A foundation for logarithmic measures of fluctuating intensity in pattern recognition

Nicholas C. Makris

Naval Research Laboratory, Washington, D.C. 20375

Received April 10, 1995

Independent analyses using Fisher information, optimal filtering, and information theory show that matched filtering images with hypothetical patterns in the logarithmic domain provides an optimal method for pattern recognition in the presence of signal-dependent noise arising from complex Gaussian fluctuations in the received fields. This provides a mathematical justification for the use of logarithmic units (i.e., decibels) in a variety of engineering applications.

The stochastic behavior of optical, radar, and acoustic fields received from both fluctuating sources and scatterers can often be well approximated with circular complex Gaussian random (CCGR) variables.[1–4] Averaged intensity from a CCGR field has a standard deviation proportional to the mean.[1] Therefore intensity images derived from CCGR fields have signal-dependent noise. Taking the logarithm of such intensity images homomorphically transforms[5] the signal-dependent noise into additive signal-independent noise.[6] It is shown that matched filtering such images with hypothetical patterns in the logarithmic domain provides an optimal method for pattern recognition according to the independent perspectives offered by minimum variance unbiased estimation with Fisher information, optimal filtering, and information theory.

Let the vector $\mathbf{W}$ contain the independent averaged intensity measurements W_k assigned to $k = 1, 2, 3 \ldots, N$ pixels in an image. Let the vector $\mathbf{a}$ contain the parameters a_i to be estimated from the image $\mathbf{W}$ for $i = 1, 2, 3 \ldots, N_a$. Assuming that the W_k are measured from CCGR fields, the conditional probability distribution for image $\mathbf{W}$ given parameter vector $\mathbf{a}$ is the product of gamma distributions[1]:

$$P(\mathbf{W}|\mathbf{a}) = \prod_{k=1}^{N} \frac{\left[\dfrac{\mu_k}{\sigma_k(\mathbf{a})}\right]^{\mu_k} (W_k)^{\mu_k - 1} \exp\left[-\mu_k \dfrac{W_k}{\sigma_k(\mathbf{a})}\right]}{\Gamma(\mu_k)} . \tag{1}$$

For pixel k, the mean intensity $\langle W_k \rangle = \sigma_k(\mathbf{a})$ explicitly depends on the parameters $\mathbf{a}$ to be estimated, and the quantity μ_k is the number of coherence cells in the intensity average.[1] This number is equal to the signal-to-noise ratio (SNR) $\langle W_k \rangle^2 / (\langle W_k^2 \rangle - \langle W_k \rangle^2)$, which is approximately equal to the time–bandwidth product of the received field for the given pixel (or the number of image-plane speckles spatially averaged[1,6]). The gamma distribution for W_k is denoted $\mathcal{G}(\langle W_k \rangle,\ \mu_k)$. The log-transformed image $\mathbf{L}$ is defined by $L_k = \ln(W_k / I_{\mathrm{ref}})$, where I_{ref} is the reference intensity. This transformed image obeys the conditional distribution[6]

$$P(\mathbf{L}|\mathbf{a}) = \prod_{k=1}^{N} \frac{\left[\dfrac{\mu_k}{\sigma_k'(\mathbf{a})}\right]^{\mu_k} \exp\left[-\mu_k \dfrac{\exp(L_k)}{\sigma_k'(\mathbf{a})} + \mu_k L_k\right]}{\Gamma(\mu_k)} , \tag{2}$$

where $\sigma_k'(\mathbf{a}) = \sigma_k(\mathbf{a}) / I_{\mathrm{ref}}$. The expectation value of L_k is $\ln[\sigma_k'(\mathbf{a})] + \psi(\mu_k) - \ln\, \mu_k$, and the variance is $\zeta(2,\, \mu_k)$. Here $\psi(\mu_k)$ is Euler's psi function[7] and $\zeta(2,\, \mu_k)$ is Riemann's zeta function.[7] For example, $\zeta(2,\, 1) = \pi^2/6$, and in the limit $\mu_k \gg 1$, $\langle L_k \rangle \approx \ln[\sigma_k'(\mathbf{a})]$ and the variance is $\zeta(2,\, \mu_k) \approx 1/\mu_k$. The exponential-gamma distribution for L_k is denoted $\mathcal{E}\mathcal{G}(\langle L_k \rangle,\ \mu_k)$.

