[7] J. L. Butler, "Comparative criteria for minicomputers," *Instrum. Technol.*, vol. 17, pp. 67-82, Oct. 1970.

[8] B. Liu, "Effect of finite word length on accuracy of digital filters–A review," *IEEE Trans. Circuit Theory*, vol. CT-18, pp. 670-677, Nov. 1971.

[9] J. B. O'Neal, Jr., "Delta modulation quantizing noise analytical and computer simulation results for Gaussian and television input signals," *Bell Syst. Tech. J.*, vol. 45, pp. 117-142, Jan. 1966.

[10] D. J. Goodman and A. Gersho, "Theory of an adaptive quantizer," *IEEE Trans. Commun.*, vol. COM-22, pp. 1037-1045, Aug. 1974.

The Generalized Correlation Method for Estimation of Time Delay

CHARLES H. KNAPP, MEMBER, IEEE, AND G. CLIFFORD CARTER, MEMBER, IEEE

Abstract–**A maximum likelihood (ML) estimator is developed for determining time delay between signals received at two spatially separated sensors in the presence of uncorrelated noise. This ML estimator can be realized as a pair of receiver prefilters followed by a cross correlator. The time argument at which the correlator achieves a maximum is the delay estimate. The ML estimator is compared with several other proposed processors of similar form. Under certain conditions the ML estimator is shown to be identical to one proposed by Hannan and Thomson [10] and MacDonald and Schultheiss [21].**

Qualitatively, the role of the prefilters is to accentuate the signal passed to the correlator at frequencies for which the signal-to-noise (S/N) ratio is highest and, simultaneously, to suppress the noise power. The same type of prefiltering is provided by the generalized Eckart filter, which maximizes the S/N ratio of the correlator output. For low S/N ratio, the ML estimator is shown to be equivalent to Eckart prefiltering.

Introduction

A SIGNAL emanating from a remote source and monitored in the presence of noise at two spatially separated sensors can be mathematically modeled as

$$x_1(t) = s_1(t) + n_1(t) \tag{1a}$$

$$x_2(t) = \alpha s_1(t + D) + n_2(t), \tag{1b}$$

where $s_1(t)$, $n_1(t)$, and $n_2(t)$ are real, jointly stationary random processes. Signal $s_1(t)$ is assumed to be uncorrelated with noise $n_1(t)$ and $n_2(t)$.

There are many applications in which it is of interest to estimate the delay D. This paper proposes a maximum likelihood (ML) estimator and compares it with other similar techniques. While the model of the physical phenomena presumes stationarity, the techniques to be developed herein are usually employed in slowly varying environments where the characteristics of the signal and noise remain stationary only for finite observation time T. Further, the delay D and attenuation α may also change slowly. The estimator is, therefore, constrained to operate on observations of a finite duration.

Another important consideration in estimator design is the available amount of *a priori* knowledge of the signal and noise statistics. In many problems, this information is negligible. For example, in passive detection, unlike the usual communications problems, the source spectrum is unknown or only known approximately.

One common method of determining the time delay D and, hence, the arrival angle relative to the sensor axis [1] is to compute the cross correlation function

$$R_{x_1 x_2}(\tau) = E[x_1(t)\, x_2(t - \tau)], \tag{2}$$

where E denotes expectation. The argument τ that maximizes (2) provides an estimate of delay. Because of the finite observation time, however, $R_{x_1x_2}(\tau)$ can only be estimated. For example, for ergodic processes [2, p. 327], an estimate of the cross correlation is given by

$$\hat{R}_{x_1 x_2}(\tau) = \frac{1}{T - \tau} \int_{\tau}^{T} x_1(t)\, x_2(t - \tau)\, dt, \tag{3}$$

where T represents the observation interval. In order to improve the accuracy of the delay estimate $\hat{D}$, it is desirable to prefilter $x_1(t)$ and $x_2(t)$ prior to the integration in (3). As shown in Fig. 1, x_i may be filtered through H_i to yield y_i for $i = 1, 2$. The resultant y_i are multiplied, integrated, and squared for a range of time shifts, τ, until the peak is obtained. The time shift causing the peak is an estimate of the true delay D. When the filters $H_1(f) = H_2(f) = 1, \forall f$, the estimate $\hat{D}$ is simply the abscissa value at which the cross-correlation function peaks. This paper provides for a generalized correlation through the introduction of the filters $H_1(f)$ and $H_2(f)$ which, when properly selected, facilitate the estimation of delay.

