $\hat{\mathcal{R}}_{\mathcal{B}}(\upsilon,\omega;\kappa)$, as given by (Fales *et al.* 1992)

$$\hat{\mathcal{R}}_{\mathcal{B}}(\upsilon,\omega;\kappa)=\sum_{\beta=1}^{\mathcal{B}}K^{-1}\tilde{s}_{\beta}(\upsilon,\omega;\kappa_{\beta})\,\hat{\Psi}_{\beta}(\upsilon,\omega;\kappa_{\beta}), \tag{4.8}$$

where $\hat{\Psi}_{\beta}(\upsilon,\omega;\kappa_{\beta})$ is the Wiener-matrix filter

$$\hat{\Psi}_{\beta}(\upsilon,\omega;\kappa_{\beta})=\hat{\Phi}'_{L}(\upsilon,\omega)\hat{\tau}^{*}(\upsilon,\omega)\sum_{\alpha=1}^{\mathcal{B}}\tilde{\tau}^{*}_{a\alpha}(\upsilon,\omega)[\tilde{\mathcal{T}}^{-1}(\upsilon,\omega;\kappa_{\beta})]_{\alpha\beta} \tag{4.9}$$

and

$$\begin{aligned}[\tilde{\mathcal{T}}(\upsilon,\omega;\kappa_{\beta})]_{\beta\alpha}=&[\hat{\Phi}'_{L}(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^{2}\tilde{\tau}_{a\beta}(\upsilon,\omega)\tilde{\tau}^{*}_{a\alpha}(\upsilon,\omega)]*\underline{\hat{|||}}_{\gamma}\\&+(K\sigma_{L}/\sigma_{p})^{-2}[\tilde{\tau}_{a\beta}(\upsilon,\omega)\tilde{\tau}^{*}_{a\alpha}(\upsilon,\omega)]*\underline{\hat{|||}}'_{\gamma}\\&+X_{\beta}Y_{\beta}(K\sigma_{L})^{-2}\left[\frac{\sigma_{s\beta}(\upsilon,\omega)}{\kappa_{\beta}(\upsilon,\omega)}\right]^{-2}\delta(\beta,\alpha).\end{aligned}$$

The corresponding minimum MSRE PSD $\hat{\epsilon}^{2}_{\mathcal{B}}(\upsilon,\omega;\kappa)$ is

$$\hat{\epsilon}^{2}_{\mathcal{B}}(\upsilon,\omega;\kappa)=\hat{\Phi}_{L}(\upsilon,\omega)[1-\hat{\Gamma}_{\mathcal{B}}(\upsilon,\omega;\kappa)], \tag{4.10}$$

where

$$\hat{\Gamma}_{\mathcal{B}}(\upsilon,\omega;\kappa)=\hat{\tau}(\upsilon,\omega)\sum_{\beta=1}^{\mathcal{B}}\tilde{\tau}_{a\beta}(\upsilon,\omega)\,\hat{\Psi}_{\beta}(\upsilon,\omega;\kappa_{\beta}) \tag{4.11}$$

is the throughput SFR. The observed image is

$$\mathcal{R}_{o\mathcal{B}}(x,y;\kappa) = \mathcal{R}_{\mathcal{B}}(x,y;\kappa) + N_r(x,y), \tag{4.12}$$

where $\mathcal{R}_{\mathcal{B}}(x,y;\kappa)$ is the IDFT of $\hat{\mathcal{R}}_{\mathcal{B}}(v,\omega;\kappa)$. The Wiener-matrix filter $\hat{\Psi}_\beta(v,\omega;\kappa_\beta)$ reduces to the Wiener filter $\hat{\Psi}(v,\omega;\kappa)$ given by equation (2.9 *b*) when $\mathcal{B} = 1$ and $\tilde{\tau}_{a\beta}(v,\omega) \equiv 1$.

(iv) *Figures of merit*

The information rate $\mathcal{H}_\beta$ for each decomposed signal $s_\beta(x,y;\kappa_\beta)$ is

$$\mathcal{H}_\beta = \frac{1}{2}\iint\limits_{\hat{B}_\beta} \log\left[1 + \frac{\hat{\Phi}_L(v,\omega)|\hat{\tau}(v,\omega)|^2|\tilde{\tau}_{a\beta}(v,\omega)|^2}{[\hat{\Phi}_L(v,\omega)|\hat{\tau}(v,\omega)|^2 * \underline{\underline{|\hat{|}|}}_s + \tilde{\Phi}_p(v,\omega)]|\tilde{\tau}_{a\beta}(v,\omega)|^2 + \tilde{\Phi}_{q\beta}(v,\omega;\kappa_\beta)}\right] dv d\omega \tag{4.13}$$

and the associated theoretical minimum data rate $\mathcal{E}_\beta$ is

$$\mathcal{E}_\beta = \frac{1}{2}\iint\limits_{\hat{B}_\beta} \log\left[1 + \frac{\tilde{\Phi}_{s\beta}(v,\omega)}{\tilde{\Phi}_{q\beta}(v,\omega;\kappa_\beta)}\right] dv d\omega. \tag{4.14}$$

The total information rate $\mathcal{H}_{\mathcal{B}}$ for the $\mathcal{B}$ decomposed signals is

$$\mathcal{H}_{\mathcal{B}} = -\frac{1}{2}\iint\limits_{\hat{B}_{\mathcal{B}}} \log[1 - \hat{\Gamma}_{\mathcal{B}}(v,\omega;\kappa)]\, dv d\omega = \frac{1}{2}\iint\limits_{\hat{B}_{\mathcal{B}}} \log \frac{\hat{\Phi}_L(v,\omega)}{\hat{\epsilon}^2_{\mathcal{B}}(v,\omega;\kappa)}\, dv d\omega, \tag{4.15}$$

where $\hat{\epsilon}^2_{\mathcal{B}}(v,\omega;\kappa)$ and $\hat{\Gamma}_{\mathcal{B}}(v,\omega;\kappa)$ are given by equations (4.10) and (4.11), respectively. Normally, $\hat{B}_{\mathcal{B}} = \hat{B}$. The associated theoretical minimum data rate $\mathcal{E}_{\mathcal{B}}$ is

$$\mathcal{E}_{\mathcal{B}} = \sum_{\beta=1}^{\mathcal{B}} \mathcal{E}_\beta = \frac{1}{2}\sum_{\beta=1}^{\mathcal{B}} \iint\limits_{\hat{B}_\beta} \log\left[1 + \frac{\tilde{\Phi}_{s\beta}(v,\omega)}{\tilde{\Phi}_{q\beta}(v,\omega;\kappa_\beta)}\right] dv d\omega. \tag{4.16}$$

The maximum-realizable fidelity $\mathcal{F}_{\mathcal{B}}$ of the digital image $\mathcal{R}_{\mathcal{B}}(x,y;\kappa)$ is

$$\mathcal{F}_{\mathcal{B}} = \sigma_L^{-2} \iint\limits_{-\infty}^{\infty} \hat{\Phi}_L(v,\omega)\hat{\Gamma}_{\mathcal{B}}(v,\omega;\kappa)\, dv d\omega. \tag{4.17}$$

Finally, the information rate $\mathcal{H}_{o\mathcal{B}}$ and fidelity $\mathcal{F}_{o\mathcal{B}}$ of the observed image $\mathcal{R}_{o\mathcal{B}}(x,y;\kappa)$ are

$$\mathcal{H}_{o\mathcal{B}} = \frac{1}{2}\iint\limits_{\hat{B}_{\mathcal{B}}} \log\left[1 + \frac{\hat{\Phi}_L(v,\omega)\hat{\Gamma}^2_{\mathcal{B}}(v,\omega;\kappa)}{\hat{\Phi}_L(v,\omega)\hat{\Gamma}_{\mathcal{B}}(v,\omega;\kappa)[1 - \hat{\Gamma}_{\mathcal{B}}(v,\omega;\kappa)] + \hat{\Phi}_r(v,\omega)}\right] dv d\omega \tag{4.18}$$

and

$$\mathcal{F}_{o\mathcal{B}} = \sigma_L^{-2} \iint\limits_{-\infty}^{\infty} [\hat{\Phi}_L(v,\omega)\hat{\Gamma}_{\mathcal{B}}(v,\omega;\kappa) - \hat{\Phi}_r(v,\omega)]\, dv d\omega. \tag{4.19}$$

These formulations reduce to the ones given in § 3 when $\mathcal{B} = 1$ and $\tilde{\tau}_{a\beta}(v,\omega) \equiv 1$.

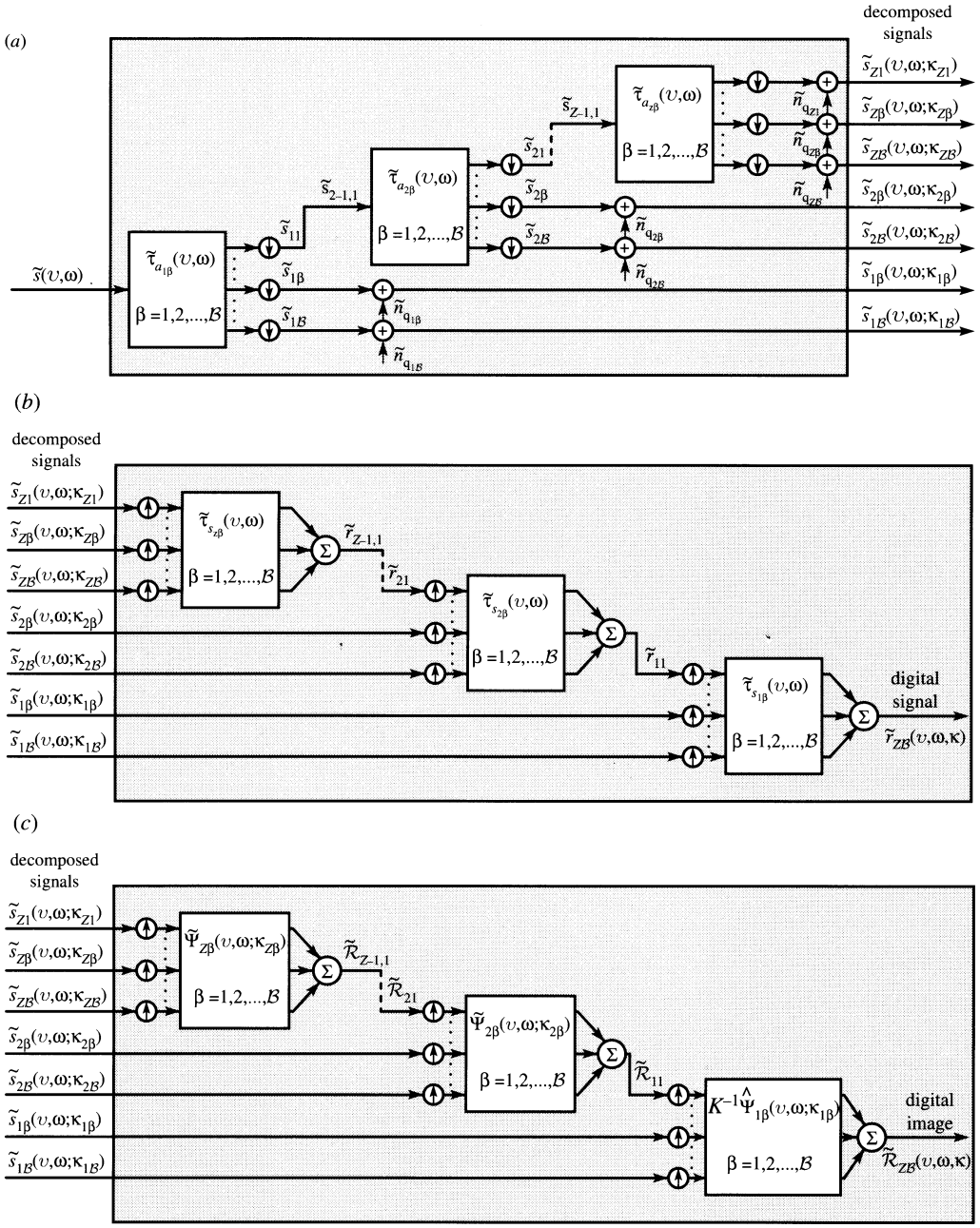



Figure 7. Multiresolution decomposition with the discrete wavelet transform. (*a*) Signal analysis, (*b*) signal synthesis, and (*c*) integrated signal synthesis and image restoration. Figure 6 depicts the associated image gathering, digital interpolation, and image reconstruction.

## (*b*) *Wavelet transform*

### (i) *Image gathering and signal analysis*

Figure 7*a* depicts the multi-level signal analysis architecture of the discrete wavelet transform. This transform decomposes the DFT of the signal $s(x, y)$ into the signals




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

$\tilde{s}_{\zeta\beta}(\upsilon,\omega;\kappa_{\zeta\beta})$ with $\boldsymbol{Z}$ levels, each with $\mathcal{B}$ bands, as defined by

$$\tilde{s}_{\zeta\beta}(\upsilon,\omega;\kappa_{\zeta\beta}) = [\tilde{s}_{\zeta-1,1}(\upsilon,\omega)\tilde{\tau}_{\mathrm{a}\zeta\beta}(\upsilon,\omega)] * \underline{\hat{|||}}_{\zeta} + \tilde{n}_{\mathrm{q}\zeta\beta}(\upsilon,\omega;\kappa_{\zeta\beta}), \tag{4.20a}$$

$$\beta = 2,3,\ldots,\mathcal{B}, \qquad \zeta = 1,2,\ldots,\boldsymbol{Z},$$

$$\tilde{s}_{\zeta 1}(\upsilon,\omega) = [\tilde{s}_{\zeta-1,1}(\upsilon,\omega)\tilde{\tau}_{\mathrm{a}\zeta 1}(\upsilon,\omega)] * \underline{\hat{|||}}_{\zeta}, \qquad \zeta = 1,2,\ldots,\boldsymbol{Z}-1, \tag{4.20b}$$

where

$$\tilde{s}_{0,1}(\upsilon,\omega) \equiv \tilde{s}(\upsilon,\omega) = [K\hat{L}(\upsilon,\omega)\hat{\tau}(\upsilon,\omega)] * \underline{\hat{|||}} + \tilde{n}_{\mathrm{p}}(\upsilon,\omega)$$

and

$$\underline{\hat{|||}}_{\zeta} = \sum_{m=0}^{X-1}\sum_{n=0}^{Y-1} \delta\left(\upsilon - \frac{m}{X^{\zeta}}, \omega - \frac{n}{Y^{\zeta}}\right)$$

is the down-sampler by $X = Y = \sqrt{\mathcal{B}}$. Hence, there are a total of $\boldsymbol{Z}(\mathcal{B}-1)+1$ decomposed signals. Compactly supported orthonormal wavelets with dyadic (octave) decimation, for which $\mathcal{B} = 4$ ($X = Y = 2$), are developed and presented in Daubechies (1988, 1990), Mallat (1989*a*, *b*), Akansu & Haddard (1992) and Vetterli & Herley (1992).

(ii) *Signal synthesis and image reconstruction*

The signal synthesis depicted in figure 7*b* reassembles the decomposed signals into the digital signal $\tilde{r}_{\boldsymbol{Z}\mathcal{B}}(\upsilon,\omega;\kappa)$, as defined by

$$\tilde{r}_{\boldsymbol{Z}\mathcal{B}}(\upsilon,\omega;\kappa) = \sum_{\beta=1}^{\mathcal{B}} \tilde{r}_{1\beta}(\upsilon,\omega;\kappa)\tilde{\tau}_{s1\beta}(\upsilon,\omega), \tag{4.21}$$

where

$$\left.\begin{aligned} \tilde{r}_{\zeta\beta}(\upsilon,\omega;\kappa) &\equiv \tilde{s}_{\zeta\beta}(\upsilon,\omega;\kappa_{\zeta\beta}), \qquad \beta = 2,3,\ldots,\mathcal{B}, \\ \tilde{r}_{\zeta 1}(\upsilon,\omega;\kappa) &= \sum_{\beta=1}^{\mathcal{B}} \tilde{r}_{\zeta+1,\beta}(\upsilon,\omega;\kappa)\tilde{\tau}_{s\zeta+1,\beta}(\upsilon,\omega), \end{aligned}\right\} \qquad \zeta = 1,2,\ldots,\boldsymbol{Z}-1$$

and

$$\tilde{r}_{\boldsymbol{Z}\beta}(\upsilon,\omega;\kappa) \equiv \tilde{s}_{\boldsymbol{Z}\beta}(\upsilon,\omega;\kappa_{\boldsymbol{Z}\beta}), \qquad \beta = 1,2,\ldots,\mathcal{B}.$$

The conditions for perfect reconstruction are analogous to those given by equation (4.6) for the single-level decomposition, and the formulation of the observed image $\hat{R}_{\mathrm{o}\boldsymbol{Z}\mathcal{B}}(\upsilon,\omega;\kappa)$ is analogous to $\hat{R}_{\mathrm{o}\mathcal{B}}(\upsilon,\omega;\kappa)$ given by equation (4.7).

(iii) *Integrated signal synthesis and image restoration*

The integrated signal synthesis and image restoration depicted in figure 7*c* transforms the decomposed signals into the digital image $\hat{\mathcal{R}}_{\boldsymbol{Z}\mathcal{B}}(\upsilon,\omega;\kappa)$, as given by

$$\hat{\mathcal{R}}_{\boldsymbol{Z}\mathcal{B}}(\upsilon,\omega;\kappa) = K^{-1}\sum_{\beta=1}^{\mathcal{B}} \tilde{\mathcal{R}}_{1\beta}(\upsilon,\omega;\kappa)\,\hat{\Psi}_{1\beta}(\upsilon,\omega;\kappa_{1\beta}), \tag{4.22}$$

where

$$\left.\begin{aligned} \tilde{\mathcal{R}}_{\zeta\beta}(\upsilon,\omega;\kappa) &\equiv \tilde{s}_{\zeta\beta}(\upsilon,\omega;\kappa_{\zeta\beta}), \qquad \beta = 2,3,\ldots,\mathcal{B}, \\ \tilde{\mathcal{R}}_{\zeta 1}(\upsilon,\omega;\kappa) &= \sum_{\beta=1}^{\mathcal{B}} \tilde{\mathcal{R}}_{\zeta+1,\beta}(\upsilon,\omega;\kappa)\,\tilde{\Psi}_{s\zeta+1,\beta}(\upsilon,\omega;\kappa_{\zeta+1,\beta}), \end{aligned}\right\} \qquad \zeta = 1,2,\ldots,\boldsymbol{Z}-1$$

and

$$\tilde{\mathcal{R}}_{\boldsymbol{Z}\beta}(\upsilon,\omega;\kappa) \equiv \tilde{s}_{\boldsymbol{Z}\beta}(\upsilon,\omega;\kappa_{\boldsymbol{Z}\beta}), \qquad \beta = 1,2,\ldots,\mathcal{B}.$$

The Wiener-matrix filters $\tilde{\Psi}_{\zeta\beta}(\upsilon,\omega;\kappa_{\zeta\beta})$ are

$$\tilde{\Psi}_{\zeta\beta}(\upsilon,\omega;\kappa_{\zeta\beta}) = \tilde{\Phi}_{s\zeta-1,1}(\upsilon,\omega)\sum_{\alpha=1}^{\mathcal{B}}\tilde{\tau}^{*}_{\mathrm{a}\zeta\beta}(\upsilon,\omega)[\tilde{\mathcal{T}}_{\zeta}^{-1}(\upsilon,\omega;\kappa_{\zeta\beta})]_{\alpha\beta}, \tag{4.23}$$

where

$$[\tilde{\mathcal{T}}_{\zeta}(\upsilon,\omega;\kappa_{\zeta\beta})]_{\beta\alpha} = [\tilde{\Phi}_{s\zeta}(\upsilon,\omega)\tilde{\tau}_{\mathrm{a}\zeta\beta}(\upsilon,\omega)\tilde{\tau}^{*}_{\mathrm{a}\zeta\alpha}(\upsilon,\omega)] * \underline{\hat{|||}}_{\zeta} + \tilde{\Phi}_{\mathrm{q}\zeta\beta}(\upsilon,\omega;\kappa_{\zeta\beta}).$$

The PSDs $\tilde{\Phi}_{s\zeta\beta}(\upsilon,\omega;\kappa_{\zeta\beta})$ are defined by the set of iterative equations

$$\tilde{\Phi}_{s\zeta\beta}(\upsilon,\omega;\kappa_{\zeta\beta}) = \tilde{\Phi}_{s\zeta}(\upsilon,\omega)|\tilde{\tau}_{\mathrm{a}\zeta\beta}(\upsilon,\omega)|^{2} * \underline{\hat{|||}}_{\zeta}, \tag{4.24}$$

and $\tilde{\Phi}_{s0,1}(\upsilon,\omega) \equiv \tilde{\Phi}_{s}(\upsilon,\omega)$ is given by equation (2.2). The final digital image restoration is performed by the Wiener-matrix filter $\hat{\Psi}_{1\beta}(\upsilon,\omega;\kappa_{1\beta})$, which is the same as $\hat{\Psi}_{\beta}(\upsilon,\omega;\kappa_{\beta})$ given by equation (4.9) simply by replacing the subscript $\beta$ by $1\beta$. If the analysis filters do not overlap, i.e. $\tilde{\tau}_{\mathrm{a}\zeta\beta}(\upsilon,\omega)\tilde{\tau}^{*}_{\mathrm{a}\zeta\alpha}(\upsilon,\omega) = |\tilde{\tau}_{\mathrm{a}\zeta\beta}|^{2}\delta(\beta,\alpha)$, then this filter reduces to

$$\hat{\Psi}_{1\beta}(\upsilon,\omega;\kappa_{1\beta}) = \frac{\hat{\Phi}_{L}(\upsilon,\omega)\hat{\tau}^{*}(\upsilon,\omega)\tilde{\tau}^{*}_{\mathrm{a}1\beta}(\upsilon,\omega)}{\tilde{\Phi}_{s}(\upsilon,\omega)|\tilde{\tau}_{\mathrm{a}1\beta}(\upsilon,\omega)|^{2} * \underline{\hat{|||}}_{1} + \tilde{\Phi}_{\mathrm{q}1\beta}(\upsilon,\omega;\kappa_{1\beta})}, \tag{4.25}$$

where $\tilde{\Phi}_{\mathrm{q}1\beta}(\upsilon,\omega;\kappa_{1\beta})$ is the PSD of the accumulated requantization noises from the $\boldsymbol{Z}$ levels of the decomposition.

