

Passive ranging through wave-front coding: information and application

Gregory E. Johnson, Edward R. Dowski, Jr., and W. Thomas Cathey

Passive-ranging systems based on wave-front coding are introduced. These single-aperture hybrid optical–digital systems are analyzed by use of linear models and the Fisher information matrix. Two schemes for passive ranging by use of a single aperture and a single image are investigated: (i) estimating the range to an object and (ii) detecting objects over a set of ranges. Theoretical limitations on estimator-error variances are given by use of the Cramer–Rao bounds. Evaluations show that range estimates with less than 0.1% error can be obtained from a single wave-front coded image. An experimental system was also built, and example results are given. © 2000 Optical Society of America
OCIS codes: 150.0150, 150.5670, 150.6910, 150.3040.

1. Introduction

Wave-front coding is a technique for uniquely coding information contained in an incoherent wave front. Optical wave fronts from an object are coded by an optical mask that is typically placed at or near the aperture stop of an otherwise general imaging system. The incoherently detected and spatially sampled wave front is decoded by a digital signal-processing algorithm. The optical–digital system is designed to communicate a portion of an object's three-dimensional information to a two-dimensional surface in an optimal manner. The particular optical masks in this ranging study are designed to preserve *range* information. Other masks may preserve or remove spatial frequencies, for example. The digital signal-processing wave-front coded ranging algorithms developed for this study were implemented in software, although other implementations can be realized. A typical wave-front coded passive-ranging system is shown in Fig. 1.

The remainder of this paper is organized as follows. Section 2 gives a brief background on optical passive-ranging systems; active systems are not discussed for brevity. Section 3 presents the optical mask used in this study. Section 4 introduces the linear system

models and signal-processing algorithms with separate sections on range estimation and object detection. Section 5 gives theoretical limitations based on the Cramer–Rao bounds (CRB's) for two example systems. Experimental validation is presented in Section 6 for one of the systems for which a prototypical mask and lens were assembled. Section 7 gives a discussion of the results and the conclusions.

2. Background: Optical Passive Ranging

Passive systems are, in general, characterized by the lack of a transmitter. As such, they enjoy advantages such as stealth, lower power, and fewer components over active systems. However, they often suffer from higher complexity and degraded performance. Example incoherent passive-ranging systems may employ multiple apertures or images such as stereo or motion ranging.¹ These techniques require multiple, displaced views of an object. Knowledge of the viewing-angle geometries and the image correspondence is used to determine the object range. Single-aperture passive systems include commercial autofocus (ATF) systems² and range-from-defocus methods.³ Range-from-defocus systems make assumptions about the object's spatial-frequency content and iterate the lens-focus position to find a best-focused lens position. In this case range can be derived from knowing or estimating the optical geometries.

A widespread and highly successful commercial ranging application is ATF systems for cameras.⁴ Active ATF is often the measurement of the angle that a transmitted beam makes when reflected off a target. Passive ATF systems come in two varieties:

The authors are with the Department of Electrical and Computer Engineering, University of Colorado, Campus Box 425, Boulder, Colorado 80303. G. E. Johnson's e-mail address is johnsog@colorado.edu.

Received 10 June 1999; revised manuscript received 14 October 1999.

0003-6935/00/111700-11\$15.00/0

© 2000 Optical Society of America

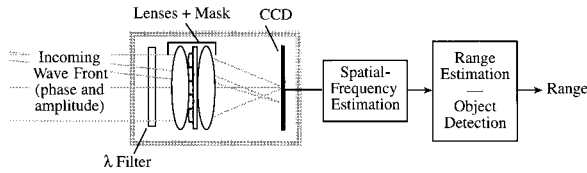


Fig. 1. Example optical–digital system for single-aperture passive ranging by use of wave-front coding.

contrast matching and phase matching. The contrast method is similar to the range-from-defocus approach in which the maximized spatial-frequency content represents the best focus. Phase ATF uses the optical concept that, for an in-focus system, the intensity across the exit pupil is uniform; this is not true for defocused systems. Note that these systems are at best crude range finders, finding focus positions that are more suitable for human photographic consumption rather than precision ranging. More information on commercial ATF systems can be found in Ref. 4.

As an alternative to the above ranging methods, wave-front coding⁵ employs a modified optical system that directly codes range information. The range information can be extracted from a single image in a straightforward manner by means of signal-processing methods. One such hybrid system was already introduced by Dowski and Cathey⁵ and Dowski.⁶ The optical mask described in Ref. 6 produced a magnitude transfer function (MTF) with periodic nulls; the periodicity is dependent on misfocus, hence range. The estimate of the periodicity of the nulls in the spatial-frequency estimate of an acquired image is related to the range.

In the null-space design overlapping objects at different ranges produce competing periodicities, making the estimate of a single range difficult. In contrast, the mask used in this ranging study and described in Refs. 7 and 8 performs a bandpass type of coding, separating objects at different ranges into different spatial-frequency regions. The encoding behavior of this mask is discussed in more detail in Section 3. Digital processing of the coded images for extracting range is discussed in Sections 4–6.

3. Optical Masks for Passive Ranging

The mask used in this passive-ranging system is designed in spatial coordinates as a $\cos(\beta\pi x)\cos(\beta\pi y)$ rectangularly separable mask. The transmissivity for the mask is shown in Fig. 2.

Because passive optical media are positive only, a properly oriented phase mask of height $\lambda/2$, where λ is the nominal center frequency (wavelength) of the incident wave, effectively provides a negative shift for the otherwise $|\cos(\beta\pi x)|$ profile. The phase profile and the spatial orientations of the phase mask and the amplitude mask are shown in Fig. 3.

The mask produces a particular MTF that changes dramatically with misfocus ψ . This theoretical MTF for an infinite aperture, the misfocus-dependent part

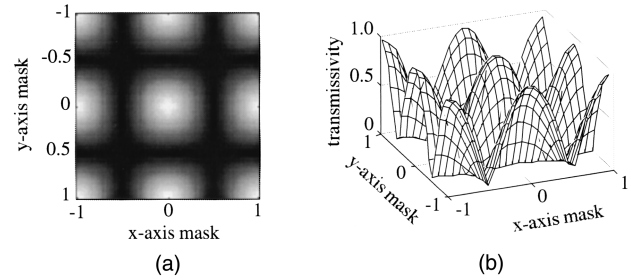


Fig. 2. (a) Transmission profile of a $\cos(\beta\pi x)\cos(\beta\pi y)$ ranging mask. (b) Contour plot of the transmission profile shown in (a).

of which is seen as a pair of shifted delta functions, is given by

$$H_\psi(u) = 1/2\delta(u) + 1/4\delta(u \pm \pi\beta/\psi),$$

$$\psi = \frac{\pi d^2}{4\lambda} \left(\frac{1}{f} - \frac{1}{z_o} - \frac{1}{z_i} \right). \quad (3.1)$$

The misfocus parameter ψ can be expressed in terms of a general optical system, as shown in Eq. (3.1), where d is the width of the aperture, z_o and z_i are the object and the image distances, respectively, f is the focal length of the system, and λ is the nominal center wavelength of the incoherent illumination. Large or small values of misfocus are translated to near or far values, respectively, for range, given the optical parameters of the imaging system. A finite aperture essentially smears the impulse response, which is given by

$$H_\psi(u) = 1/2(1 - |u|/d)\cos(\beta\pi u)\text{sinc}[\psi u(d - |u|)]$$

$$+ 1/4(1 - |u|/d)\text{sinc}[(\psi u \pm \pi\beta)(d - |u|)]. \quad (3.2)$$

Example MTF's $H(u)$ for the cosine mask are shown in Fig. 4. The plots of the MTF's are clearly seen as having a bandpass region as well as a common zero-frequency subspace. Both the shifting and the broadening of the MTF passband regions is clear for the two values of misfocus.