The cross correlation between $x_1(t)$ and $x_2(t)$ is related to

Manuscript received July 24, 1975; revised November 21, 1975 and February 23, 1976.

C. H. Knapp is with the Department of Electrical Engineering and Computer Science, University of Connecticut, Storrs, CT 06268.

G. C. Carter is with the Naval Underwater Systems Center, New London Laboratory, New London, CT 06320.

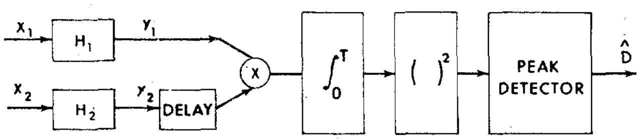

Fig. 1. Received waveforms filtered, delayed, multiplied, and integrated for a variety of delays until peak output is obtained.

the cross power spectral density function by the well-known Fourier transform relationship

$$R_{x_1 x_2}(\tau) = \int_{-\infty}^{\infty} G_{x_1 x_2}(f) e^{j2\pi f \tau}\, df. \tag{4}$$

When $x_1(t)$ and $x_2(t)$ have been filtered as depicted in Fig. 1, then the cross power spectrum between the filter outputs is given by [3, p. 399]

$$G_{y_1 y_2}(f) = H_1(f) H_2^*(f) G_{x_1 x_2}(f), \tag{5}$$

where $*$ denotes the complex conjugate. Therefore, the generalized correlation between $x_1(t)$ and $x_2(t)$ is

$$R_{y_1 y_2}^{(g)}(\tau) = \int_{-\infty}^{\infty} \psi_g(f) G_{x_1 x_2}(f) e^{j2\pi f \tau}\, df, \tag{6a}$$

where

$$\psi_g(f) = H_1(f) H_2^*(f) \tag{6b}$$

and denotes the general frequency weighting.

In practice, only an estimate $\hat{G}_{x_1 x_2}(f)$ of $G_{x_1 x_2}(f)$ can be obtained from finite observations of $x_1(t)$ and $x_2(t)$. Consequently, the integral

$$\hat{R}_{y_1 y_2}^{(g)}(\tau) = \int_{-\infty}^{\infty} \psi_g(f) \hat{G}_{x_1 x_2}(f) e^{j2\pi f \tau}\, df \tag{6c}$$

is evaluated and used for estimating delay. Indeed, depending on the particular form of $\psi_g(f)$ and the *a priori* information, it may also be necessary to estimate $\psi_g(f)$ in (6a)-(6b). For example, when the role of the prefilters is to accentuate the signal passed to the correlator at those frequencies at which the signal-to-noise (S/N) ratio is highest, then $\psi_g(f)$ can be expected to be a function of signal and noise spectra which must either be known *a priori* or estimated.

The selection of $\psi_g(f)$ to optimize certain performance criteria has been studied by several investigators. (See, for example, [4]-[12].) This paper will derive the ML estimator for delay D in the mathematical model (1a) and (1b), given signal and noise spectra. The results will be shown to be equivalent to (6a)-(6c) with an appropriate $\psi(f)$. This weighting turns out to be equivalent to that proposed in [12] and under simplifying assumptions to that proposed in [21]. The development presented here does not presume initially that the estimator has the form (6c). Rather, it is shown that the ML estimator may be realized by choosing τ that maximizes (6c) with proper weighting, $\psi_g(f)$, and proper estimate, $\hat{G}_{x_1 x_2}(f)$. The weighting $\psi_g(f)$ yielding the ML estimate will be compared to other weightings that have been proposed. Under certain conditions the ML estimator is shown to be equivalent to other processors.

