01 Jan 1976

Detection In Image Dependent Noise

David R. Cunningham
Missouri University of Science and Technology, drc@mst.edu

Robert D. Laramore

Eamon Barrett

Recommended Citation

D. R. Cunningham et al., "Detection In Image Dependent Noise," *IEEE Transactions on Information Theory*, vol. 22, no. 5, pp. 603 - 610, Institute of Electrical and Electronics Engineers, Jan 1976.
The definitive version is available at https://doi.org/10.1109/TIT.1976.1055591

V. Scalar Parameter

Assume that $S = s =$ scalar and let

$$A = a = \text{scalar}$$

$$M(S) = \mu(s) = \text{scalar}$$

$$\nabla(\cdot) = \frac{\vartheta(\cdot)}{\vartheta s}.$$

Then, Theorem 1 simplifies to the following Corollary.

Corollary: For an arbitrary natural number p and $P(X|S)$ statistics satisfying (4) and (5) in Lemma 1, the following inequality holds:

$$E\{[\hat{s}(X) - s]^{2p} | S\} \geq \frac{[\vartheta/\vartheta s \, E\{\hat{s}(X)|s\}]^{2p}}{\left[\int dX \, P(X|s) \left[\frac{\vartheta}{\vartheta s} \log P(X|s)\right]^{2p/2p-1}\right]^{2p-1}}; \quad (49)$$

(49) is satisfied with equality if and only if

$$\frac{\vartheta}{\vartheta s} \log p(X/s) = C(s)[\hat{s}(X) - s]^{2p-1} \quad (50)$$

where $C(s)$ is an arbitrary function.

Remarks

a) The most general family of distributions that satisfy the bound (49) is obtained from (50) and is given by

$$P(X|s) = R(X) \cdot \exp\{\int ds \, C(s)[\hat{s}(X) - s]^{2p-1}\}. \quad (51)$$

If $C(s)$ and $R(X)$ are chosen to be constant, the following subfamily of distributions satisfying (49) with equality is obtained:

$$P(X|s) = R \exp\{-C_0[\hat{s}(X) - s]^{2p}\} \quad (52)$$

where C_0 is a positive constant and R is the normalizing one. The distributions in (52) are of exponential nature.

b) The result in (48) for the model

$$S = F(s) + N$$

described in the last paragraph of Section IV takes the form

$$E\{[\hat{s}(X) - s]^{2p} | s\} \geq \frac{\pi(2p-1)/2 [(\vartheta/\vartheta s) E\{\hat{s}(X)|s\}]^{2p}}{\|W(\vartheta/\vartheta s)F(s)\|_2^{2p} 2^p [\Gamma(\tfrac{1}{2}(2p-1))]^{2p-1}}. \quad (53)$$

For $p = 1$, the well known Cramér–Rao expression follows from (53).

VI. Conclusions

The main objective of this correspondence has been the presentation and analysis of a generalized evaluation criterion in estimation. As a case in which the mean-square criterion is not adequate, the nonlinear modulation problem should be mentioned. Indeed, if s is a scalar modulating parameter and $F(s)$ is the modulated m-dimensional signal, an evaluation criterion for the estimate $\hat{s}$ that penalizes heavily the large deviations between transmitted signal s and $\hat{s}$ will help the designer to forsee *highly folding* effects and to avoid them in advance. One such evaluation criterion is that of minimizing $E\{[\hat{s} - s]^{2p} | s\}$, $p > 1$.

Finally, we want to emphasize that evaluation criteria vary as much as real problems. Persistence in using the mean-square error criterion in every problem is unjustified and, in many cases, misleading.

References

[1] H. Cramér, "A contribution to the theory of statistical estimation," *Skandinavisk Actuarientidskrift,* vol. 29, pp. 85–94, 1946.

[2] J. L. Hodges, Jr., and E. L. Lehman, "Some applications of the Cramér–Rao inequality," *Proc. Second Berkeley Symp. Math. Stat. Prob.*, pp. 13–22, 1951.

[3] C. R. Rao, "Information and accuracy attainable in the estimation of statistical parameters," *Bull. Calcutta Math. Soc.*, vol. 37, pp. 81–91, 1945.

[4] E. W. Barankin, "Locally best unbiased estimates," *Ann. Math. Stat.*, vol. 29, pp. 447–501, 1949.

[5] D. Slepian, "Estimation of signal parameters in the presence of noise," *IRE Trans. Inform. Theory,* vol. IT-3, pp. 68–89, Mar. 1954.

[6] J. Ziv and M. Zakai, "Some lower bounds on signal parameter estimation," *IEEE Trans. Inform. Theory,* vol. IT-15, pp. 386–391, May 1969.

[7] R. S. McAulay and E. M. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Trans. Inform. Theory,* vol. IT-17, pp. 669–676, Nov. 1971.

[8] D. C. Rife, M. Goldstein, and R. R. Boorstyn, "A unification of Cramér–Rao type bounds," *IEEE Trans. Inform. Theory,* vol. IT-21, pp. 330–332, May 1975.

[9] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression.* Englewood Cliffs, NJ: Prentice-Hall, 1971.

[10] M. Tasto and P. A. Wintz, "Picture bandwidth compression by adaptive block quantization," Tech. Rep. TR-EE 70-14, Sch. Elec. Eng., Purdue Univ., Lafayette, IN, July 1970.

[11] B. G. Haskell, "The computation and bounding of rate-distortion functions," *IEEE Trans. Inform. Theory,* vol. IT-15, pp. 525–531, Sept. 1969.

[12] M. Tasto and P. A. Wintz, "A bound on the rate distortion function and application to images," *IEEE Trans. Inform. Theory,* vol. IT-18, pp. 150–160, Jan. 1972.

[13] D. S. Sakrison, "The rate distortion function for a class of sources," *Inform. Contr.*, vol. 15, pp. 165–195, Jan. 1969.

[14] D. L. Wenhoff, R. M. Gray, and L. D. Davisson, "Fixed-rate universal source block coding with a fidelity criterion," submitted to the *IEEE Trans. Inform. Theory.*

[15] R. M. Gray, D. L. Neuhoff, and D. S. Ornstein, "Nonblock source coding with a fidelity criterion," submitted to *Ann. Prob.*

[16] O. Bryngdahl, "Visual transfer characteristics from Mach band measurements," *Kybernetik,* vol. 2, pp. 71–77, 1964.

[17] J. A. Whiteside and M. L. Davidson, "Symmetrical appearance of bright and dark Mach bands from an exponential illumination gradient," *J. Opt. Soc. Am.*, vol. 61, pp. 530–536, July 1971.

[18] T. G. Stockham, Jr., "Image processing in the context of a visual model," *Proc. IEEE*, vol. 60, pp. 828–842, July 1972.

[19] J. L. Mannos and D. S. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," *IEEE Trans. Inform. Theory,* vol. IT-20, pp. 525–536, July 1974.

[20] D. Kazakos, "New lower bounds to the Bayes probability of error," to be published.

[21] P. Papantoni-Kazakos, "p-efficient estimators," Tech. Rep. #7512, Elec. Eng. Dep., Rice Univ., Houston, TX, Aug. 1975.

