

Image Detection Under Low-Level Illumination

Raúl E. Sequeira, *Student Member, IEEE*, John A. Gubner, *Member, IEEE*, and Bahaa E. A. Saleh, *Fellow, IEEE*

Abstract—Image detection under low-light-level conditions is treated as a hypothesis-testing problem in which the observations are modeled as a shot-noise process. Since computing the likelihood ratio for shot-noise processes is not feasible, we propose the use of a one-dimensional test statistic obtained by filtering and sampling the observations. The filter is chosen to maximize a generalized signal-to-noise ratio. The likelihood ratio for the one-dimensional test statistic is evaluated numerically by inverting the corresponding characteristic function under each hypothesis.

Index Terms—Hypothesis testing, generalized signal-to-noise ratio, shot noise, and filtered point process.

I. INTRODUCTION

WHEN images are produced under high-level illumination, they are often modeled as the sum of a “signal” image plus signal-independent Gaussian noise. However, when an image is formed under low-light-level illumination, it can be better modeled as a filtered point process, also known as shot noise, described as follows.

The process of image acquisition consists of the measurement of arriving photons in the image plane. In practical situations, it is difficult to measure the exact location of the photons since these are filtered by the finite response of the imaging device. What is measured, instead, is the superposition of the responses of the imaging system to each arriving photon. This superposition, measured at a point $x \in \mathbb{R}^2$, can be described by the random variable

$$Z(x) = \sum_{\nu} h(x - X_{\nu}) \quad (1.1)$$

where $\{X_{\nu}\}$ denotes the set of positions at which photons are detected, and h represents the impulse response or point-spread function of the imaging device. A block diagram of the imaging system is shown in Fig. 1.

We consider the following hypothesis-testing problem. Under hypothesis H_i , $i = 0, 1$, the $\{X_{\nu}\}$ are points of a two-dimensional Poisson process with nonnegative intensity $\lambda_i(x)$, $x \in \mathbb{R}^2$.

Special cases of this problem that have previously been addressed include the following.

1) The exact locations of the photons are available. This is equivalent to the point-spread function $h(x)$ being equal to a

Manuscript received February 5, 1992; revised August 11, 1992. This work was supported in part by the Air Force Office of Scientific Research under Grant AFOSR-90-0181 and in part by the University of Wisconsin—Madison Graduate School. This work was presented in part at the Optical Society of America Annual Meeting, San Jose, CA, Nov. 1991. The associate editor coordinating the review of this paper and approving it for publication was Dr. Rama Chellappa.

The authors are with the Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI 53706.

IEEE Log Number 9205217.

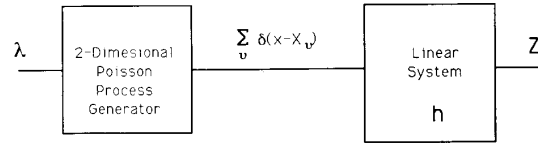


Fig. 1. Imaging system model.

Dirac delta function $\delta(x)$, in which case one has a detection problem with Poisson-process observations. When one has Poisson-process observations, the likelihood ratio (LR) is, of course, well known [9, p. 94]. In [4], a suboptimal detection scheme that could be easily implemented was considered. This led to the consideration of a correlation detector in which the Poisson-process observations were passed through a linear filter (taken to be one of the intensity functions) and sampled. This led to a hypothesis test based on a single shot-noise random variable.

2) Counts of photons in disjoint regions are available, and hence one is faced with a Poisson counting process detection problem. This case can be regarded as a filtered Poisson process with a special form of h , followed by sampling. The LR is well known for this special case also [9, p. 94]. In [10], a correlation scheme was used for classification. The Poisson counting process observations (with counts being either 0 or 1 with high probability) were cross-correlated with various reference functions. Three reference functions were considered, one of which was constructed so that the value of the correlation between that function and the observed image approximates the value of the logarithm of the LR.

Little work has been done for the more general case of filtered point-process observations due to the difficulty of computing the density and distribution functions involved [3].

II. HYPOTHESIS TESTING WITH SHOT-NOISE OBSERVATIONS

A. Preliminary Considerations

Using the mathematical model described in Section I, we would like to decide whether λ_0 or λ_1 is the true intensity of the underlying Poisson process that gives rise to our observations $\{Z(x)\}$. Clearly, a likelihood ratio test (LRT) is called for [7, p. 11]. Unfortunately, a formula for the LR of the continuum of observations $\{Z(x)\}$ does not seem to be available in the literature. In fact, even if we based our decision in a finite number of samples of $\{Z(x)\}$, say $Z(x_1), \dots, Z(x_K)$, the LR would be obtained by inverting the K -variate characteristic function of $Z(x_1), \dots, Z(x_K)$ under each hypothesis to obtain the K -variate density under each hypothesis. The quotient of these densities would yield the

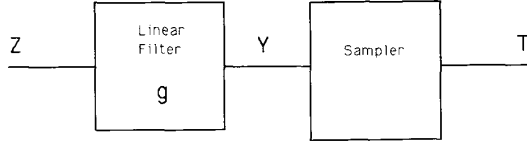


Fig. 2. Detector structure.

LR. This approach is impractical unless K is very small. It is because of these considerations that we introduce the following suboptimal approach.

B. A Suboptimal Detector

As a suboptimal detector scheme, we propose that the received image $\{Z(x)\}$ be passed through a linear filter g (to be chosen later in Section II-E) and then sampled, as shown in Fig. 2. The final discrimination of hypotheses will be based on the sample T . More precisely, let

$$Y(x) \triangleq \int g(x-u)Z(u)du \quad (2.1)$$

where $Z(\cdot)$ is given by (1.1). (Here and in the sequel, integrals are understood as being over all of \mathbb{R}^2 unless otherwise indicated. One-dimensional integrals over \mathbb{R} are indicated by $\int_{-\infty}^{\infty} \cdot$.) Observe that if we set

$$\tilde{g}(x) \triangleq \int g(x-u)h(u)du \quad (2.2)$$

then

$$Y(x) = \sum_{\nu} \tilde{g}(x - X_{\nu}). \quad (2.3)$$

Clearly, (2.3) has the same form as (1.1). In other words, when the shot-noise process $\{Z(x)\}$ is passed through the linear system g , the output $\{Y(x)\}$ is also a shot-noise process. The final processing step shown in Fig. 2 is sampling. We set

$$T \triangleq Y(0) = \sum_{\nu} \tilde{g}(-X_{\nu}) \quad (2.4)$$

and perform an LRT based on T .

