AN INFORMATION THEORY APPROACH TO THREE INCOHERENT INFORMATION PROCESSING SYSTEMS

Edward R. Dowski, Jr.

Imaging Systems Laboratory
Department of Electrical & Computer Engineering
University of Colorado
Boulder, Colorado
80309-0425

1. INTRODUCTION

Many incoherent optical/digital systems can be used for non-imaging purposes, such as passive ranging. These systems cannot effectively be analyzed or designed in terms of traditional image-forming systems. Instead, such systems should be analyzed in terms of information theory. Through mathematical modelling of the sampled image, information theory can be used to optimize a given system.

The Fisher Information matrix and the Cramer-Rao bound are two widespread and tractable measures of the information content of a signal. By analyzing information processing systems in terms of these measures of information, illuminating conclusions that are related to optimum design can be found. Such conclusions lead to necessary conditions that particular types of information processing systems must possess. We consider these necessary conditions for three incoherent optical systems, namely passive range estimation, extended depth of field, and passive range detection.

2. FISHER INFORMATION AND THE CRAMER-RAO BOUND

Let the parameters desired from a given estimation system be denoted by the length p vector $\underline{\Phi}$

$$\underline{\Phi} = [\phi_1, \phi_2, \cdots, \phi_p]^T \qquad (1)$$

where the noiseless measurement is some vector function of these parameters, say $\underline{x}(\underline{\Phi})$. The actual measurement in any real system will always be corrupted by noise. The limit of this noise will be signal dependent shot noise or detector quantization noise. We assume a zero mean white gaussian noise with variance σ^2. Our ability, on the average, to estimate $\underline{\Phi}$ is bounded by the Cramer-Rao bound [1, 2, 3]. This bound can describe both biased and unbiased estimators. We consider here only unbiased estimators. The variance of any unbiased estimator of one component of $\underline{\Phi}$, say ϕ_i, is bounded below as

$$var(\hat{\phi_i}) \geq J_{ii}^{-1}(\underline{\Phi}) \qquad (2)$$

where $J(\underline{\Phi})$ is the Fisher Information matrix of the parameter $\underline{\Phi}$ [2, 3]. Let $p(\underline{y};\underline{\Phi})$ be the probability density function for the observed noisy data $\underline{y}$. The Fisher information matrix is then given by

$$J(\underline{\Phi}) = E\left[\frac{\partial}{\partial \underline{\Phi}}\, ln\ p(\underline{y};\underline{\Phi})\right]\left[\frac{\partial}{\partial \underline{\Phi}}\, ln\ p(\underline{y};\underline{\Phi})\right]^T \qquad (3)$$

where E denotes expected value. Under the zero mean white gaussian noise assumption (3) reduces [4] to

$$J(\underline{\Phi}) = \frac{1}{\sigma^2} G^T G; \quad G^T = \frac{\partial \underline{x}^T(\underline{\Phi})}{\partial \underline{\Phi}} \qquad (4)$$

The matrix G is called a sensitivity matrix. Assume that the parameter $\underline{\Phi}$ is partitioned into two sets so that $\underline{\Phi} = [\underline{\phi}_1^T, \underline{\phi}_2^T]^T$ By partitioning the matrix G of (4) as

$$G = [G_1\ G_2]; \quad G_1 = \frac{\partial \underline{x}(\underline{\Phi})}{\partial \underline{\phi}_1}; \quad G_2 = \frac{\partial \underline{x}(\underline{\Phi})}{\partial \underline{\phi}_2} \qquad (5)$$

we can show that the inverse of the Fisher Information matrix of (3) is given by [4]

$$J^{-1}(\underline{\Phi}) = \sigma^2 \begin{bmatrix} [G_1^T\ P_{G_2}^{\perp}\ G_1]^{-1} & \star \\ \star & [G_2^T\ P_{G_1}^{\perp}\ G_2]^{-1} \end{bmatrix} \qquad (6)$$

where

$$P_{G_i}^{\perp} = I - G_i(G_i^T G_i)^{-1} G_i^T$$

is a projection matrix projecting onto the space orthogonal to the space spanned by the matrix G_i, or

$< G_i >$. The identity matrix is given by I. Application of (6) to specific information processing systems leads to estimates of the theoretical performance limit for specific systems, while also determining necessary the conditions for such systems.