Detection in Image Dependent Noise

DAVID R. CUNNINGHAM, ROBERT D. LARAMORE, AND EAMON BARRETT

***Abstract*—The detection of two-dimensional optical signals which have been corrupted by noise is considered. Discussion is limited to the detection of a known object in a known location. The problem is approached from the classical statistical technique of hypothesis testing. Initially the solution is formulated in very general terms. The decision rule is derived for a signal distorted by noise of an unspecified type which may include signal dependent noise. Once the decision rule is obtained, the probabilities of false alarm and detection are evaluated from *a priori* knowledge of the noise and imaging system. The general results are applied to Poisson noise, signal dependent Gaussian noise, and binomial noise.**

Manuscript received September 2, 1975; revised February 3, 1976. This work was supported in part by ESL, Inc., Sunnyvale, CA. This correspondence was presented at the National Telecommunications Conference, New Orleans, LA, December 1–3, 1975.

D. R. Cunningham is with the Electrical Engineering Department, University of Missouri-Rolla, Rolla, MO 65401.

R. D. Laramore is with the Missouri Engineering Company and Associates, Rolla, MO 65401.

E. Barrett is with ESL, Inc., Sunnyvale, CA 94086.

I. Introduction

Signal dependent noise can occur in many signal processing problems, in particular the two-dimensional image processing problem treated in this correspondence. The application of the principles of communication science to optics is not a recent development. Optical imaging has been examined in the light of information theory [1], [2], and the modern theory of image formation is clearly expressed in terms of Fourier analysis and linear systems theory [3]–[5].

Several authors have evaluated photographic and electronic imaging on an absolute scale [6]–[12]. Signal dependent models for the grain noise in photographic film have been suggested by Walkup and Choens [13] and others. The material available on restoration of degraded optical signals is vast and is not reviewed here. Harris [14] treated the detection problem under the assumption of additive Gaussian noise, while Helstrom solved the problem for the case of signals in photon noise detected by an ideal receptor [15] and for a random field formulation of the problem [16].

A set of operating characteristics applicable to detection in a broad class of noises is developed in this correspondence.

II. The Decision Rule

The problem of detecting a known object in an image degraded by arbitrary noise and sampling is now considered. The object location, size, and intensity in the image plane, as well as the background, are assumed known *a priori*. A likelihood ratio test is used as suggested by the Neyman–Pearson criterion [17].

Let $R_k(x,y)$ be the available signal at the point (x,y) in the image plane. The signal $R(x,y)$ obviously depends upon whether or not the object in question is present; we write

$$H_0: R(x,y) = R_0(x,y) \tag{1}$$

and

$$H_1: R(x,y) = R_1(x,y) \tag{2}$$

where H_0 and H_1 are the hypotheses that the signal is absent and present, respectively. Assume that the image is sampled to form an $M \times N$ array of data samples using a sampling function $g_{ij}(x,y)$ to form the sample R_{ij}. Then

$$R_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R(x,y) g_{ij}(x,y)\, dx\, dy. \tag{3}$$

R_{ij} will be a random variable with probability density function

$$p_{r_{ij}|H_k}(R_{ij}|H_k). \tag{4}$$

The likelihood ratio test is

$$\Lambda(R) = \frac{p_{\boldsymbol{r}|H_1}(\boldsymbol{R}|H_1)}{p_{\boldsymbol{r}|H_0}(\boldsymbol{R}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{5}$$

where η is a threshold value determined by the maximum acceptable false alarm probability and $\boldsymbol{R}$ is the sample vector.

Although adjacent picture elements are highly correlated in most images, the elements can be considered conditionally independent, given H_k, in many systems. Consider the following two cases.

If the image is sensed by an ideal photon counter, the samples are degraded by Poisson noise with mean specified by H_k. Thus the samples are conditionally independent, given H_k.

Film used in aerial photography was tested for correlation between picture elements. Microdensitometer measurements using a nominal spot diameter of two micrometers indicated conditional correlation coefficients, given H_k, of about five tenths if the measurements are separated by one spot diameter, but only about one tenth if the separation is increased to two spot diameters. The same relative relation also holds for larger spot diameters, suggesting that the conditional correlation coefficient is caused by the overlapping of the spot skirts. Measurements also indicate that the film noise is approximately Gaussian with mean and variance determined by H_k. Thus, for photographic film, if the sampling functions $g_{ij}(x,y)$ are nonzero, or of significant value, only over disjoint areas in the x,y-plane, the individual samples can reasonably be assumed conditionally independent.

Assuming the individual samples are conditionally independent, the decision rule is

$$\Lambda(R) = \frac{\prod_{j=1}^{N} \prod_{i=1}^{M} p_{r_{ij}|H_1}(R_{ij}|H_1)}{\prod_{j=1}^{N} \prod_{i=1}^{M} p_{r_{ij}|H_0}(R_{ij}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{6}$$

Taking the natural logarithm of the likelihood ratio defines the random variable

$$L(R) = \sum_{j=1}^{N} \sum_{i=1}^{M} \ln \left[p_{r_{ij}|H_1}(R_{ij}|H_1)\right] - \sum_{j=1}^{N} \sum_{i=1}^{M} \ln \left[p_{r_{ij}|H_0}(R_{ij}|H_0)\right]. \tag{7}$$

Now the decision rule may be stated as

$$L(R) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma = \ln \eta. \tag{8}$$

The probabilities of false alarm and detection are given by

$$P_F = \int_{\gamma}^{\infty} p_{l|H_0}(L|H_0)\, dL, \tag{9}$$

and

$$P_D = \int_{\gamma}^{\infty} p_{l|H_1}(L|H_1)\, dL. \tag{10}$$

Notice that $L(R)$ is formed by the addition and subtraction of conditionally independent random variables. If the object is large enough to require a reasonably large number of samples, let us say 50, in the image plane, the central limit theorem implies that $L(R)$ will be approximately normally distributed. If the conditional probability density functions in (9) and (10) are assumed to be normal, it is necessary only to determine the conditional means and variances of $L(R)$. These are designated by μ_k and σ_k^2, respectively. The relation of P_F and P_D to the conditional density functions of L can be clearly illustrated by sketching $p_{l|H_0}(L|H_0)$ and $p_{l|H_1}(L|H_1)$. However, it is more useful to standardize one of the curves by making the following change of variables:

$$U = \frac{L - \mu_0}{\sigma_0}. \tag{11}$$

Also, let

$$\mu = \frac{\mu_1 - \mu_0}{\sigma_0}, \tag{12}$$

$$\sigma = \frac{\sigma_1}{\sigma_0}, \tag{13}$$

and

$$\gamma' = \frac{\gamma - \mu_0}{\sigma_0}. \tag{14}$$

Then

$$P_F = \operatorname{erfc}(\gamma'), \tag{15}$$

or

$$\gamma' = \operatorname{erfc}^{-1}(P_F), \tag{16}$$

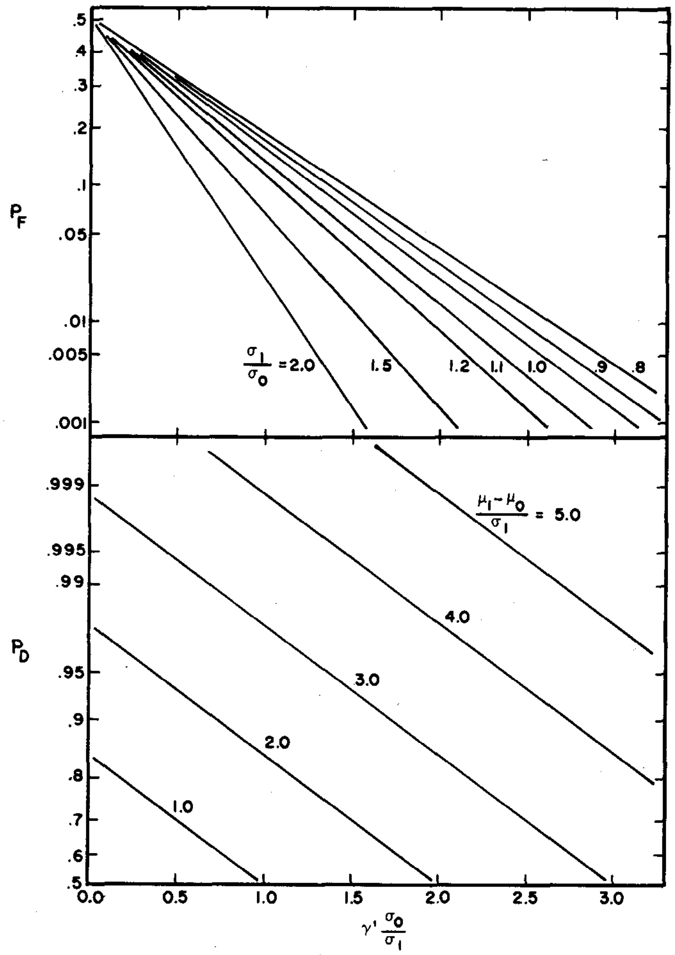

Fig. 1. Receiver operating characteristics.

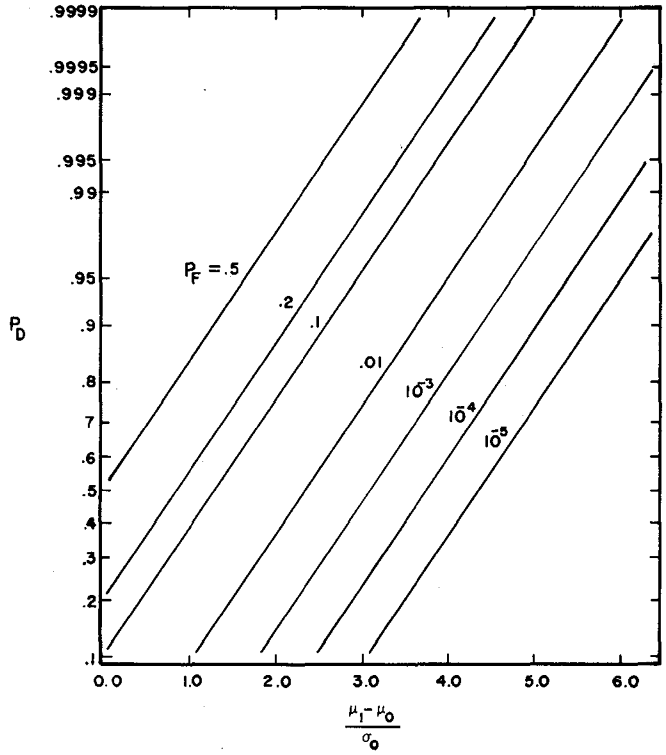

Fig. 2. Detection at low contrast, $\sigma_1/\sigma_0 = 1$.

where

$$\operatorname{erfc}(u) = \frac{1}{\sqrt{2\pi}} \int_u^{\infty} e^{-(x^2/2)}\, dx. \tag{17}$$

Therefore, if a maximum allowable p_F is specified, γ' is fixed and

$$P_D = \operatorname{erfc}\left(\frac{\gamma' - u}{\sigma}\right) = \operatorname{erfc}\left(\frac{\sigma_0}{\sigma_1}\gamma' - \frac{\mu_1 - \mu_0}{\sigma_1}\right). \tag{18}$$

A decision rule based on U which is equivalent to that based on L of (8) is

$$U \underset{H_0}{\overset{H_1}{\gtrless}} \gamma'. \tag{19}$$

Thus, P_D has been expressed as a function of γ' (or P_F) and the two parameters σ_1/σ_0 and $(\mu_1 - \mu_0)/\sigma_0$. These equations are presented in graphical form in Fig. 1. It should be stressed that these results are quite general. Other than the requirement that the samples R_{ij} be conditionally independent and that the number of samples be reasonably large, no restrictions have been placed on the noise model used or the sampling technique. If the values μ_k and σ_k^2 can be found, the curves of Fig. 1 and the equations on which they are based may be used.

A. Example

Assume that $\sigma_1/\sigma_0 = 1.2$ and $(\mu_1 - \mu_0)/\sigma_1 = 3.0$. If the maximum allowable $P_F = 0.1$ (i.e., on 10 percent of tests made with the object absent, the detector will indicate it is present), the problem proceeds as follows.

From the intersection of the $P_F = 0.1$ line and the curve corresponding to $\sigma_1/\sigma_0 = 1.2$ in Fig. 1, draw a vertical line intersecting the $\gamma'\sigma_0/\sigma_1$ axis and the curve for $(\mu_1 - \mu_0)/\sigma_1 = 3.0$ on the P_D graph. The values $\gamma'\sigma_0/\sigma_1 = 1.06$ and $P_D = 0.973$ are found. Solving for γ', we find that the decision rule is

$$U \underset{H_0}{\overset{H_1}{\gtrless}} 1.27. \tag{20}$$

If it is required that the maximum $P_F = 0.01$, $\gamma'\sigma_0/\sigma_1$ is found to be 1.93 and $P_D = 0.86$. Thus the more stringent requirement on P_F reduces significantly the probability of detecting the object.

B. Low Contrast Case

Detection when the contrast between object and background is low is of particular interest. Low contrast implies that the signals have a small variation about some larger value. This implies that the signal dependent noise will be nearly the same regardless of which hypothesis is true. For this case, the ratio σ_1/σ_0 is approximately one. Under this constraint,

$$P_F = \operatorname{erfc}(\gamma'), \tag{21}$$

and

$$P_D = \operatorname{erfc}\left(\gamma' - \frac{\mu_1 - \mu_0}{\sigma_0}\right). \tag{22}$$

P_D is plotted as a function of the parameter $(\mu_1 - \mu_0)/\sigma_1$ in Fig. 2.

III. Noise Models

The general results of the preceding section will now be applied to detection problems using different models for the noise content of the received signal $R(x,y)$. In each case, the conditional means μ_k and variances σ_k^2 necessary to use the curves are calculated.

A. Poisson Noise

An ideal diode array is basically a photon counter. Photon emission is by its nature a discrete random process in which the

noise level depends on the level of intensity [8] and is often modeled by a Poisson process.