C. The Likelihood Ratio for T

Let $F_i(t) \triangleq \mathcal{P}_i(T \leq t)$, $i = 0, 1$, be the cumulative probability distribution of T given that λ_i is the true intensity of the underlying Poisson process that models the location at which arriving photons strike the imaging system. Clearly, if no photons arrive, $T = 0$. The probability of this event is

$$\mathcal{P}_i(T = 0) = e^{-\Lambda_i}$$

where

$$\Lambda_i \triangleq \int \lambda_i(x) dx.$$

In our applications, we have $\Lambda_i < \infty$. We thus expect $F_i(t)$ to have the form

$$F_i(t) = \begin{cases} \int_{-\infty}^t f_i(\tau) d\tau, & t < 0 \\ e^{-\Lambda_i} + \int_{-\infty}^t f_i(\tau) d\tau, & t \geq 0 \end{cases} \quad (2.5)$$

for some nonnegative f_i satisfying

$$\int_{-\infty}^{\infty} f_i(\tau) d\tau = 1 - e^{-\Lambda_i}.$$

Remark: In general, $F_i(t)$ could have jumps at points $t \neq 0$, depending on the behavior of \tilde{g} . In our examples, this does not occur.

Assuming equally likely hypotheses, if we observe $T = t$, we perform the LRT:

$$\begin{cases} \Lambda_0 - \Lambda_1 & > 0, & t = 0 \\ & < 0, & t = 0 \\ & & H_0 \\ & & H_1 \\ f_1(t)/f_0(t) & > 1, & t \neq 0. \\ & < 1, & t \neq 0. \\ & & H_0 \end{cases} \quad (2.6)$$

The numerical calculation of f_0 and f_1 is discussed in Appendix A.

D. The Probability of Error

The probability of error incurred using the LRT (2.6) is denoted by P_e ; we can write an expression for P_e as follows. First, let D_0 denote the set of t such that we decide in favor of H_0 . Clearly, $0 \in D_0$ if and only if $\Lambda_0 < \Lambda_1$. For $t \neq 0$, $t \in D_0$ if and only if $f_0(t) < f_1(t)$. Let D_1 denote the complement of D_0 . We write $D_1 = D_0^c$. Then, under equally likely hypotheses,

$$\begin{aligned} P_e &= \frac{1}{2} \{ \mathcal{P}_1[T \in D_0] + \mathcal{P}_0[T \in D_1] \} \\ &= \frac{1}{2} \left\{ 1 + \int_{D_0} [dF_1(t) - dF_0(t)] \right\}. \end{aligned} \quad (2.7)$$

Let $I_A(t)$ denote the indicator function of a set A . In other words, $I_A(t) = 1$ if $t \in A$ and $I_A(t) = 0$ otherwise. Using (2.7) and (2.5), we can then write

$$\begin{aligned} P_e &= \frac{1}{2} \left\{ 1 + [e^{-\Lambda_1} - e^{-\Lambda_0}] I_{D_0}(0) \right. \\ &\quad \left. + \int_{D_0 \cap \{0\}^c} (f_1(t) - f_0(t)) dt \right\} \\ &= \frac{1}{2} \left\{ 1 + [e^{-\Lambda_1} - e^{-\Lambda_0}] I_{D_0}(0) \right. \\ &\quad \left. + \int_{D_0} (f_1(t) - f_0(t)) dt \right\}. \end{aligned} \quad (2.8)$$

In our applications, D_0 will turn out to be an interval or a union of disjoint intervals. Hence, the last term in (2.8) can easily be computed if we have a simple way to evaluate

$$G_i(t) \triangleq \int_{-\infty}^t f_i(\tau) d\tau. \quad (2.9)$$

We discuss this further in Appendix B.

E. Selecting the Filter g

Ideally, we would like to select g to minimize the probability of a decision error P_e . Since the dependence of P_e on g is not readily apparent, we introduce the following ad hoc criterion for selecting g . We would like to choose g to maximize the generalized “signal-to-noise ratio” (cf. [1, eq. (18)], [2, eq. (6)]):

$$\frac{(\mu_1 - \mu_0)^2}{\sigma_1^2 + \sigma_0^2} \quad (2.10)$$

where μ_i and σ_i^2 are the mean and variance of T under hypothesis $i = 0, 1$. Since Y in (2.3) is a shot-noise process, it follows from (2.4) and [5, pp. 382–383] that

$$\begin{aligned} \mu_i &\triangleq \mathbb{E}_i[T] \\ &= \int \tilde{g}(-x) \lambda_i(x) dx \end{aligned} \quad (2.11)$$

and

$$\begin{aligned} \sigma_i^2 &\triangleq \mathbb{E}_i[(T - \mu_i)^2] \\ &= \int \tilde{g}(-x)^2 \lambda_i(x) dx. \end{aligned} \quad (2.12)$$

Using (2.11) and (2.12), it follows from the Cauchy–Schwarz inequality that \tilde{g} maximizes (2.10) if and only if

$$\tilde{g}(-x) = c \frac{\lambda_1(x) - \lambda_0(x)}{\lambda_1(x) + \lambda_0(x)} \quad (2.13)$$

where c is an arbitrary constant. Using (2.2), (2.13) then becomes

$$\int g(-x - u) h(u) du = c \frac{\lambda_1(x) - \lambda_0(x)}{\lambda_1(x) + \lambda_0(x)}. \quad (2.14)$$