(iv) *Figures of merit*

The information rate $\mathcal{H}_{\zeta\beta}$ for each level $\zeta$ and band $\beta$ is

$$\mathcal{H}_{\zeta\beta} = \frac{1}{2}\iint\limits_{\hat{B}_{\beta}} \log\left[1 + \frac{\tilde{\Phi}_{s\zeta-1,1}(\upsilon,\omega)|\tilde{\tau}_{\mathrm{a}\zeta\beta}(\upsilon,\omega)|^{2}}{\tilde{\Phi}_{s\zeta-1,1}(\upsilon,\omega)|\tilde{\tau}_{\mathrm{a}\zeta\beta}(\upsilon,\omega)|^{2} * \underline{\hat{|||}}_{\zeta s} + \tilde{\Phi}_{\mathrm{q}\zeta\beta}(\upsilon,\omega;\kappa_{\zeta\beta})}\right]\mathrm{d}\upsilon\mathrm{d}\omega, \tag{4.26}$$

where

$$\underline{\hat{|||}}_{\zeta s} = \sum_{\substack{m=0\\ m=n\neq 0}}^{X-1}\sum_{n=0}^{Y-1}\delta\left(\upsilon - \frac{m}{X^{\zeta}}, \omega - \frac{n}{Y^{\zeta}}\right)$$

are the sidebands of the down-sampler, and the PSDs $\tilde{\Phi}_{s\zeta1}(\upsilon,\omega)$ are given by equation (4.24). The associated theoretical minimum data rate $\mathcal{E}_{\zeta\beta}$ is

$$\mathcal{E}_{\zeta\beta} = \frac{1}{2}\iint\limits_{\hat{B}_{\beta}} \log\left[1 + \frac{\tilde{\Phi}_{s\zeta-1,1}(\upsilon,\omega)|\tilde{\tau}_{\mathrm{a}\zeta\beta}(\upsilon,\omega)|^{2} * \underline{\hat{|||}}_{\zeta}}{\tilde{\Phi}_{\mathrm{q}\zeta\beta}(\upsilon,\omega;\kappa_{\zeta\beta})}\right]\mathrm{d}\upsilon\mathrm{d}\omega. \tag{4.27}$$

The total theoretical minimum data rate $\mathcal{E}_{\boldsymbol{Z}\mathcal{B}}$ of the $\boldsymbol{Z}(\mathcal{B}-1)+1$ decomposed signals is

$$\mathcal{E}_{\boldsymbol{Z}\mathcal{B}} = \sum_{\zeta=1}^{\boldsymbol{Z}-1}\sum_{\beta=2}^{\mathcal{B}}\mathcal{E}_{\zeta\beta} + \sum_{\beta=1}^{\mathcal{B}}\mathcal{E}_{\boldsymbol{Z}\beta}. \tag{4.28}$$

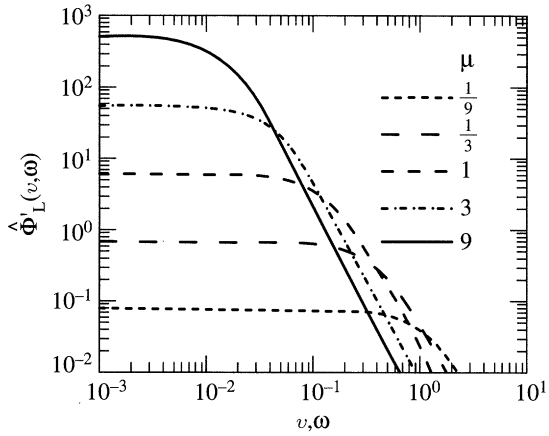



Figure 8. Normalized PSD $\hat{\Phi}'_L(v,\omega)$ of the random radiance field with mean spatial detail $\mu$ relative to the sampling lattice.

If the analysis filters $\tilde{\tau}_{\mathrm{a}\zeta\beta}(v,\omega)$ do not overlap, then the total information rate $\mathcal{H}_{\boldsymbol{Z}\mathcal{B}}$ has the same summation as $\mathcal{E}_{\boldsymbol{Z}\mathcal{B}}$, i.e.

$$\mathcal{H}_{\boldsymbol{Z}\mathcal{B}} = \sum_{\zeta=1}^{\boldsymbol{Z}-1}\sum_{\beta=2}^{\mathcal{B}}\mathcal{H}_{\zeta\beta} + \sum_{\beta=1}^{\mathcal{B}}\mathcal{H}_{\boldsymbol{Z}\beta}. \tag{4.29}$$

## 5. Quantitative and qualitative assessments

This final section characterizes the upper bounds of performance of visual communication in terms of the figures of merit developed in the above two sections and correlates these quantitative assessments with experimental results produced by computer simulations. The actual data rate of the visual communication can, of course, be directly correlated with its lower bound, the theoretical minimum data rate. However, the visual quality of the restored image is too elusive to be captured entirely by a single quantitative measure or even by any existing model of human vision. In practice, the preferred visual representation varies with target and application. Therefore, our emphasis is on correlating the information rate with the restorability (Frieden 1970) (ability to restore) of images with high visual quality. This emphasis leads us to a close examination of the perceptible effects of the perturbations that occur in the visual communication channel.

### (*a*) *Simulation*

(i) *Targets*

The PSD $\hat{\Phi}_L(v,\omega)$ of the radiance field $L(x,y)$ of natural scenes commonly approaches the shape given by (Itakura *et al.* 1974; Kass & Hughes 1983)

$$\hat{\Phi}_L(v,\omega) = \frac{2\pi\mu^2\sigma_L^2}{[1+(2\pi\mu\rho)^2]^{3/2}} \tag{5.1}$$

and shown in figure 8. Figure 9*a* illustrates a target of random polygons that has this PSD. The target is generated by a Poisson process with mean distance $\mu$ between the radiance-field transitions and an independent Gaussian intensity distribution with variance $\sigma_L^2$ at the transitions (Modestino & Fries 1980). The mean distance $\mu$

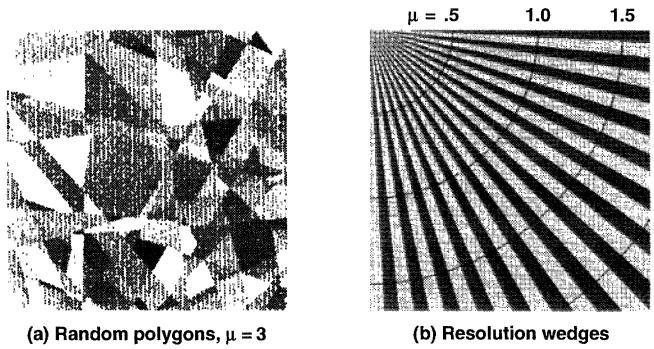



Figure 9. Targets.

between the edges of the polygons is measured relative to the sampling interval of the image-gathering device (i.e. the mean spatial detail is equal to the sampling interval when $\mu = 1$). The target of resolution wedges in figure 9*b* is included to facilitate the assessment of resolution and clarity (i.e. the absence of visually annoying defects).

These targets and the subsequent images are produced with $256 \times 256$ display elements, which is (at least nearly) sufficient to give them the appearance of being continuous relative to the sampling lattice of the image-gathering process. The sampling lattice contains $64 \times 64$ picture elements (pixels) for the random polygons and $32 \times 32$ pixels for the resolution wedges. Therefore, to produce images of equal size to the target, the polygons and wedges are reconstructed with an interpolation lattice, respectively, of $4 \times 4$ and $8 \times 8$ display elements. The implementation of this interpolation varies with the method of image reproduction: For the traditional image reconstruction (given in § 2 *b*), the interpolation is simply used to magnify the image (using cubic splines (Fales *et al.* 1996)); and for the restoration with the Wiener filter (given in § 2 *c*), it is used to suppress the blurring and raster effects of the reconstruction process. The difference in the interpolation density used for the restoration of the two targets appears as a magnification (by a factor of 2) of the wedges relative to the polygons. This magnification can have a significant effect on the visual appearance of some image degradations. Halftone prints of this size which contain more than $64 \times 64$ pixels (as is often found in the digital processing literature) hide many defects that would immediately become visible in a wide range of applications, from medical diagnosis and military or planetary reconnaissance to (digital) high-definition television (HDTV).

(ii) *Image gathering and display*

The SFRs of image gathering and display devices, $\hat{\tau}(\upsilon, \omega)$ and $\hat{\tau}_{\mathrm{r}}(\upsilon, \omega)$, respectively, commonly approach the Gaussian shape given by (Schade 1951, 1952, 1953, 1955; Fales *et al.* 1984; Schreiber 1993; Huck *et al.* 1995)

$$\hat{\tau}(\upsilon, \omega) = \exp[-(\rho/\rho_c)^2], \tag{5.2}$$

and shown in figure 10, where $\rho_c$ is the optical-design index for which $\hat{\tau}(\upsilon, \omega) = 1/\mathrm{e} \approx 0.37$. The reconstruction response $\tau_{\mathrm{r}}(x, y)$ is made just large enough to minimize the visibility of the display raster. This minimization occurs when the optical-design index of the SFR $\hat{\tau}_{\mathrm{r}}(\upsilon, \omega)$ is $\rho_c = 0.45$ (Schreiber 1993). Schade (1951, 1952, 1953, 1955) and Schreiber (1993) found that the corresponding image-gathering response $\tau(x, y)$ should be approximately half as large (i.e. $\rho_c \approx 0.8$) to provide the generally

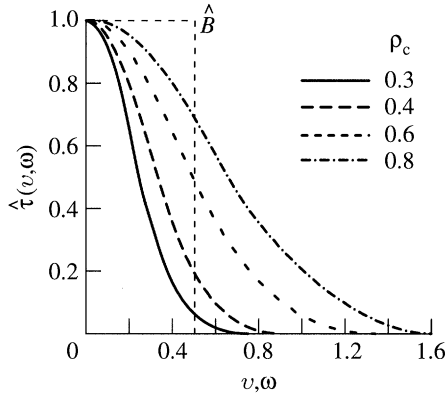



Figure 10. SFRs $\hat{\tau}(v,\omega)$ of the image-gathering device relative to the sampling passband $\hat{B}$ for unit sampling intervals.

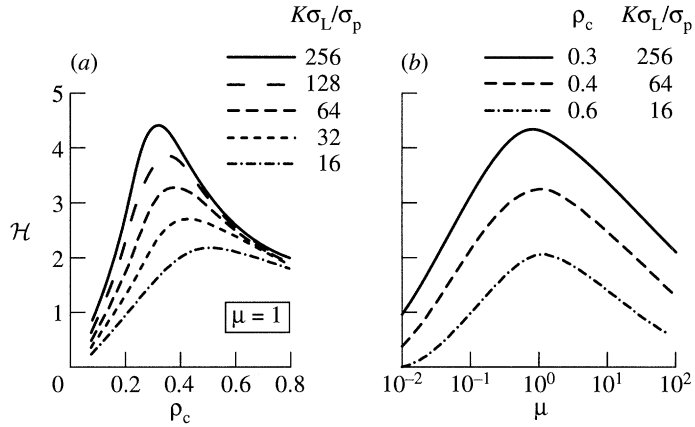



Figure 11. Information rate $\mathcal{H}$ versus (*a*) the optical design index $\rho_c$ and (*b*) the mean spatial $\mu$, both for several SNRs $K\sigma_L/\sigma_p$.

most favourable trade-off between blurring and aliasing for image reconstruction without the aid of digital processing.

The simulation does not account for the compression of the normally wide dynamic range of natural terrestrial radiance fields (Barlow 1981) into the relatively narrow dynamic range of the image on film. Instead, it deals only with the dynamic range of the reflectance of natural scenes (Bowker *et al.* 1985) and of film and halftone prints (Ulichney 1987), which range from *ca.* 1% to 80%. This small signal analysis is justifiable because the information-theoretic assessment deals largely with fine detail which normally has relatively low contrast. In practice, of course, it also becomes necessary to deal with dynamic-range compression and nonlinear tone-scale transformation (Fales *et al.* 1996).

### (*b*) *Image gathering and transmission*

#### (i) *Maximum-realizable information rate* $\mathcal{H}_m$

Figure 11 characterizes the information rate $\mathcal{H}$ as a function of the electro-optical design of the image-gathering device and of the mean spatial detail of the radiance field. The number of quantization levels is assumed to be sufficiently high so that

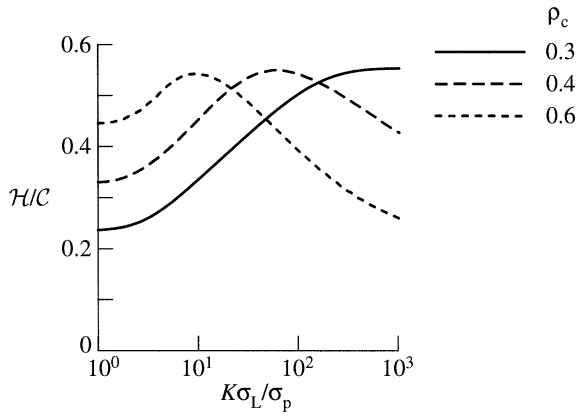



Figure 12. Ratio of information rate $\mathcal{H}$ to channel capacity $\mathcal{C}$ versus the SNR $K\sigma_L/\sigma_{\mathrm{p}}$. $\mathcal{H}$ is given for three optical-design indexes $\rho_c$ and a mean spatial detail $\mu = 1$.

the quantization noise is negligible. These results demonstrate that $\mathcal{H}$ reaches its maximum-realizable value $\mathcal{H}_m$ when both of the following two conditions are met:

1. The SFR $\hat{\tau}(\upsilon, \omega)$ is appropriately matched to the sampling passband $\hat{B}$ for the available SNR $K\sigma_L/\sigma_{\mathrm{p}}$. This condition is intuitively appealing for image restoration. In one extreme, when the SNR is low, one would prefer to avoid substantial blurring because the noise constrains the enhancement of fine spatial detail. In the other extreme, when the SNR is high, one would prefer to avoid substantial aliasing because then the noise no longer constrains this enhancement.

2. The sampling passband $\hat{B}$ is appropriately matched to the radiance-field PSD $\hat{\Phi}_L(\upsilon, \omega)$. For the PSDs $\hat{\Phi}_L(\upsilon, \omega)$ typical of natural scenes, this match occurs when the sampling interval is close to the mean spatial detail of the radiance field (i.e. $\mu \approx 1$). This condition is also intuitively appealing for image restoration. Ordinarily, one cannot expect to restore spatial detail that is finer than the sampling interval, whereas one can expect to restore detail that is much coarser from fewer samples.

(ii) *Information rate $\mathcal{H}$ and channel capacity $\mathcal{C}$*

Figure 12 compares $\mathcal{H}$ with $\mathcal{C}$ as a function of the SNR $K\sigma_L/\sigma_{\mathrm{p}}$ for three optical-design indexes $\rho_c$. The curves show that the constraints of the image-gathering process limit $\mathcal{H}_m$ to approximately one-half of $\mathcal{C}$ (i.e. $\mathcal{H}_m \approx 0.55\mathcal{C}$). The upper bound of the ratio $\mathcal{H}/\mathcal{C}$ is independent of the SNR; it requires only that the two conditions that lead to $\mathcal{H}_m$ (as specified above) are met. Therefore, we can treat these two conditions analogous to the role that white and bandwidth-limited signals play in classical communication theory (namely, as the conditions that establish the upper bound on the information rate through the communication channel). Although $\mathcal{H}_m$ cannot be specified without qualifications (i.e. radiance-field spectrum and image-gathering response), the value given here seems to be approximately valid for a wide range of PSDs $\hat{\Phi}_L(\upsilon, \omega)$ of natural scenes and of SFRs $\hat{\tau}(\upsilon, \omega)$ of realizable image-gathering devices (Huck *et al.* 1988).

(iii) *Information-entropy $\mathcal{H}(\mathcal{E})$ plot*

Figure 13 presents an information-entropy $\mathcal{H}(\mathcal{E})$ plot that characterizes the information rate $\mathcal{H}$ versus the associated theoretical minimum data rate, or entropy, $\mathcal{E}$. This plot serves as an alternative to the familiar rate-distortion function (Shan-

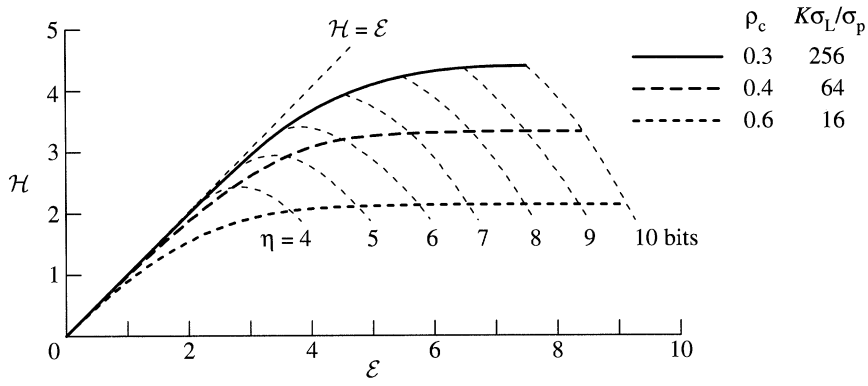



Figure 13. The information-entropy $\mathcal{H}(\mathcal{E})$ plot that characterizes the information rate $\mathcal{H}$ versus the associated theoretical minimum data rate $\mathcal{E}$ for $\eta$-bit quantization. The three curves represent informationally optimized designs specified by the optical-design index $\rho_c$ and SNR $K\sigma_L/\sigma_p$. The mean spatial detail $\mu = 1$.

non 1959), which is based on the premise that the encoder has unperturbed access to the original source and, therefore, directly controls the trade-off between distortion and data rate. The basic difference between these two methods of assessment is as follows: The *rate-distortion function* accounts for the perturbations in coding and transmission in terms of a criterion, usually fidelity, that directly measures the resemblance between the original message and its reproduction. The *information-entropy plot*, instead, accounts for the perturbations before coding as well as in coding and transmission in terms of a criterion, here information, that measures the ability to reproduce an approximate representation of the original message with an appropriate filter. In our assessment of visual communication, the maximum-realizable fidelity represents a particular image (namely, the one restored with the Wiener filter) and the information rate represents the ability to restore images, which Frieden (1970) described as restorability. We use the WIGE filter (given in § 2 *d*) because it allows us to interactively trade among fidelity, sharpness, and clarity (i.e. absence of visually annoying defects) for the best visual quality. Thus, our reason for choosing information rate over fidelity as the pivotal measure of performance is not based on a fundamental theoretical consideration, but, instead, it is based on the entirely pragmatic realization that the perceived quality of an image is too elusive to be captured by a single criterion, such as fidelity, or even by any existing model of human vision.

Obviously, the merit of the information-theoretic assessment of visual communication depends on the correlation that exists between the information rate and the visual quality with which images can be restored. The following § 5 *c* demonstrates that this correlation is, in fact, very high. It shows that any loss in information rate, regardless of where it occurs in the visual communication channel, reduces the best possible image quality. (However, it should be noted that our assessment accounts only for perturbations in image gathering, coding, and display, but not for those in data transmission.)