The expected value $S_k(x,y)$ of the received signal is now defined as the photon flux per unit area per second. The output R_{ij} of the diode which senses area A_{ij} in the image plane will be a Poisson random variable with mean

$$H_k: m_{ijk} = \beta_{ij} T_0 \int\int_{A_{ij}} S_k(x,y)\, dx\, dy, \tag{23}$$

where β_{ij} is the quantum efficiency (the probability with which an incident photon is counted) and T_0 is the exposure time. Because R_{ij} is Poisson, its conditional probability density function is

$$p_{r_{ij}|H_k}(R_{ij}|H_k) = \frac{m_{ijk}^{R_{ij}}}{R_{ij}!} e^{-m_{ijk}}, \qquad R_{ij} = 0,1,2,\cdots. \tag{24}$$

From (8), the decision rule becomes

$$\sum_{j=1}^{N}\sum_{i=1}^{M} R_{ij} \ln\left(\frac{m_{ij1}}{m_{ij0}}\right) - \sum_{j=1}^{N}\sum_{i=1}^{M} (m_{ij1} - m_{ij0}) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma. \tag{25}$$

If $L(R)$ is assumed to be normally distributed, its conditional mean and variance are

$$\mu_k = \sum_{j=1}^{N}\sum_{i=1}^{M} m_{ijk} \ln\left(\frac{m_{ij1}}{m_{ij0}}\right) - \sum_{j=1}^{N}\sum_{i=1}^{M} (m_{ij1} - m_{ij0}), \tag{26}$$

and

$$\sigma_k^2 = \sum_{j=1}^{N}\sum_{i=1}^{M} m_{ijk}\left[\ln\left(\frac{m_{ij1}}{m_{ij0}}\right)\right]^2. \tag{27}$$

These values of μ_k and σ_k^2 may be used with the curves of Fig. 1.

The upper bound on diode array performance for the case of Poisson noise can be found by examining μ_k and σ_k^2 when the diodes have unity quantum efficiency, are contiguous, and the diode area shrinks to zero as the number of diodes grows without bound to cover the image area I. For these conditions,

$$\mu_k = T_0 \int\int_I S_k(x,y) \ln\left[\frac{S_1(x,y)}{S_0(x,y)}\right] dx\, dy - T_0 \int\int_I [S_1(x,y) - S_0(x,y)]\, dx\, dy \tag{28}$$

and

$$\sigma_k^2 = T_0 \int\int_I S_k(x,y)\left[\ln\left(\frac{S_1(x,y)}{S_0(x,y)}\right)\right]^2 dx\, dy. \tag{29}$$

This is the same result obtained by Helstrom [15]. These values of μ_k and σ_k^2 can be used with the graphs in Fig. 1, and the results compared with those obtained for a discrete array. The degradation in detectability of a specified object caused by sampling and diode quantum efficiency can thus be determined.

B. Signal Dependent Gaussian Noise

One possible noise model is that the signal $R(x,y)$ will be normally distributed about the expected value $S_k(s,y)$ with the variance deterministically related to the signal [13].

The above model may be formulated as follows:

$$H_k: R(x,y) = S_k(x,y) + Q_k(x,y)N(x,y), \tag{30}$$

where $N(x,y)$ is a zero-mean white Gaussian noise with autocorrelation $\delta(x,y)$, which is independent of $S_k(x,y)$, and where $Q_k(x,y)$ is a deterministic function of $S_k(x,y)$ and modulates the noise. An individual sample R_{ij} is assumed to be the average intensity $R(x,y)$, weighted by the sampling function $g_{ij}(x,y)$ and normalized by the volume under the sampling function curve:

$$R_{ij} = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} R(x,y) g_{ij}(x,y)\, dx\, dy}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_{ij}(x,y)\, dx\, dy}. \tag{31}$$

For convenience, define

$$A = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_{ij}(x,y)\, dx\, dy.$$

The random variables R_{ij} will be normal with means

$$m_{ijk} = \frac{1}{A}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S_k(x,y) g_{ij}(x,y)\, dx\, dy \tag{32}$$

and variances

$$\sigma_{ijk}^2 = \frac{1}{A^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} Q_k^2(x,y) g_{ij}^2(x,y)\, dx\, dy. \tag{33}$$

After simplification, the decision rule reduces to

$$L(R) = \sum_{j=1}^{N}\sum_{i=1}^{M} [a_{ij}R_{ij}^2 + b_{ij}R_{ij} + c_{ij}] \underset{H_0}{\overset{H_1}{\gtrless}} \gamma, \tag{34}$$

where

$$a_{ij} = \frac{1}{2\sigma_{ij0}{}^2} - \frac{1}{2\sigma_{ij1}{}^2}, \tag{35}$$

$$b_{ij} = \frac{m_{ij1}}{\sigma_{ij1}{}^2} - \frac{m_{ij0}}{\sigma_{ij0}{}^2}, \tag{36}$$

and

$$c_{ij} = \frac{m_{ij0}{}^2}{2\sigma_{ij0}{}^2} - \frac{m_{ij1}{}^2}{2\sigma_{ij1}{}^2} + \ln\left(\frac{\sigma_{ij0}}{\sigma_{ij1}}\right). \tag{37}$$

As before, L will be assumed normal by the central limit theorem and only the conditional means and variances of L are needed to use the curves. The conditional mean and variance are found to be

$$\mu_k = \sum_{j=1}^{N}\sum_{i=1}^{M} [a_{ij}(\sigma_{ijk}{}^2 + m_{ijk}{}^2) + b_{ij}m_{ijk} + c_{ij}], \tag{38}$$

and

$$\sigma_k^2 = \sum_{j=1}^{N}\sum_{i=1}^{M} [a_{ij}^2(4m_{ijk}{}^2\sigma_{ijk}{}^2 + 2\sigma_{ijk}{}^4) + (4a_{ij}b_{ij}m_{ijk} + b_{ij}^2)\sigma_{ijk}{}^2]. \tag{39}$$

A particular type of signal dependent Gaussian noise has been used to model the grain noise of photographic film. For this case, let $Q_k(x,y) = \sqrt{\alpha' S_k(x,y)}$ and assume either that $S_k(x,y)$ is essentially constant over a sample region or that $g_{ij}(x,y)$ takes on only two values, zero and a positive constant. In this situation, the following results are obtained:

$$\sigma_{ijk}{}^2 = \frac{\alpha}{A} m_{ijk}, \tag{40}$$

with α determined from α' and $g_{ij}(x,y)$. Also

$$a_{ij} = \frac{A}{2\alpha m_{ij0}} - \frac{A}{2\alpha m_{ij1}}, \tag{41}$$

$$b_{ij} = 0, \tag{42}$$

and

$$c_{ij} = \frac{A}{2\alpha}(m_{ij0} - m_{ij1}) + \frac{1}{2}\ln\left(\frac{m_{ij0}}{m_{ij1}}\right). \tag{43}$$

Therefore, for this case,

$$L(R) = \sum_{j=1}^{N}\sum_{i=1}^{M}\left[\frac{A}{2\alpha}\left(\frac{1}{m_{ij0}} - \frac{1}{m_{ij1}}\right)R_{ij}^2 + \frac{A}{2\alpha}(m_{ij0} - m_{ij1}) + \frac{1}{2}\ln\left(\frac{m_{ij0}}{m_{ij1}}\right)\right] \tag{44}$$

with mean

$$\mu_k = \sum_{j=1}^{N}\sum_{i=1}^{M}\left[\frac{A}{2\alpha}\left(\frac{1}{m_{ij0}} - \frac{1}{m_{ij1}}\right)\left(\frac{\alpha}{A} m_{ijk} + m_{ijk}{}^2\right) + \frac{A}{2\alpha}(m_{ij0} - m_{ij1}) + \frac{1}{2}\ln\left(\frac{m_{ij0}}{m_{ij1}}\right)\right] \tag{45}$$