The behavior shown in Fig. 4 is seen as a passband-filtering operation, the center frequency of which is highly range dependent. The object's spatial frequencies that are coded in the image will vary de-

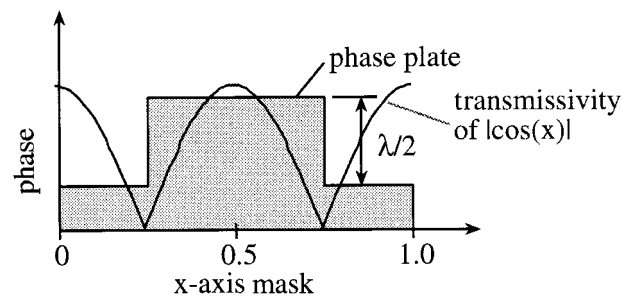


Fig. 3. Spatial orientation of the phase-mask profile with respect to the transmission-mask profile.

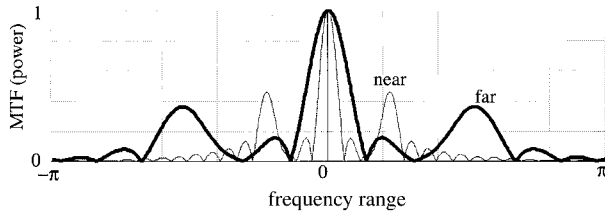


Fig. 4. MTF's for the two example misfocus values. Note the optical system essentially produces power (the MTF amplitude) as a function of range (the peak-power location).

pending on the range between the object and the imaging system.

4. Range Decoding by Means of Signal Processing

Range information is coded in the spatial frequencies of a sampled image. To evaluate the spatial-frequency content of a wave-front coded image, one implements Fourier transform techniques. The Fourier method leads to a maximum-likelihood estimate of the sinusoidal frequency content and is valuable in the signal-estimation and the signal-detection tasks described below. For the cosinusoidal mask, we can perform range estimation by implementing a peak detector on the spectral estimate. Alternatively, we can perform object detection by measuring the energy content in the spatial-frequency estimate. Note that the entire image can be processed at once, or we can process only a small region of interest (ROI) within the image. By processing smaller ROI's, we can produce a two-dimensional range image (or range map) by displacing the ROI across a coded image and repeating the spectral estimation and the range estimation or the object-detection routines.

A. Spectral Estimation

The spatial-frequency estimation is implemented by use of Fourier analysis. Equation (4.1) is a well-known equation for averaging complex-valued Fourier coefficients. The Fourier coefficients are f_k , where g_q ($q = 0, \dots, 2N - 1$) is one element of the $2N$ -length (zero-padded, if necessary) column (say, the m th column) of the ROI (or the input image):

$$y_k = |\mathbf{E}[f_k]| = \left| \frac{1}{M} \sum_{m=1}^M f_{k,m} \right| \\ = \left| \frac{1}{M} \sum_{m=1}^M \sum_{q=0}^{2N-1} w_q u_{q,m} \exp(-j2\pi qn/2N) \right|, \quad (4.1)$$

where \mathbf{E} is the expected-value operator.

We can evaluate horizontal spectra by processing rows in the ROI or the image. Note the presence of a smoothing window, w_q ($q = 0, \dots, 2N - 1$), that is applied to the data prior to transforming. This helps to reduce the sidelobes in the estimate. Averaging of the complex-valued coefficients reduces random noise in the data.

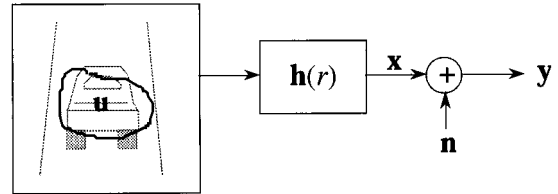


Fig. 5. Typical noisy image-formation model for a single object.

B. Range Estimation

Here the range to a single unknown object is desired. A noiseless image \mathbf{x} , where $\mathbf{x} = [x_1 x_2 \dots x_N]^T$, is modeled as the convolution of an unknown object \mathbf{u} with the range-dependent impulse response $\mathbf{h}(r)$. A noisy version of the image \mathbf{y} is modeled as \mathbf{x} through a noisy channel with additive noise, \mathbf{n} , $N[0, \sigma^2 \mathbf{I}]$, without loss of generality. These relations are shown in Eq. (4.2), where $\mathbf{F}(r)$ is a convolution matrix based on $\mathbf{h}(r)$ and the asterisk denotes convolution:

$$\mathbf{x} = \mathbf{h}(r) * \mathbf{u} = \mathbf{F}(r)\mathbf{u}, \quad \mathbf{y} = \mathbf{x} + \mathbf{n} \quad (4.2)$$

(this linear system is also illustrated graphically in Fig. 5).

In the case of passive ranging the system impulse response $\mathbf{h}(r)$ and its convolution matrix $\mathbf{F}(r)$ can contain the lens, the mask, and the propagation effects as well as the optical aberrations. The goal is to estimate the parameter r for unknown objects \mathbf{u} , given $\mathbf{F}(r)$ and the data \mathbf{y} with unknown noise \mathbf{n} . Note that the assumption of a noise-covariance matrix $\sigma^2 \mathbf{I}$ does not detract from the generality of these expressions. If the data \mathbf{y} have a nondiagonal noise-covariance matrix \mathbf{R} (i.e., the noise is not uncorrelated), they can be prewhitened as $\mathbf{R}^{-1/2} \mathbf{y}$, and the system response in Eq. (4.2) becomes $\mathbf{R}^{-1/2} \mathbf{F}(r_i)$.

On the basis of the optical system behavior described in Section 2, one can expect a sinusoidal-like response to a spectrally flat object located at a single range. The estimation of range is therefore the task of estimating the spatial frequencies in the image and locating the maximum energy in the spectra (not including the zero-frequency response); essentially, a simple peak detector is used on the non-zero-frequency portion of the maximum-likelihood spatial-frequency estimate described in Subsection 4.A. Section 5 shows that the error in range estimates from such a system can be less than 0.1% for a point-source object.

C. Object Detection

Here the goal is to detect objects within a given set of ranges. Thus we are interested in the sum of objects \mathbf{u} that make up an overall image \mathbf{x} . A noiseless image \mathbf{x} can be modeled as the sum of the convolutions of the unknown objects \mathbf{u}_i with their respective range-dependent impulse responses $\mathbf{h}(r_i)$, where r_i is the range parameter of interest for the i th object. Again, a noisy version of the image \mathbf{y} is modeled through an additive-noise channel with noise \mathbf{n} .

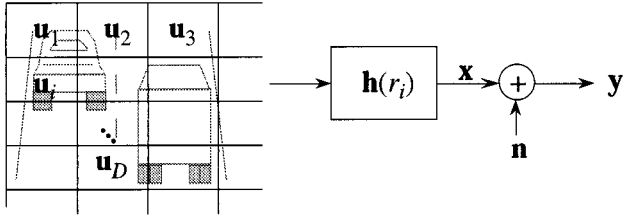


Fig. 6. Typical noisy image-formation model for multiple objects. Note that the partitioning of the object volume into regions is arbitrary.