Unfortunately, (2.14) may not have a solution. For example, if g and h are square integrable, the left-hand side of (2.14) will be a continuous function of x , while the right-hand side need not be, as is the case in our examples in Section III-C. In order to avoid this problem, we *a priori* constrain g to be of the form

$$g(x) = \sum_{k=1}^K w_k \delta(x + x_k) \quad (2.15)$$

where K and the locations x_1, \dots, x_K are chosen in some heuristic fashion, perhaps as a uniform grid, and the weights are chosen to maximize (2.10). Observe that if (2.15) holds, then it follows from (2.1) that

$$Y(x) = \sum_{k=1}^K w_k Z(x + x_k),$$

and thus

$$T \triangleq Y(0) = \sum_{k=1}^K w_k Z(x_k),$$

i.e., the test statistic T is a weighted superposition of the measurements $Z(x_1), \dots, Z(x_K)$. Now, let $\mathbf{w} \triangleq [w_1, \dots, w_K]'$

and $\mathbf{Z} \triangleq [Z(x_1), \dots, Z(x_K)]'$. Let $\mathbf{m}_i \triangleq \mathbb{E}_i[\mathbf{Z}]$ and $\mathbf{\Gamma}_i \triangleq \mathbb{E}_i[(\mathbf{Z} - \mathbf{m}_i)(\mathbf{Z} - \mathbf{m}_i)']$. Then $T = \mathbf{w}'\mathbf{Z}$, and

$$\mu_i = \mathbf{w}'\mathbf{m}_i \quad \text{and} \quad \sigma_i^2 = \mathbf{w}'\mathbf{\Gamma}_i\mathbf{w} \quad (2.16)$$

where the k th entry of the vector \mathbf{m}_i is [cf. (2.11)]

$$\int h(x_k - x) \lambda_i(x) dx \quad (2.17)$$

and the $k\ell$ entry of the matrix $\mathbf{\Gamma}_i$ is [cf. (2.12)]

$$\int h(x_k - x) h(x_\ell - x) \lambda_i(x) dx. \quad (2.18)$$

Letting $\mathbf{m} = \mathbf{m}_1 - \mathbf{m}_0$ and $\mathbf{\Gamma} = \mathbf{\Gamma}_0 + \mathbf{\Gamma}_1$, we see that under the constraint (2.15), (2.10) becomes

$$\frac{|\mathbf{w}'\mathbf{m}|^2}{\mathbf{w}'\mathbf{\Gamma}\mathbf{w}}. \quad (2.19)$$

By the Cauchy–Schwarz inequality, \mathbf{w} maximizes (2.19) if and only if $\mathbf{\Gamma}\mathbf{w} = c\mathbf{m}$, where c is an arbitrary constant. For the numerical examples in Section III-C, we take $c = 1$, and $\mathbf{\Gamma}\mathbf{w} = \mathbf{m}$ is easily solved using the NAG routine F04ASF. The NAG routine D01FCF is used to compute (2.17) and (2.18).

Remark: One of the reviewers has suggested selecting g so that

$$\tilde{g}(-x) = \int g(-x - u) h(u) du = \ln \frac{\lambda_1(x)}{\lambda_0(x)} \quad (2.20)$$

the idea being that (2.4) would then be equal (up to an additive constant) to the logarithm of the LR of the point process itself. A related idea was used in [10]. Unfortunately, (2.20) may not have a solution for the same reasons given following (2.14).

III. HYPOTHESIS TESTING AND ERROR PERFORMANCE

In this section, we apply the preceding ideas to several examples. For comparison, we also discuss the consequences of assuming that T has a Gaussian distribution under each hypothesis; this assumption was used, under somewhat higher light levels, in [4], [10] based on the Central Limit Theorem. Under certain conditions, this approximation is adequate [6], and it avoids the burden of computing $f_i(t)$ and $G_i(t)$ by the numerical evaluation of inverse Fourier transforms. However, under the low-light-level conditions considered here, we do not expect the Gaussian approximation to work well, and this is indeed the case.

A. Likelihood Ratio Test

Following the observation $T = t$, the LRT is given by (2.6). In the numerical examples discussed below, we plot $f_0(t)$ and $f_1(t)$ (Figs. 4–6), and we see that the equation

$$\frac{f_1(t)}{f_0(t)} = 1$$

has at most one solution of interest, denoted by η . Hence, if $t \neq 0$, the LRT reduces to the single threshold test

$$\begin{aligned} &H_1 \\ &t > \eta. \\ &H_0 \end{aligned}$$

B. Gaussian Test (GT)

A simple test to write is the following. Let

$$p_i(t) \triangleq \frac{1}{\sqrt{2\pi}\sigma_i} e^{-(t-\mu_i)^2/2\sigma_i^2} \quad (3.1)$$

where μ_i and σ_i^2 are given by (2.16). In other words, we are assuming that T is normal, but with the correct mean and variance. We consider the test

$$\frac{p_1(t)}{p_0(t)} > 1. \\ H_0$$

In the examples discussed below, $\sigma_1^2 \geq \sigma_0^2$, and the decision regions for this test are easily shown to be

$$D_0 = \left\{ t \in \mathbb{R} : -\gamma - \frac{b}{a} \leq t \leq \gamma - \frac{b}{a} \right\}$$

and

$$D_1 = D_0^c$$

where $a \triangleq \sigma_1^2 - \sigma_0^2$, $b \triangleq \mu_1\sigma_0^2 - \mu_0\sigma_1^2$, and

$$\gamma \triangleq \left[\frac{1}{a} \left\{ \sigma_0^2\sigma_1^2 \ln \frac{\sigma_1^2}{\sigma_0^2} + \mu_1^2\sigma_0^2 - \mu_0^2\sigma_1^2 \right\} + \frac{b^2}{a^2} \right]^{\frac{1}{2}}.$$

In the examples discussed below, the integrals $\int_{-\infty}^{-\gamma-b/a} p_i(t) dt$ are negligible, and so we consider the following single-threshold test. We set $\eta = \gamma - b/a$ and use $D_0 = (-\infty, \eta]$. Even with this choice of D_0 , the probability of error is given by (2.8), which requires complicated numerical integration as discussed in Appendix B. However, we also consider P_{Ge} , the ‘‘Gaussian approximation’’ of P_e , that we define by

$$P_{Ge} \triangleq \frac{1}{2} \left\{ \int_{-\infty}^{\eta} p_1(t) dt + \int_{\eta}^{\infty} p_0(t) dt \right\}.$$

In the examples below, P_{Ge} was computed easily with the NAG routine S15ABF for evaluating the cumulative distribution of the standard normal density.