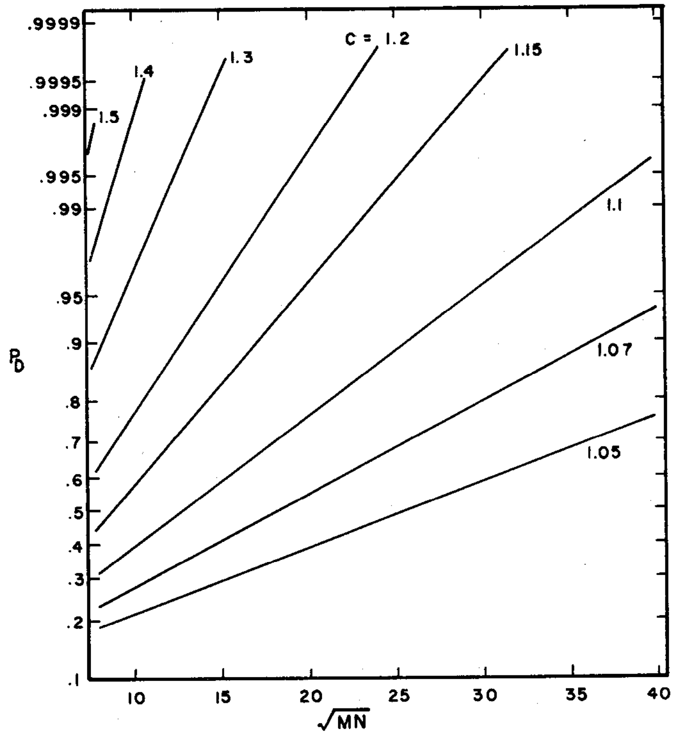

Fig. 3. Detection in binomial noise for various contrast ratios, C, and $P_F = 0.1$.

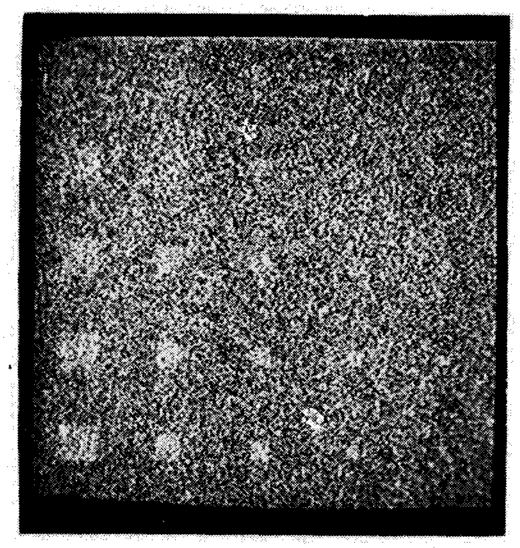

Fig. 5. Photograph illustrating visual detectability of uniform object against uniform background in binomial noise $m_{ij0} = 0.5$.

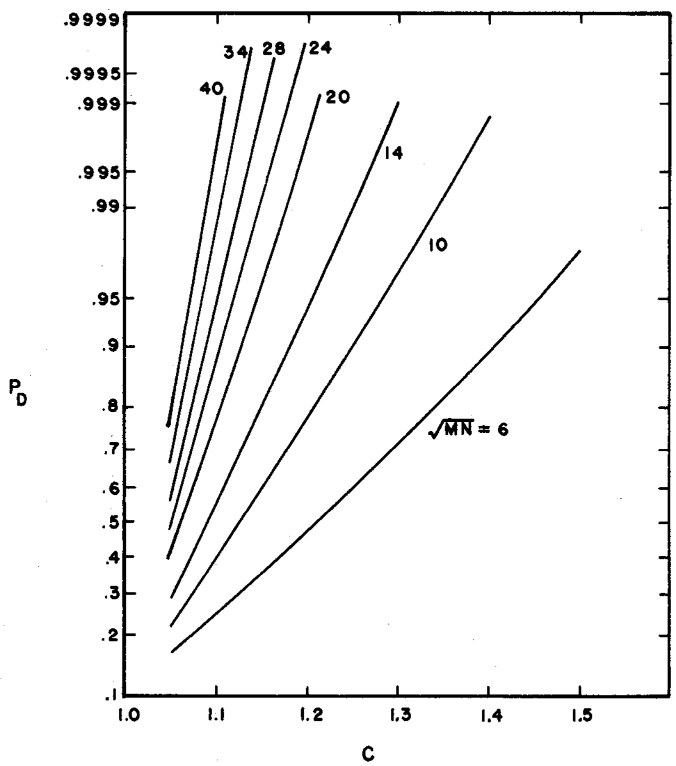

Fig. 4. Detection in binomial noise for various numbers of samples and $P_F = 0.1$.

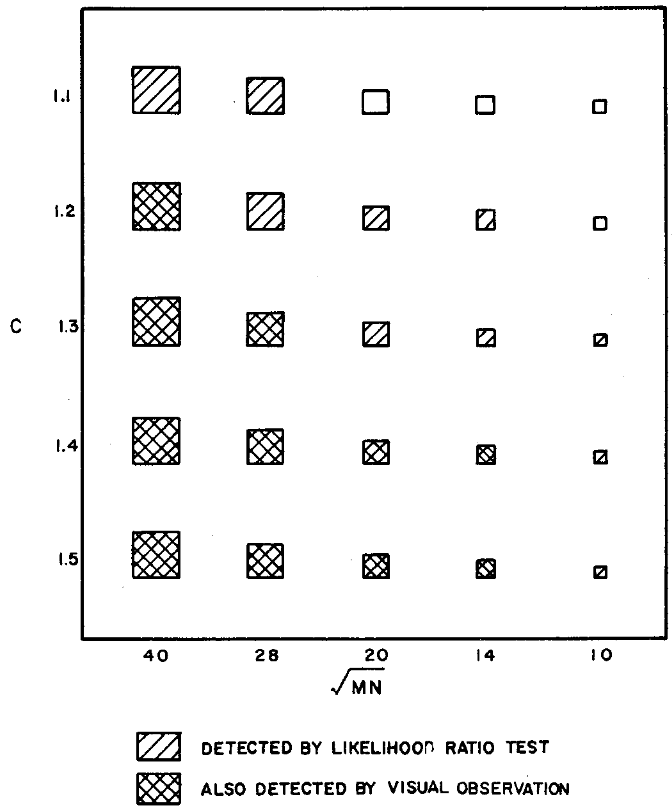

Fig. 6. Detectability of uniform object against uniform background in binomial noise $m_{ij0} = 0.5$, $C = m_{ij1}/m_{ij0}$.

and variance

$$\sigma_k^2 = \sum_{j=1}^{N}\sum_{i=1}^{M}\left[\left(\frac{1}{m_{ij0}} - \frac{1}{m_{ij1}}\right)^2\left(\frac{A}{\alpha}m_{ijk}{}^3 + \frac{1}{2}m_{ijk}{}^2\right)\right]. \quad (46)$$

Thus the parameters needed to use Fig. 1 can be calculated.