The relations for uncorrelated noise are indicated mathematically in Eq. (4.3) and graphically in Fig. 6:

$$\mathbf{x} = \sum_i^p \mathbf{h}(r_i) * \mathbf{u}_i = \sum_i^p \mathbf{F}(r_i) \mathbf{u}_i, \quad \mathbf{y} = \mathbf{x} + \mathbf{n}. \quad (4.3)$$

Equation (4.3) is now rephrased through linear algebra to bring matrix tools to bear. A linear system \mathbf{B} is assumed to be an $N \times p$ matrix whose columns represent the modes of the system. In the case of object detection the columns represent the positive frequency coefficients from the modulation transfer function at a discrete selection of ranges [thus N equals half of the fast Fourier transform (FFT) length, and p represents the number of MTF's]. For example, $\mathbf{B} = [\mathbf{b}_1 \mathbf{b}_2 \dots \mathbf{b}_p]$ contains eight MTF's, where \mathbf{b}_i is the MTF at range r_i . An input image \mathbf{u} is assumed to be an $(N \times 1)$ -length vector and is formed by a linear combination of terms \mathbf{b}_i , i.e., $\mathbf{u} = \mathbf{B}\mathbf{a}$, where \mathbf{a} is a $(p \times 1)$ -length vector of weights, one for each \mathbf{b}_i , as shown in Eq. (4.4):

$$\mathbf{u} = \mathbf{B}\mathbf{a} = \sum_{i=1}^p \mathbf{b}_i a_i, \quad \mathbf{a} = [a_1 \dots a_p]^T, \quad \mathbf{B}: N \times p, \quad \mathbf{u}: N \times 1, \quad \mathbf{a}: p \times 1. \quad (4.4)$$

The detection methodology seeks to find a representation of \mathbf{a} from noisy observations \mathbf{y} of the object \mathbf{u} , where $\mathbf{y} = \mathbf{u} + \mathbf{n}$, with \mathbf{n} as a noise vector that is independent of the data, $\mathbf{n}: N[0, \sigma^2 \mathbf{I}]$. An estimate $\hat{\mathbf{a}}$ (the minimum mean-squared-error estimate) of \mathbf{a} can be formed by means of the pseudoinverse of \mathbf{B} , denoted $\mathbf{B}^\#$. From $\mathbf{u} = \mathbf{B}\mathbf{a}$ and $\mathbf{y} = \mathbf{u} + \mathbf{n}$, we estimate \mathbf{u} through $\hat{\mathbf{u}} = \mathbf{P}_B \mathbf{y} = \mathbf{B}\mathbf{B}^\# \mathbf{y}$ to get $\hat{\mathbf{a}} = \mathbf{B}^\# \hat{\mathbf{u}}$ or $\hat{\mathbf{a}} = \mathbf{B}^\# \mathbf{B}\mathbf{B}^\# \mathbf{y} = \mathbf{B}^\# \mathbf{y}$, where $\mathbf{B}^\# = (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T$. Note that we can equivalently estimate the contributions of the Fourier modes (the weights on the values of \mathbf{b}_i) or the contributions of the impulse responses [point-spread functions (PSF's)]. In the following discussion the system is denoted as $\mathbf{u} = \mathbf{B}\mathbf{a}$, and \mathbf{B} is assumed to be composed of MTF's and not of PSF's.

The detectors in this study are based on binary hypothesis testing (H_0 versus H_1). If one assumes that the signal \mathbf{y} is generated from exactly one mode (the i th mode) the parameter vector \mathbf{a} is all zeros except for a 1 in the i th position: $\mathbf{a} = [0 \dots 0 \ 1 \ 0 \dots 0]^T$. Thus one can think of a set of vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_p$, where \mathbf{a}_i indicates that the i th value is a 1 and that the rest are zeros. A simple hypotheses

test H_0 (with a composite alternative H_1) would then appear as

$$\begin{aligned} H_0 \text{ (Object at } r_i) \\ H_1 \text{ (No object at } r_i) \end{aligned} \Rightarrow \mathbf{y} = \mathbf{B}\mathbf{a}, \quad \begin{cases} \mathbf{a} = \mathbf{a}_i \text{ under } H_0 \\ \mathbf{a} \neq \mathbf{a}_i \text{ under } H_1 \end{cases} \quad (4.5)$$

In this detection approach the signal is reformulated as the combination of the interference modes and a signal mode. The interference terms may now be rejected with the techniques of subspace projections. Here the notations for the $N \times p$ system matrix \mathbf{B} and the $p \times 1$ parameter vector \mathbf{a} are replaced with \mathbf{B}_0 and \mathbf{a}_0 , respectively. The data \mathbf{y} are described as having a subspace signal, with subspace interference and noise of an unknown level, or $\mathbf{y} = \mathbf{B}_0 \mathbf{a}_0 + \mathbf{n}$ can also be modeled as $\mathbf{y} = \mathbf{b}_i^T \mathbf{a}_i + \mathbf{B}\mathbf{a} + \mathbf{n}$. This gives \mathbf{y} as the combination of a single-mode subspace signal $\mathbf{b}_i^T \mathbf{a}_i$ and an interference signal $\mathbf{B}\mathbf{a}$. The $N \times (p - 1)$ -sized \mathbf{B} matrix is \mathbf{B}_0 with the i th column missing, and the $(p - 1)$ -length \mathbf{a} vector is \mathbf{a}_0 with the i th element removed.⁹ The projectors of interest are defined as

for

$$\mathbf{B}_0 = [\mathbf{b}_1 \dots \mathbf{b}_p],$$

for the i th element define $\mathbf{b} = \mathbf{b}_i$,

$$\mathbf{B} = [\mathbf{b}_1 \dots \mathbf{b}_{i-1} \quad \mathbf{b}_{i+1} \dots \mathbf{b}_p],$$

i.e., \mathbf{B}_0 is missing the i th column,

now

$$\begin{aligned} \mathbf{P}_B &= \mathbf{B}\mathbf{B}^\#, & \mathbf{P}_B^\perp &= \mathbf{I} - \mathbf{P}_B, & \mathbf{P}_G^\perp &= \mathbf{I} - \mathbf{P}_G, \\ \mathbf{P}_G &= \mathbf{P}_{\mathbf{P}_B^\perp \mathbf{b}} = (\mathbf{P}_B^\perp \mathbf{b}\mathbf{b}^T \mathbf{P}_B^\perp)^\# = \mathbf{P}_B^\perp \mathbf{b}\mathbf{b}^T \mathbf{P}_B^\perp (\mathbf{b}^T \mathbf{P}_B^\perp \mathbf{b})^{-1}, \end{aligned} \quad (4.6)$$

by use of the ranging-system variables.