C. Examples

We compare the error performance of the two tests in three examples. In each example, we consider three cases, sampling at $K =$ one, five, and ten points. For this purpose, let

$$\begin{aligned} \mathcal{R} &= \{x = (u, v) : 0 \leq u \leq 1 \text{ and } 0 \leq v \leq 1\} \\ \lambda_0(x) &= \begin{cases} c_0 I_{\mathcal{O}}(x) + n. & x \in \mathcal{R} \\ 0. & x \notin \mathcal{R} \end{cases} \\ \lambda_1(x) &= \begin{cases} c_1 I_{\mathcal{E}}(x) + n. & x \in \mathcal{R} \\ 0. & x \notin \mathcal{R} \end{cases} \\ h(x) &= e^{-50(u^2+v^2)} \end{aligned}$$

where $I_{\mathcal{O}}(x)$ and $I_{\mathcal{E}}(x)$ are the indicator functions of the shaded regions in Fig. 3, and n is a constant background level. The sampling points, x_1, \dots, x_{10} , are shown in Fig. 4, and are

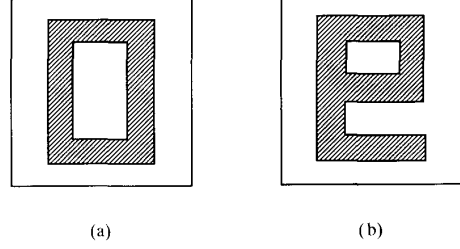


Fig. 3. (a) Region for hypothesis 0. (b) Region for hypothesis 1.

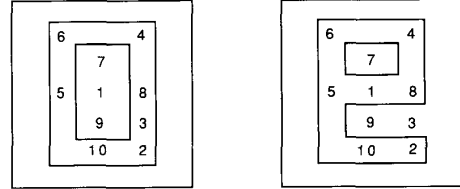


Fig. 4. The sampling points.

TABLE I
THE SAMPLING POINTS

k	x_k	k	x_k
1	(0.5, 0.5)	6	(0.275, 0.825)
2	(0.7, 0.175)	7	(0.5, 0.65)
3	(0.725, 0.325)	8	(0.725, 0.5)
4	(0.725, 0.825)	9	(0.5, 0.325)
5	(0.275, 0.5)	10	(0.5, 0.175)

TABLE II
PARAMETERS FOR EXAMPLE 1

	$K = 1$	$K = 5$	$K = 10$
μ_0	1.52	-1.27	-1.38
σ_0^2	0.422	1.53	1.61
μ_1	4.40	3.25	3.21
σ_1^2	2.58	2.98	2.99

explicitly listed in Table I. When $K = 1$, we set $w_1 = 1$, and hence $T = Z(x_1)$.

Example 1: In this example, $c_0 = c_1 = 100$ and $n = 10$. Then, $\Lambda_0 = 43.00$ and $\Lambda_1 = 44.87$. The weights $\{w_k\}$ for the case $K = 5$ are 1.100019 , 2.603646×10^{-1} , -8.424318×10^{-1} , -1.044526×10^{-2} , -8.713632×10^{-2} , and for $K = 10$, 1.191592 , 3.038038×10^{-1} , -9.131236×10^{-1} , 4.417913×10^{-3} , -7.798749×10^{-2} , 1.564578×10^{-2} , -2.080950×10^{-1} , 9.888259×10^{-2} , -9.350529×10^{-2} , 4.023416×10^{-3} . Table II contains the means and variances of the filtered point process for the three cases. Table III contains the thresholds for the two tests under consideration, the corresponding decision regions D_0 and the probabilities of error. The value of P_{Ge} is also included. A plot of f_0 and f_1 for each case is shown in Fig. 5. From Table III, we see that a reduction of 44.1% in P_e is obtained for the LRT when using $K = 5$ instead of $K = 1$, and a further reduction of 3.25% is obtained by using $K = 10$.