C. Binomial Noise

The form of the sampling will be altered for this case. Previously the sampling has been assumed linear. Now, suppose it is desired to detect an object in an image composed of an $M \times N$ array of samples. The individual sample values r_{ij} must be either one (white) or zero (black). This model is often used in simulated images for digital computer processing. The probability of the sample in the ij position having value R_{ij} is given by

$$\Pr\{r_{ij} = R_{ij} | H_k\} = m_{ijk}R_{ij} + (1 - m_{ijk})(1 - R_{ij}), \qquad R_{ij} = 0,1, \quad (47)$$

where

$$m_{ijk} = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_k(x,y) g_{ij}(x,y)\, dx\, dy}{\max\{S_k(x,y)\} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_{ij}(u,v)\, du\, dv}. \tag{48}$$

It is clear that m_{ijk} is the expected value of R_{ij}, given that hypothesis H_k is true. Also, m_{ijk} is the normalized value of the "true image" at the ij position weighted by the sampling function $g_{ij}(x,y)$. If the r_{ij} are conditionally independent, the log likelihood ratio test becomes

$$L(R) = \sum_{j=1}^{N} \sum_{i=1}^{M} \ln \left[\frac{m_{ij1}R_{ij} + (1 - m_{ij1})(1 - R_{ij})}{m_{ij0}R_{ij} + (1 - m_{ijp})(1 - R_{ij})} \right] \underset{H_0}{\overset{H_1}{\gtrless}} \gamma. \tag{49}$$

Again assume that enough samples have been taken to make $L(R)$ approximately normal and that the conditional probability density functions $p_{l|H_k}(L(R)|H_k)$ are determined by the means $E\{L(R)|H_k\}$ and variances var $\{L(R)|H_k\}$. It is a simple matter to find

$$\mu_k = \sum_{j=1}^{N} \sum_{i=1}^{M} \left\{ \ln \left(\frac{1 - m_{ij1}}{1 - m_{ij0}} \right) - m_{ijk} \ln \left[\frac{m_{ij0}(1 - m_{ij1})}{m_{ij1}(1 - m_{ij0})} \right] \right\} \tag{50}$$

and

$$\sigma_k^2 = \sum_{j=1}^{N} \sum_{i=1}^{M} m_{ijk}(1 - m_{ijk}) \left\{ \ln \left[\frac{m_{ij0}(1 - m_{ij1})}{m_{ij1}(1 - m_{ij0})} \right] \right\}^2. \tag{51}$$

D. *An Example*

As a specific example, the detection of an area of known size, location, and contrast in an area of uniform brightness was considered. With P_F held constant at 0.1 and $m_{ij0} = 0.5$, the probability of detection was found for squares with various numbers of elements, MN, and various values of $C = m_{ij1}/m_{ij0}$. The results are shown in Figs. 3 and 4. A display consisting of five sizes of squares and five values of C was generated on a computer graphics terminal and is shown in Fig. 5. When $P_D \geq 0.9$ is considered sufficient for likelihood ratio detection, Fig. 6 compares the performance of the likelihood ratio detector against visual detection on the unrestored image. Since on the average no restoration scheme can equal the likelihood ratio detector, all of the test squares are not restorable.

IV. Detection of Degraded Images

Consider the case where the ideal image predicted by geometric optics, $S_k'(x,y)$, is degraded by a point spread function $h(x,y;u,v)$. In the absence of noise, the received signal is, given that hypothesis H_k is true,

$$H_k\colon S_k(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S_k'(u,v) h(x,y;u,v)\, du\, dv. \tag{52}$$

In the presence of noise, the received signal is

$$H_k\colon R(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_k'(u,v) h(x,y;u,v)\, du\, dv \tag{53}$$

where $R_k'(x,y)$ would be received if the point spread function were an impulse.

As before, assume the degraded image is sampled by a set of sampling functions $\{g_{ij}(x,y)\}$ to form an $M \times N$ array of samples $\{R_{ij}\}$. Thus the problem is simply to detect the degraded image $S(x,y)$ rather than $S'(x,y)$. The detection procedure is depicted by Fig. 7(b) while in the absense of degradation the procedure would be as in Fig. 7(a).

If the point spread function is unknown, the likelihood ratio test must be based on the undegraded image, and the detection procedure is depicted by Fig. 7(c). In either case, the detection of an object in a degraded image will be more difficult than the undegraded case, and the resulting decision rule will yield poorer performance.

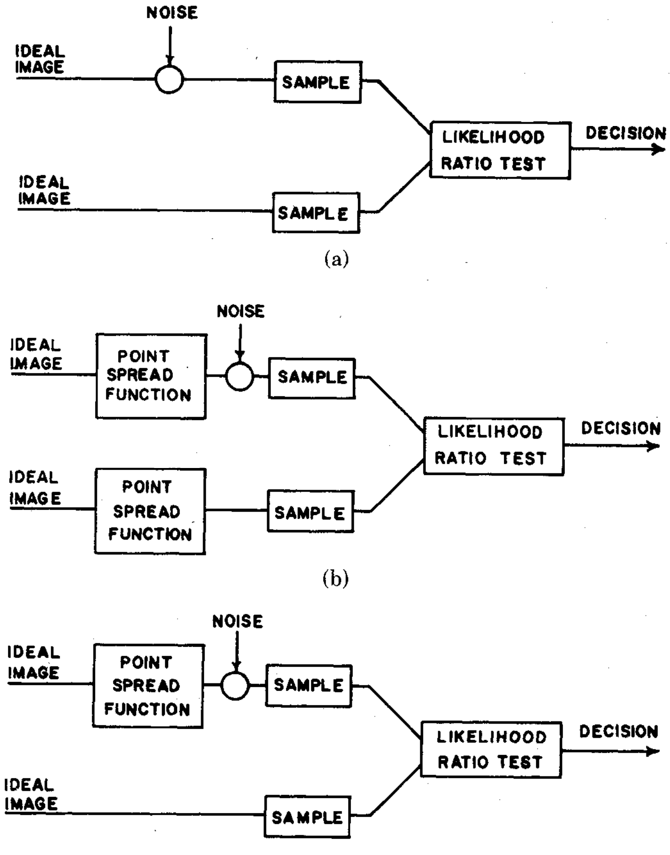

Fig. 7. Various detection schemes.

An Example

A specific example of point spread function, object size, contrast, etc., was examined to provide insight into the effect of system degradation on detectability. The detection of a uniform object against a uniform background in the presence of Poisson noise was selected. Although this example was worked out in only one dimension, the two-dimensional problem is no more difficult conceptually.

Two basic assumptions were made concerning the system. The imaging system was assumed to be diffraction-limited with the area under the point spread function independent of the width, and an ideal diode array was assumed with the diode spacing sufficiently small in relation to object size to find μ_k and σ_k^2 by means of the equations derived for the limiting case ($T \to 0$, $M \to \infty$, $N \to \infty$, $\beta = 1$).

Define R to be the ratio of the object width to the width of the main lobe of the point spread function, and the contrast to be $[S_1'(x) - s_0'(x)]/S_0'(x)$. The degradation of the image by the point spread function is shown in Fig. 8, and Fig. 9 shows the degradation in detectability caused by diffraction. For curve a, the detection rule is based on the degraded image as in Fig. 7(b); for curve b, the detection rule is based on the undegraded image when in fact the image is actually degraded as in Fig. 7(c). The difference between curve a and b is a measure of the cost of ignorance.