A matched-subspace detector⁹ provides an F -distributed statistic $L_2 - 1$, hereafter denoted simply as L , as shown by

$$L = \frac{\mathbf{y}^T \mathbf{P}_B^\perp \mathbf{P}_G \mathbf{P}_B^\perp \mathbf{y}}{\mathbf{y}^T \mathbf{P}_B^\perp \mathbf{P}_G^\perp \mathbf{P}_B^\perp \mathbf{y}} \left(\frac{s-p}{p} \right). \quad (4.7)$$

Equation (4.7) essentially gives the ratio of the signal power in subspace $\langle \mathbf{b} \rangle$ to the signal power in the noise

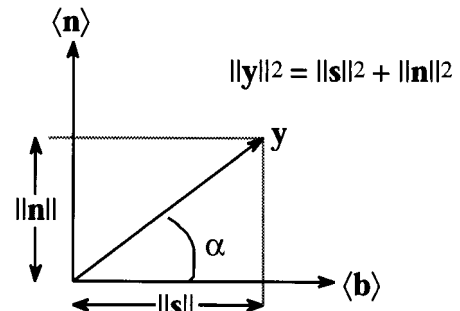


Fig. 7. $L(\mathbf{y})$ and the angular relations in the \mathbf{P}_B^\perp space.

subspace $\langle \mathbf{n} \rangle$ plus the interference subspace $\langle \mathbf{B} \rangle$, after projecting \mathbf{y} through $\mathbf{P}_{\mathbf{B}}$. This concept is slightly modified in Eqs. (4.8) and (4.9), below. Note the scaling by the dimensions of the subspaces in Eq. (4.7), where in this case p is the number of modes in \mathbf{B} (not in \mathbf{B}_0) and $s = (N - \dim(\mathbf{b})) = (N - 1)$.

Note that the design of the interference subspace $\langle \mathbf{B} \rangle$ can have a tremendous impact on the quality of the detector.¹⁰ A subspace rich in interference can actually scale the noise beyond what an orthogonal projection system might, providing degraded results compared with a simple orthogonal projection. Optimal subspace designs are largely application and environment dependent.

In essence the statistic L in Eq. (4.7) measures the ratio of the length of \mathbf{y} in $\langle \mathbf{b} \rangle$ to the length of \mathbf{y} in $\langle \mathbf{n} \rangle$, as illustrated in Fig. 7. This ratio is the cotangent squared of the angle α between the two vectors. Thus $L = \|\mathbf{s}\|^2 / \|\mathbf{n}\|^2 \rightarrow \cot^2(\alpha)$, which has values ranging from zero to infinity. An alternative statistic β is developed in Eqs. (4.8) and (4.9) in which the cosine of the angle is measured. This alternative statistic has the advantage that it is scaled from 0.0 to 1.0 and is beta distributed.¹¹ This distribution is helpful in determining the detector thresholds for each mode in a particular application:

$$L : \cot^2(\alpha) = \frac{\cos^2(\alpha)}{\sin^2(\alpha)} = \frac{\cos^2(\alpha)}{1 - \cos^2(\alpha)} = \frac{\|\mathbf{s}\|^2 / \|\mathbf{y}\|^2}{1 - \|\mathbf{s}\|^2 / \|\mathbf{y}\|^2} : \frac{\beta}{1 - \beta}. \quad (4.8)$$

Rewriting the statistic L of Eq. (4.7) in the form of Eq. (4.8) yields

$$\beta = \cos^2(\alpha) = \frac{\|\mathbf{s}\|^2}{\|\mathbf{y}\|^2} = \frac{\mathbf{y}^\top \mathbf{P}_{\mathbf{B}}^\perp \mathbf{P}_{\mathbf{G}} \mathbf{P}_{\mathbf{B}}^\perp \mathbf{y}}{\mathbf{y}^\top \mathbf{P}_{\mathbf{B}}^\perp \mathbf{y}}, \quad (4.9)$$

where the statistic is denoted β . Note that \mathbf{y} still passes through $\mathbf{P}_{\mathbf{B}}^\perp$ in both the numerator and the denominator in Eq. (4.9). One can think of this projector as placing \mathbf{y} in the correct signal plane (i.e., in Fig. 7 all the vectors are assumed to lie in the $\mathbf{P}_{\mathbf{B}}^\perp$ subspace).

5. Cramer–Rao Bounds

For many systems the data \mathbf{y} can be modeled as functions of the parameters \mathbf{q} , $\mathbf{y}(\mathbf{q})$, where \mathbf{q} is the unknown range vector. As such, it is desirable to form an estimate of \mathbf{q} , say, $\hat{\mathbf{q}}$, by use of the data \mathbf{y} . The Fisher information matrix $\mathbf{J}(\mathbf{q})$ provides an analysis of a system's sensitivities to the model parameters \mathbf{q} . When we estimate \mathbf{q} , the inverse of the Fisher matrix $\mathbf{J}^{-1}(\mathbf{q})$ provides the Cramer–Rao lower bound¹² on an unbiased estimator's error-covariance matrix $E[\hat{\mathbf{q}} - \mathbf{q}][\hat{\mathbf{q}} - \mathbf{q}]^\top$, where $E[\dots]$ is the expected-value operator. The trace of $\mathbf{J}^{-1}(\mathbf{q})$ lower-

bounds the estimator's error variance, $E[\hat{\mathbf{q}} - \mathbf{q}]^\top [\hat{\mathbf{q}} - \mathbf{q}] \geq \text{tr}[\mathbf{J}^{-1}(\mathbf{q})]$, where $\text{tr}[\dots]$ represents the trace of the elements (in square brackets) of the matrix, or the sum of the diagonal elements. These theoretical relations can be used to guide application-specific system design and provide a basis for accepting or rejecting a particular estimator during the development and the testing phases.

If the model for a linear system has unknown deterministic parameters, $\mathbf{q} = [q_1 q_2 \dots q_P]^\top$, then the Fisher matrix $\mathbf{J}(\mathbf{q})$ is defined as

$$\mathbf{J}(\mathbf{q}) = E \left[\frac{\partial}{\partial \mathbf{q}} \ln p_{\mathbf{q}}(\mathbf{y}) \right] \left[\frac{\partial}{\partial \mathbf{q}} \ln p_{\mathbf{q}}(\mathbf{y}) \right]^\top, \quad (5.1)$$

where $p_{\mathbf{q}}(\mathbf{y})$ is the probability-density function of the data \mathbf{y} that are parameterized by \mathbf{q} .^{12,13}

A. Cramer–Rao Bounds for Range Estimation

For the range-estimation model the unknown parameters are the range r and the object \mathbf{u} , such that $\mathbf{q} = [r \mathbf{u}^\top]^\top$. Furthermore, if we assume that the noise is distributed as $\mathbf{n} : N[\mathbf{0}, \sigma^2 \mathbf{I}]$, the Fisher matrix $\mathbf{J}(\mathbf{q})$

may be written as

$$\begin{aligned} \mathbf{J}(\mathbf{q}) &= \frac{1}{\sigma^2} [\mathbf{G}_r \quad \mathbf{G}_u]^\top [\mathbf{G}_r \quad \mathbf{G}_u], \\ \mathbf{G}_r &= \frac{\partial}{\partial r} \ln p_{\mathbf{q}}(\mathbf{y}) \Rightarrow \frac{\partial}{\partial r} \mathbf{F}(r) \mathbf{u}, \\ \mathbf{G}_u &= \frac{\partial}{\partial \mathbf{u}} \ln p_{\mathbf{q}}(\mathbf{y}) \Rightarrow \mathbf{F}(r), \end{aligned} \quad (5.2)$$

where \mathbf{G} expresses the sensitivities to the respective model parameters. In this form the inverse of $\mathbf{J}(\mathbf{q})$ provides the lower bound on the estimator error, as shown in

$$\mathbf{J}^{-1}(\mathbf{q}) = \sigma^2 \begin{bmatrix} [\mathbf{G}_r^\top \mathbf{P}_{\mathbf{G}_u}^\perp \mathbf{G}_r]^{-1} & * \\ * & [\mathbf{G}_u^\top \mathbf{P}_{\mathbf{G}_r}^\perp \mathbf{G}_u]^{-1} \end{bmatrix}, \quad (5.3)$$

where the projector, $\mathbf{P}_{\mathbf{G}_u}^\perp = \mathbf{I} - \mathbf{G}_u (\mathbf{G}_u^\top \mathbf{G}_u)^{-1} \mathbf{G}_u^\top$, is projecting the sensitivity matrix \mathbf{G}_r onto a subspace that is orthogonal to the subspace spanned by the sensitivity matrix \mathbf{G}_u . The terms represented by the asterisks in Eq. (5.3) are not needed for the CRB derivation.