The plot $\mathcal{H}(\mathcal{E})$ confirms quantitatively what the equations for $\mathcal{H}$ and $\mathcal{E}$ suggest directly (see § 3 *c*), namely, that the electro-optical design that increases $\mathcal{H}$ also decreases the associated $\mathcal{E}$. Together, these two effects substantially increase $\mathcal{H}/\mathcal{E}$ as long as the quantization noise remains smaller than the aliasing and photodetector noises. If the quantization noise grows larger, as the result of coarse quantization to




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Table 1. *Characteristics of traditional and informationally optimized image gathering and transmission* ($\mu = 1$)

| image gathering for | design | $\rho_c$ | $K\sigma_L/\sigma_p$ | $\eta$, bits | $\mathcal{H}$, bits | $\mathcal{E}$ bits | $\mathcal{H}/\mathcal{E}$ |
|---|---|---|---|---|---|---|---|
| reconstruction | 1 | 0.8 | 16 | 8 | 1.8 | 7.5 | 0.24 |
| high information rate | 2a | 0.4 | 64 | 7 | 3.3 | 5.6 | 0.59 |
| | 3a | 0.3 | 256 | 8 | 4.2 | 5.6 | 0.75 |
| high information efficiency | 2b | 0.4 | 64 | 5 | 2.9 | 3.6 | 0.81 |
| | 3b | 0.3 | 256 | 6 | 3.4 | 3.7 | 0.92 |

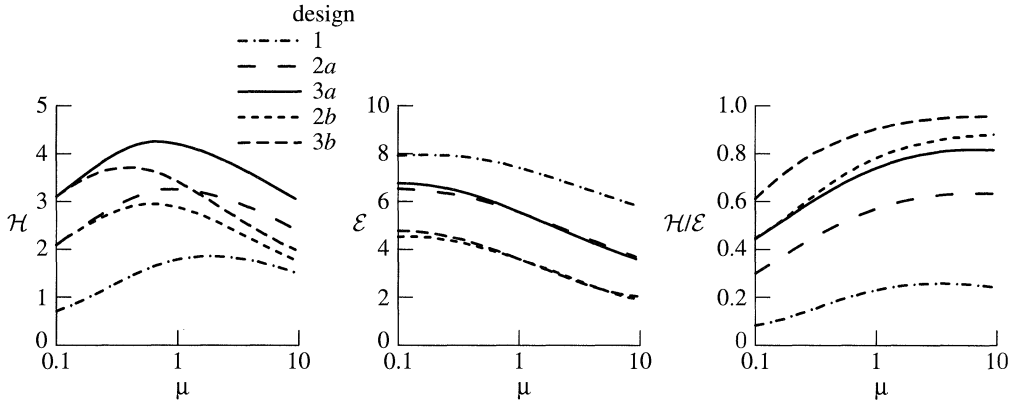



Figure 14. Information and data transmission versus mean spatial detail $\mu$. The designs are specified in table 1.

gain high data compression, then the selection of the number of quantization levels leads to a trade-off between $\mathcal{H}$ and $\mathcal{H}/\mathcal{E}$.

Table 1 summarizes the characteristics of visual communication for three electro-optical designs of the image-gathering device. Design 1 is the traditional design, and designs 2 and 3 are informationally optimized designs obtained from the $\mathcal{H}(\mathcal{E})$ plot of figure 13. The $\eta$-bit quantizations for the latter designs favour either a high $\mathcal{H}$ or a high $\mathcal{H}/\mathcal{E}$.

(iv) *Performance versus mean spatial detail*

Figure 14 characterizes $\mathcal{H}$, $\mathcal{E}$ and $\mathcal{H}/\mathcal{E}$ as functions of the mean spatial detail $\mu$. As $\mu$ grows coarser than the sampling interval ($\mu > 1$), both $\mathcal{H}$ and $\mathcal{E}$ decrease, but they do so with slightly different slopes so that $\mathcal{H}/\mathcal{E}$ increases modestly. It may be noted that the loss in $\mathcal{H}$ is much more severe for the coarsely quantized signals (designs 2b and 3b) than for the more finely quantized signals (designs 2a and 3a). Conversely, as $\mu$ grows finer than the sampling interval ($\mu < 1$), $\mathcal{H}$ decreases substantially while $\mathcal{E}$ increases modestly. The resultant decrease in $\mathcal{H}/\mathcal{E}$ obviously gives rise to a strong motivation for data compression in which only those features in the scene are encoded for transmission that are large enough to be resolved in the restored image. This selective encoding process (or irrelevancy reduction) usually entails some loss of information.

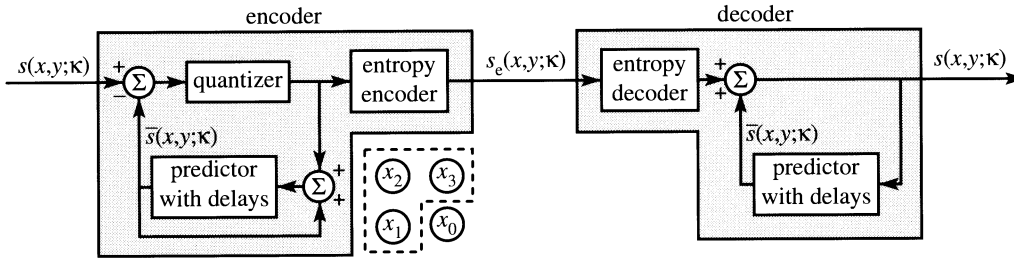



Figure 15. Differential pulse code modulation (DPCM) with entropy coding.

(v) *Redundancy reduction*

Figure 15 depicts a common redundancy reduction scheme that consists of differential pulse code modulation (DPCM) with entropy coding (Jain 1989). The DPCM encoder first predicts the acquired sample $x_0$ based on some previous samples $x_1$, $x_2$, ..., $x_n$, and it then subtracts this prediction from the actual value. The decoder reverses this process by adding the prediction to the received signal. The entropy (Huffman) coding, which follows the DPCM, deals with the efficient assignment of binary code words. The efficiency is gained by letting the length of the binary code word for a quantization level be inversely related to its frequency of occurrence.

The $n$th-order conditional entropy $E_n$ of the encoded data is

$$E_n = -\sum_{x_0=0}^{\kappa-1}\sum_{x_1=0}^{\kappa-1}\cdots\sum_{x_n=0}^{\kappa-1} p(x_0, x_1 \ldots, x_n)\log p(x_0|x_1, x_2, \ldots, x_n). \tag{5.3}$$

The lower bound of $E_n$ (for a Gaussian signal) is $\mathcal{E}$ given by equations (3.5), and the upper bound of $E_n$ is

$$E_o = -\sum_{x_0=0}^{\kappa-1} p(x_0)\log p(x_0),$$

where $p(x_0)$ is the probability of a sample having the quantized amplitude $x_0 \in (0, 1, 2, \ldots, \kappa-1)$. The values $p(x_0)$ are usually obtained from the probability distribution (histogram) of the quantized signal. If the probability distribution is uniform so that all quantization levels are equally likely, then $p(x_0) = 1/\kappa$ and $E_o = \mathcal{E}_m = \eta$ is the maximum possible entropy. The conditional entropy $E_1$ uses the past nearest sample (i.e. $x_1$), and $E_3$ uses the past three nearest samples (i.e. $x_1$, $x_2$, and $x_3$). The prediction minimizes the linear mean-square error estimation $E\{[s(x, y; \kappa) - \bar{s}(x, y; \kappa)]^2\}$.

Figure 16*a* characterizes the entropy $E_3$ and its theoretical lower bound $\mathcal{E}$ as functions of the mean spatial detail $\mu$, and figure 16*b* similarly characterizes the 'decorrelation efficiency' $\mathcal{E}/E_3$. The plot of $\mathcal{E}/E_3$ demonstrates that the DPCM with Huffman coding can attain *ca.* 75–100% decorrelation. The actual decorrelation within this range depends on the statistics of the signal, which, in turn, depends on both the statistics of the radiance field and the design of the image-gathering device. In general, the decorrelation efficiency tends to be high (*ca.* 90%) when neighbouring pixels are highly uncorrelated because (*a*) the mean spatial detail of the radiance field is finer than the sampling interval ($\mu < 1$) and (*b*) the perturbations in the image-gathering process due to aliasing and photodetector noise are high. Conversely, the decorrelation efficiency tends to be low (*ca.* 65%) when neighbouring pixels are correlated because the mean spatial detail is coarse and the perturbations are small.

The DPCM with entropy coding attains its maximum-realizable compression only

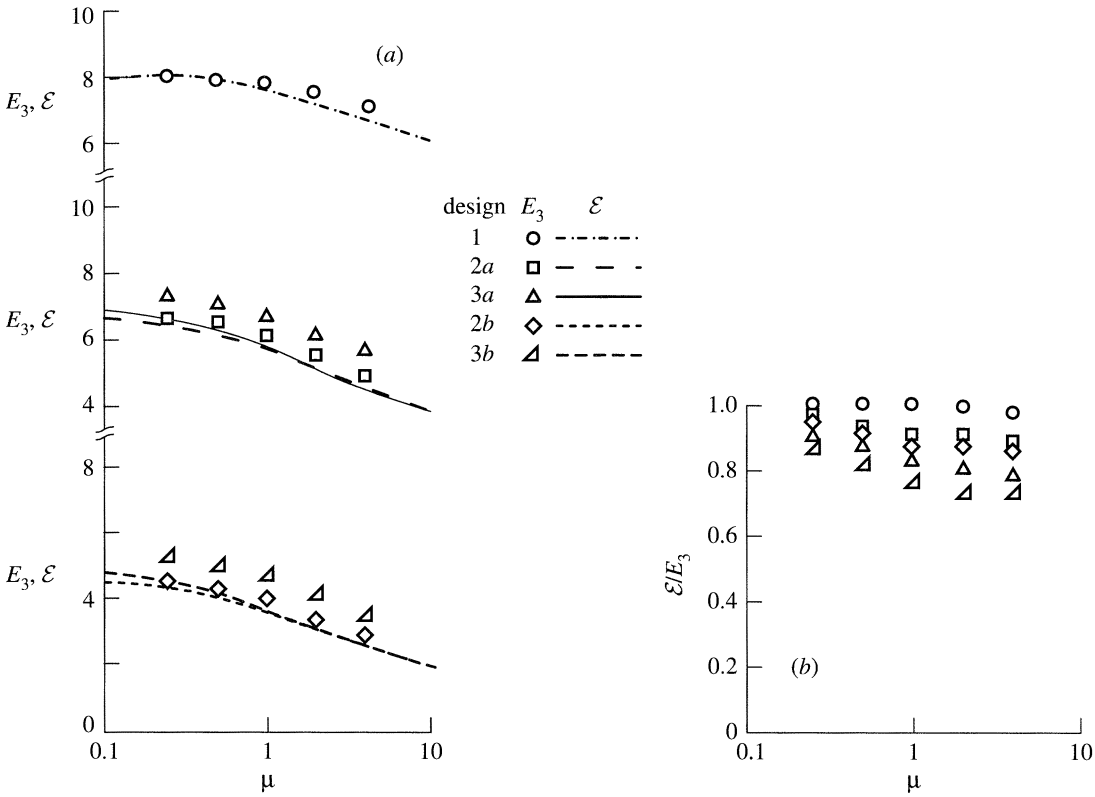



Figure 16. Data rate ($a$) and decorrelation efficiency ($b$) versus the mean spatial detail $\mu$. The data rate is characterized by the entropy $E_3$ for DPCM with Huffman coding and the theoretical lower bound $\mathcal{E}$. The designs are specified in table 1.

when the statistics of the acquired signal actually correspond to the statistics for which the mean-square error estimation is minimized. Because the signal statistics tend to vary from one image to the next, the Rice Machine (Yeh *et al.* 1991), for example, uses the following scheme to closely approach the maximum-realizable compression for a wide range of signal statistics. First, it simultaneously processes each frame, or segment, of the acquired signal with a number of identical encoders, of which each is tuned to a different signal statistic. Subsequently, it selects the most compressed (i.e. shortest) signal for transmission.

### (*c*) *Image gathering and reproduction*

#### (i) *Transition from reconstruction to restoration*

Figure 17 presents images that illustrate the transition from traditional telephotography and television in which the images are produced without digital processing (i.e. the reconstruction in § 2 *b*) to modern visual communication in which the images are produced with digital processing (e.g. the Wiener restoration in § 2 *c*). The transition is illustrated in the following three steps:

1. Figure 17*a* characterizes the traditional image gathering (design 1) and reconstruction. The images are perceptively degraded by blurring, aliasing, and raster effects. The aliasing and raster effects appear as stair-step distortions of the diagonal sharp edges for spatial detail that is larger than the sampling interval ($\mu > 1$), and

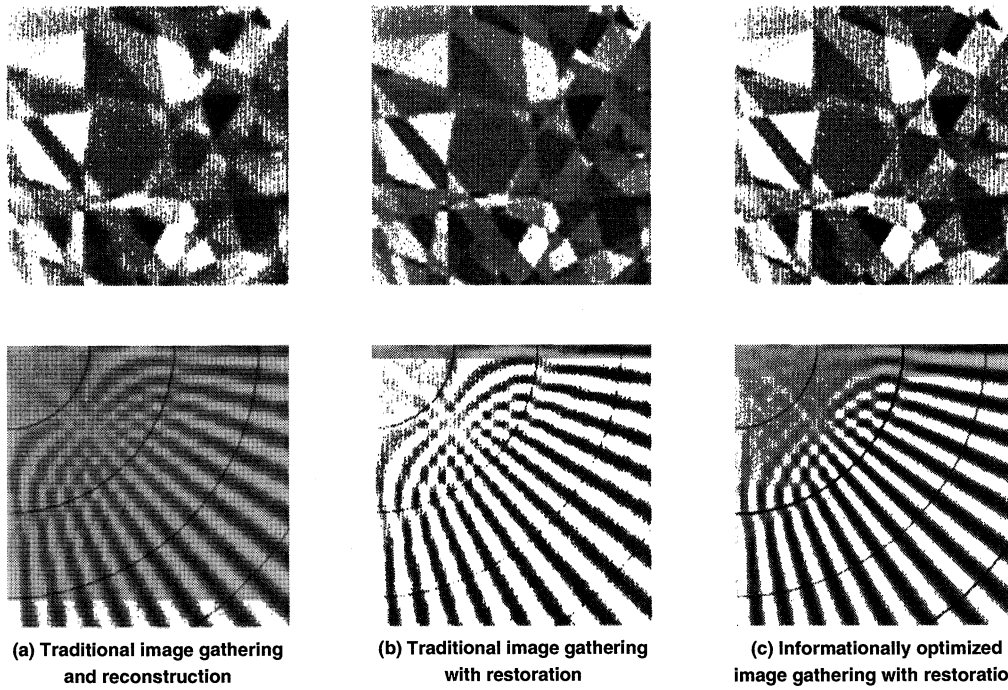



Figure 17. Reconstructions and Wiener restorations for two designs of the image-gathering device, the traditional design (design 1) and an informationally optimized design (design 2a) as specified in table 1.

the aliasing appears also as moiré patterns for detail of the resolution wedges that is finer ($\mu < 1$).

2. Figure 17*b* again characterizes the traditional image gathering (design 1) but with restoration. The images now are sharper and resolve fine detail with higher contrast. However, the moiré patterns are also stronger.

3. Figure 17*c* characterizes informationally optimized image gathering (design 2a) with restoration. These images resolve detail as fine as the sampling lattice without severe aliasing artifacts. However, this improvement is slightly marred by the emergence of ringing near sharp edges (Gibbs phenomenon) and coloured photodetector noise.

The last two figures clearly demonstrate that the visual quality with which images can be restored with the aid of digital processing depends critically on the design of the image-gathering device. Figures 18 and 19 further characterize salient differences between the traditional image gathering and reconstruction and the informationally optimized image gathering with restoration.

The degradations due to aliasing and raster effects are immediately apparent in the images of the resolution wedges. However, these degradations also exist, but more subtly, in the images of the random polygons. To illustrate this, figure 20 presents the reconstructed image shown in figure 17*a* together with its blurred and aliased components (see equation (2.5 *c*)). These images reveal that (*a*) the blurred component is substantially sharper than the observed image and (*b*) the aliased components are closely correlated with the sharp edges of the polygons. At a first glance, the aliased components even mimic the appearance of an edge enhanced

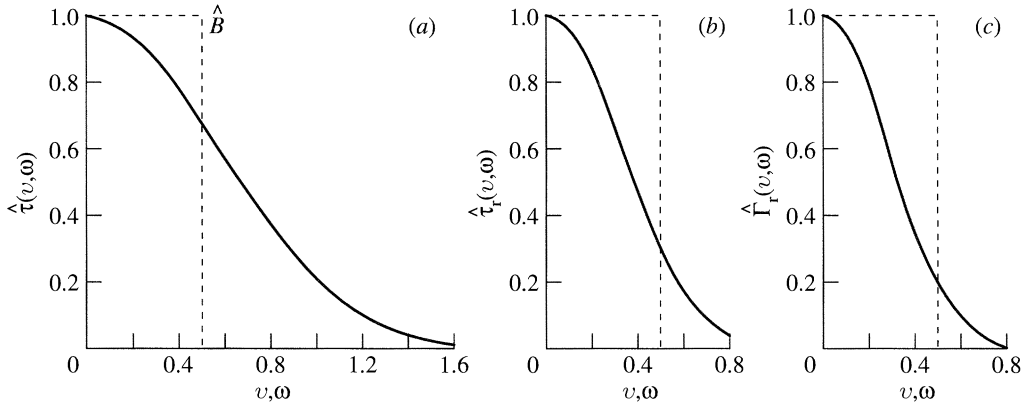



Figure 18. SFRs for traditional image gathering and reconstruction (design 1). (*a*) Image gathering, (*b*) image reconstruction, and (*c*) throughput.

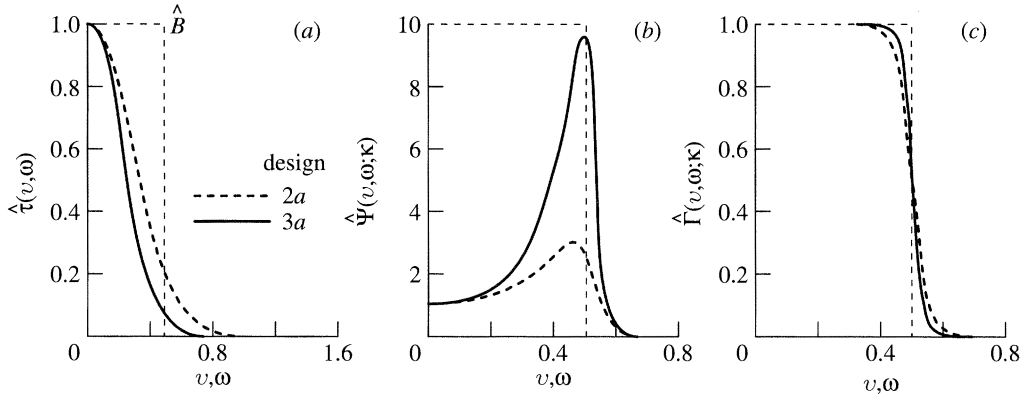



Figure 19. SFRs for informationally optimized image gathering with restoration (designs 2a and 3a). (*a*) Image gathering, (*b*) Wiener filter, and (*c*) throughput.

image. However, the visual effect of the aliased components, when added to the blurred component, is that of increased blurring. The reason is simply that the image representation of the polygons obscures what the magnified representation of the wedges clearly reveals, namely, that the apparent increase in blurring is actually due to the (unresolved) stair-step distortions that aliasing and raster effects cause.

(ii) *Information rate $\mathcal{H}$ and fidelities $\mathcal{F}$ and $\mathcal{F}_e$*

Figure 21 compares $\mathcal{H}$ with $\mathcal{F}$ and $\mathcal{F}_e$ as a function of the optical-design index $\rho_c$ for three SNRs $K\sigma_L/\sigma_p$. $\mathcal{F}$ and $\mathcal{F}_e$ represent the maximum-realizable fidelities, respectively, with which the radiance field $L(x,y)$ and the radiance-field transitions $L(x,y) * \nabla^2 G(x,y)$ can be restored. Whereas $\mathcal{F}$ emphasizes structure, shape, and shading, $\mathcal{F}_e$ emphasizes sharp edges and fine detail. It may appear, at first glance, that the design that maximizes $\mathcal{F}$ differs from the design that maximizes $\mathcal{F}_e$. Yet, a close comparison of the information and fidelity curves reveals that the design that maximizes $\mathcal{H}$ also maximizes both fidelities $\mathcal{F}$ and $\mathcal{F}_e$ when the SNR is sufficiently high, above *ca.* 64, and provides an intuitively attractive compromise between $\mathcal{F}$ and $\mathcal{F}_e$ when the SNR is lower.