V. Conclusions

It is clear that statistical techniques can be applied to the detection of two-dimensional signals in signal dependent noise. The

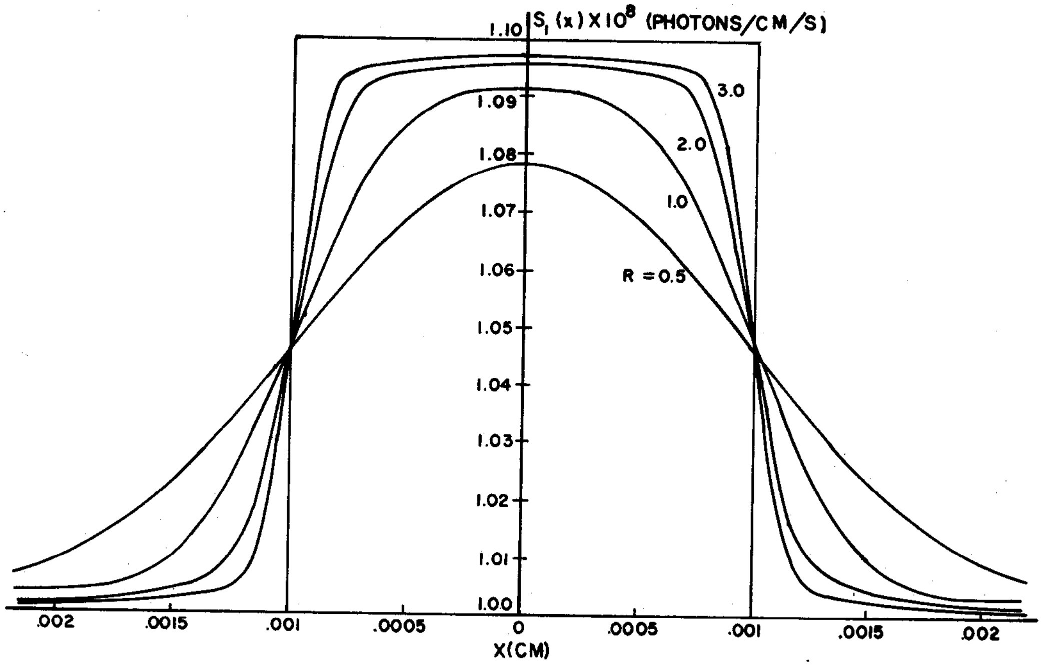

Fig. 8. Degraded images resulting from point spread function.

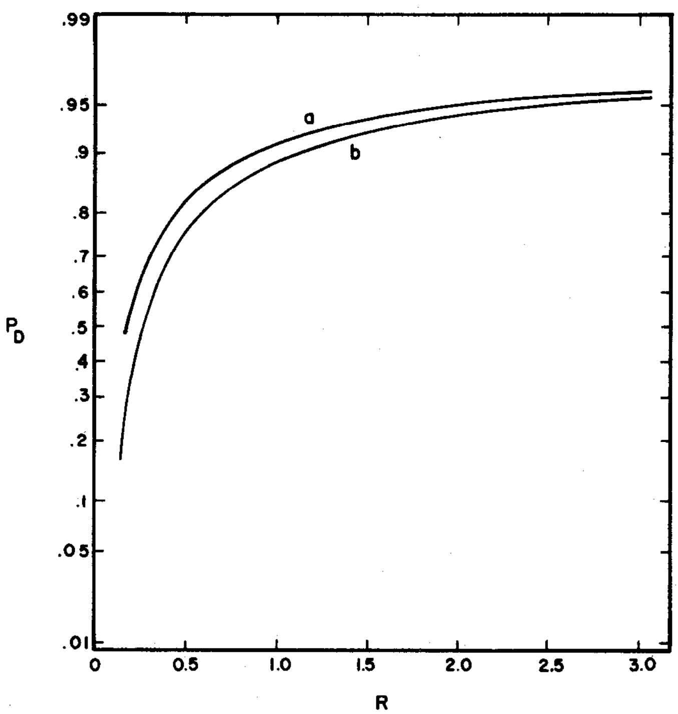

Fig. 9. Performance of detection schemes shown in Fig. 7(b) and (c). Curve a shows detectability when the decision rule is based on the degraded image. Decision rule for curve b is based on undegraded image. Background intensity = 10^8 photons/cm/s, contrast = 0.1, width = 0.002 cm, P_F = 0.01.

method appears to be useful in many applications. The added *a priori* knowledge necessary for detection improves performance beyond visual detection of an unrestored image, and the performance of no restoration scheme can be expected to equal that of the likelihood ratio test. Extensive use of such detection techniques may be limited by the accurate knowledge required of the signal and image sensing system. The method, however, may be used quite effectively to evaluate the relative performance of various imaging systems.

Throughout this correspondence, it has been assumed that sufficient samples of the image have been taken for L to be considered normal. Therefore, caution should be observed in applying these results when the available samples are few and when dealing with extremely small probabilities of false alarm.

REFERENCES

[1] A. E. Saunders, "On the application of information theory to photography," *The Journal of Photographic Science,* vol. 21, pp. 257–261, 1973.

[2] N. J. Bershad, "Resolution, optical-channel capacity and information theory," *Journal of the Optical Society of America,* vol. 59, pp. 157–175, Feb. 1969.

[3] J. W. Goodman, *Introduction to Fourier Optics.* San Francisco: McGraw-Hill, 1968.

[4] R. J. Becherer and J. D. Geller, "Optimum shaded apertures for reducing photographic grain noise," *Proc. of SPIE Seminar-in-Depth,* Image Information Recovery, vol. 16, pp. 89–96, 1969.

[5] B. R. Frieden and J. J. Burke, "Restoring with maximum entropy II: Superresolution of photographs of diffraction-blurred impulses," *Journal of the Optical Society of America,* vol. 62, pp. 1202–1210, Oct. 1972.

[6] O. H. Schade, Sr., "An evaluation of photographic image quality and resolving power," *Journal of the Society of Motion Picture and Television Engineers,* vol. 73, pp. 81–119, Feb. 1964.

[7] A. G. Millikan, "Image detection at the telescope," *American Scientist,* vol. 62, pp. 324–333, May–June 1974.

[8] A. Rose, *Vision: Human and Electronic.* New York: Plenum, 1973.

[9] R. Shaw, "Multilevel grains and the ideal photographic detector," *Photographic Science and Engineering,* vol. 16, pp. 192–200, May–June 1972.

[10] H. J. Zweig, "The relation of quantum efficiency to energy- and contrast-detectivity for photographic materials," *Photographic Science and Engineering,* vol. 5, pp. 142–148, May–June 1961.

[11] S. Karp and J. R. Clark, "Photon counting: A problem in classical noise theory," *IEEE Trans. Inform. Theory,* vol. IT-16, pp. 672–680, Nov. 1970.

[12] C. W. Helstrom and L. Wang, "Optimum detection of an optical image on a photoelectric surface," *IEEE Trans. Aerosp. Electron. Syst.,* vol. AES-9, pp. 561–570, July 1973.