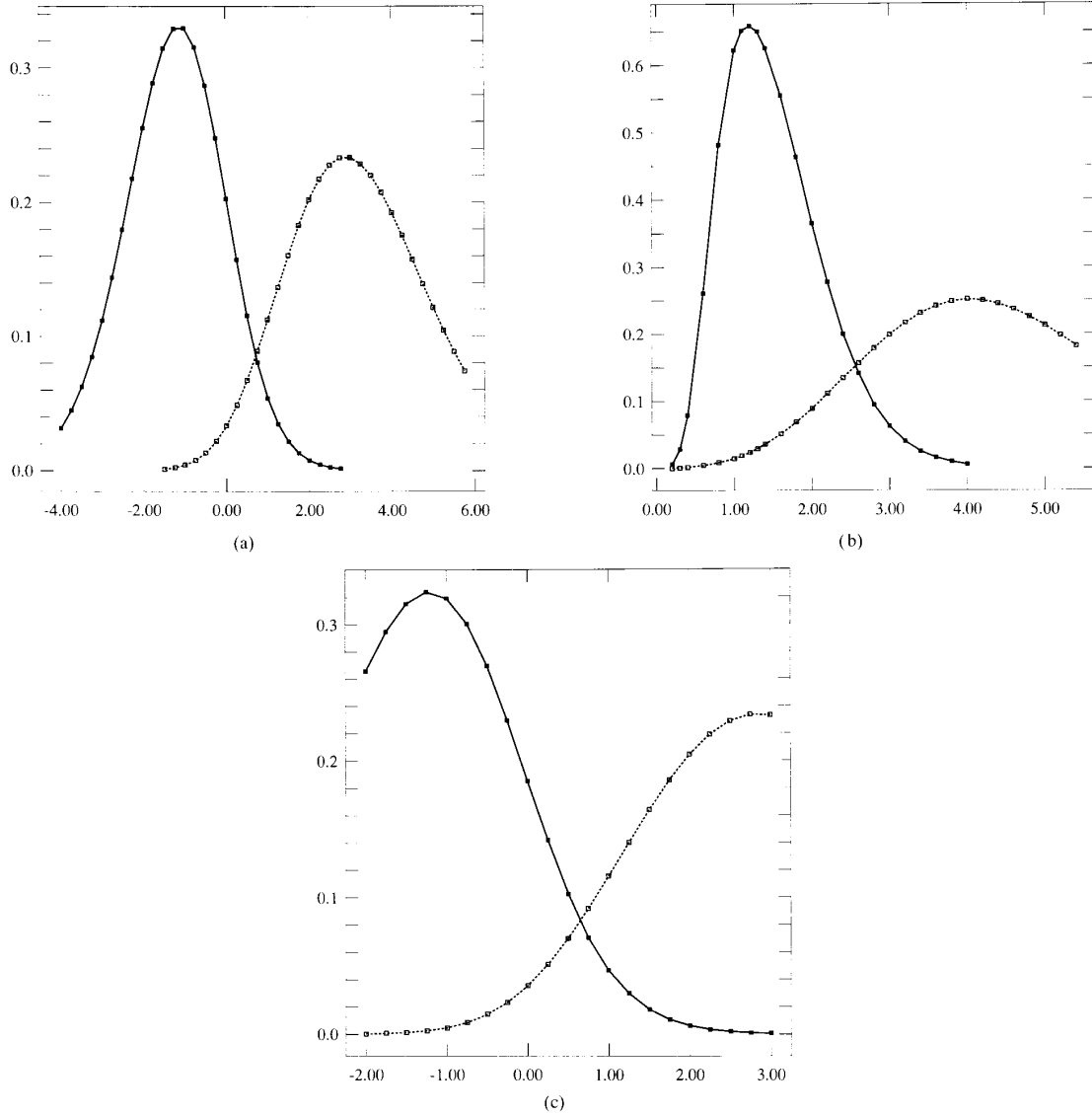


Fig. 5. Probability "density" function of T under H_0 (solid squares) and under H_1 (open squares) for Example 1. (a) $K = 1$. (b) $K = 5$. (c) $K = 10$.

TABLE III
TEST PERFORMANCE IN EXAMPLE 1

K	Test	η	D_0	P_i	P_{Gi}
1	LRT	2.57	$[0, \eta)$	0.0936	—
	GT	2.65	$[0, \eta)$	0.0949	0.0887
5	LRT	0.720	$(-\infty, \eta)$	0.0523	—
	GT	0.773	$(-\infty, \eta)$	0.0525	0.0628
10	LRT	0.650	$(-\infty, \eta)$	0.0506	—
	GT	0.706	$(-\infty, \eta)$	0.0507	0.0617

TABLE IV
PARAMETERS FOR EXAMPLE 2

	$K = 1$	$K = 5$	$K = 10$
m_0	0.759	-0.63	-0.691
σ_0^2	0.211	0.760	0.800
m_1	2.20	1.62	1.601
σ_1^2	1.29	1.49	1.50

Example 2: In this example, $c_0 = c_1 = 50$ and $n = 5$. Then, $\Lambda_0 = 21.5$ and $\Lambda_1 = 22.4$. The weights $\{w_k\}$ for the case $K = 5$ are 1.100019 , 2.603646×10^{-1} , -8.424318×10^{-1} , -1.044526×10^{-2} , -8.713632×10^{-2} , and for $K = 10$,

1.191592 , 3.038038×10^{-1} , -9.131236×10^{-1} , 4.417913×10^{-3} , -7.798749×10^{-2} , 1.564578×10^{-2} , -2.080950×10^{-1} , 9.888259×10^{-2} , -9.350529×10^{-2} , 4.023416×10^{-3} . Table IV contains the means and variances of the filtered point process for the three cases. Table V contains the thresholds for the two tests under consideration, the corresponding decision