In Eq. (5.3) the upper left-hand term provides the lower bound on the error variance for the best esti-

mator of r with the unknown object \mathbf{u} , as summarized in

$$\text{var}(\hat{r}) \geq \frac{\sigma^2}{[\mathbf{G}_r^T \mathbf{P}_{\mathbf{G}_u}^\perp \mathbf{G}_r]}. \quad (5.4)$$

Note that relation (5.4) is a rank-1 evaluation ($\mathbf{G}_r^T \mathbf{P}_{\mathbf{G}_u}^\perp \mathbf{G}_r$ is a scalar) because there is only one parameter to be estimated, r .

Expression (5.4) leads to the variance bounds in relation (5.5) as expressed in terms of the frequency response of the system components. Note that the projector $\mathbf{P}_{\mathbf{G}_u}^\perp$ is replaced with a frequency-domain equivalent $P[\exp(j\theta)]$, which takes on nonzero values only when the nuisance sensitivities [defined in Eq. (5.2)] are zero:

$$\begin{aligned} \text{var}(\hat{r}) &\geq \frac{\sigma^2}{\int_{-\infty}^{\infty} P[\exp(j\theta)] \left| \frac{\partial}{\partial r} H[\exp(j\theta); r] U[\exp(j\theta)] \right|^2 d\theta}, \\ P[\exp(j\theta)] &= \begin{cases} 1 & |H[\exp(j\theta); r]| = 0 \\ 0 & \text{else} \end{cases}. \end{aligned} \quad (5.5)$$

From relation (5.5) it is evident that the estimator's performance depends on the spectral content of the object, $U[\exp(j\theta)]$. Although at first relation (5.5) represents a design difficulty because no *a priori* information about the object is assumed, general approximations may be made regarding anticipated objects. As an alternative to range estimation [which is dependent on the object spectra, according to relation (5.5)], the CRB's on object detection are developed in the Subsection 5.B. There it is evident that the object spectra do not play into the system equations.

B. Cramer–Rao Bounds for Object Detection

From Eq. (4.4) and Fig. 6, we can see that the partitioned model parameters for this system are expressed as $\mathbf{q} = [\mathbf{u}_D^T \mathbf{u}^T]^T$, where \mathbf{u}_D is the desired unknown object at some range r_D and \mathbf{u} represents all other unknown objects. Note that, in this situation, the range r_D is assumed to be known. This assumption essentially is the act of looking at a given range and estimating an object or the presence (or the absence) of an object. The Fisher information for such a system may be written as

$$\begin{aligned} \mathbf{J}(\mathbf{q}) &= \frac{1}{\sigma^2} [\mathbf{G}_{\mathbf{u}_D} \quad \mathbf{G}_{\mathbf{u}}]^T [\mathbf{G}_{\mathbf{u}_D} \quad \mathbf{G}_{\mathbf{u}}], \\ \mathbf{G}_{\mathbf{u}_D} &= \frac{\partial}{\partial \mathbf{u}_D} \ln p_q(\mathbf{y}) \Rightarrow \mathbf{F}(r_D), \\ \mathbf{G}_{\mathbf{u}} &= \frac{\partial}{\partial \mathbf{u}} \ln p_q(\mathbf{y}) \Rightarrow \mathbf{F}(r). \end{aligned} \quad (5.6)$$

The inverse of the Fisher matrix in Eq. (5.6) is given by

$$\mathbf{J}^{-1}(\mathbf{q}) = \sigma^2 \begin{bmatrix} [\mathbf{G}_{\mathbf{u}_D}^T \mathbf{P}_{\mathbf{G}_{\mathbf{u}}}^\perp \mathbf{G}_{\mathbf{u}_D}]^{-1} & \\ & * \\ * & [\mathbf{G}_{\mathbf{u}}^T \mathbf{P}_{\mathbf{G}_{\mathbf{u}_D}}^\perp \mathbf{G}_{\mathbf{u}}]^{-1} \end{bmatrix}, \quad (5.7)$$

where the projector $\mathbf{P}_{\mathbf{G}_{\mathbf{u}}}^\perp = \mathbf{I} - \mathbf{G}_{\mathbf{u}}(\mathbf{G}_{\mathbf{u}}^T \mathbf{G}_{\mathbf{u}})^{-1} \mathbf{G}_{\mathbf{u}}^T$ is projecting the sensitivity matrix $\mathbf{G}_{\mathbf{u}_D}$ onto a subspace that is orthogonal to the subspace spanned by the sensitivity matrix $\mathbf{G}_{\mathbf{u}}$. For this system, then, the lower bound on the error variance for the best estimator of the unknown object \mathbf{u}_D is given by

$$\text{var}(\hat{\mathbf{u}}_D) \geq \frac{\sigma^2}{\text{tr}[\mathbf{G}_{\mathbf{u}_D}^T \mathbf{P}_{\mathbf{G}_{\mathbf{u}}}^\perp \mathbf{G}_{\mathbf{u}_D}]} = \frac{\sigma^2}{\text{tr}[\mathbf{F}(r_D)^T \mathbf{P}_{\mathbf{F}(r)}^\perp \mathbf{F}(r_D)]}, \quad (5.8)$$

where $\text{tr}[\dots]$ represents the trace of the terms in square brackets, which represent the sum of the diagonal elements. Expression (5.8) leads to the variance bounds

$$\begin{aligned} \text{var}(\hat{\mathbf{u}}_D) &\geq \frac{\sigma^2}{\int P[\exp(j\theta)] H[\exp(j\theta); r_D]^2 d\theta}, \\ P[\exp(j\theta)] &= \begin{cases} 1 & |H[\exp(j\theta); r]| = 0 \\ 0 & \text{else} \end{cases}, \end{aligned} \quad (5.9)$$

expressed in terms of the frequency response of the system components. Note that the projector $\mathbf{P}_{\mathbf{G}_{\mathbf{u}}}^\perp$ is replaced with a frequency-domain equivalent $P[\exp(j\theta)]$, which takes on values of 1 only when the MTF for the nuisance parameter [defined in Eq. (5.6)] is zero. In relation (5.9) note that the estimator-error performance does *not* depend on the spectral content of the object.