According to estimation theory,[8,9] the Cramer–Rao lower bound[8,9] (CRLB) $E[(\hat{a}_i - a_i)^2] \geq [\mathbf{J}^{-1}]_{ii}$ limits the minimum mean-square estimation error for any unbiased estimate $\hat{a}_i$ of the true parameter value a_i from measurements $\mathbf{Y}$, where $\mathbf{J}$ is the Fisher information matrix[8,9] with elements

$$J_{ij} = -E\left[\frac{\partial^2}{\partial a_i \partial a_j} \ln P(\mathbf{Y}|\mathbf{a})\right]. \tag{3}$$

The Fisher information matrices for image $\mathbf{W}$ and its log transform $\mathbf{L}$ are identical and found to be equal to

$$J_{ij} = \sum_{k=1}^{N} \left[\mu_k \frac{\partial \ln \sigma_k(\mathbf{a})}{\partial a_i} \frac{\partial \ln \sigma_k(\mathbf{a})}{\partial a_j}\right], \tag{4}$$

where $\sigma_k(\mathbf{a})$ can be replaced by $\sigma_k'(\mathbf{a}) = \sigma_k(\mathbf{a}) / I_{\mathrm{ref}}$ to reference the logarithmic measure to physical units without altering the Fisher information. Some useful applications of Eq. (4) are given in Ref. 4. The Fisher information matrix of Eq. (4) may equivalently be written as

$$J_{ij} = \sum_{k=1}^{N} \left(\mu_k \frac{\partial \langle L_k \rangle}{\partial a_i} \frac{\partial \langle L_k \rangle}{\partial a_j}\right). \tag{5}$$

Equation (5) shows that the CRLB can be computed directly from the expectation value of logarithmic intensity measures. It is significant because it shows that Fisher information is contained in (1) the variation of the expectation value of a logarithmic measure of image intensity with respect to the parameters to be estimated and (2) the number of coherence cells averaged for each pixel. It is interesting that the Fisher information matrix can be obtained directly from the Gaussian form of $P(\mathbf{L}|\mathbf{a})$ or $P(\mathbf{W}|\mathbf{a})$ in the asymptotic limit $\mu_k \gg 1$ because $\mathbf{a}$ is not a function of μ_k.[4] This is consistent with Fisher's use of the central limit theorem in his initial derivation[10] of what became known as Fisher information.

The efficiency of logarithmic measures is particularly evident in the estimation of a single parameter a from a single measurement L. Here the minimum estimation error is $\{E[(\hat{a} - a)^2]\}^{1/2} = [\sqrt{\mu}|(\partial\langle L\rangle/\partial a)|]^{-1}$. The optimal resolution of the parameter is inversely proportional to the parameter's slope magnitude over the expectation value of a logarithmic measurement of intensity. For example, if the parameter to be estimated is the expected intensity level $\langle L\rangle$, as is often the case in engineering applications, the bound is dependent only on the number of coherence cells in the average. This situation is desirable in comparing measurements with different intensity expectation values, because the resolution can be kept constant. Conversely, if the parameter to be estimated is the intensity expectation value σ, the bound is directly dependent on both the number of coherence cells in the average and the local expectation value of intensity. This makes comparison of measurements with different intensity expectation values more difficult, because the resolution varies in direct proportion to the parameter to be estimated.