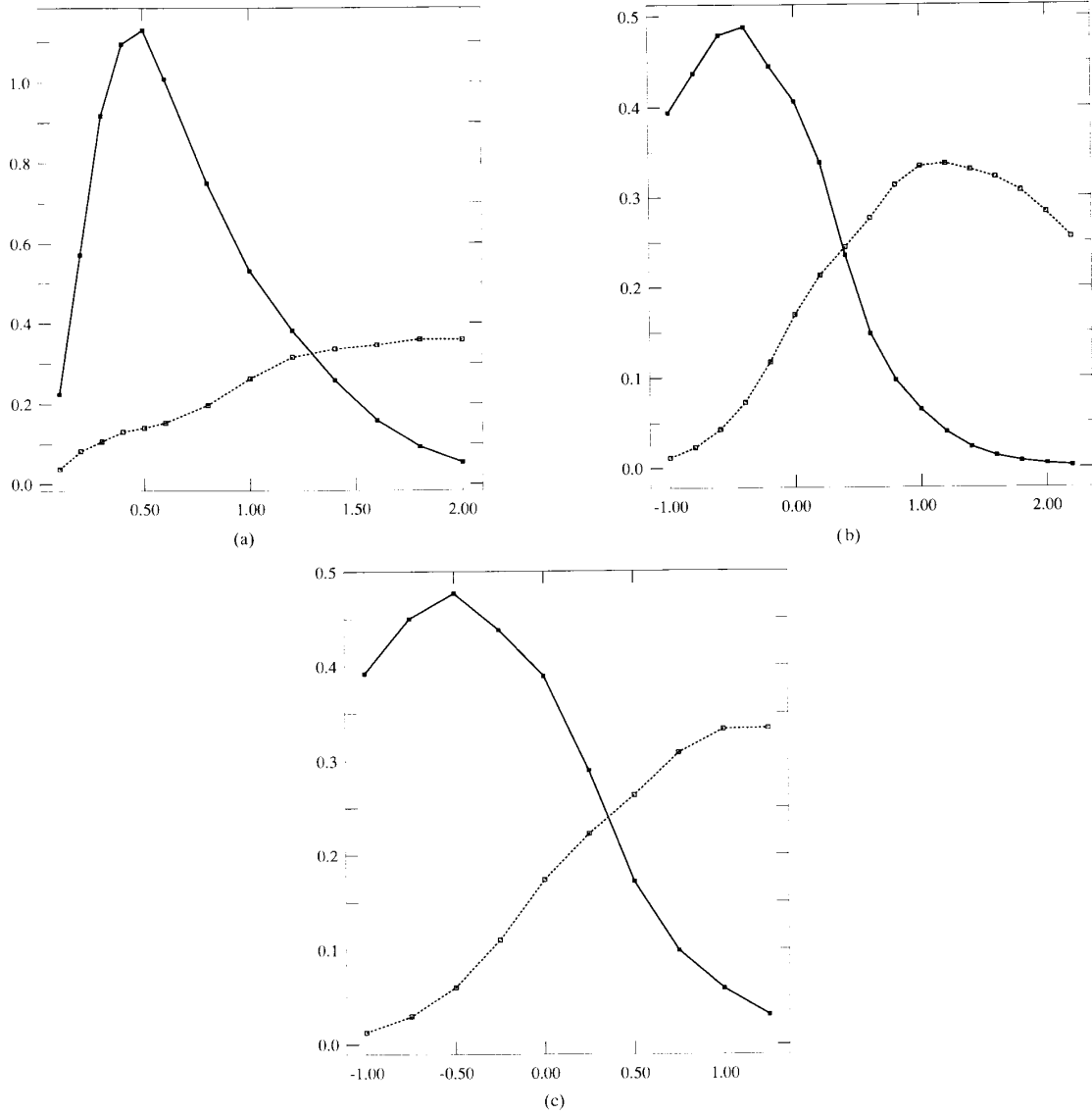


Fig. 6. Probability “density” function of T under H_0 (solid squares) and under H_1 (open squares) for Example 2. (a) $K = 1$. (b) $K = 5$. (c) $K = 10$.

TABLE V
TEST PERFORMANCE IN EXAMPLE 2

K	Test	η	D_0	P_e	P_{Gr}
1	LRT	1.29	$[0, \eta)$	0.177	—
	GT	1.44	$[0, \eta)$	0.182	0.157
5	LRT	0.389	$(-\infty, \eta)$	0.129	—
	GT	0.464	$(-\infty, \eta)$	0.129	0.138
10	LRT	0.360	$(-\infty, \eta)$	0.126	—
	GT	0.423	$(-\infty, \eta)$	0.127	0.137

regions D_0 and the probabilities of error. The value of P_{Gr} is also included. A plot of f_0 and f_1 for each case is shown in Fig. 6. From Table V, we see that a reduction of 27.1% in P_e is obtained for the LRT when using $K = 5$ instead of $K = 1$,

and a further reduction of 2.32% is obtained by using $K = 10$.

Example 3: In this example, $c_0 = c_1 = 16$ and $n = 5$. Then, $\Lambda_0 = 10.28$ and $\Lambda_1 = 10.58$. The weights $\{w_k\}$ for the case $K = 5$ are 7.590372×10^{-1} , 1.791687×10^{-1} , -5.971229×10^{-1} , -8.206410×10^{-3} , -5.451747×10^{-2} , and for $K = 10$, 8.767560×10^{-1} , 2.073143×10^{-1} , -6.402233×10^{-1} , 7.654726×10^{-3} , -4.801134×10^{-2} , 1.534052×10^{-2} , -1.978471×10^{-1} , 7.396542×10^{-2} , -1.243330×10^{-1} , 2.736930×10^{-2} . Table VI contains the means and variances of the filtered point process for the three cases. Table VII contains the thresholds for the two tests under consideration, the corresponding decision regions D_0 and the probabilities of error. The value of P_{Gr} is also included. A plot of f_0 and f_1 for each case is shown in Fig. 7. From Table VII, we see that a reduction of 6.87% in P_e is obtained for the LRT