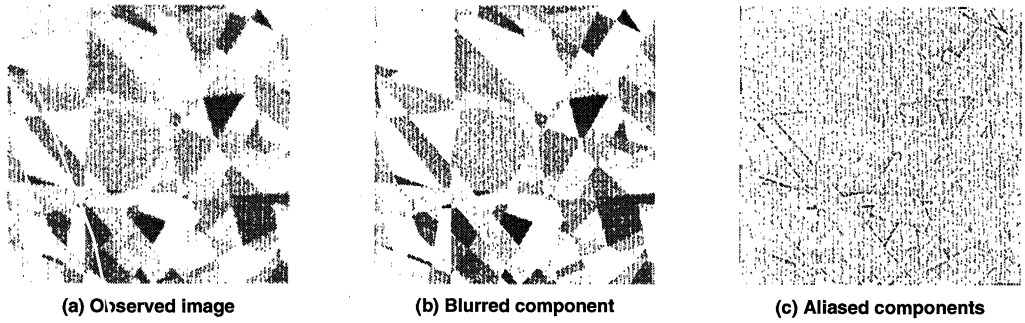



Figure 20. Observed image and its blurred and aliased components for the traditional image gathering and reconstruction.

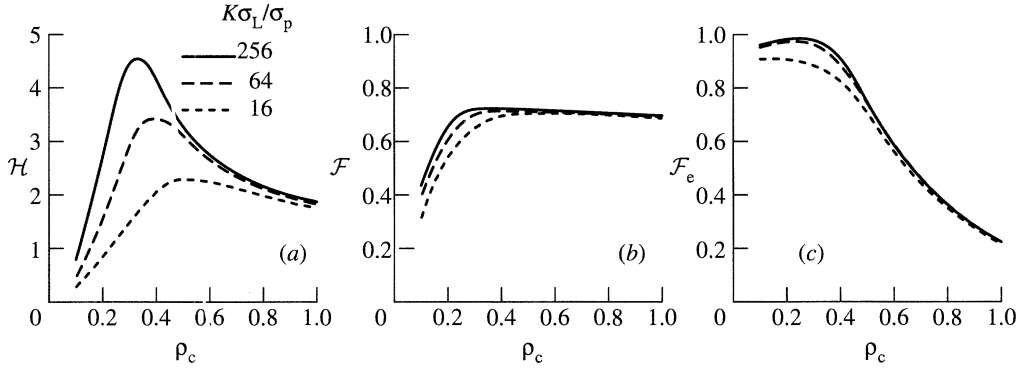



Figure 21. Information rate and maximum-realizable fidelities versus the optical-design index $\rho_c$ for three SNRs $K\sigma_L/\sigma_p$. (*a*) Information rate, (*b*) fidelity to $L(x, y)$, and (*c*) fidelity to $L(x, y) * \nabla^2 G(x, y)$. The mean spatial detail $\mu = 1$.

(iii) *Information rate $\mathcal{H}$ and robustness of image restoration*

The radiance-field PSD $\hat{\Phi}_L(v, \omega)$ is seldom known *a priori*. Even if the actual PSD closely approaches the shape shown in figure 8, the mean $\mu$ and variance $\sigma_L^2$ often remain uncertain because they depend not only on the scene but also on the viewing distance and geometry. Therefore, to optimally restore images, it ordinarily is necessary to rely on estimates of the statistical properties of the captured radiance field. The tolerance of the restoration to errors in these estimates is commonly referred to as robustness.

Figure 22 compares $\mathcal{H}$ with $\mathcal{F}$ of matched and mismatched Wiener restorations as functions of the optical-design index $\rho_c$ for three SNRs $K\sigma_L/\sigma_p$. The mismatched restorations use erroneous estimates of $\mu$, either 0.1 or 10 instead of 1. The results show that the design that maximizes $\mathcal{H}$ provides the most robust restoration and that the robustness increases with increasing $\mathcal{H}$. In practice, one cannot go far wrong simply by assuming that the mean spatial detail is the finest detail that can be resolved (i.e. $\mu = 1$).

(iv) *Information rate $\mathcal{H}_o$, fidelity $\mathcal{F}_o$, and visual quality*

Figure 23 characterizes $\mathcal{H}_o$ and $\mathcal{F}_o$ of the observed images of the random polygons. It can be seen immediately that $\mathcal{H}_o$, unlike $\mathcal{F}_o$, is highly sensitive to both the design of the image-gathering device and the number of distinguishable grey levels

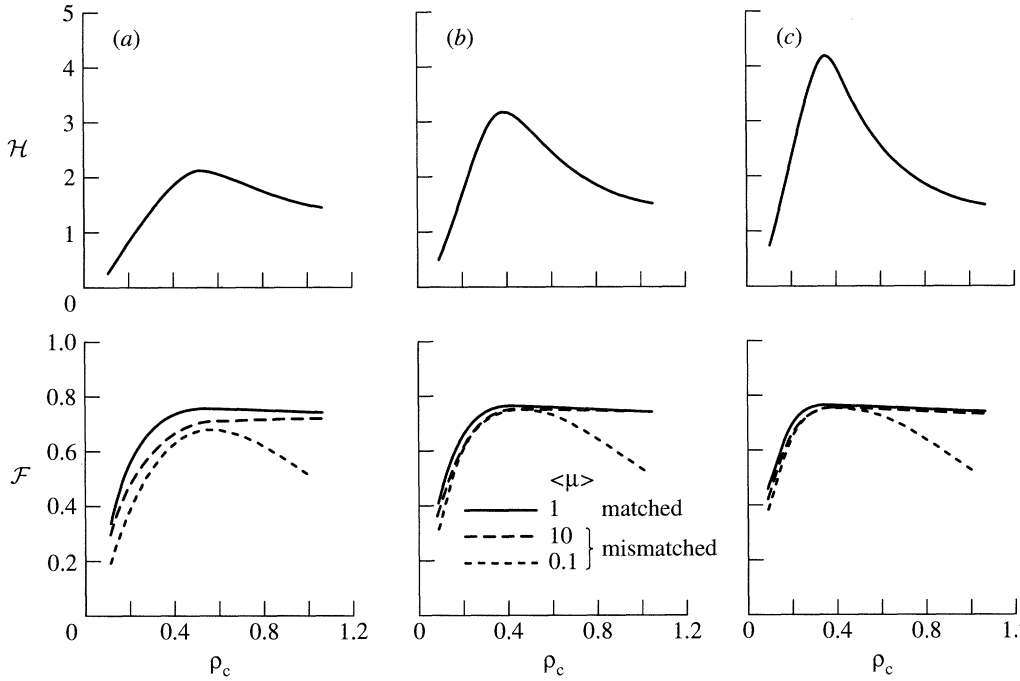



Figure 22. Information rate $\mathcal{H}$ and fidelity $\mathcal{F}$ versus the optical-design index $\rho_c$ for three SNRs $K\sigma_L/\sigma_p$. $\mathcal{F}$ is given for the matched and two mismatched Wiener restorations. (*a*) $K\sigma_L/\sigma_p = 16$, (*b*) $K\sigma_L/\sigma_p = 64$, and (*c*) $K\sigma_L/\sigma_p = 256$. The estimate $\langle\mu\rangle = 1$ is correct, whereas the estimates $\langle\mu\rangle = 0.1$ and 10 are incorrect.

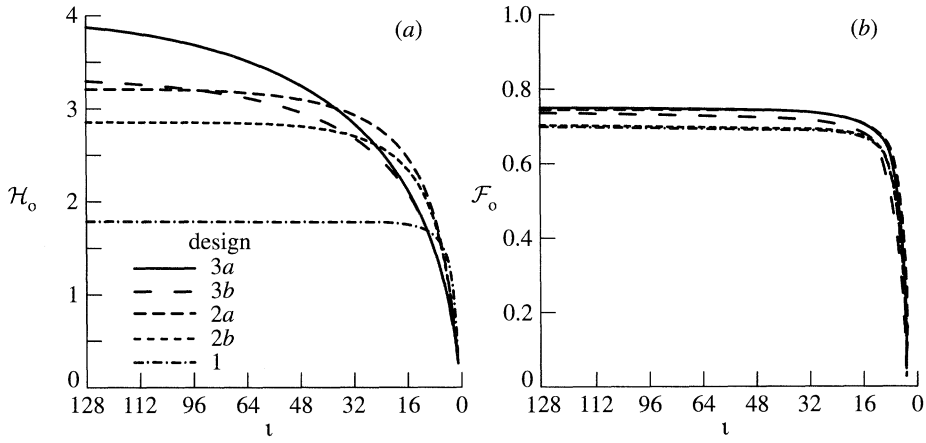



Figure 23. Quantitative characteristics of the observed image versus number of distinguishable grey levels $\iota$. (*a*) Information rate and (*b*) maximum-realizable fidelity. The designs are specified in table 1 and the mean spatial detail $\mu = 1$.

$\iota$. The halftone prints (Ulichney 1987) in this paper are limited to $\iota \approx 44$. This limit constrains the highest information rate of these images to $\mathcal{H}_o \approx 3.3$ bits. Thus, the high information rate $\mathcal{H}$ of the data transmission for design 3a does not carry over to the information rate $\mathcal{H}_o$ of the observed image. In fact, if the image-display medium were limited to even fewer distinguishable grey levels (e.g. $\iota \approx 10$), then design 2 would actually appear to be preferable to design 3.

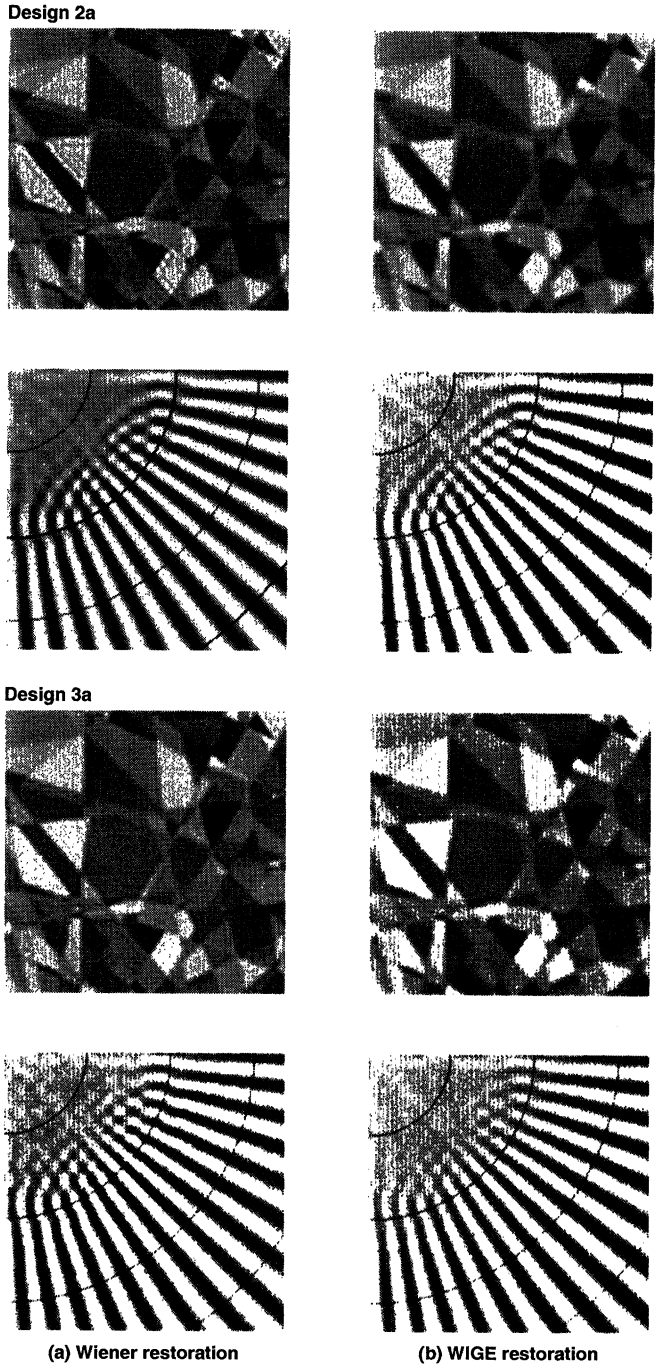



Figure 24. Wiener and WIGE restorations for informationally optimized image gathering. The designs are specified in table 1.

Figure 24 presents images restored with the Wiener filter and interactively with the WIGE filter for designs 2a and 3a. Both restorations produce images that are sharp and resolve detail as fine as the sampling lattice. However, the WIGE restorations




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

improve visual quality over the Wiener restorations by suppressing aliasing artifacts and ringing at the cost of a modest loss in sharpness.

In general, these results indicate that the visual quality with which images can be restored increases with increasing information rate $\mathcal{H}_o$, even after the maximum-realizable fidelity $\mathcal{F}_o$ has essentially been reached. Thus, if the goal is to restore images with the best possible visual quality, regardless of the complexity of the digital processing, then information rate is a more appropriate criterion than fidelity for assessing the requirements of the visual communication channel. This conclusion was anticipated by Linfoot (1956, 1958) nearly 40 years ago. It was he who first correlated Shannon's measure of the amount of information with the visual quality of optical (photographic) images in which the perturbations were limited to those caused by blurring (diffraction and aberrations) and noise. In summarizing his results, Linfoot observed in his customary elegant style:

> An optical system can properly be said to be of high quality only if the amount of information contained in its image approaches the maximum possible..., and it is an agreeable consequence... that those which are efficient according to this criterion also form images which are sharp and clear in the usual sense of the words. (Linfoot 1956)

> If the arithmetical recording of optical images were a standard practice today, instead of a prospect for the future opened up by the advent of the fast computing machines, we would go on to add that informationally optimized designs were always to be preferred. (Linfoot 1958)

In visual communication, however, in which image gathering and display are combined with digital transmission, the correlation between information rate and image quality is complicated by the fact that the communication process is perturbed not only by blurring and noise but also by sampling. Because of this additional perturbation, visual quality depends not only on the information rate that reaches the observer, but also, for a given information rate (especially if this rate approaches the maximum that is possible), by the compromise that must be made between enhancing sharpness and suppressing visually annoying defects due to aliasing and ringing. Therefore, to arrive at a suitable trade-off between performance and requirement in the design of a visual communication channel, it remains necessary to relate quantitative and qualitative assessments to each other. This need obviously carries over to data compression in which some information is lost. Although any loss of information will impair the visual quality with which images can be restored, the degree of this impairment can be ascertained only by qualitative assessments.

### (*d*) *Multiresolution decomposition*

Figure 25 presents images of the resolution wedges restored with the Wiener-matrix filter from data produced by an image-gathering system that combines design 3 with four analysis filters. The filters are a set of shifted Gaussian SFRs that are linear on a log scale, as given by

$$\tilde{\tau}_{a\beta}(v,\omega) = \exp\left[-\left(\frac{\rho-\rho_\beta}{\Delta\rho_\beta}\right)^2\right]$$

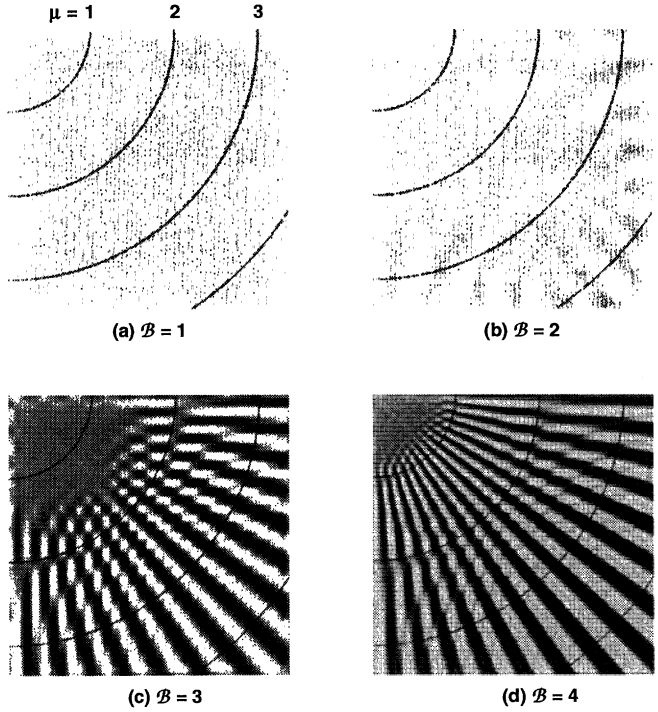



Figure 25. Wiener-matrix restorations for design 3 with single-level transformation. The four analysis filters and requantization map produce $\mathcal{H}_{\mathcal{B}} = 3.8$ bits at $\mathcal{E}_{\mathcal{B}} = 5.1$ bits. (The resolution wedges are coarser than elsewhere.)

for $\beta = 2, 3, 4$ together with the low-pass filter

$$\tilde{\tau}_{a_1}(\upsilon, \omega) = \begin{cases} \exp\left[-\left(\dfrac{\rho - \rho_1}{\Delta\rho_1}\right)^2\right], & \rho \geqslant \rho_1, \\ 1, & 0 \leqslant \rho \leqslant \rho_1, \end{cases}$$

where $\rho_\beta = 2^{\beta-1}\rho_1$ is the centre frequency and $\Delta\rho_\beta = 2^{\beta-1}\Delta\rho_1$ is the width of the $\beta$th filter. Figure 26 illustrates the corresponding image gathering and throughput SFRs for $\rho_1 = 0.06$ and $\Delta\rho_1 = 0.02$. This signal analysis may be interpreted either as a single-level decomposition ($\boldsymbol{Z} = 1$) with four bands ($\mathcal{B} = 4$), as marked in figures 25 and 26, or as a three-level decomposition ($\boldsymbol{Z} = 3$) in which (see figure 7) $(\zeta\beta) = (3,1)$, (3,2), (2,2), and (1,2) in order of the increasing resolution with which the images are restored. Either way, this example illustrates that the location of the filters is, for best performance, closely linked with the electro-optical design of the image-gathering device (i.e. the trade-off between blurring and aliasing as a function of the SNR).

The associated requantization map may be developed in several different ways. Using the trade-off between preserving $\mathcal{H}$ and reducing $\mathcal{E}$ as the general objective, we developed a map that produces $\mathcal{H}_{\mathcal{B}} = 3.8$ bits at $\mathcal{E}_{\mathcal{B}} = 5.1$ bits for random polygons with $\mu = 1$. Hence, compared with design 3a (table 1), there is a small loss in the information rate together with a small reduction in the theoretical minimum data rate, which yields again an information efficiency $\mathcal{H}/\mathcal{E} = 0.75$.

The resolution wedges in figure 25 are expanded in scale by a factor of two in

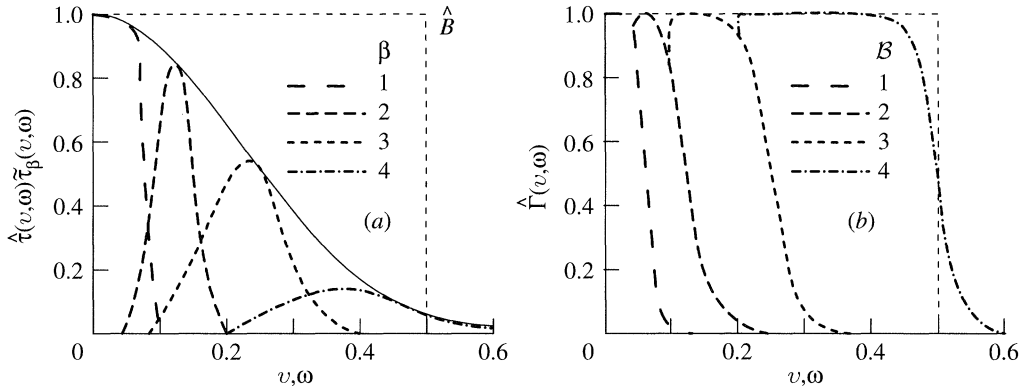



Figure 26. SFRs for multiresolution image gathering and restoration for design 3 with Gaussian analysis filters. (*a*) Image gathering and analysis and (*b*) throughput.

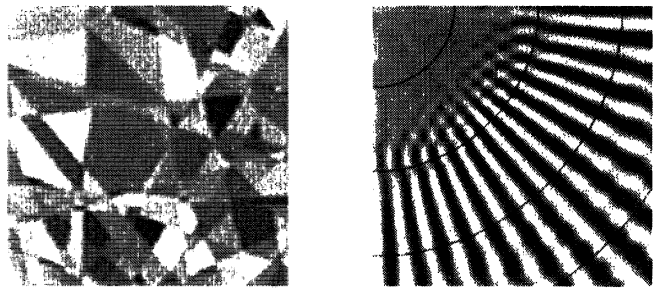


Figure 27. Final Wiener-matrix restorations for the multiresolution image gathering and restoration characterized in figures 25 and 26.

order to illustrate the improvement that is gained in resolution with the addition of increasingly higher spatial frequency bands. Figure 27 presents images of the final restoration at the same scale as that shown in figure 24 for easy comparison. As can be seen, the decomposition together with the loss in information rate allowed for the requantization causes a small but perceptible loss in sharpness and clarity compared with the restored images in figure 24*a* for design 3a. However, the visual quality remains substantially superior to that for the traditional image gathering, either with reconstruction or restoration, as shown in figures 17*a* and 17*b*, respectively.