3. SINGLE-IMAGE, SINGLE-LENS, PASSIVE RANGE ESTIMATION SYSTEMS

Incoherent Single-image, single-lens, passive range estimation systems code the object range information into unambiguous spatial information at the image. Such systems should be insensitive to the characteristics of the particular object. The noiseless sampled image model of this system is

$$\underline{x} = \underline{h}(\psi) * \underline{u} = F(\psi)\underline{u} \qquad (7)$$

where the system point spread function (PSF) or impulse response is given by $\underline{h}(\psi)$. This PSF is characterized by the misfocus or normalized range parameter ψ. The symbol $*$ denotes convolution. The matrix $F(\psi)$ is a convolution matrix containing $\underline{h}(\psi)$ as elements. The unknown object is given by $\underline{u}$.

The unknown parameters of the noiseless sampled data are the normalized range ψ and the object $\underline{u}$. The system, described by $\underline{h}(\psi)$ and $F(\psi)$, is assumed to be known. The unknown parameters can be grouped as

$$\underline{\Phi} = [\phi_1, \underline{\phi}_2^T]^T, \quad \phi_1 = \psi, \quad \underline{\phi}_2 = \underline{u}$$

With this partitioning of the parameters, the sensitivity matrices from (5) are found to be

$$G_1 = \frac{\partial}{\partial \psi} F(\psi)\underline{u}, \quad G_2 = F(\psi) \qquad (8)$$

From (6) the Cramer-Rao bound on estimating the normalized range ψ with unknown $\underline{u}$ is given by

$$var(\hat{\psi}) \geq \frac{\sigma^2}{G_1^T\ P^{\perp}_{F(\psi)}\ G_1} \qquad (9)$$

where $P^{\perp}_{F(\psi)}$ is a projection matrix projecting onto the subspace orthogonal to $< F(\psi) >$. From (9) we conclude that for passive range estimation to be possible $P^{\perp}_{F(\psi)}$ must not be a rank-zero projection, or equivalently $F(\psi)$ must not be full rank. The Cramer-Rao bound of (9) also shows that the variation of $F(\psi)$ with ψ, in the intersection of the rank-one subspace $< \underline{u} >$ and the subspace orthogonal to $< F(\psi) >$, must be large for accurate range estimation.

Since the matrix $F(\psi)$ can be approximated by a circulant matrix, the eigenvalues of $F(\psi)$ are approximately the DFT values of $\underline{h}(\psi)$ [5]. The denominator of (9) can then be approximated as

$$G_1^T\ P^{\perp}_{F(\psi)}\ G_1 \approx \int P(e^{j\theta}) |\frac{\partial}{\partial \psi} H(\psi, e^{j\theta})|^2 |U(e^{j\theta})|^2 \, d\theta \qquad (10)$$

where

$$P(e^{j\theta}) = \begin{cases} 1 & \text{if } |H(\psi, e^{j\theta})| = 0 \\ 0 & \text{otherwise} \end{cases} \qquad (11)$$

and where $H(\psi, e^{j\theta})$ and $U(e^{j\theta})$ are the DFTs of the sampled vectors $\underline{h}(\psi)$ and $\underline{u}$ respectively. Notice that $H(\psi, e^{j\theta})$ is the optical transfer function (OTF) of the sampled system with misfocus ψ, while $U(e^{j\theta})$ is the spatial frequency spectrum of the sampled object.

From (9) and (10), a necessary condition for a passive range estimation systems is that the OTF must contain zeros that are a function of misfocus or normalized range ψ. It is impossible to build a general passive range estimation system whose OTF does *not* contain zeros as a function of misfocus.

With *a priori* information of the object spatial frequency spectrum, from (10) a desired condition on the system is that the magnitude squared variation of the OTF should be a matched filter for the expected object spatial frequency power spectrum.

4. EXTENDED DEPTH OF FIELD SYSTEMS

In an incoherent extended depth of field system, unknown objects are estimated without knowledge of misfocus or object range. The model for the noiseless sampled image is identical to the model of the passive ranging system

$$\underline{x} = \underline{h}(\psi) * \underline{u} = F(\psi)\underline{u} \qquad (12)$$

The difference between the passive ranging problem and the extended depth of field problem is that instead of estimating object range ψ without knowledge of the object $\underline{u}$, the unknown object $\underline{u}$ is estimated without knowledge of the range ψ.

Partition the parameter vector as in the passive ranging problem. The sensitivity matrices G_1 and G_2 are then the same as in (8). The Cramer-Rao bound on the estimate of the unknown object $\underline{u}$ from (6) is then

$$var(\hat{\underline{u}}) \geq \frac{\sigma^2}{\text{trace}(F(\psi)^T\ P^{\perp}_{G_1}\ F(\psi))} \qquad (13)$$

where $P_{G_1}^{\perp}$ is the projection matrix projecting onto the subspace orthogonal to $< G_1 >$. The relationship of (13) then implies that the variation of $F(\psi)$ with ψ, in the rank-one subspace $< \underline{u} >$, should be zero. Additionally, the eigenvalues of $F(\psi)^T F(\psi)$ should be large.