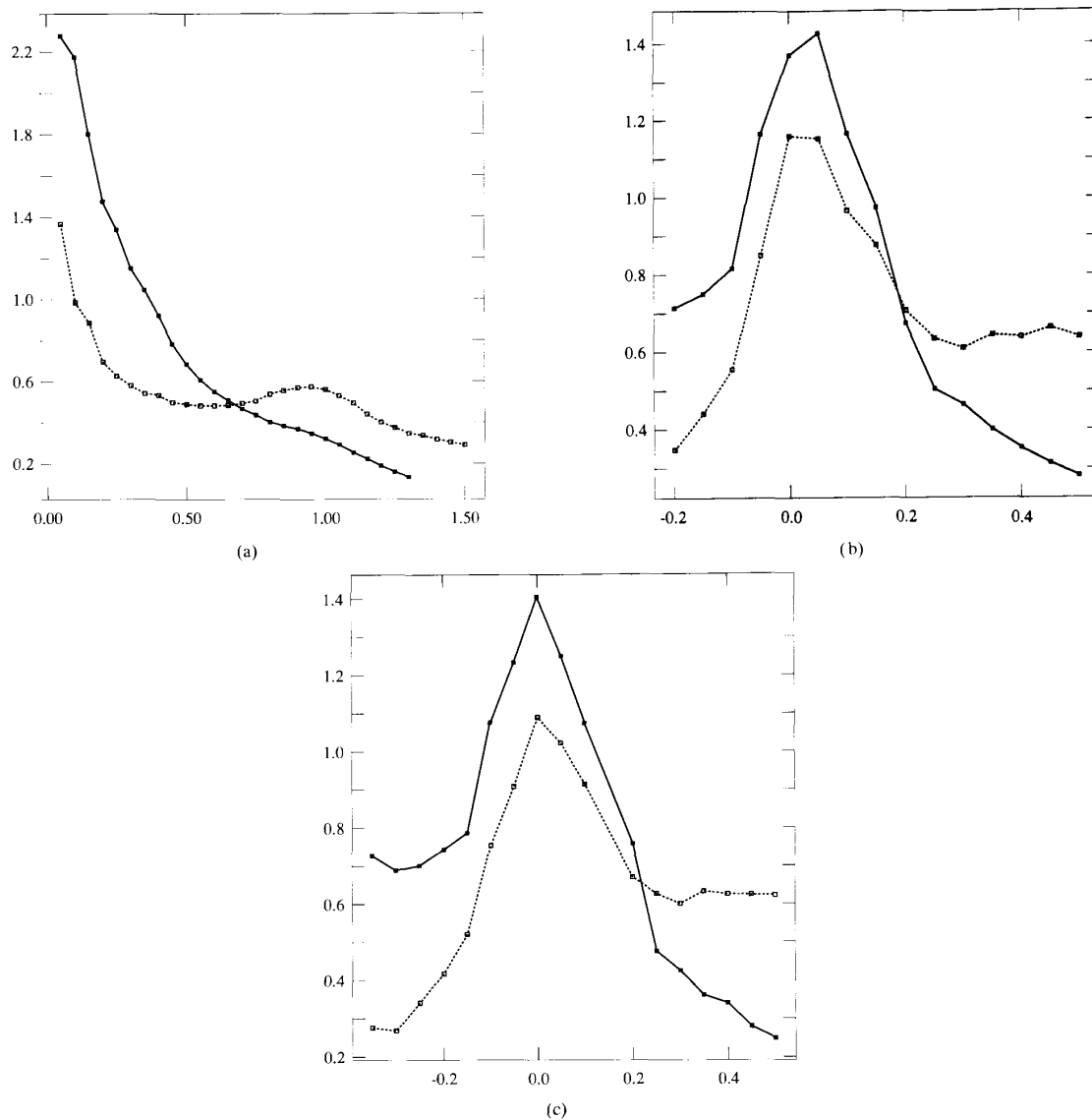


Fig. 7. Probability "density" function of T under H_0 (solid squares) and under H_1 (open squares) for Example 3. (a) $K=1$. (b) $K=5$. (c) $K=10$.

TABLE VI
PARAMETERS FOR EXAMPLE 3

	$K=1$	$K=5$	$K=10$
m_0	0.457	-0.0894	-0.122
σ_0^2	0.174	0.199	0.204
m_1	0.918	0.414	0.398
σ_1^2	0.518	0.305	0.316

when using $K=5$ instead of $K=1$, and a further reduction of 0.962% is obtained by using $K=10$.

IV. DISCUSSION AND CONCLUSION

We have treated image detection at low light levels as a binary hypothesis-testing problem based on a one-dimensional

TABLE VII
TEST PERFORMANCE IN EXAMPLE 3

K	Test	η	D_0	P_i	P_{G_i}
1	LRT	0.675	$[0, \eta)$	0.335	—
	GT	0.893	$[0, \eta)$	0.348	0.310
5	LRT	0.188	$(-\infty, \eta)$	0.312	—
	GT	0.236	$(-\infty, \eta)$	0.313	0.303
10	LRT	0.172	$(-\infty, \eta)$	0.309	—
	GT	0.212	$(-\infty, \eta)$	0.311	0.300

test statistic. This statistic was obtained by filtering the received image and then sampling at one point. The filter we used was obtained by maximizing an ad hoc signal-to-noise ratio. In the examples we considered, we found that the largest

weights of the filter (for the cases $K = 5$ and $K = 10$) correspond to the locations where the “o” and the “e” do not overlap. This makes intuitive sense: at x_1 , the “e” is present and w_1 is the largest positive weight; at x_3 , the “o” is present and w_3 is the largest-magnitude negative weight. Since x_1 and x_3 are the most important points for discrimination between H_0 and H_1 , we observed little improvement in performance when using $K = 10$ instead of $K = 5$. The lower the intensity of the point process, the harder it is to discriminate between the hypotheses. The samples in Example 3 bear less information for discrimination than those in the other two examples. This resulted in a much smaller improvement in performance in Example 3 than in the other two examples when going from $K = 1$ to $K = 5$. We compared our LRT with a test that uses Gaussian densities. From the results of the examples, it is clear that the distribution function of T is not Gaussian under either hypothesis (see Figs. 5–7). It is interesting to observe, though, that the probability of error P_e is very similar for the two tests, i.e., P_e is not very sensitive to the value of η . We note that the Gaussian approximation P_{G_e} of the probability of error is neither an upper nor a lower bound for P_e since sometimes it overpredicts (by 21.9% in Example 1, $K = 10$) and sometimes it underpredicts (by 11.3% in Example 2, $K = 1$) the value of P_e . Hence, P_{G_e} is not a reliable quantity for estimating P_e .

Further research should be devoted to studying the behavior of P_e as a function of the weights $\{w_k\}$, as a function of the sampling points $\{x_k\}$, and as a function of the total number of weights of the filter g . It is very important that fast and accurate methods be found in order to compute the functions $f_i(t)$ and $G_i(t)$ since straightforward applications of numerical integration are rather time consuming.

APPENDIX A EVALUATION OF THE LIKELIHOOD RATIO

In order to perform the test (2.6), we need to compute the “density” function of T under each hypothesis. Loosely speaking, this can be accomplished by computing the inverse Fourier transform of the characteristic function of T . Several methods have been proposed (see [3] and the references therein) for carrying out this computation. Our numerical solution relies on the use of the quadrature subroutines D01FCF and D01AMF from the NAG library.

For the purpose of computing the LR function and the probability of error for the test (2.6), we introduce the moment-generating function of T , denoted by $M_i(s)$; it is given by [5, p. 381]

$$M_i(s) \triangleq \mathbb{E}_i[e^{sT}] = \exp\left\{\int \lambda_i(x) [e^{s\hat{g}(-x)} - 1] dx\right\}. \quad (\text{A.1})$$

From (2.5), it follows that

$$M_i(s) = e^{-\Lambda_i} + K_i(s) \quad (\text{A.2})$$

where

$$K_i(s) \triangleq \int_{-\infty}^{\infty} e^{st} f_i(t) dt. \quad (\text{A.3})$$

With $s = \sigma + j\omega$, let

$$C_i^\sigma(\omega) \triangleq \operatorname{Re}\left\{\int \lambda_i(x) e^{s\hat{g}(-x)} dx\right\} = \int \lambda_i(x) e^{\sigma\hat{g}(-x)} \cos(\omega\hat{g}(-x)) dx$$

and

$$S_i^\sigma(\omega) \triangleq \operatorname{Im}\left\{\int \lambda_i(x) e^{s\hat{g}(-x)} dx\right\} = \int \lambda_i(x) e^{\sigma\hat{g}(-x)} \sin(\omega\hat{g}(-x)) dx.$$

We compute $C_i^\sigma(\omega)$ and $S_i^\sigma(\omega)$ numerically using the NAG routine D01FCF. If we set $s = j\omega$ in (A.3) and take inverse Fourier transforms, we obtain, since $\operatorname{Re}\{K_i(j\omega)e^{j\omega t}\}$ is an even function of ω ,

$$f_i(t) = \frac{1}{\pi} \int_0^\infty \left\{ e^{-(\Lambda_i - C_i(\omega))} \cos(S_i(\omega) - \omega t) - e^{-\Lambda_i} \cos(\omega t) \right\} d\omega \quad (\text{A.4})$$

where, for convenience of notation, we write C_i and S_i instead of C_i^σ and S_i^σ when $\sigma = 0$.