As these results show, the proper integration of signal analysis and synthesis with image gathering and restoration is a prerequisite for producing sharp and clear pictures of fine detail. This integration does not preclude, of course, the need for adaptive requantization to particular scene statistics and features to obtain high data compression. However, the problem of optimizing analysis filters and requantization maps is not quantitatively addressed in this paper.

## 6. Conclusions

In extending information theory to the assessment of visual communication, this paper united two disciplines, the electro-optical design of image gathering and display devices and the digital processing for image coding and restoration. The mathematical development led to six intuitively attractive figures of merit for assessing visual communication as a function of its critical limiting factors. Two basic problems were




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

dealt with that previous assessments of visual communication have not addressed. One problem is that the limitations inherent in the realizability of the spatial frequency response of optical apertures and the sampling passband of photodetection mechanisms require a trade-off between blurring and aliasing. This precludes the maximum-realizable information rate $\mathcal{H}_m$ that the image-gathering system can produce from closely approaching Shannon's channel capacity $\mathcal{C}$ for a bandwidth-limited system. The other problem is that the image-gathering process bars the encoder from unperturbed access to the scene. This precludes the theoretical minimum data rate $\mathcal{E}$ from approaching its lower bound, the information rate $\mathcal{H}$, unless $\mathcal{H}$ is allowed to be substantially reduced by coarse quantization.

Another problem, which is intrinsic to any quantitative assessment of visual communication, is that of measuring image quality. The best visual quality with which an image can be restored is inevitably limited by a trade-off among blurring, aliasing, and ringing. It is not possible, regardless of the complexity of digital processing, to maximize sharpness without either aliasing or ringing artifacts. The perceptually preferred trade-off varies with target and application. For example, in medical diagnosis and military or planetary reconnaissance, sharpness is often enhanced until artifacts and noise begin to interfere with the interpretation; whereas in entertainment, visually annoying defects are usually suppressed even at the cost of some loss of sharpness. Therefore, the visual quality that can be attained is too elusive to be measured by a single criterion. Instead, our assessment focused on the correlation between information rate and the restorability of images. To do so, we used the WIGE filter which combines the Wiener filter with an interactive enhancement function that permits the user to control the trade-off among maximum-realizable fidelity, sharpness, and clarity (i.e. absence of visually annoying defects).

Quantitative and qualitative assessments show that:

1. The information rate $\mathcal{H}$ that the image-gathering device can produce reaches its maximum-realizable value $\mathcal{H}_m$ when both (*a*) the relationship between the SFR and the sampling passband (i.e. the trade-off between blurring and aliasing) is informationally optimized for the available SNR and (*b*) the sampling passband of the image-gathering device best matches the radiance-field spectrum. For natural scenes this match occurs when the sampling interval is near the mean spatial detail in the scene. The constraints of the image-gathering process limit the maximum-realizable information rate $\mathcal{H}_m$ to approximately one-half of Shannon's channel capacity for a bandwidth-limited system with an average power limitation.

2. The image-gathering device that is designed to produce the maximum-realizable information rate $\mathcal{H}_m$ ordinarily maximizes (*a*) the *efficiency* of the information transmission (i.e. the ratio of the information rate $\mathcal{H}$ to the theoretical minimum data rate $\mathcal{E}$), (*b*) the *quality* of the image restoration (i.e. the restorability of images for fidelity, resolution, sharpness, and clarity), and (*c*) the *robustness* of the image restoration (i.e. the tolerance of the restoration to errors in estimates of the radiance-field statistics). This critical dependence of the efficiency, quality, and robustness of visual communication on the design of the image-gathering device is largely independent of the statistical properties of random scenes.

3. Informationally optimized designs of the image-gathering device differ substantially from the traditional design (table 1). Typically, for the traditional design, the information rate and efficiency are limited to $\mathcal{H}_m \approx 1.8$ bits and $\mathcal{H}_m/\mathcal{E} \approx 0.24$; whereas for informationally optimized designs, they may reach $\mathcal{H}_m \approx 4.2$ bits and

$\mathcal{H}_m/\mathcal{E} \approx 0.75$. The higher information rates permit images to be restored with substantially sharper and clearer detail.

4. The decorrelation attained with redundancy reduction algorithms depends on the statistics of the radiance field and the design of the image-gathering device. For the conditions that are often encountered, the common DPCM with entropy (Huffman) coding can attain *ca.* 75–100% decorrelation.

5. Visual communication with multiresolution decomposition can approach the best possible performance only if both (*a*) the signal analysis filters are integrated with the design of the image-gathering device to maximize the information rate and efficiency and (*b*) the signal synthesis filters are integrated with the image-restoration algorithm to maximize the fidelity and visual quality of the restored image.

6. No loss of information (either in image gathering, coding, transmission, or display) is entirely irrelevant in the sense that it does not impair the visual quality with which images can be restored. However, the degree of this impairment may be difficult to ascertain because the perceived quality of the image is too elusive to be captured by a figure of merit or even by any existing model of human vision.

The emphasis of these assessments were on characterizing practical upper bounds on the information rate and efficiency of image gathering and transmission and on the associated fidelity and visual quality of image restoration. We did not assess the constraints that the blurring and raster effects of image reconstruction often impose on the realizable image quality. However, Fales *et al.* (1996) fully account for these constraints, and Appendix D includes them into the formulations of information rate and fidelity to pave the way for such further assessments.

For example, in the evolution of commerical television to (digital) high-definition television (HDTV), the frame rates of transmission and display will be separated from each other and a buffer will be added to the receiver to decode and process the received signal. This event will open up the opportunity to digitally interpolate and enhance the observed image. When this occurs, then the potential that HDTV offers for improving image quality can be fully realized, for producing the best possible picture, only if the traditional electro-optical design of the image-gathering system (for image reconstruction without the aid of digital processing) is changed to an informationally optimized design. As the results in this paper have shown, this design will not only improve the visual quality with which images can be restored, but it will also improve the efficiency with which encoded information can be transmitted.

## Appendix A. Electro-optical design

The electro-optical design of image-gathering devices ordinarily revolves around two interdependent trade-offs. One trade-off, in terms of geometrical optics (see figure 28), deals with instantaneous field of view (IFOV) versus SNR as controlled largely by the aperture size and the focal length of the objective lens together with the aperture size and the responsivity of the photodetection mechanism (Boyd 1983). Additional factors that may enter into this trade-off are field of view, depth of field, and exposure time. The IFOV and associated sampling lattice that evolve from this trade-off establish the angular resolution, or spatial scale, of the visual communication channel. The other trade-off, in terms of physical optics, deals with the relationship between the SFR of the optical (objective lens and photodetector) apertures and the sampling passband of the photodetection mechanism as a function of the SNR. The




PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

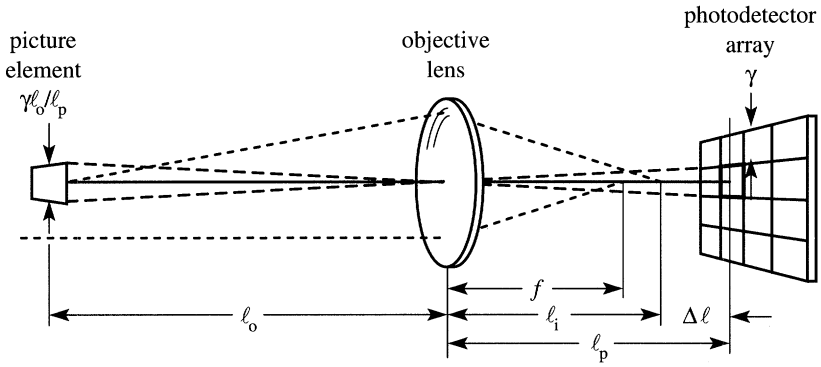



Figure 28. Optical configuration of the image-gathering device with a photodetector array.

Table 2. *Electro-optical design parameters*

| parameter | symbol | unit |
|---|---|---|
| lens aperture diameter | $D$ | m |
| lens focal length | $f$ | m |
| photodetector aperture width | $\gamma$ | m |
| photodetector aperture distance from lens | $\ell_p$ | m |
| steady-state gain | $K$ | A |
| wavelength | $\lambda$ | μm |
| photodetector responsivity | $r(\lambda)$ | A $W^{-1}$ |
| spectral radiance | $L(\lambda)$ | W $m^{-2}$ $sr^{-1}$ $\mu m^{-1}$ |

design that evolves from this trade-off (as analysed in this paper) establishes the upper bounds on the information rate and efficiency of the visual communication channel at the spatial scale fixed by the sampling lattice.

The steady-state gain $K$ in the model of image gathering given by equations (2.1) can be expressed in terms of the geometrical design by (Boyd 1983)

$$K = k\Theta\Omega \int_0^\infty L(\lambda)r(\lambda)\,\mathrm{d}\lambda \approx \frac{\pi k\gamma^2}{4F^2} \int_0^\infty L(\lambda)r(\lambda)\,\mathrm{d}\lambda,$$

where $k$ is the optical transmittance, $\Theta = \pi D^2/4$ is the area of the objective lens aperture, $\Omega = (\gamma/\ell_p)^2 \approx (\gamma/f)^2$ steradian (sr) is the solid angle of the IFOV, and $F = f/D$ is the lens $f$-number. Other design parameters are identified in figure 28 and specified in table 2. Fales *et al.* (1984) and Huck *et al.* (1995) assessed the design of image-gathering devices in terms of these parameters as well as of lens diffraction and transmittance shading, defocus blur, line-scan and photodetector-array mechanisms, square and hexagonal sampling lattices, and image-plane (charge-domain) lateral inhibition.

## Appendix B. Insufficient sampling

This appendix examines the statistical properties of insufficiently sampled signals to delineate their proper treatment in the information-theoretic assessment of the

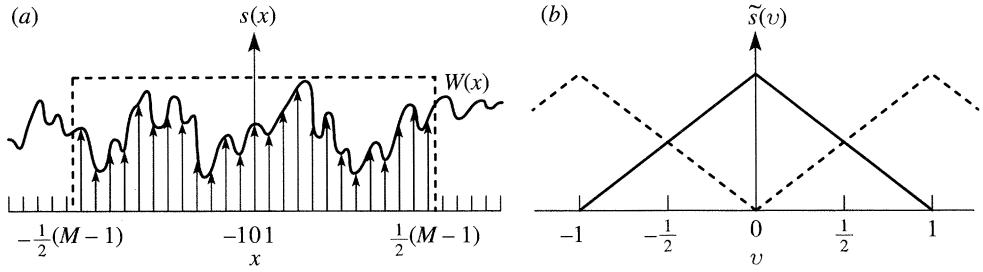



Figure 29. Representation of the insufficiently sampled signal in terms of a finite number of discrete samples. (*a*) The function $s(x)$ sampled at unit intervals within the window $W(x)$ and (*b*) the DFT of $s(x)$.

image-gathering process. For brevity, we present the analysis for a one-dimensional signal, let the gain constant $K = 1$, and omit the photodetector and quantization noises. Hence, equations (2.1) for the sampled signal $s(x)$ simplify to

$$s(x) = [L(x) * \tau(x)]\, \underline{|||}(x),$$

$$\tilde{s}(\upsilon) = [\hat{L}(\upsilon)\ \hat{\tau}(\upsilon)] * \hat{\underline{|||}}(\upsilon) = \hat{L}(\upsilon)\hat{\tau}(\upsilon) + [\hat{L}(\upsilon)\hat{\tau}(\upsilon)] * \hat{\underline{|||}}_s(\upsilon),$$

where $\hat{\underline{|||}}(\upsilon) = \delta(\upsilon) + \hat{\underline{|||}}_s(\upsilon)$. In the next two paragraphs, we manipulate these two equations into a finite dimensional framework, and in the remainder of this appendix, we use this framework to examine the statistical properties of the sampled signal.

First, we represent the signal $s(x)$ by the finite set of discrete signals $s(m)$ at locations $-\frac{1}{2}(M-1), \ldots, -1, 0, 1, \ldots \frac{1}{2}(M-1)$, where $M$ is an odd integer (see figure 29). Therefore,

$$\tilde{s}'(\upsilon) = \sum_{m=-(M-1)/2}^{(M-1)/2} s(m) \mathrm{e}^{-\mathrm{i}2\pi m\upsilon},$$

where the prime on $\tilde{s}(\upsilon)$ indicates the finite set obtained by letting $s(m) = 0$ for $|m| > \frac{1}{2}(M-1)$. Moreover, we choose the finite set of spatial frequencies $\upsilon = k/M$, where $k = -\frac{1}{2}(M-1), \ldots, -1, 0, 1, \ldots \frac{1}{2}(M-1)$. Therefore,

$$\tilde{s}'(k) \equiv \tilde{s}'(k/M) = \sum_{m=-(M-1)/2}^{(M-1)/2} s(m) \mathrm{e}^{-\mathrm{i}2\pi mk/M}.$$

The corresponding inverse DFT is

$$s(m) = \frac{1}{M} \sum_{k=-(M-1)/2}^{(M-1)/2} \tilde{s}'(k) \mathrm{e}^{\mathrm{i}2\pi mk/M}.$$

Thus this initial manipulation replaces $s(x)$ and $\tilde{s}(\upsilon)$ by $s(m)$ and $\tilde{s}'(k)$, respectively, for a finite set of discrete signals $s(m)$ and a finite set of spatial frequencies $k$. A simple interpolation of the discrete set of amplitudes provides all other frequencies for a continuous spatial frequency representation.

Second, we substitute the inverse DFT

$$s(m) = \int_{-1/2}^{1/2} \tilde{s}(\upsilon) \mathrm{e}^{\mathrm{i}\pi m\upsilon}\, \mathrm{d}\upsilon$$

into the above expression for $\tilde{s}'(k)$ and let $u = v - k/M$. Therefore,

$$\tilde{s}'(k) = \int_{-1/2}^{1/2} \tilde{s}(v) \sum_{m=-(M-1)/2}^{(M-1)/2} e^{i2\pi mv}\, dv.$$

Through Fourier manipulations involving the window function $W(x)$ depicted in figure 29*a*, it can be shown that

$$\tilde{s}'(k) = \int_{-\infty}^{\infty} \tilde{s}(v)\hat{W}\left(v - \frac{k}{M}\right) dv,$$

where

$$\hat{W}(v) = M\mathrm{sinc}\, Mv.$$

For large $M$, $\hat{W}(v - k/m)$ possesses a strong peak at and is highly localized near $v = k/m$, and for $M$ approaching infinity, $\hat{W}(v - k/m)$ becomes the impulse function at $v = k/m$, i.e.

$$\lim_{M\to\infty} M\mathrm{sinc}\, Mv = \delta(v).$$

Thus the difference between $\tilde{s}'(k)$ and $\tilde{s}(k/m)$ is merely that the amplitudes of the frequency components of $\tilde{s}'(k)$ are slightly blurred.

Now we proceed with searching for the statistical properties of the set of $M$ variables $\tilde{s}'(k)$. Because the radiance field is Gaussian and all the transformations in the image-gathering process are linear, the signal components $\tilde{s}'(k)$ are also Gaussian by nature of the linear transformation between $\tilde{s}'(k)$ and $s(m)$. Therefore, to determine the statistical properties of $\tilde{s}'(k)$, we need to examine only the covariance $E\{\tilde{s}'(k)\tilde{s}'^*(k')\}$ given by

$$E\{\tilde{s}'(k)\tilde{s}'^*(k')\} = \int_{-\infty}^{\infty}\int E\{\tilde{s}(v)\tilde{s}^*(v')\}M^2\mathrm{sinc}\, M\left(v - \frac{k}{M}\right)\mathrm{sinc}\, M\left(v' - \frac{k'}{M}\right) dv dv'. \tag{B 1}$$

Toward this end, for the statistical properties of the spatially unbounded radiance field, we first note the well-known result that the Fourier amplitudes, $\hat{L}(v)$, are uncorrelated for distinct frequencies, i.e.

$$E\{\hat{L}(v)\hat{L}^*(v')\} = \hat{\Phi}_L(v)\delta(v - v').$$

It follows, then, almost immediately from the introductory relations defining $\tilde{s}(v)$ that the Fourier amplitudes of the sampled signal exhibit the correlation property

$$E\{\tilde{s}(v)\tilde{s}^*(v')\} = \tilde{\Phi}_s(v)\sum_m \delta(v - v' - m),$$

where the PSD $\tilde{\Phi}_s(v)$ of the sampled signal is given by

$$\tilde{\Phi}_s(v) = [\hat{\Phi}_L(v)|\hat{\tau}(v)|^2] * \underline{|\hat{|}|}(v).$$

Therefore,

$$E\{\tilde{s}'(k)\tilde{s}'^*(k')\} = \int_{-\infty}^{\infty} \tilde{\Phi}_s(v)\Delta(v; k, k')\, dv,$$

where

$$\Delta(\upsilon;k,k') = M^2 \operatorname{sinc} M\left(\upsilon - \frac{\kappa}{M}\right) \sum_m \operatorname{sinc} M\left(\upsilon - \frac{k'}{M} - m\right).$$

For large $M$, $\Delta(\upsilon;k,k')$ exhibits strong peaks under the conditions $\upsilon - k'/M = \upsilon - k/M - m = 0$ or $\upsilon = k/M$ and $(k-k')/M = m$. But because $|k| \leqslant (M-1)/2$, only the $m = 0$ term contributes measurably. This leaves the integral property

$$\int_{-\infty}^{\infty} \Delta(\upsilon;k,k')\,\mathrm{d}\upsilon = M\delta_{kk'}.$$

Assuming the PSD $\tilde{\Phi}_s(\upsilon)$ possesses satisfactory smoothness properties relative to $\Delta(\upsilon;k,k')$, we can finally conclude that for sufficiently large $M$ the (finite set of) sampled signal Fourier amplitudes are uncorrelated, i.e.

$$E\{\tilde{s}'(k)\tilde{s}'^*(k')\} \cong \tilde{\Phi}_s(\upsilon)M\delta_{kk'}.$$

Thus we have established the conditions that allow us to write the equality

$$\tilde{s}'(k) = \tilde{s}(k/M) = \sum_m \hat{L}\left(\frac{k}{M} - m\right)\hat{\tau}\left(\frac{k}{M} - m\right)$$

in which the individual components of $\tilde{s}'(k)$ are uncorrelated. For example, because of the restriction $|k| \leqslant (M-1)/2$, the condition $k/M - m = k'/M - m'$ can be satisfied only when $(k,m) = (k',m')$. This result establishes the basic premise by which the aliased signal components, which before sampling were a meaningful part of the input signal, are relegated to the status of noise in tandem with the photodetector noise.

## Appendix C. Quantization

This appendix examines the relationship between the information rate through the analogue-to-digital transformation and the entropy of the digital output signal. We begin the discussion of this relationship with a general model of the quantization process and then continue with the assumed statistical model of this process given in § 2 *a*. To distinguish between the general (or exact) model of quantization and its statistical approximation, we denote the digital output signal as $s_{\mathrm{q}}(x,y)$ for the former and, consistent with the notation elsewhere, as $s(x,y;\kappa)$ for the latter.