Because objects are to be detected rather than estimated, we modify relation (5.8) to evaluate the variance in the detectable object energy, $\mathbf{u}^T \mathbf{u}$. Note that the subscript D has been dropped from the \mathbf{u} terms. As such, the threshold detection of an estimated object $\hat{\mathbf{u}}$ by use of a measure of the estimated object's energy $\hat{\mathbf{u}}^T \hat{\mathbf{u}}$ from the image data \mathbf{y} is of interest. This evaluation evolves as a modification of the Fisher matrix derived in the Subsection 5.A, as summarized below in relations (5.10) and (5.11). When we compute the variance of a function of a parameter ($\hat{\mathbf{u}}^T \hat{\mathbf{u}}$ is a function of the original parameter $\hat{\mathbf{u}}$) the Fisher matrix is modified by consideration of the sensitivity of the function to the original parameter.¹³ This modification is manifest in the Fisher matrix as a linear transformation \mathbf{A} , as shown in

$$\mathbf{J}(\mathbf{u}^T \mathbf{u}) = \mathbf{A}^{-1} \mathbf{J}(\mathbf{u}) \mathbf{A}^{-T}, \quad \mathbf{A}_{ij} = \left[\frac{\partial}{\partial u_i} u_j^2 \right]. \quad (5.10)$$

Note that the elements of the object \mathbf{u} (i.e., u_i and u_j) are now included in the CRB evaluation through

Table 1. Simulation Parameter Values

Simulation Parameter	Resolution R	
	4 m	20 mm
Focal length f	60 mm	10 mm
Aperture width d	5 mm	0.35 mm
Mask beta β	2	2
Illumination λ_{eff}	550 nm	550 nm
Image distance $f = z_i$	60 mm	10 mm
Object distance $r = z_o$	1–4 m	5–20 mm

the transformation \mathbf{A} , bringing the object content into question. The variance on estimating $\mathbf{u}^T \mathbf{u}$ is bounded as shown in

$$\text{var}(\hat{\mathbf{u}}^T \hat{\mathbf{u}}) \geq \frac{\sigma^2}{\text{tr}[\mathbf{A}\mathbf{F}(r_D)^T \mathbf{P}_{F(r)}^+ \mathbf{F}(r_D)\mathbf{A}]}. \quad (5.11)$$

Effects of the system parameters on the variance bound can be seen in relation (5.11) by consideration of a spectrally white object \mathbf{u} , where $u_k = 1/N$, $k = -N/2, \dots, N/2$ for an N -length object \mathbf{u} . The transformation \mathbf{A} becomes $\partial/\partial u_i(u_j^2)$, which is $2u$ when $i = j$ and zero otherwise. Recall that $\mathbf{u} = \mathbf{B}\mathbf{a}$ and that \mathbf{B} is a matrix of MTF's, so \mathbf{u} is the object's spectral representation, which was defined to be white. Thus the terms u_i are independent. For an N -length object \mathbf{u} , the matrix \mathbf{A} therefore becomes a diagonal matrix with $2/N$ on the diagonal, $\mathbf{A} = (2/N)\mathbf{I}$.

C. Simulated Optical Systems

Two optical systems are simulated, one that is capable of ranging from 1 to 4 m and another that is capable of ranging from 5 to 20 mm. The former system is similar to that which might be employed in general proximity-sensing applications, and the latter is similar to a miniature inspection system or a medical scope design. The optical parameters for each system are given in Table 1.

Both designs contain a simple lens, an aperture, an illumination bandpass filter, and the cosine wavefront coding mask as introduced in Section 3. The generalized optical system is shown in Fig. 8. Example MTF's for the 4-m and the 20-mm optical sys-

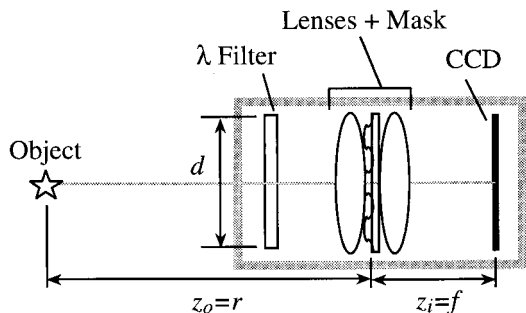


Fig. 8. Schematic of the object and the simulated optical system. Note that r is measured from the first principle plane of the optical system. The λ filter is used to reduce infrared wavelengths.

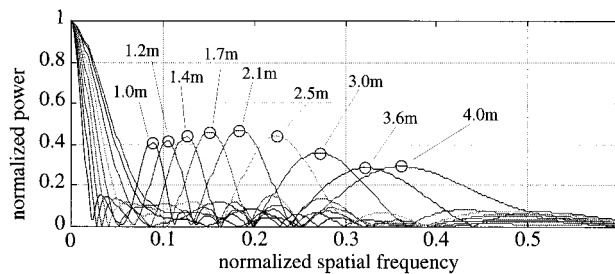


Fig. 9. Example MTF's for the simulated 1–4-m system. The peaks are marked with the range for the simulated MTF.

tems described above are shown in Figs. 9 and 10, respectively, in which the MTF's for several ranges have been plotted.

The peak locations marked in Fig. 9 show the ranges at which each MTF was simulated. The relation of the range (in physical units) to the peak location (in normalized frequency coordinates) forms a calibration for the system. This calibration is used to determine the most likely range from the maximum peak location in a given spectral estimate. Observations to be made between Figs. 9 and 10 include the fact that the MTF's look remarkably similar for two distinctly different optical systems. As a design rule, one can shuffle both the position of the close-range peak and the position of the far-range peak by adjusting both the focal length and the focus of the system. Changes in the aperture size and the mask value β can also be used to help vary the sharpness of the peaks for a given frequency location.

For object detection, the $N \times p$ model \mathbf{B}_0 contains a discrete number of modes p , and N is a conveniently large FFT length. In this case p is related to the range resolution of the object-detection system. For a set of object distances (i.e., 1 to 4 m) the displacement between modes in the model is chosen to be a percentage of the distance. If the detector resolution is 10%, or $R_{\text{pct}} = 0.10$, the detection space is divided as $R_{\text{next}} = (1 + R_{\text{pct}})R_{\text{this}}$. Thus R_{next} is the range at which the next MTF (i.e., the next column of \mathbf{B}_0) will be evaluated. One example system \mathbf{B}_0 is evaluated with a range resolution of 4%, or $R_{\text{pct}} = 0.04$. At lower percentages (higher resolutions) the farther range-estimate variance exceeds the closer estimates by several orders of magnitude. This difference in-

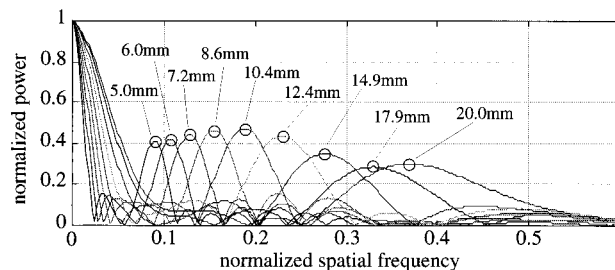


Fig. 10. Example MTF's for the simulated 5–20-mm system. Several peaks are marked with the range for the simulated MTF.