[13] J. R. Walkup and R. C. Choens, "Image processing in signal dependent noise," *Optical Engineering,* vol. 13, May–June, 1974.

[14] J. L. Harris, "Resolving power and decision theory," *Journal of the Optical Society of America,* vol. 54, pp. 606–611, May 1964.

[15] C. W. Helstrom, "The detection and resolution of optical signals," *IEEE Trans. Inform. Theory,* vol. IT-10, pp. 275–287, Oct. 1964.

[16] ——, "Detection and resolution of incoherent objects by a background-limited optical system," *Journal of the Optical Society of America*, vol. 59, pp. 164–175, Feb. 1969.
[17] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, Part I. New York: Wiley, 1968.

A Likelihood Ratio Formula for Spherically Invariant Processes

A. F. GUALTIEROTTI

Abstract—**This correspondence provides a likelihood ratio formula for equivalent spherically invariant measures on a real and separable Hilbert space.**

Spherically invariant stochastic processes are of interest in communication theory because they allow one to relax somewhat the assumption of normality, while keeping many of its most useful consequences. In [1], we considered their equivalence, here we provide a likelihood ratio formula.

An extensive study of the properties of the "marginal" distributions of spherically invariant processes can be found in [2]. These processes contain the Gaussian ones and are in fact mixtures of the latter. It then immediately follows from a result of Skorohod [3] that the likelihood ratio formula is also a mixture of Gaussian likelihoods.

A few definitions are necessary to indicate the domain of validity of the formula. We start with a real and separable Hilbert space H with inner product $(\cdot,\cdot)_H$ and norm $\|\cdot\|_H$. Denote its Borel subsets by $\mathcal{B}$. On H, consider random variables of the form $X_h(x) = (x,h)_H$ and, when the variance $V(X_h)$ exists, write $v(h)$ for $V(X_h)$. We say that a probability measure P on $\mathcal{B}$ is spherically invariant if $v(h) = v(k)$ implies that X_h and X_k have the same distribution. In a typical example, we would take $H = L_2[0,T]$, $X_h = \int_0^T h(t)\, X_t\, dt$, and X_t a zero-mean measurable process with paths in $L_2[0,T]$. In [1], we described how such measures P are characterized by couples (F,R), where F is a probability measure on $\boldsymbol{R}$, with support in $\boldsymbol{R}_+$, and possessing a finite second moment, and R is a covariance operator (that is, a nonnegative, selfadjoint operator with finite trace). Gaussian measures are obtained by taking the F concentrated at one point, and operators R are obtained typically as operators determined by covariances of stochastic processes. If P_t is the Gaussian measure determined by the covariance t^2R, then the spherically invariant measure associated with (F,R) is given by

$$P(B) = \int_0^\infty P_t(B)\, dF(t), \qquad \text{for } B \text{ in } \mathcal{B}. \tag{1}$$

Let us now state Skorohod's result [3, p. 99]. Let $(H,\mathcal{B})$ be the previously stated measurable Hilbert space and let $(X,\mathcal{X})$ be some measurable space. For x in X, let P_x and Q_x be measures on $(H,\mathcal{B})$ such that:

a) for fixed B in $\mathcal{B}$, $x \to P_x(B)$ and $x \to Q_x(B)$ are $\mathcal{X}$-measurable
b) there exists sets B_x in $\mathcal{B}$ such that: i) $P_y(B_x) = Q_y(B_x) = \delta_{x,y}$, ii) for every Y in $\mathcal{X}$, $\bigcup_{x\in Y} B_x$ is in $\mathcal{B}$, iii) $B_x \cap B_y = \phi$, $x \neq y$.

Let M and N be measures on $\mathcal{X}$. Define, for B in $\mathcal{B}$,

$$P(B) = \int_X P_x(B)\, dM(x) \qquad Q(B) = \int_X Q_x(B)\, dN(x).$$

Then $P \ll Q$, if and only if i) $M \ll N$, ii) $P_x \ll Q_x$, M almost everywhere x. Furthermore, when $P \ll Q$,

$$\frac{dP}{dQ}(h) = \frac{dP_x}{dQ_x}(h)\,\frac{dM}{dN}(x), \qquad \text{for } h \text{ in } B_x. \tag{2}$$

In order to use the likelihood ratio formula of (2) for spherically invariant measures given by $P(B)$ (or $Q(B)$) defined in (1), choose P_x (or Q_x) to be Gaussian, with mean zero and covariance operator x^2R (or x^2S). Let $X = \boldsymbol{R}_+$. In [3], it is also shown that

$$f(h) \triangleq \lim_n \frac{1}{n} \sum_{k=1}^n \frac{(h,e_k)_H^2}{a_k}$$

exists Q_x almost everywhere and has the value x^2 where e_k is the eigenvector of S with associated eigenvalue a_k. It is further well known (e.g., [3, p. 95]) that, if P_R is the Gaussian measure corresponding to R, and Q_S is the one associated with S, then $P_R \equiv Q_S$ if and only if

$$R = S^{1/2}(I + U)S^{1/2},$$

where $I + U$ invertible U selfadjoint, and Hilbert–Schmidt. If then b_k is the kth eigenvalue of U,

$$\frac{dP_R}{dQ_S}(h) = \exp(A),$$

where

$$A = -\frac{1}{2}\sum_{k,l} (U(I + U)^{-1}e_k,e_l)_H \left(\frac{(h,e_k)_H(h,e_l)_H}{(a_k a_l)^{1/2}} - \delta_{k,l}\right) + \frac{1}{2}\sum_k \left(\ln(1 + b_k) - \frac{b_k}{1 + b_k}\right).$$

Thus, when computing dP_x/dQ_x, the eigenvalues are multiplied by x^2, which in turn can be written as $f(h)$. The latter function provides the "disintegration" B_x. It may be worth noticing that $P_x \perp P_y$, $x \neq y$, and $Q_x \perp Q_y$ also, which "explains" the result!

ACKNOWLEDGMENT

I wish to thank Dr. C. R. Baker, of the University of North Carolina, Chapel Hill, as well as the reviewer, for advice on the writing of this correspondence.

REFERENCES

[1] A. F. Gualtierotti, "Some remarks on spherically invariant distributions," *J. Multivariate Anal.*, vol. 4, pp. 347–349, 1974.
[2] K. Yao, "A representation theorem and its applications to spherically invariant random processes," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 600–608, Sept. 1973.
[3] A. V. Skorohod, *Integration in Hilbert Space*. New York: Springer Verlag, 1974.

Manuscript received July 7, 1975; revised January 21, 1976.
The author is with the Département de Mathématiques, Ecole Polytechnique Fédérale, Lausanne, Switzerland.

An Upper Bound on the Mean-Square Error for Bayesian Parameter Estimators

R. M. HAWKES AND J. B. MOORE, MEMBER, IEEE

Abstract—**The estimation of unknown parameters of deterministic time-varying signals contaminated with additive Gaussian white noise is considered, and an upper bound is obtained for the mean-square estimation error. As an example, the theory is applied to the synchronous demodulation of pulse frequency-modulated (PFM) signals in Gaussian white noise.**

Manuscript received July 3, 1975; revised January 5, 1976. This work was supported in part by the Australian Research Grants Committee and Radio Research Board.
The authors are with the Department of Electrical Engineering, University of Newcastle, N.S.W., Australia.