Using a one-dimensional signal for brevity, the mutual information $\mathcal{H}_{\mathrm{q}}$ between the analogue input signal $s(x)$ and the digital output signal $s_{\mathrm{q}}(x)$ is

$$\mathcal{H}_{\mathrm{q}} = \mathcal{E}[s_{\mathrm{q}}(x)] - \mathcal{E}[s_{\mathrm{q}}(x)|s(x)],$$

where $\mathcal{E}[s_{\mathrm{q}}(x)]$ is the entropy of $s_{\mathrm{q}}(x)$ and $\mathcal{E}[s_{\mathrm{q}}(x)|s(x)]$ is the conditional entropy that represents the uncertainty about $s_{\mathrm{q}}(x)$ when $s(x)$ is known. To give explicit expressions for these entropies, we first define the input and output signals as vectors in an image frame of $N$ pixels. In particular, we let the vector $\bar{s}$ represent the analogue input signal $\bar{s} = (\ldots, s(n), s(n+1), \ldots)$, and $\bar{s}_{\mathrm{q}}$ represent the digital output signal $\bar{s}_{\mathrm{q}} = (\ldots, s_{\mathrm{q}}(n), s_{\mathrm{q}}(n+1), \ldots)$ with the quantized levels $\bar{s}(\bar{j}) = (\ldots, s_{j_{\mathrm{n}}}, s_{j_{n+1}}, \ldots)$, where the vector $\bar{j} = (\ldots, j_{\mathrm{n}}, j_{n+1}, \ldots)$ represents the frame of quantization integers. Moreover, we let $s^-_{j_{\mathrm{n}}}$ and $s^+_{j_{\mathrm{n}}}$ $[s^-_{j_{\mathrm{n}}} \leqslant s(n) < s^+_{j_{\mathrm{n}}}]$ represent the range of analogue amplitudes to be quantized as level $s_{j_{\mathrm{n}}}$ (for the $n$th pixel). When this range of analogue amplitudes is arranged as the vectors $\bar{s}^-(\bar{j})$ and $\bar{s}^+(\bar{j})$, then this arrangement defines

a frame space with elemental quantization cells associated with the vector $\bar{s}(\bar{j})$ that is represented by the conditional probability

$$P_{s_q|s}[\bar{s}(\bar{j})|\bar{s}] = \begin{cases} 1, & \text{if } \bar{s}^-(\bar{j}) \leqslant \bar{s} < \bar{s}^+(\bar{j}) \\ 0, & \text{elsewhere.} \end{cases}$$

With these definitions, the entropy $\mathcal{E}[s_q(x)]$ of the digital output signal is explicitly given by

$$\mathcal{E}[s_q(x)] = -\frac{1}{N}\sum_{\bar{j}=\bar{0}}^{\overline{\kappa-1}} P_{s_q}[\bar{s}(\bar{j})] \log P_{s_q}[\bar{s}(\bar{j})],$$

where $\kappa$ is the number of quantization levels. Using the defining relation,

$$P_{ss_q}[\bar{s}, \bar{s}_q] = P_s[\bar{s}]P_{s_q|s}[\bar{s}_q|\bar{s}],$$

the probability $P_{s_q}[\bar{s}(\bar{j})]$ in the above summation can be expressed as

$$P_{s_q}[\bar{s}(\bar{j})] = \int_{-\bar{\infty}}^{\bar{\infty}} \mathrm{d}^N s\, P_s[\bar{s}]P_{s_q|s}[\bar{s}(\bar{j})|\bar{s}],$$

where $\mathrm{d}^N s$ is the differential volume element of the analogue pixel frame amplitudes. Assuming that $s^+_{j_{n-1}} = s^-_{j_n}$, the integration over the entire analogue amplitude space can be partitioned into the set of all quantization cells as

$$P_{s_q}[\bar{s}(\bar{j})] = \int_{\bar{s}^-(\bar{j})}^{\bar{s}^+(\bar{j})} \mathrm{d}^N s\, P_s[\bar{s}]P_{s_q|s}[\bar{s}(\bar{j})|\bar{s}] + \sum_{\substack{\bar{k}=\bar{0} \\ (\bar{k}\neq\bar{j})}}^{\overline{\kappa-1}} \int_{\bar{s}^-(\bar{k})}^{\bar{s}^+(\bar{k})} \mathrm{d}^N s\, P_s[\bar{s}]P_{s_q|s}[\bar{s}(\bar{j})|\bar{s}].$$

The conditional probability is unity in the first integral, and it vanishes in the second integral. Therefore, as expected,

$$P_{s_q}[\bar{s}(\bar{j})] = \int_{\bar{s}^-(\bar{j})}^{\bar{s}^+(\bar{j})} \mathrm{d}^N s\, P_s[\bar{s}].$$

Similarly, the conditional entropy $\mathcal{E}[s_q(x)|s(x)]$ is explicitly given by

$$\begin{aligned}\mathcal{E}[s_q(x)|s(x)] &= -\frac{1}{N}\int_{-\bar{\infty}}^{\bar{\infty}} \mathrm{d}^N s \sum_{\bar{j}=\bar{0}}^{\overline{\kappa-1}} P_s[\bar{s}]P_{s_q|s}[\bar{s}(\bar{j})|\bar{s}] \log P[\bar{s}(\bar{j})|\bar{s}] \\ &= -\frac{1}{N}\left\{\sum_{\bar{j}=\bar{0}}^{\overline{\kappa-1}} \int_{\bar{s}^-(\bar{j})}^{\bar{s}^+(\bar{j})} \mathrm{d}^N s + \sum_{\substack{\bar{j},\bar{k}=\bar{0} \\ (\bar{j}\neq\bar{k})}}^{\overline{\kappa-1}} \int_{\bar{s}^-(\bar{k})}^{\bar{s}^+(\bar{k})} \mathrm{d}^N s\right\} \\ &\quad \times P_s[\bar{s}]P_{s_q|s}[\bar{s}(\bar{j})|\bar{s}] \log P_{s_q|s}[\bar{s}(\bar{j})|\bar{s}].\end{aligned}$$

The conditional probability $P_{s_q|s}[\bar{s}(\bar{j})|\bar{s}]$ leads to $(1)\log(1) = 0$ in the first integral and to $(0)\log(0) = 0$ (in the sense of a limit) in the second integral. Hence, $\mathcal{E}[s_q(x)|s(x)]$ vanishes, and

$$\mathcal{H}_q = \mathcal{E}[s_q(x)].$$

This pivotal result states that the entropy $\mathcal{E}[s_q(x)]$ of the digital output signal $s_q(x)$ is precisely given by the information rate $\mathcal{H}_q$ through the analogue-to-digital transformation or, equivalently, by the mutual information between the analogue input

and digital output signals. This result is intuitively acceptable when one considers that the uncertainty about $s_q(x)$ vanishes when $s(x)$ is known because the path from analogue to digital is deterministic. This does not hold for the reverse from digital to analogue.

It is reasonable to assume that the conclusion about the equivalence of the information rate through the analogue-to-digital transformation and the entropy of the digital output signal carries over to the approximate statistical model of the quantization process that § 2 *a* presents. Accordingly, we use the entropy symbol $\mathcal{E}$ to denote the mutual information between the analogue input signal $s(x)$ and the digital output signal $s(x;\kappa)$, as defined by

$$\mathcal{E} = \mathcal{E}[s(x;\kappa)] - \mathcal{E}[s(x;\kappa)|s(x)].$$

The digital signal $s(x;\kappa)$ can be decomposed as the sum of the analogue signal $s(x)$ (or $\bar{s}$ in the general model) and a correction signal $n_q(x;\kappa)$ (or $\bar{n}_q$), so that $s(x;\kappa) = s(x) + n_q(x;\kappa)$ (or $\bar{s}_q = \bar{s} + \bar{n}_q$). The correction signal $\bar{n}_q$ did not appear explicitly in the general model of quantization since $\bar{s}$ uniquely determines $\bar{s}_q$. Hence, $\bar{n}_q$ simply becomes the deterministic signal that is required, when added to $\bar{s}$, to give the already known digital signal $\bar{s}_q$. In the statistical model, $n_q(x;\kappa)$ is an additive random process that is statistically independent of the analogue input signal $s(x)$ and forces a continuous characteristic on the output signal $s(x;\kappa)$. Therefore, the conditional entropy $\mathcal{E}[s(x;\kappa)|s(x)]$ no longer vanishes but, instead, becomes the entropy $\mathcal{E}[n_q(x;\kappa)]$ of the quantization noise $n_q(x;\kappa)$. Correspondingly, the entropy $\mathcal{E}[s(x;\kappa)]$ of the approximate output signal $s(x;\kappa)$ ceases to be equal to the entropy $\mathcal{E}[\bar{s}_q]$ of the discrete output signal $\bar{s}_q$. The entropies $\mathcal{E}[s(x;\kappa)]$ and $\mathcal{E}[n_q(x;\kappa)]$ of the approximate signal and noise take on a physical meaning only as the difference

$$\mathcal{E} = \mathcal{E}[s(x;\kappa)] - \mathcal{E}[n_q(x;\kappa)]$$

that § 3 *b* uses to develop quantitative expressions for $\mathcal{E}$.

The equivalence between the information rate through the analogue-to-digital transformation and the entropy of the digital output signal is further supported by the following consideration. As above, we let $N$ be the number of pixels in an image frame. Now, however, we model the amplitude of each pixel of the analogue input signal $s(x)$ as a quantized version. Here, the number of quantization levels $\kappa_a$ is arbitrarily large to closely represent the analogue signal, but it remains finite for counting purposes so that $M[s(x)] \leqslant \kappa_a^N$ represents the number of distinguishable images (messages) that $s(x)$ contains. Hence, the entropy of $s(x)$ is

$$\mathcal{E}[s(x)] = \frac{1}{N}\log M[s(x)] \leqslant \frac{1}{N}\log \kappa_a^N = \log \kappa_a.$$

Now, let $M[s(x)|s_q(x)]$ be the number of images that the analogue-to-digital transformation (channel noise) causes to be erroneous. Then the entropy of the analogue input signal $s(x)$ given the digital output signal $s_q(x)$ is

$$\mathcal{E}[s(x)|s_q(x)] = \frac{1}{N}\log M[s(x)|s_q(x)],$$

and the corresponding information rate is

$$\begin{aligned}\mathcal{H}_q &= \mathcal{E}[s(x)] - \mathcal{E}[s(x)|s_q(x)] \\ &= \frac{1}{N}\log M[s(x)] - \frac{1}{N}\log M[s(x)|s_q(x)] = \frac{1}{N}\log M_q[s(x)],\end{aligned}$$

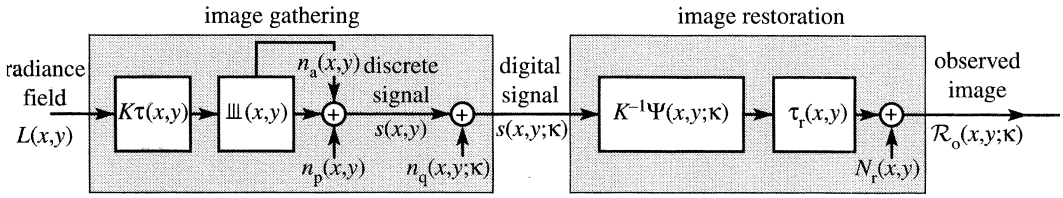



Figure 30. General model of image gathering and restoration.

where

$$M_{\mathrm{q}}[s(x)] \equiv \frac{M[s(x)]}{M[s(x)|s_{\mathrm{q}}(x)]}.$$

This definition of $M_{\mathrm{q}}[s(x)]$ may be interpreted as the number of distinguishable images that the digital output signal $s(x;\kappa)$ contains. Hence, $\mathcal{H}_{\mathrm{q}}$ is the ordinary well-defined entropy that characterizes the set of $M_{\mathrm{q}}[s(x)]$ images. Note that as $\kappa_{\mathrm{a}} \rightarrow \infty$, the upper bound of the entropy $\mathcal{E}[s(x)]$ of the analogue input signal increases without limit (which is an intrinsic characteristic of the entropy of continuous signals). However, the information rate $\mathcal{H}_{\mathrm{q}}$ of the digital output signal remains finite.

Finally, let $\kappa_{\mathrm{e}}^N = M_{\mathrm{q}}[s(x)]$ be the set of equal length code words to represent the $M_{\mathrm{q}}[s(x)]$ distinguishable images. Then $\mathcal{H}_{\mathrm{q}}$ becomes

$$\mathcal{H}_{\mathrm{q}} = \frac{1}{N} \log \kappa_{\mathrm{e}}^N = \log \kappa_{\mathrm{e}}.$$

For the simplifying assumption in § 2 *a* that the analogue-to-digital transformation consists of a signal $s(x)$ and additive, independent noise $n_{\mathrm{q}}(x)$, both of which are white, band-limited, and Gaussian with variances $\sigma_s^2$ and $\sigma_{\mathrm{q}}^2$, respectively, $\mathcal{H}_q$ is

$$\mathcal{H}_{\mathrm{q}} = \tfrac{1}{2} \log(1 + \sigma_s^2/\sigma_{\mathrm{q}}^2).$$

Equating these last two expressions for $\mathcal{H}$ yields the relationship

$$\kappa_{\mathrm{e}} = (1 + \sigma_s^2/\sigma_{\mathrm{q}}^2)^{1/2} \approx \sigma_s/\sigma_{\mathrm{q}}.$$

This result suggests that $\kappa_{\mathrm{e}}$ has the physical equivalence of uniformly quantizing the analogue signal amplitude ($\sim \sigma_s$) into intervals of width of the quantization noise ($\sim \sigma_{\mathrm{q}}$).

## Appendix D. Image restoration without interpolation

Figure 30 depicts a model of image gathering and restoration without the digital interpolation assumed throughout the paper. The image-restoration process now transforms the digital signal $s(x, y; \kappa)$ into the observed image $\mathcal{R}_{\mathrm{o}}(x, y; \kappa)$, as defined by the expressions

$$\mathcal{R}_{\mathrm{o}}(x, y; \kappa) = K^{-1} s(x, y; \kappa) * \Psi(x, y; \kappa) * \tau_{\mathrm{r}}(x, y) + N_{\mathrm{r}}(x, y),$$
$$\hat{\mathcal{R}}_{\mathrm{o}}(v, \omega; \kappa) = K^{-1} \tilde{s}(v, \omega; \kappa)\, \tilde{\Psi}(v, \omega; \kappa) \hat{\tau}_{\mathrm{r}}(v, \omega) + \hat{N}_{\mathrm{r}}(v, \omega),$$

where $s(x, y; \kappa)$ and $\tilde{s}(v, \omega; \kappa)$ are given by equations (2.1 *a*) and (2.1 *b*), respectively. The corresponding discrete Wiener filter $\tilde{\Psi}(v, \omega; \kappa)$ is (Fales *et al.* 1996)

$$\tilde{\Psi}(v, \omega; \kappa) = \frac{[\hat{\Phi}_L(v, \omega) \hat{\tau}^*(v, \omega) \hat{\tau}_{\mathrm{r}}^*(v, \omega)] * \underline{|\hat{|}|}}{[\hat{\Phi}_L(v, \omega) |\hat{\tau}(v, \omega)|^2 * \underline{|\hat{|}|} + K^{-2} \tilde{\Phi}_{\mathrm{p}}(v, \omega) + K^{-2} \tilde{\Phi}_{\mathrm{q}}(v, \omega; \kappa)][|\hat{\tau}_{\mathrm{r}}(v, \omega)|^2 * \underline{|\hat{|}|}]},$$

the throughput SFR $\hat{I}(\upsilon,\omega;\kappa)$ is

$$\hat{I}(\upsilon,\omega;\kappa) = \hat{\tau}(\upsilon,\omega)\,\tilde{\Psi}(\upsilon,\omega;\kappa)\hat{\tau}_{\mathrm{r}}(\upsilon,\omega),$$

and the minimum MSRE $\epsilon_{\mathrm{o}}^2$ is

$$\epsilon_{\mathrm{o}}^2 = \iint_{-\infty}^{\infty} \hat{\epsilon}_{\mathrm{o}}^2(\upsilon,\omega;\kappa)\,\mathrm{d}\upsilon\mathrm{d}\omega,$$

where

$$\hat{\epsilon}_{\mathrm{o}}^2(\upsilon,\omega;\kappa) = \hat{\Phi}_L(\upsilon,\omega)[1-\hat{I}(\upsilon,\omega;\kappa)] + \hat{\Phi}_{\mathrm{r}}(\upsilon,\omega).$$

The corresponding information rate $\mathcal{H}_{\mathrm{o}}$ and maximum-realizable fidelity $\mathcal{F}_{\mathrm{o}}$ of the observed image are given as follows. The information rate is

$$\mathcal{H}_{\mathrm{o}} = \frac{1}{2}\iint_{\hat{B}} \log\left[1 + \frac{\hat{\Phi}_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2}{\hat{\Phi}_L(\upsilon,\omega)|\hat{\tau}(\upsilon,\omega)|^2 * \underline{\hat{|||}}_s + K^{-2}\tilde{\Phi}_{\mathrm{n}}(\upsilon,\omega;\kappa)}\right]\mathrm{d}\upsilon\mathrm{d}\omega,$$

where

$$\tilde{\Phi}_{\mathrm{n}}(\upsilon,\omega;\kappa) = \tilde{\Phi}_{\mathrm{p}}(\upsilon,\omega) + \tilde{\Phi}_{\mathrm{q}}(\upsilon,\omega;\kappa) + \left[\frac{|\tilde{\Psi}(\upsilon,\omega;\kappa)|\hat{\tau}_{\mathrm{r}}(\upsilon,\omega)|^2}{K^2\hat{\Phi}_{\mathrm{r}}(\upsilon,\omega)} * \underline{\hat{|||}}\right]^{-1}.$$

This expression reduces to equation (3.9 *a*) in the limits

$$\tilde{\Psi}(\upsilon,\omega;\kappa)\hat{\tau}_{\mathrm{r}}(\upsilon,\omega) \to \hat{\Psi}(\upsilon,\omega;\kappa)$$

and

$$\left[\frac{|\tilde{\Psi}(\upsilon,\omega;\kappa)\hat{\tau}_{\mathrm{r}}(\upsilon,\omega)|^2}{\hat{\Phi}_{\mathrm{r}}(\upsilon,\omega)} * \underline{\hat{|||}}\right]^{-1} \to \frac{\hat{\Phi}_{\mathrm{r}}(\upsilon,\omega)}{|\hat{\Psi}(\upsilon,\omega;\kappa)|^2}$$

for which it is evident that the effects of the small response of $\hat{\Psi}(\upsilon,\omega;\kappa)$ beyond the sampling passband are neglected. And the maximum-realizable fidelity is

$$\mathcal{F}_{\mathrm{o}} = \sigma_L^{-2}\iint_{-\infty}^{\infty}[\hat{\Phi}_L(\upsilon,\omega)\hat{I}(\upsilon,\omega;\kappa) - \hat{\Phi}_{\mathrm{r}}(\upsilon,\omega)]\,\mathrm{d}\upsilon\mathrm{d}\omega.$$

In the limit $\tilde{\Psi}(\upsilon,\omega;\kappa)\hat{\tau}_{\mathrm{r}}(\upsilon,\omega) \to \hat{\Psi}(\upsilon,\omega;\kappa)$,

$$\mathcal{F}_{\mathrm{o}} = \sigma_L^{-2}\iint_{-\infty}^{\infty}\{\hat{\Phi}_L(\upsilon,\omega)[1-2^{-\hat{\mathcal{H}}(\upsilon,\omega;\kappa)}] - \hat{\Phi}_{\mathrm{r}}(\upsilon,\omega)\}\,\mathrm{d}\upsilon\mathrm{d}\omega,$$

where $\hat{\mathcal{H}}(\upsilon,\omega;\kappa)$ is the spectral information density given by the integrand of equation (3.2 *f*) for $\mathcal{H}$.

## References

Akansu, A. N. & Haddad, R. A. 1992 *Multiresolution signal decomposition.* Boston: Academic Press.

Alter-Gartenberg, R., Fales, C. L., Huck, F. O. & McCormick, J. A. 1992 Image gathering and processing for high-resolution edge detection. In *Computer vision and image understanding* (ed. L. Shapiro & A. Rosenfeld). New York: Academic Press.

Andrews, H. C. & Hunt, B. R. 1977 *Digital image restoration.* Englewood Cliffs, NJ: Prentice-Hall.

Barlow, H. B. 1981 Critical limiting factors in the design of the eye and visual cortex. *Proc. R. Soc. Lond.* B **212**, 1–34.

Bowker, D. E., Davis, R. E., Myrick, D. L., Stacy, K. & Jones, W. T. 1985 Spectral reflectances of natural targets for use in remote sensing studies. *NASA Ref. Publ.* **1139**.

Boyd, R. W. 1983 *Radiometry and the detection of optical radiation.* New York: Wiley.

Burt, P. J. & Adelson, E. H. 1983 The Laplacian pyramid as a compact image code. *IEEE Trans. Commun.* **COM-31**, 532–540.

Daubechies, I. 1988 Orthonormal bases of compactly supported wavelets. *Commun. Pure Appl. Math.* **41**, 909–996.

Daubechies, I. 1990 The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inform. Theory* **36**, 961–1005.

Ebrahimi, T. & Kunt, M. 1991 Image compression by Gabor expansion. *Opt. Engng* **30**, 873–880.

Fales, C. L., Huck, F. O. & Samms, R. W. 1984 Imaging system design for improved information capacity. *Appl. Opt.* **23**, 872–888.

Fales, C. L. & Huck, F. O. 1991 An information theory of image gathering. *Inform. Sci.* **57–58**, 245–285.

Fales, C. L., Huck, F. O., Alter-Gartenberg, R. & Rahman, Z. 1992 Multiresolution image-gathering and restoration. *J. Visual Commun. Image Represent.* **3**, 356–363.

Fales, C. L., Huck, F. O., Alter-Gartenberg, R. & Rahman, Z. 1996 Image gathering and digital restoration. *Phil. Trans. R. Soc. Lond.* A **354**, 2249–2287. (Following paper.)

Fellgett, P. B. & Linfoot, E. H. 1955 On the assessment of optical images. *Phil. Trans. R. Soc. Lond.* A **247**, 369–407.

Frieden, B. R. 1970 Information, and the restorability of images. *J. Opt. Soc. Am.* **60**, 575–576.

Gabor, D. 1946 Theory of communication. *Proc. IEEE* **93**, 429–457.