The minimum error possible in the unbiased position estimation of an object in an intensity image is now derived. This resolution bound is also a bound on recognition of the object because an object that cannot be resolved cannot be recognized. Let the intensity measurements $\Psi[x - x_0]$ describe an object measurement vector $\mathbf{\Psi}$ of finite length X and unknown position x_0 in a one-dimensional image $\mathbf{W}$ of spatial pixel index x such that $-X/2 \le x - x_0 \le X/2$. The $\Psi[x - x_0]$ are gamma distributed according to $\mathcal{G}(\langle\Psi[x - x_0]\rangle, \mu)$, where the expectation value of the object is defined by $\langle\Psi[x - x_0]\rangle$ and the number of coherence cells in the measurement average is a constant μ throughout the object. With the definition $q[x - x_0] = \ln(\langle\Psi[x - x_0]\rangle/I_{\mathrm{ref}})$ and Δx as a pixel increment, the logarithmic form of the Fisher information given in Eq. (4) can be used to yield the CRLB for estimation of the position parameter x_0. This Fisher information is

$$J = \frac{\mu}{\Delta x} \int_{x=-X/2}^{X/2} \left| \frac{\partial q[x - x_0]}{\partial x} \right|^2 \mathrm{d}x \,. \tag{6}$$

Defining $Q[\xi]$ as the Fourier transform of $q[x]$,

$$E \equiv \int_{-\infty}^{\infty} |Q[\xi]|^2 \mathrm{d}\xi = \int_{-X/2}^{X/2} |q[x - x_0]|^2 \mathrm{d}x \tag{7}$$

as the energy, and

$$B_{\mathrm{rms}}^2 \equiv \frac{(2\pi)^2}{E} \int_{-\infty}^{\infty} \xi^2 |Q[\xi]|^2 \mathrm{d}\xi \tag{8}$$

as the rms bandwidth characterizing the natural log of the expected value of object intensity, yields the CRLB for position estimation of the object:

$$E[(\hat{x}_0 - x_0)^2] \ge J^{-1} = \frac{\Delta x}{\mu B_{\mathrm{rms}}^2 E} \,\cdot \tag{9}$$

This result and the development from relations (6)–(9) are analogous to those used in bounding time-delay resolution for radar returns in uncorrelated noise.[11] This is due to the similar form of the Fisher information matrix in both problems. However, it must be stressed that in this case B_{rms} and E are descriptors of the log-transformed object expectation value, which can span a full range of values. Also, the noise is from signal-dependent fluctuations that are only signal independent and additive in the logarithmic domain, whereas in the radar/sonar time-delay problem similar methods are used to describe the rms bandwidth and energy of a coherent waveform (which also can span a full range of values) in independent additive white noise. With these definitions, Fisher information for object recognition and position resolution increases with the number of coherence cells and (log-domain) bandwidth and energy and decreases with pixel increment. For example, it is apparent from relations (7)–(9) that occlusion of an object can decrease the energy or bandwidth and cause the Fisher information to decrease and the CRLB to increase.

The matched filter estimator has been shown to attain the CRLB for position resolution of a signal in independent additive white noise when its output SNR is high.[2,11] For example, matched filtering an image with the expected object in the logarithmic domain provides the minimum variance unbiased estimator for object localization because it attains the CRLB of relation (9) when $E \gg \zeta(2, \mu)\Delta x$, where $\zeta(2, \mu) \approx 1/\mu$ always holds approximately but becomes more accurate as μ increases. Furthermore, the output of such a log-transformed matched filter always has maximum SNR. One can see this by first defining the log-transformed object measurement $\mathbf{\Phi}$ such that $\Phi[x] = \ln(\Psi[x]/I_{\mathrm{ref}})$ and then employing the decomposition $\Phi[x] = S[x] + N[x]$, where the signal component is deterministic and defined as $S[x] = q[x] + \psi(\mu) - \ln \mu$, so that the noise component $N[x]$ is zero mean and exponential-gamma distributed according to $\mathcal{E}\mathcal{G}(\langle N[x]\rangle = 0, \mu)$. The linear time-invariant filter that maximizes the output SNR for such zero-mean uncorrelated noise is the matched filter.[2] The output of the matched filter for the log-transformed image is