APPENDIX B EVALUATION OF THE PROBABILITY OF ERROR

In order to compute P_e , we first need to compute (2.9). To this end, let $\hat{G}_i(s)$ denote the Laplace transform of G_i . More precisely,

$$\hat{G}_i(s) \triangleq \int_{-\infty}^{\infty} G_i(t) e^{st} dt, \quad \operatorname{Re}\{s\} < 0. \quad (\text{B.1})$$

Assuming that

$$\lim_{t \rightarrow -\infty} e^{st} G_i(t) = 0, \quad \operatorname{Re}\{s\} < 0,$$

which is true in our applications, integration by parts yields

$$\hat{G}_i(s) = -\frac{1}{s} K_i(s). \quad (\text{B.2})$$

Combining (B.2) with (A.2) and (A.1), we can substitute in (B.1) and take inverse Fourier transforms (after writing $s = \sigma + j\omega$, with $\sigma < 0$) to obtain

$$G_i(t) = -\frac{e^{-\sigma t}}{\pi} \int_0^\infty \frac{Q_i^\sigma(t, \omega)}{\sigma^2 + \omega^2} d\omega \quad (\text{B.3})$$

where

$$Q_i^\sigma(t, \omega) = e^{-(\Lambda_i - C_i^\sigma(\omega))} \cdot [\sigma \cos(S_i^\sigma(\omega) - \omega t) + \omega \sin(S_i^\sigma(\omega) - \omega t)] - e^{-\Lambda_i} (\sigma \cos(\omega t) - \omega \sin(\omega t)).$$

The integrals (A.4) and (B.3) are computed numerically using the NAG routine D01AMF.

REFERENCES

- [1] W. A. Gardner, "A unifying view of second-order measures of quality for signal classification," *IEEE Trans. Commun.*, vol. COM-28, pp. 807-816, June 1980.
- [2] E. A. Geraniotis and H. V. Poor, "Minimax discrimination for observed Poisson processes with uncertain rate functions," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 660-669, Sept. 1985.
- [3] J. A. Gubner, "On the computation of shot-noise probability distributions," submitted to *IEEE Trans. Inform. Theory*.
- [4] G. M. Morris, "Scene matching using photon-limited images," *J. Opt. Soc. Amer. A*, vol. 1, pp. 482-488, May 1984.
- [5] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1984.
- [6] ———, "High density shot noise and Gaussianity," *J. Appl. Prob.*, vol. 8, pp. 118-127, 1971.
- [7] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1988.
- [8] R. E. Sequeira, "Binary hypothesis testing for filtered point processes," M.S. Thesis, Univ. Wisconsin—Madison, May 1990.
- [9] D. L. Snyder, *Random Point Processes*. New York: Wiley, 1975.
- [10] M. N. Wernick and G. M. Morris, "Image classification at low light levels," *J. Opt. Soc. Amer. A*, vol. 3, pp. 2179-2187, Dec. 1986.



Raúl E. Sequeira (S'88) was born in San José, Costa Rica, on July 2, 1966. He received the B.S.E.E. degree from the Universidad de Costa Rica in 1988, and the M.S.E.E. degree from the University of Wisconsin—Madison in 1990.

He is currently working on the Ph.D. degree in electrical engineering at the University of Wisconsin—Madison, where he has been a Research Assistant since 1989. His main research interests are in detection and estimation theory and in statistical signal processing.

Mr. Sequeira is a member of the IEEE Signal Processing Society.



John A. Gubner (S'83-M'83) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1983, 1985, and 1988, respectively.

In 1988 he joined the faculty of the University of Wisconsin—Madison as an Assistant Professor in the Department of Electrical and Computer Engineering. In 1985 he received an IEEE Frank A. Cowan Scholarship for graduate study in communications, and from 1986 to 1988 he was a Graduate Fellow with the Systems Research Center at the University of Maryland. His research interests include distributed estimation, shot-noise processes, wavelet transforms, and information theory.



Bahaa E. A. Saleh (M'73-SM'86-F'91) received the B.S. degree from Cairo University, Cairo, Egypt, in 1966 and the Ph.D. degree from the Johns Hopkins University, Baltimore, MD, in 1971, both in electrical engineering.

From 1971 to 1974 he was an Assistant Professor at the University of Santa Catarina, Brazil. Thereafter, he joined the Max Planck Institute, Goettingen, Germany, where he was involved in research of laser light scattering and photon correlation spectroscopy. He is presently professor and Chairman

of the Department of Electrical and Computer Engineering at the University of Wisconsin—Madison, where he has been since 1977. He held visiting appointments at the University of California, Berkeley, in 1977, and the Columbia Radiation Laboratory of Columbia University, New York, NY, in 1983. He is currently involved in research of image processing, optical information processing, statistical optics, optical communication, and vision. He is the author of *Photoelectron Statistics* (Springer, 1978) and a coeditor of *Transformations in Optical Signal Processing* (SPIE, 1981). During 1980-1988 he was an Associate and Topical Editor of the *Journal of the Optical Society of America*. He is currently Editor of the same journal. He is a member of the Board of Editors of *Quantum Optics*.

Dr. Saleh is a Fellow of the Optical Society of America and a member of Phi Beta Kappa and Sigma Xi. He received the Wisconsin Romnes Award in 1981 and was appointed a Guggenheim Fellow in 1984.