Gonzalez, R. C. & Wintz, P. 1987 *Digital image processing*, 2nd edn. Reading, MA: Addison-Wesley.

Huang, T. S. (ed.) 1979 *Picture processing and digital filtering.* Berlin: Springer.

Huck, F. O., Fales, C. L., Halyo, N., Samms, R. W. & Stacy, K. 1985 Image gathering and processing: information and fidelity. *J. Opt. Soc. Am.* A **2**, 1644–1666.

Huck, F. O., Fales, C. L., McCormick, J. A. & Park, S. K. 1988 Image-gathering system design for information and fidelity. *J. Opt. Soc. Am.* A **5**, 285–299.

Huck, F. O., Fales, C. L., Alter-Gartenberg, R., Rahman, Z. & Reichenbach, S. E. 1993 Visual communication: information and fidelity. *J. Visual Commun. Image Represent.* **4**, 62–78.

Huck, F. O., Fales, C. L., Alter-Gartenberg, R. & Rahman, Z. 1994 Visual communication: Information and data transmission. *J. Visual Commun. Image Represent.* **5**, 282–300.

Huck, F. O., Fales, C. L., Jobson, D. J. & Rahman, Z. 1995 Electro-optical design for efficient visual communication. *Opt. Engng* **34**, 795–813.

Itakura, Y., Tsutsumi, S. & Takagi, T. 1974 Statistical properties of the background noise for the atmospheric windows in the intermediate infrared region. *Infrared Phys.* **14**, 17–29.

Jain, A. K. 1989 *Fundamentals of digital image processing.* Englewood Cliffs, NJ: Prentice-Hall.

Kass, M. & Hughes, J. 1983 A stochastic image model for AI. In *Proc. IEEE Int. Conf. on Systems, Man, and Cybernetics*, pp. 369–372. New York: Institute of Electrical and Electronics Engineers.

Linfoot, E. H. 1955 Information theory and optical images. *J. Opt. Soc. Am.* **45**, 808–819.

Linfoot, E. H. 1956 Transmission factors and optical design. *J. Opt. Soc. Am.* **46**, 740–752.

Linfoot, E. H. 1958 Quality evaluations of optical systems. *Optica Acta* **5**, 1–14.

McCormick, J. A., Alter-Gartenberg, R. & Huck, F. O. 1989 Image gathering and restoration: information and visual quality. *J. Opt. Soc. Am.* A **6**, 987–1005.

Mallat, S. 1989 Multifrequency channel decompositions of images and wavelet models. *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-37**, 2901–2110.





PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Mallat, S. 1989 A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Analysis Mach. Intell.* **11**, 674–693.

Marr, D. & Hildreth, E. 1980 Theory of edge detection. *Proc. R. Soc. Lond.* B **207**, 187–217.

Marr, D. 1982 *Vision.* San Francisco, CA: Freeman.

Mertz, P. & Gray, F. 1934 Theory of scanning and its relation to the characteristics of the transmitted signal in telephotography and television. *Bell Syst. Technol.* **13**, 494–515.

Modestino, J. W. & Fries, R. W. 1980 Construction and properties of a useful two-dimensional random field. *IEEE Trans. Inform. Theory* **IT-26**, 44–50.

Pratt, W. K. 1978 *Digital image processing.* New York: Wiley.

Rosenfeld, A. & Kak, A. C. 1982 *Digital picture processing.* New York: Academic Press.

Schade Sr, O. H. 1951 Image gradation, graininess and sharpness in television and motion-picture systems. I. Image structure and transfer characteristics. *J. Soc. Motion Pict. Telev. Engng* **56**, 137–171.

Schade Sr, O. H. 1952 Image gradation, graininess and sharpness in television and motion-picture systems. II. The grain structure of motion picture images. *J. Soc. Motion Pict. Telev. Engng* **58**, 181–222.

Schade Sr, O. H. 1953 Image gradation, graininess and sharpness in television and motion-picture systems. III. The grain structure of television images. *J. Soc. Motion Pict. Telev. Engng* **61**, 97–164.

Schade Sr, O. H. 1955 Image gradation, graininess and sharpness in television and motion-picture systems. IV. Image analysis in photographic and television systems. *J. Soc. Motion Pict. Telev. Engng* **64**, 593–617.

Schreiber, W. F. 1993 *Fundamentals of electronic imaging systems*, 3rd edn. New York: Springer.

Shannon, C. E. 1948 A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423; **28**, 623–656.

Shannon, C. E. 1959 Coding theorems for a discrete source with a fidelity criterion. *IRE Natn. Conv. Rec.*, part 4, pp. 142–164. (Also in *Information and decision process* (ed. R. A. Machol), pp. 93–126 (1960). New York: McGraw-Hill.)

Shannon, C. E. & Weaver, W. 1964 *The mathematical theory of communication.* Urbana: University of Illinois Press.

Ulichney, R. 1987 *Digital halftoning.* Cambridge, MA: MIT Press.

Vetterli, M. & Herley, C. 1992 Wavelets and filter banks: theory and design. *IEEE Trans. Signal Process.* **10**, 2207–2232.

Wiener, N. 1949 *Extrapolation, interpolation, and smoothing of stationary time series.* New York: Wiley.

Woods, J. W. 1991 *Subband image coding.* Boston: Kluwer.

Yeh, P.-S., Rice, R. F. & Miller, W. 1991 On the optimality of code options for a universal noiseless coder. *JPL Publ.* **91-2**.

*Received 8 June 1994; revised 18 September 1995; accepted 22 January 1996*





PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

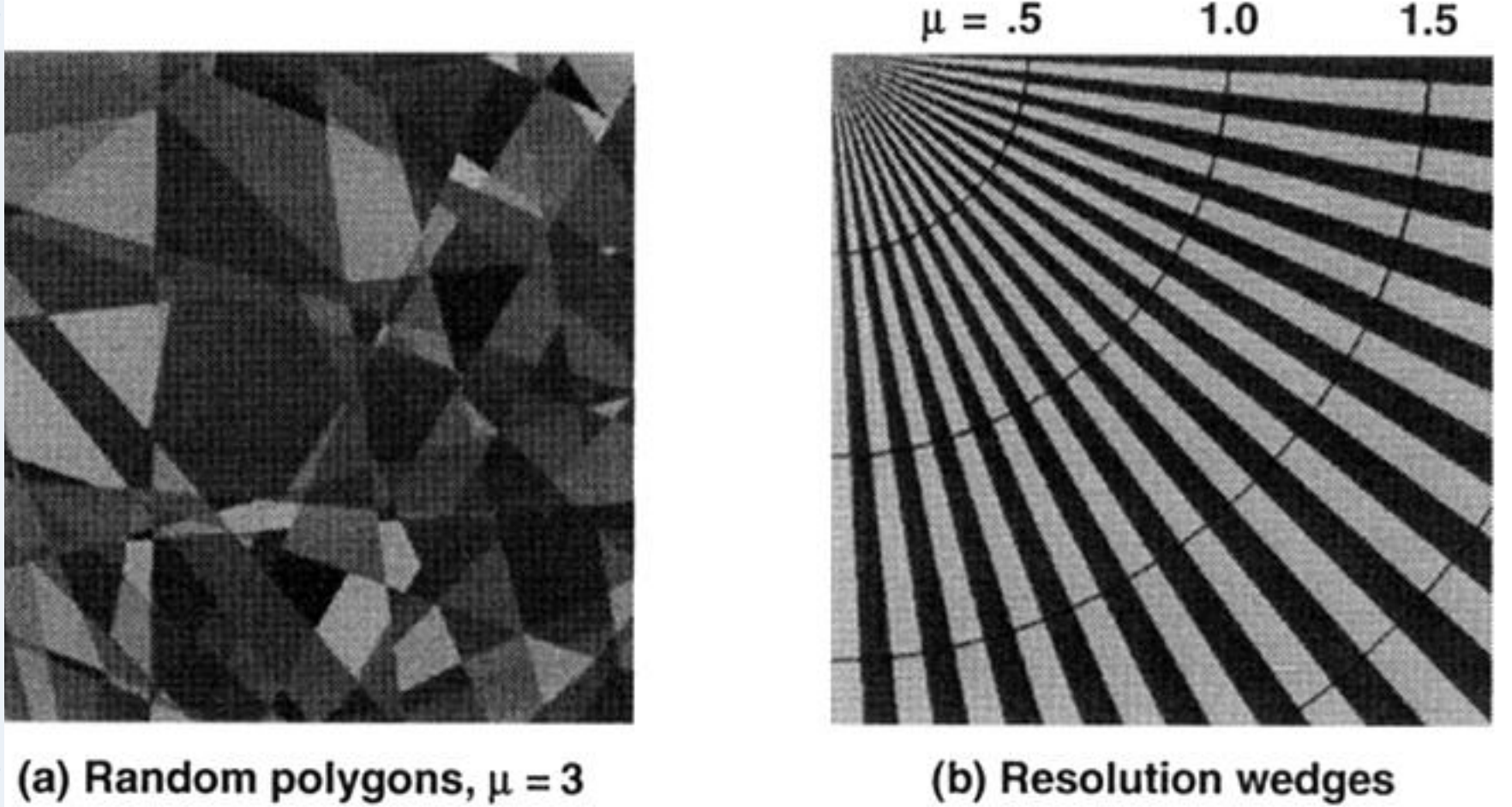
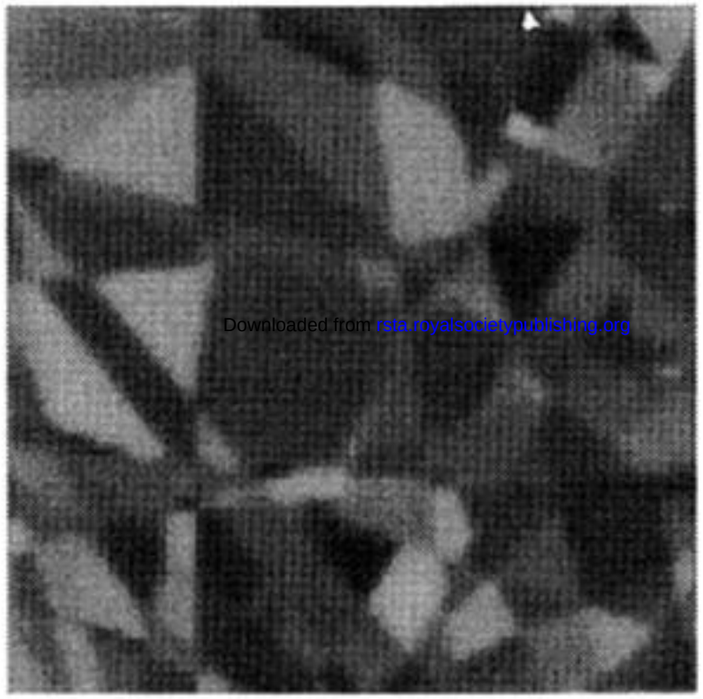
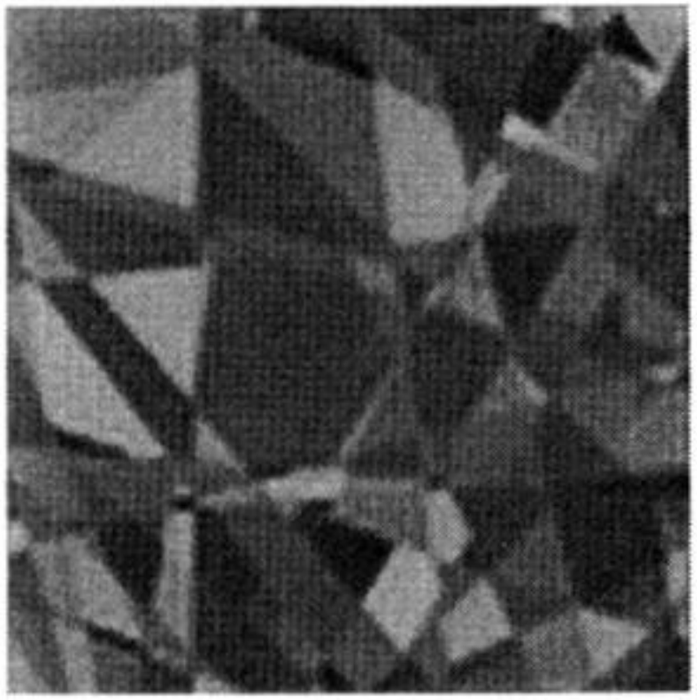
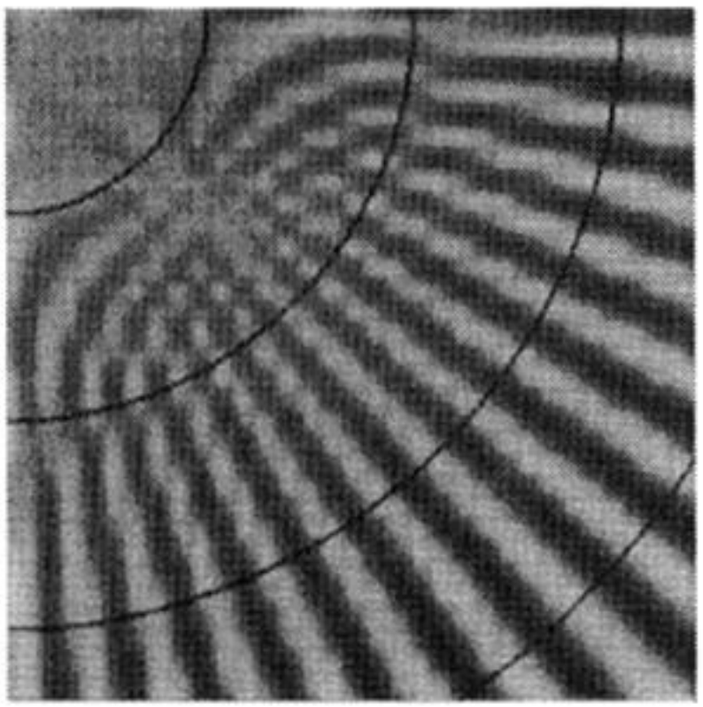
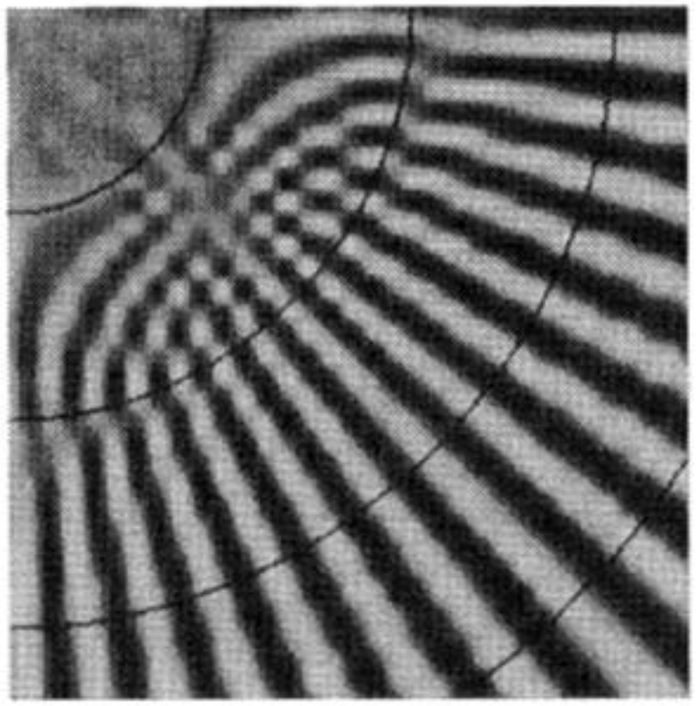


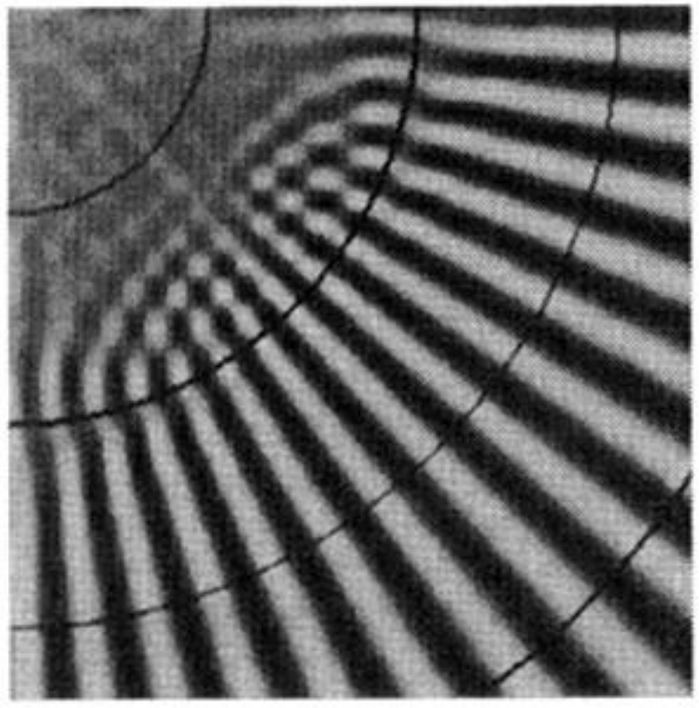

Figure 9. Targets.









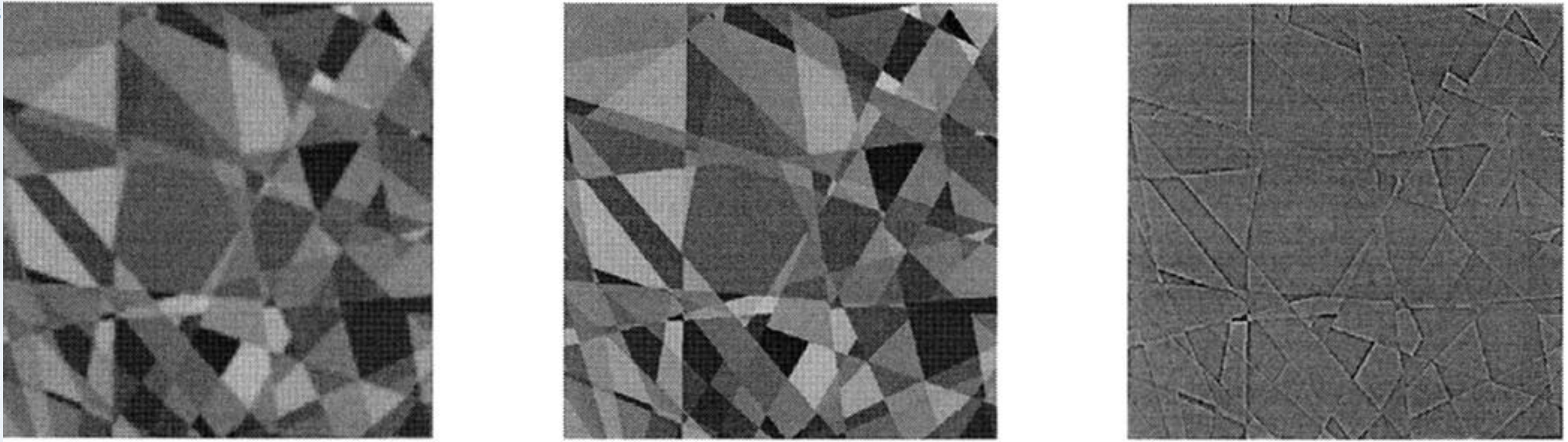


(a) Traditional image gathering and reconstruction

(b) Traditional image gathering with restoration

(c) Informationally optimized image gathering with restoration

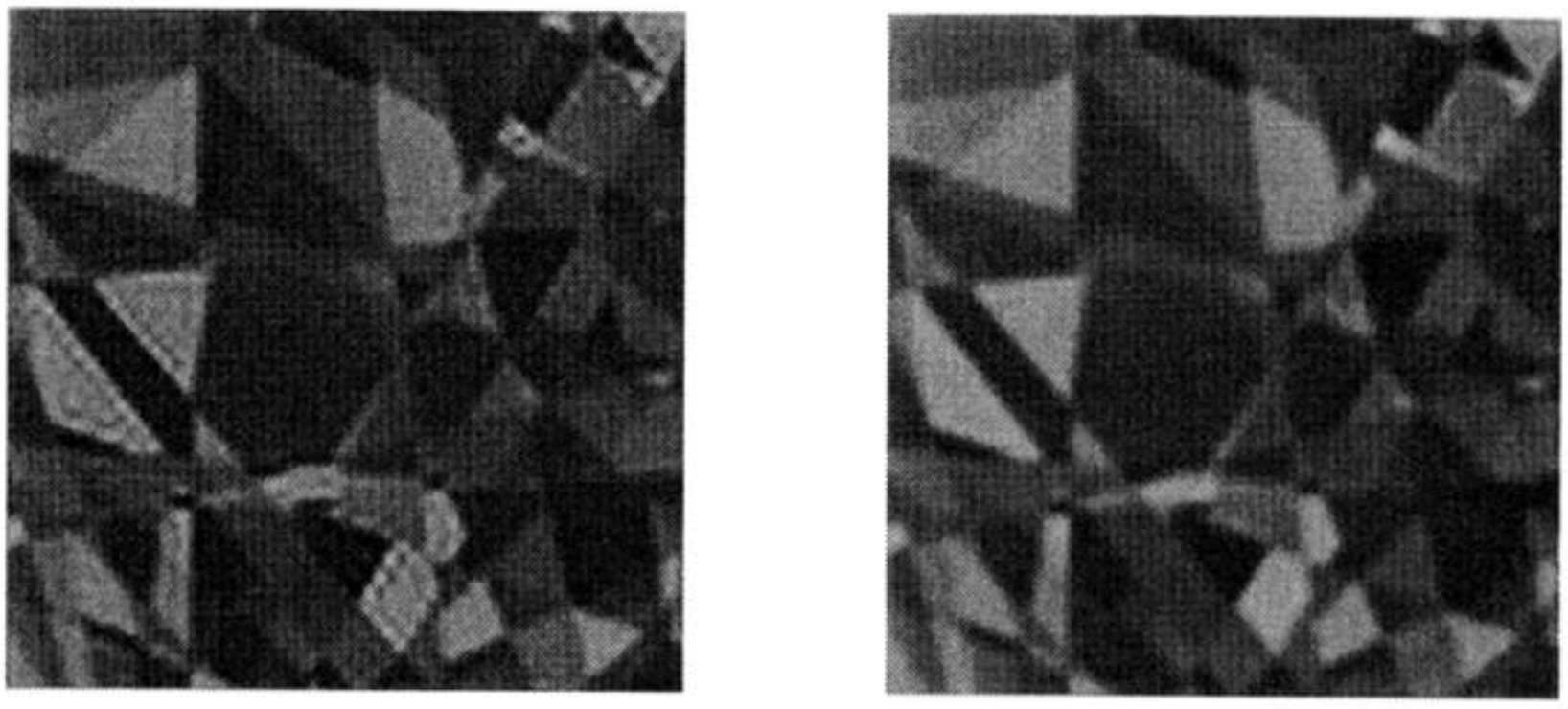
igure 17. Reconstructions and Wiener restorations for two designs of the image-gathering evice, the traditional design (design 1) and an informationally optimized design (design 2a) as pecified in table 1.