$$R[l] = \frac{k_0}{\Delta x} \int_{x=x_1}^{x_2} S[x]\Phi[x - l]\mathrm{d}x \tag{10}$$

for $-X \le l \le X$, where $S[x]$ is the optimal filter, the noise covariance $\langle N[x]N[x']\rangle$ is $\zeta(2, \mu)$ for $x = x'$ and zero otherwise, k_0 is a normalization constant, and the integration is only over the overlap of $S[x]$ and $\Phi[x - l]$. The output SNR is defined as $\mathrm{SNR}\ (R[l]) \equiv \langle R[l]\rangle^2/(\langle R[l]^2\rangle - \langle R[l]\rangle^2)$. It has a maximum value at zero-lag $\mathrm{SNR}\ (R[0]) = E_S/[\zeta(2, \mu)\Delta x]$, where the signal energy is $E_S = \int_{-X/2}^{X/2} |S[x]|^2 \mathrm{d}x$. The signal energy is approximated by E, and the maximum SNR ($R[0]$) is approximated by $\mu E/\Delta x$ with increasing accuracy as μ increases. An optimal logarithmic matched filter for the special case of $\mu = 1$ was recently derived by Downie and Walkup.[12] Their filter also differs from that above because it has not been defined so that the additive noise is zero mean.

To investigate the information theory perspective, the conditional probability distribution for the log-transformed object measurement can be written as

$$P(\mathbf{\Phi}|\mathbf{a}) = \left[\frac{\mu^{\mu}}{\Gamma(\mu)}\right]^{M} \times \exp\left(\mu \sum_{x=-X/2}^{X/2} \{(\Phi[x] - q[x;\, \mathbf{a}]) - \exp(\Phi[x] - q[x;\, \mathbf{a}])\}\right) \tag{11}$$

by straightforward algebraic manipulation of Eq. (2), where M is the number of pixels in $\mathbf{\Phi}$ and μ is constant over x. The argument of the interior exponential can be expanded into a Taylor series[6] such that

$$P(\mathbf{\Phi}|\mathbf{a}) \approx \left[\frac{\mu^{\mu}}{\Gamma(\mu)} \exp(-\mu)\right]^{M} \times \exp\left\{-\frac{\mu}{2} \sum_{x=-X/2}^{X/2} (\Phi[x] - q[x;\, \mathbf{a}])^2\right\} \tag{12}$$

for $|\Phi[x] - q[x;\, \mathbf{a}]| \ll 3$. (This can be a reasonable approximation even for $\mu = 1$ because the corresponding standard deviation of $\Phi[x]$ is then $\pi\sqrt{6} \approx 1.3$. As μ increases, the approximation becomes better and Stirling's formula leads to a Gaussian density.[6]) It is assumed that there is *a priori* probability $P(\nu)$ that the νth object is present in an image, where $\langle\Psi_{\nu}[x]\rangle$ defines the νth object's expectation value such that $q_{\nu}[x] = \ln(\langle\Psi_{\nu}[x]\rangle/I_{\mathrm{ref}})$. This *a priori* probability $P(\nu)$ should be set by the context of the recognition problem. According to the information theory of Woodward and Davies,[2,13] the ideal or *a posteriori* receiver for the νth object is one that computes the *a posteriori* probability $P(\nu|\mathbf{\Phi}) = P(\nu)P(\mathbf{\Phi}|\nu)/P(\mathbf{\Phi})$, where $P(\mathbf{\Phi})$ is a known constant fixed by the measurement $\mathbf{\Phi}$. From relation (12) it follows that

$$P(\mathbf{\Phi}|\nu) = \left[\frac{\mu^{\mu}}{\Gamma(\mu)} \exp(-\mu)\right]^{M} \exp\left(-\frac{\mu}{2}\left\{\left(\sum_{x=-X/2}^{X/2} \Phi^2[x]\right) + \left(\sum_{x=-X/2}^{X/2} q_{\nu}^2[x]\right) - 2\left(\sum_{x=-X/2}^{X/2} \Phi[x] q_{\nu}[x]\right)\right\}\right), \tag{13}$$