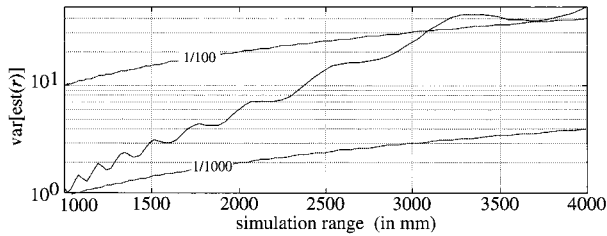


Fig. 11. Plot of the CRB range-estimate variance as a function of range for the 4-m system design. Note that the variance is near 1/1000 at 1 m and approximately 1/100 at 4 m.

dicates that, for this detector, the MTF peaks are too broad at farther ranges to resolve reliably objects that are separated by less than 4% of the range, or approximately 16 cm in the 4-m design and approximately 0.8 mm in the 20-mm design, at the farthest ranges.

D. Cramer–Rao Bound Results for Range Estimation

For the system parameters given in Subsection 5.C and a signal-to-noise ratio of 45 dB (typical of a noisy CCD camera) Fig. 11 shows that one can reasonably estimate the range to a spectrally white object with errors of the order of 1/100 to 1/1000 at the closer ranges in the 4-m system. The variance begins to exceed 1/100 as the range approaches 4 m. Figure 12 shows that one can reasonably estimate the range to a spectrally white object with errors of less than 1/1000 in the 20-mm system.

In this subsection only the single-range estimation task is explored by use of spectrally white objects. Although the results are compelling (in general, range errors are expected to be less than 0.1%), most practical applications will not be detecting point sources or such broadband objects. As such, a more robust object-detection technique is employed in Subsection 5.E., in which the detector is based on subspace-projection methods and is designed to work with sources of varying spectral distributions.

E. Cramer–Rao Bound Results for Object Detection

For the system parameters given in Subsection 5.D and a signal-to-noise ratio of 45 dB Fig. 13 shows that one may reasonably detect objects with errors of the order of 1/10⁶ by using the 4-m system. Figure 14 shows that one can reasonably detect a spectrally

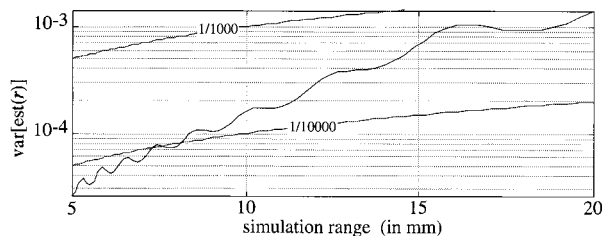


Fig. 12. Plot of the CRB range-estimate variance as a function of range for the 20-mm system design. Note that the variance is less than 1/1000 for the entire simulated range.

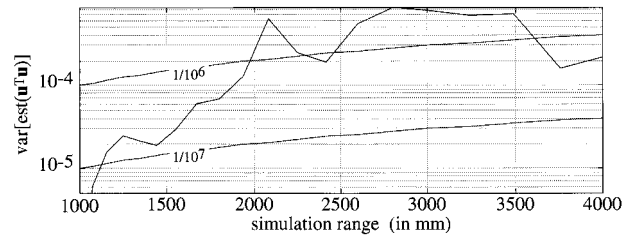


Fig. 13. Plot of the CRB object-detection variance as a function of range for the 4-m system design by use of a 4% range resolution. The variance is much less than 1/1000.

white object with errors of less than 1/1000 by using the 20-mm system. Although at first these values may appear to be extremely low, it should be reiterated that the systems are operating at only a 4% resolution. At higher resolutions the variance for the farther ranges will exceed 1/10 and higher. In general, object-detection errors are expected to be less than 0.1%, similar to the range estimates for the same point source.

6. Experimental Results

To validate the point-source object results provided in the CRB evaluations, we use a 15- μ m-diameter hole in an opaque material as a transmission source. The experimental setup allows for the testing of ranges nominally from 1 to 2 m, considerably shorter than the ranges at which the simulations were performed. Furthermore, the system is not focused to infinity to stretch the MTF responses over this range. Rather, it is focused to just beyond the experimental region. The diagram shown in Fig. 15 shows the basic setup.

A. Experimental Data Quality

The size of the pinhole, detector noise, and difficulties with controlling camera exposure time and source

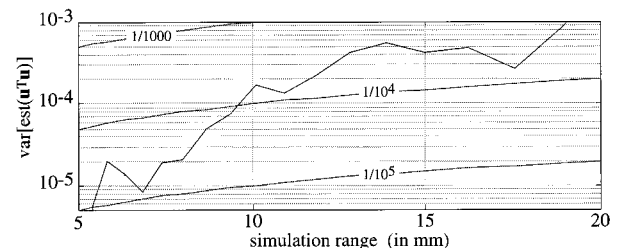


Fig. 14. Plot of the CRB object-detection variance as a function of range for the 20-mm system design with a 4% range resolution. Note that the variance is still less than 1/1000.

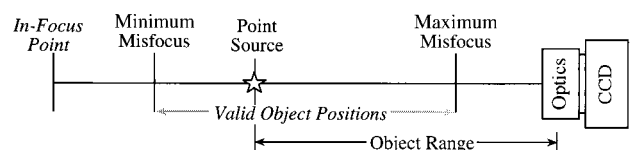


Fig. 15. Experimental setup.

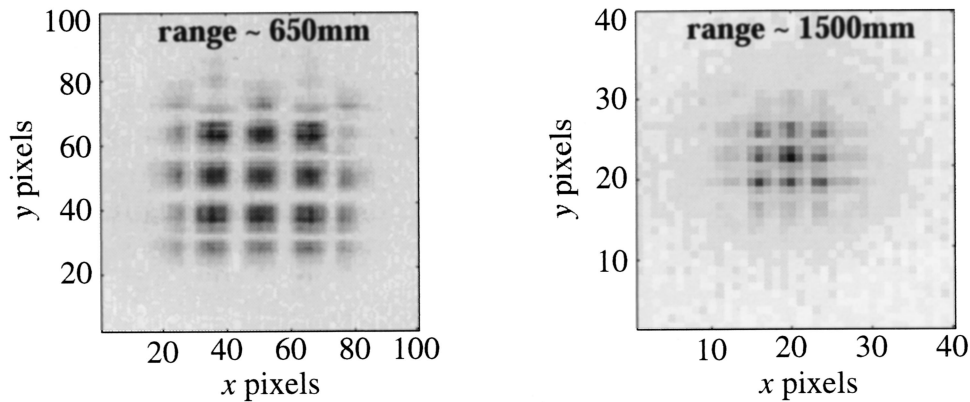


Fig. 16. Images of a point-source object located approximately 650 mm and 1.5 m away from the principle plane of the experimental optical system.

illumination combine to provide images with a small dynamic range: 14 to 33 levels out of an 8-bit camera (i.e., 5% to 13% of the dynamic range). As such, the signal-to-noise ratio was quite low, and the image contrast was considerably low. Example images of a point source at 650 mm and 1.5 m are given in Fig. 16.

Because the mask is cosinusoidal, it consists of both a phase piece and an amplitude piece. Cementing the assembly under a microscope produced alignment and other quality issues. Optical aberrations and misalignment of the lenses further contribute to degraded performance of the experimental system. The nonuniformity of the peaks in the captured PSF in the contour plots of Fig. 17 highlight these issues as well as the low illumination (i.e., the high noise) level. A comparison of the spatial-frequency estimates of Fig. 17 is shown in Fig. 18. Note that there are two estimates for each image, a horizontal and a vertical plot. The CCD pixels are rectangular ($11 \mu\text{m} \times 13 \mu\text{m}$), so there are two spatial sampling frequencies available.