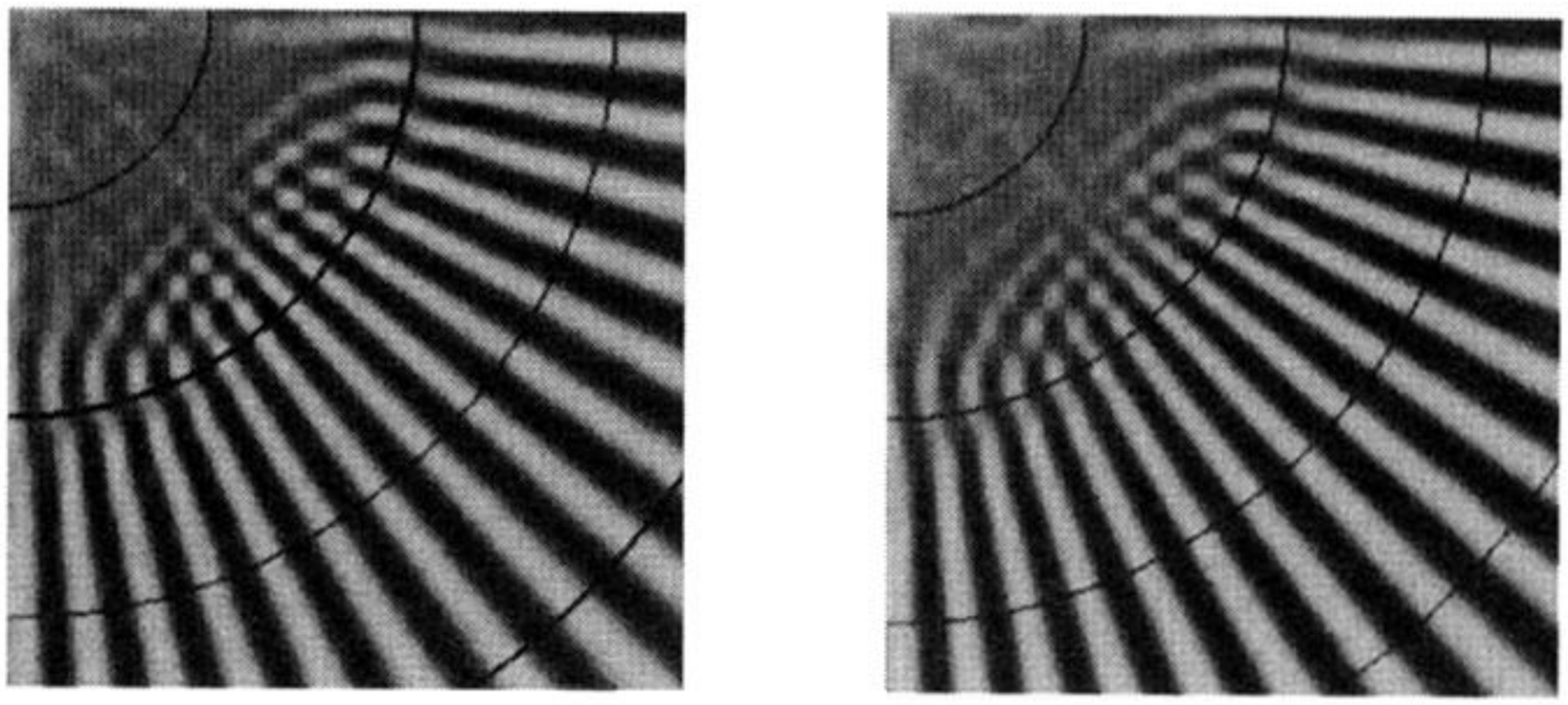


(a) Observed image (b) Blurred component (c) Aliased components

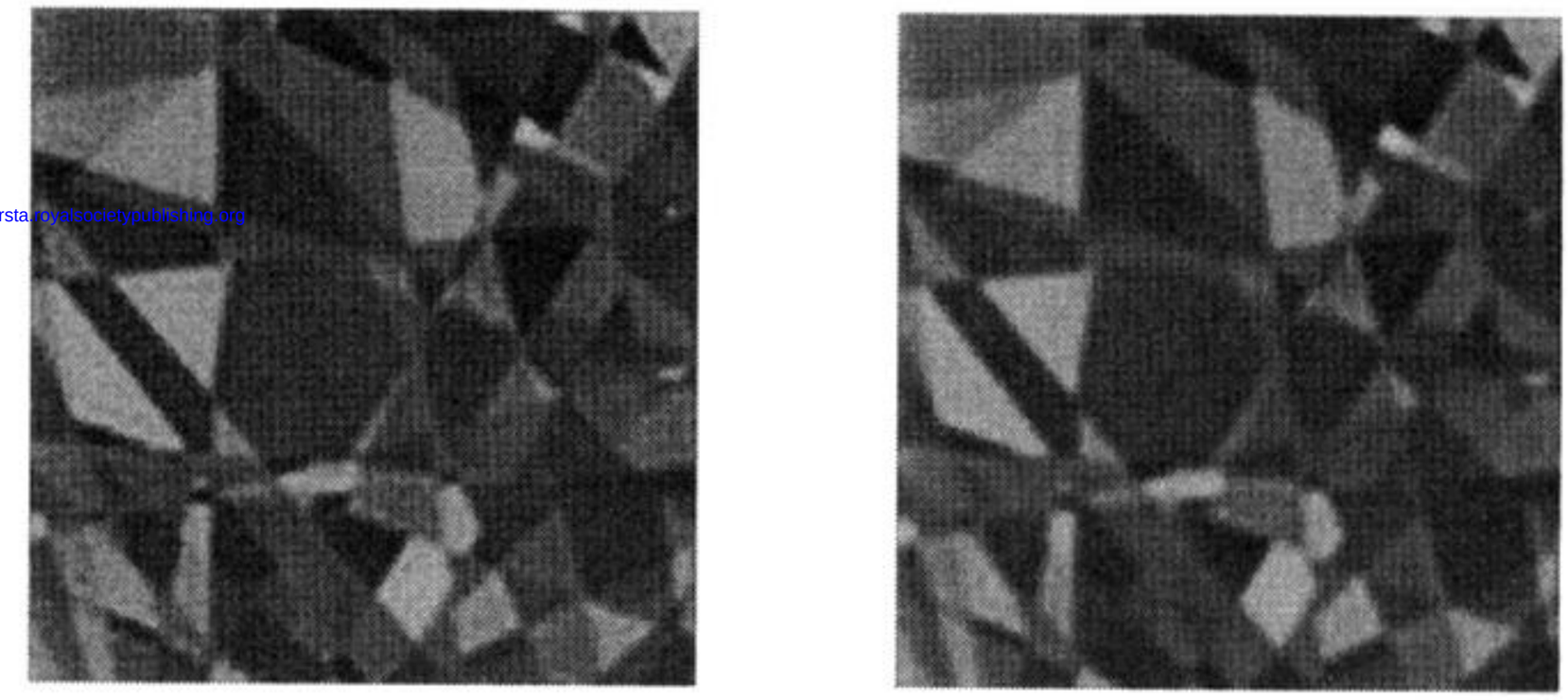
gure 20. Observed image and its blurred and aliased components for the traditional image gathering and reconstruction.

**Design 2a**



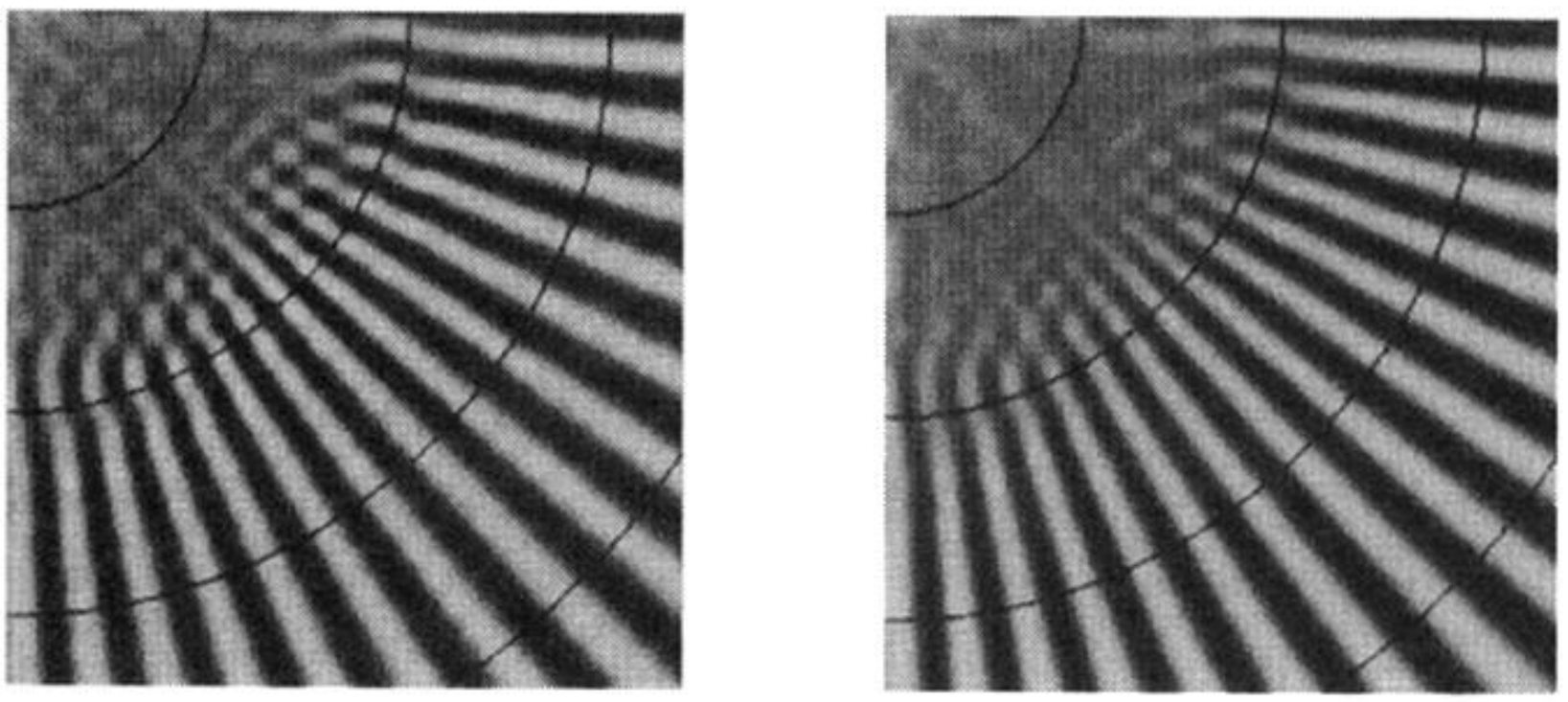


**Design 3a**





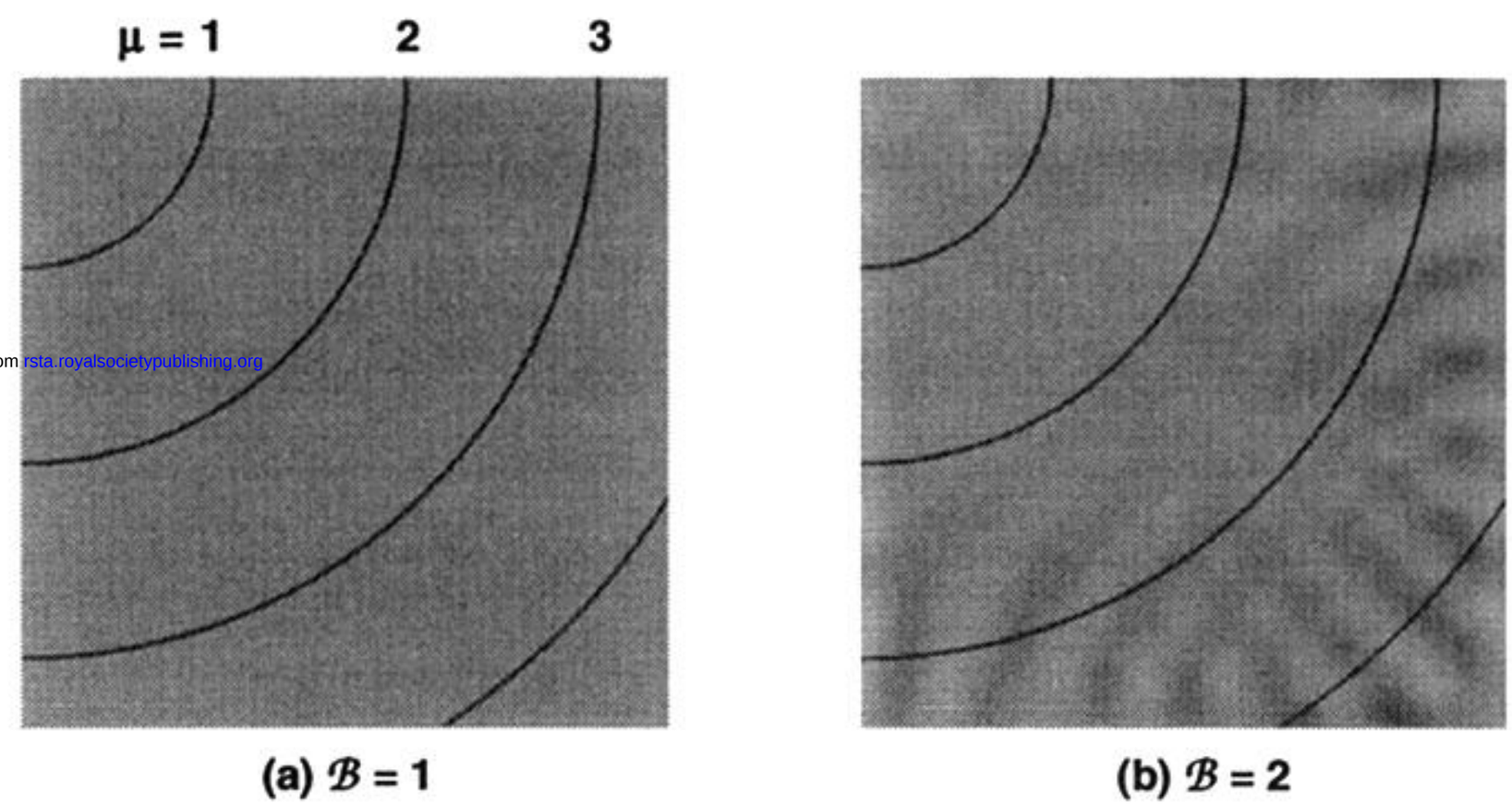




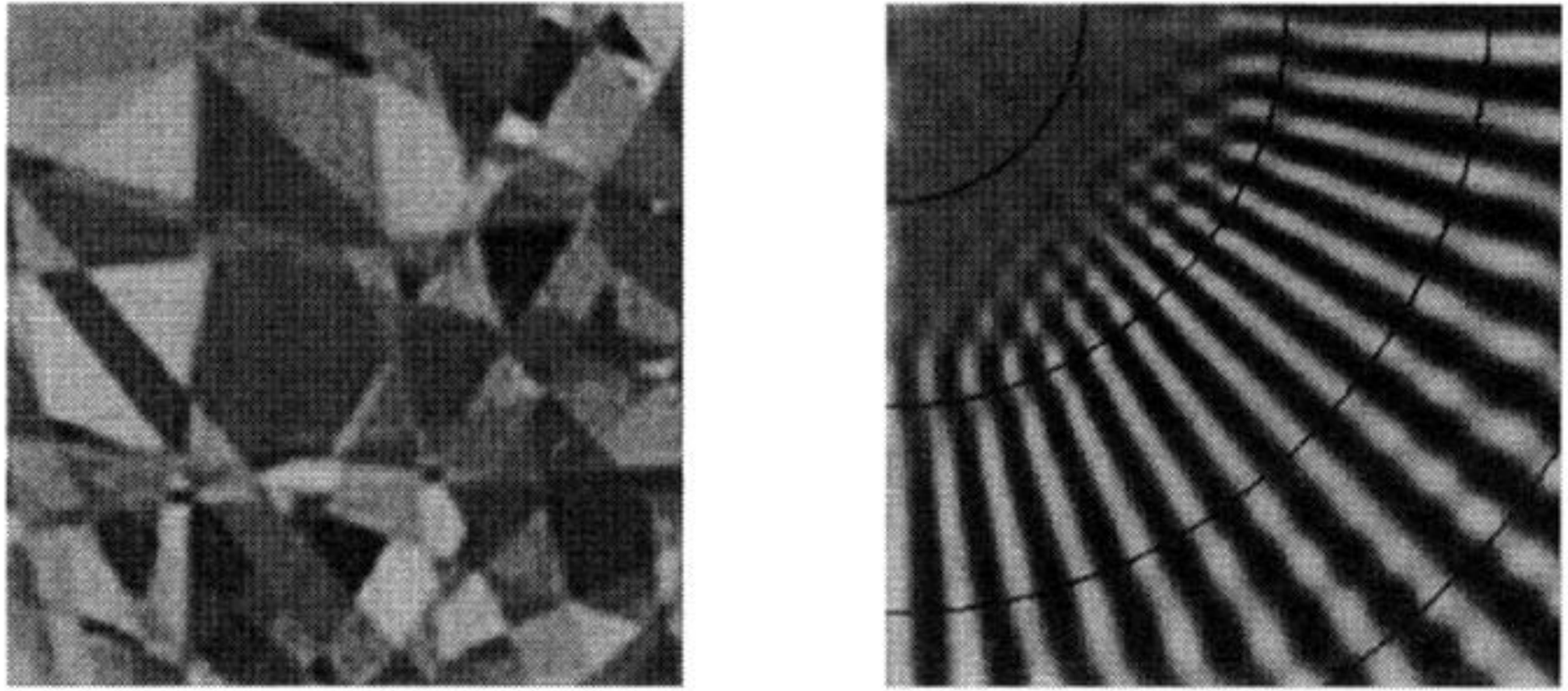
(a) Wiener restoration

(b) WIGE restoration

gure 24. Wiener and WIGE restorations for informationally optimized image gathering. The designs are specified in table 1.




igure 25. Wiener-matrix restorations for design 3 with single-level transformation. The four nalysis filters and requantization map produce $\mathcal{H}_{\mathcal{B}} = 3.8$ bits at $\mathcal{E}_{\mathcal{B}} = 5.1$ bits. (The resolution edges are coarser than elsewhere.)



gure 27. Final Wiener-matrix restorations for the multiresolution image gathering and restoration characterized in figures 25 and 26.