where the general parameter set $\mathbf{a}$ is replaced by the specific object index ν. Because the $\Phi[x]$ are measured, the first summation term in the exponent is a known constant. Following the analogous radar problem, it is common to assume that the $q_{\nu}[x]$ have the same energy over ν,[2,13] so that the second summation is a constant. Therefore the correlation receiver of the log-transformed intensities

$$C(\nu) = \frac{\mu}{\Delta x} \int_{x=-X/2}^{X/2} \Phi[x] q_{\nu}[x] \mathrm{d}x \tag{14}$$

is a sufficient statistic[2,8–10,13] from which the *a posteriori* density $P(\nu|\mathbf{\Phi})$ may be fully reconstructed and all the received information about the presence of the νth object may be fully recovered in an optimum receiver.[2,13] Clearly, there is a simple linear relationship between the correlation receiver $C(\nu)$ and the zero-lag value of the optimal filter output $R[0]$, which involves known additive and multiplicative constants. For example, when $P(\nu)$ is uniformly distributed, $C(\nu)$ attains a maximum with respect to ν for the object that is most likely to be present.

The *a posteriori* probability $P(\nu|\mathbf{\Psi})$ for the presence of the νth object given the intensity measurement $\mathbf{\Psi}$ is equal to the *a posteriori* probability $P(\nu|\mathbf{\Phi})$ given the log measurement $\mathbf{\Phi}$, because $\mathbf{\Phi}$ is completely specified by $\mathbf{\Psi}$ and vice versa. However, the sufficient statistic of Eq. (14), the optimal filter of Eq. (10), and the Fisher information of Eq. (9) indicate that matched filtering images with hypothetical patterns in the logarithmic domain provides an optimal method for pattern recognition in the presence of signal-dependent noise arising from CCGR fluctuations in the received fields. Because CCGR fields are commonly measured in optical, radar, and acoustic imaging, this result provides a mathematical justification for the use of logarithmic intensity measurements to efficiently convey information for pattern recognition in a variety of engineering applications. These results may also be useful in interpreting the apparent logarithmic response of human auditory and visual perception to intensity stimulus as exhibited in the Weber–Fechner laws.[14]

References

1. J. W. Goodman, *Statistical Optics* (Wiley, New York, 1985), Chap. 6.
2. J. V. Difranco and W. L. Rubin, *Radar Detection* (Prentice-Hall, Englewood Cliffs, N.J., 1968).
3. P. Swerling, Trans. IRE Inform. Theory **IT-6,** 269 (1960).
4. N. C. Makris, "A parameter estimation bound for images," J. Acoust. Soc. Am. (to be published).
5. A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, N.J., 1975).
6. H. H. Arsenault and G. April, J. Opt. Soc. Am. **66,** 1160 (1976).
7. I. S. Gradshteyn and I. M. Ryshik, *Tables of Integrals, Series and Products* (Academic, San Diego, Calif., 1980), pp. 943, 1073.
8. C. R. Rao, *Linear Statistical Inference and Its Applications* (Wiley, New York, 1966), Chap. 5.
9. S. M. Kay, *Fundamental of Statistical Signal Processing: Estimation Theory* (Prentice-Hall, Englewood Cliffs, N.J., 1993), Chaps. 3 and 5.
10. R. A. Fisher, Philos. Trans. R. Soc. London **222,** 309 (1922).
11. N. Levanov, *Radar Principles* (Wiley, New York, 1988).
12. J. D. Downie and J. F. Walkup, J. Opt. Soc. Am. A **11,** 1599 (1994).
13. P. M. Woodward and I. L. Davies, Proc. Inst. Electr. Eng. Pt. III **99,** 37 (1952).
14. W. Wong and H. Norwich, J. Acoust Soc. Am. **97,** 3761 (1995).