The denominator of (13) can be placed in terms of the OTF and spatial frequency spectrum of $\underline{u}$, as in (10), by

$$\text{trace}(F(\psi)^T \ P_{G_1}^{\perp} \ F(\psi)) \approx \int P(e^{j\theta})|H(\psi, e^{j\theta})|^2 d\theta \tag{14}$$

where

$$P(e^{j\theta}) = \begin{cases} 1 & \text{if } |\frac{\partial}{\partial \psi} H(\psi, e^{j\theta}) U(e^{j\theta})| = 0 \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

From (13) and (14), the necessary condition on a reliable extended depth of field system is that the variation of the OTF with misfocus should be zero over the spatial frequencies where the expected object has non-zero power. A desired condition is that the power of the OTF should be maximized where its variation with misfocus, or the object spatial frequency spectrum, is zero. From (15), the performance of this type of system is independent of the unknown object if the object spatial frequency spectrum contains no zero values in the passband of the system.

5. PASSIVE RANGE DETECTION SYSTEMS

In an incoherent passive range detection system the sum of images of unknown objects from a number of known ranges is sampled. This can be modelled as

$$\underline{x} = \sum_i \underline{h}(\psi_i) * \underline{u}_i = \sum_i F(\psi_i) \underline{u}_i \tag{16}$$

The normalized range values ψ_i are assumed known while the objects at the different ranges, $\underline{u}_i$, are assumed unknown. For simplicity consider the imaging at two ranges, $i = 1, 2$. The results can be generalized to any number of ranges.

Partition the parameter vector for the sampled image as

$$\underline{\Phi} = [\underline{\phi}_1^T, \underline{\phi}_2^T]^T; \ \ \underline{\phi}_1 = \underline{u}_1, \ \ \underline{\phi}_2 = \underline{u}_2 \tag{17}$$

The sensitivity matrices of (5) then become

$$G_1 = F(\psi_1), \ \ G_2 = F(\psi_2) \tag{18}$$

The Cramer-Rao bound on the estimation of $\underline{u}_1$ with unknown $\underline{u}_2$ is then

$$var(\underline{\hat{u}}_1) \geq \frac{\sigma^2}{\text{trace}(F(\psi_1)^T \ P_{F(\psi_2)}^{\perp} \ F(\psi_1))} \tag{19}$$

The projection matrix of (19) projects onto the subspaces orthogonal to $< F(\psi_2) >$. For reliable operation at distinct ranges ψ_1 and ψ_2, the system should possess orthogonal subspaces $< F(\psi_1) >$ and $< F(\psi_2) >$. And, the eigenvalues of $F(\psi_1)^T F(\psi_1)$, within the subspace $< P_{F(\psi_2)}^{\perp} >$, should be maximized. Parallel statements to those above can be made for the estimation of object $\underline{u}_2$.

Again, this bound can be placed in terms of the OTF and spatial frequency spectrum of the object as

$$\text{trace}(F(\psi_1)^T \ P_{F(\psi_2)}^{\perp} \ F(\psi_1)) \approx \int P(e^{j\theta})|H(\psi_1, e^{j\theta})|^2 \ d\theta \tag{20}$$

where

$$P(e^{j\theta}) = \begin{cases} 1 & \text{if } |H(\psi_2, e^{j\theta})| = 0 \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

From (19) and (20), the necessary condition for passive range detection systems is that the expected OTFs, as a function of misfocus, should form an orthogonal set. A desired condition is that the total power of each expected OTF should be maximized and equalized among all expected OTFs. Notice that the performance of this type of system is independent of the unknown objects.

6. REFERENCES

[1] H. W. Sorenson, *Parameter Estimation.* New York: Marcel Dekker, Inc., 1980.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: John Wiley & Sons, Inc, 1991.

[3] L. L. Scharf, *Statistical Signal Processing.* Reading, Massachusetts: Addison-Wesley, 1991.

[4] L. L. Scharf and L. McWhorter, "Geometry of the Cramer-Rao Bound," *Signal Processing,* vol. 31, pp. 301–311, April 1993.

[5] B. Noble and J. W. Daniel, *Aplied Linear Algebra.* Englewood Cliffs, NJ: Prentice-Hall, 1988.