Figure 18 highlights the broadened (and the lowered) peak responses of the experimental system. The broadening is due in part to the high noise level and, more importantly, to the optical problems men-

tioned above and the use of broadband illumination in the experiments (the CRB investigations described in Section 5 used monochromatic illumination). The shift in peak locations between experiment and simulation is due to the defocused (from infinity) experimental system; the image simulator was focused at infinity.

B. Experimental Results for Range Estimation

The experiments described in this subsection evaluate the spectra for point-source objects with nominal distances of 650 to 1500 mm. At each range tested 10 instances of an image are acquired and processed. Note that each instance is an average of five images, which helps to reduce noise because the illumination is low. The center region (sample instance) of the acquired image is Fourier processed, as explained in Subsection 4.A, to produce spectral estimates for that image.

Figure 19 shows that the standard deviation in the range estimates for 10 instances is less than 0.50% of the object range. The standard deviation shown in Fig. 19 is close to the 0.10% variance bound (i.e., $\text{var} = 0.1^{1/2}$ is a standard deviation of 0.316%), as described in Section 5 for the 4-m design.

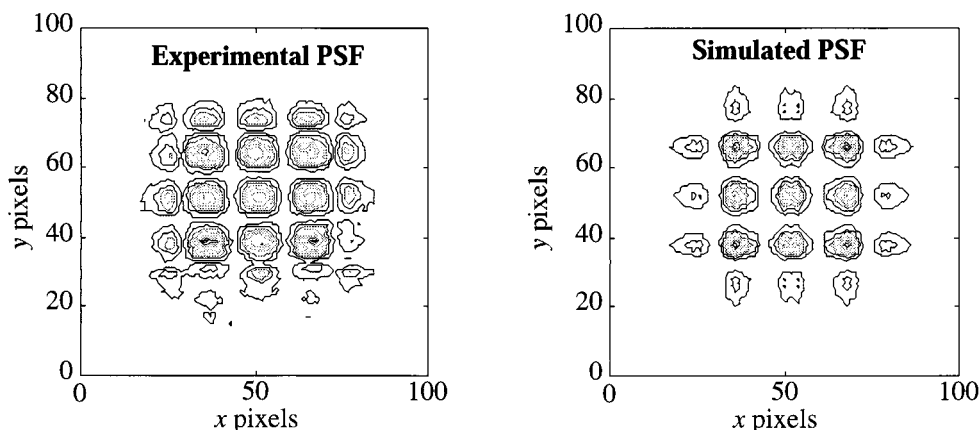


Fig. 17. Experimental and simulated PSF's at nominally 1 m.

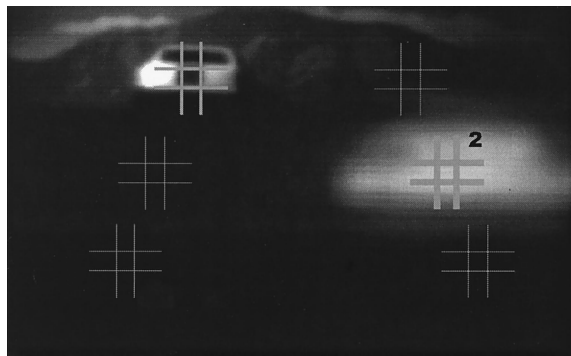


Fig. 22. Detector locations for traffic monitoring.

treat the data shown in Fig. 21 differently. The categories may be translated into more meaningful items such as tracking, warning, proximal, and the like.

Figure 22 provides another type of proximity-sensing situation, this time with traffic monitoring. Six regions are analyzed to detect the presence of a vehicle. The center rightmost sensor detected a vehicle and estimated its range to be near bin 2. The upper leftmost sensor detected a vehicle but cannot assign it a bin number because the threshold level for proper mode identification was not reached.

7. Conclusions

Two schemes for passive ranging by use of a single wave-front coded image from a single aperture have been discussed: (i) estimating the range to an object and (ii) detecting objects over a given set of ranges. CRB evaluations showed that range estimates based on optical system parameters with less than 0.1% error are possible. Experimental range estimates with approximately 0.2% error support the CRB evaluations. Furthermore, the experimental system has also been used in a detector mode in which detected objects were placed in a proximity map. This qualitative example of a detection system shows the capability of building a range map that is based on wave-front coding. Proximity detectors are in widespread use in simple applications. Many such detectors are Boolean, sensing if something is there or not. Wave-front coding offers an attractive passive alternative for both proximity-sensing and ranging systems.

Future research is directed at the design of other

optical masks for passive ranging as well as enhancing the digital signal-processing algorithms currently in use. Masks that are nonabsorbing (the cosinusoidal mask has 50% absorption) are desired for light-sensitive applications. Fast and adaptive algorithms for proximity detection and range mapping are desired for systems with limited processing power.

This material is based on research that was supported in part by U.S. Army Research Office grant DAAG 55-97-1-0348 and by the Colorado Advanced Technology Institute.

References

1. B. Bhanu, S. Das, P. Symosek, S. Snyder, and B. Roberts, "Synergism of binocular and motion stereo for passive ranging," *IEEE Trans. Aerosp. Electron. Sys.* **30**, 709–721 (1994).
2. M. Subbarao, T. Choi, and A. Nikzad, "Focusing techniques," *Opt. Eng.* **32**, 2824–2836 (1993).
3. M. Subbarao and Y.-F. Liu, "Analysis of defocused image data for 3D shape recovery using a regularization technique," in *Three-Dimensional Imaging and Laser-Based Systems for Metrology and Inspection III*, A. G. Harding and D. J. Svetkoff, eds., Proc. SPIE **3204**, 24–73 (1997).
4. N. Goldberg, *Camera Technology: The Dark Side of the Lens* (Academic, New York, 1992).
5. E. R. Dowski and W. T. Cathey, "Single-lens, single-image, incoherent passive ranging systems," *Appl. Opt.* **33**, 6762–6773 (1994).
6. E. R. Dowski, "Passive ranging with an incoherent optical system," Ph.D. dissertation (Department of Electrical and Computer Engineering, University of Colorado, Boulder, Colo., 1993).
7. J. Curlander, E. R. Dowski, and R. McCoy, "Passive three-dimensional location and tracking for autonomous rendezvous," SBIR Phase I Final Rep. (Vexcel Corporation, 4909 Nautilus Court, Boulder, Colo. 80301, 1995).
8. E. R. Dowski, "An information theory approach to incoherent information processing systems," in *Digest of the Topical Meeting on Signal Recovery and Synthesis V* (Optical Society of America, Washington, D.C., 1995), pp. 106–108.
9. L. L. Scharf and B. Friedlander, "Matched subspace detectors," *IEEE Trans. Signal Process.* **42**, 2146–2157 (1994).
10. R. T. Beherens and L. L. Scharf, "Signal processing applications of oblique projection operators," *IEEE Trans. Signal Process.* **42**, 1413–1423 (1994).
11. S. Kraut and L. L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. Signal Process.* **47**, 2538–2541 (1999).
12. L. L. Scharf and L. McWhorter, "Geometry of the Cramer–Rao bound," *Signal Process.* **31**, 301–311 (1993).
13. L. L. Scharf, *Statistical Signal Processing* (Addison-Wesley, Reading, Mass., 1991), pp. 209–276.