Processor Interpretation

It is informative to examine the effect of processor weightings on the shape of $R_{y_1 y_2}(\tau)$ under ideal conditions. For models of the form of (1), the cross correlation of $x_1(t)$ and $x_2(t)$ is

$$R_{x_1 x_2}(\tau) = \alpha R_{s_1 s_1}(\tau - D) + R_{n_1 n_2}(\tau). \tag{7}$$

The Fourier transform of (7) gives the cross power spectrum

$$G_{x_1 x_2}(f) = \alpha G_{s_1 s_1}(f) e^{-j2\pi f D} + G_{n_1 n_2}(f). \tag{8}$$

If $n_1(t)$ and $n_2(t)$ are uncorrelated ($G_{n_1 n_2}(f) = 0$), the cross power spectrum between $x_1(t)$ and $x_2(t)$ is a scaled signal power spectrum times a complex exponential. Since multiplication in one domain is a convolution in the transformed domain (see, for example, [13]), it follows for $G_{n_1 n_2}(f) = 0$ that

$$R_{x_1 x_2}(\tau) = \alpha R_{s_1 s_1}(\tau) \circledast \delta(t - D), \tag{9}$$

where $\circledast$ denotes convolution.

One interpretation of (9) is that the delta function has been spread or "smeared" by the Fourier transform of the signal spectrum. If $s_1(t)$ is a white noise source, then its Fourier transform is a delta function and no spreading takes place. An important property of autocorrelation functions is that $R_{ss}(\tau) \leqslant R_{ss}(0)$. Equality will hold for certain τ for periodic functions (see, for example, [3, pp. 323-326]). However, for most practical applications, equality does not hold for $\tau \neq 0$, and the true cross correlation (9) will peak at D regardless of whether or not it is spread out. The spreading simply acts to broaden the peak. For a single delay this may not be a serious problem. However, when the signal has multiple delays, the true cross correlation is given by

$$R_{x_1 x_2}(\tau) = R_{s_1 s_1}(\tau) \circledast \sum_i \alpha_i \delta(\tau - D_i). \tag{10}$$

In this case, the convolution with $R_{s_1 s_1}(\tau)$ can spread one delta function into another, thereby making it impossible to distinguish peaks or delay times. Under ideal conditions where $\forall f$, $\hat{G}_{x_1 x_2}(f) \cong G_{x_1 x_2}(f)$, $\psi_g(f)$ should be chosen to ensure a large sharp peak in $R_{y_1 y_2}(\tau)$ rather than a broad one in order to ensure good time-delay resolution. However, sharp peaks are more sensitive to errors introduced by finite observation time, particularly in cases of low S/N ratio. Thus, as with other spectral estimation problems, the choice of $\psi_g(f)$ is a compromise between good resolution and stability.

The preceding discussion sets the background for the role that $\psi_g(f)$ is to play. Now the five generalizations of the cross-correlation function listed in Table I will be examined individually.

The Roth Processor

The weighting proposed by Roth [9],

$$\psi_R(f) = \frac{1}{G_{x_1 x_1}(f)} \tag{11}$$

[where the subscript R is to distinguish the choice of $\psi_g(f)$], yields[1]

$$\hat{R}^{(R)}_{y_1 y_2}(\tau) = \int_{-\infty}^{\infty} \frac{\hat{G}_{x_1 x_2}(f)}{G_{x_1 x_1}(f)} e^{j2\pi f\tau} \, df. \tag{12}$$

Equation (12) estimates the impulse response of the optimum linear (Wiener-Hopf) filter

$$H_m(f) = \frac{G_{x_1 x_2}(f)}{G_{x_1 x_1}(f)}, \tag{13}$$

which "best" approximates the mapping of $x_1(t)$ to $x_2(t)$ (see, for example, [14], [15]). If $n_1(t) \neq 0$, as is generally the case for (1), then

$$G_{x_1 x_1}(f) = G_{s_1 s_1}(f) + G_{n_1 n_1}(f), \tag{14}$$

and

$$R^{(R)}_{y_1 y_2}(\tau) = \delta(\tau - D) \circledast \int_{-\infty}^{\infty} \frac{\alpha G_{s_1 s_1}(f)}{\{G_{s_1 s_1}(f) + G_{n_1 n_1}(f)\}} \cdot e^{j2\pi f\tau} \, df. \tag{15}$$

Therefore, except when $G_{n_1 n_1}(f)$ equals any constant (including zero) times $G_{s_1 s_1}(f)$, the delta function will again be spread out. The Roth processor has the desirable effect of suppressing those frequency regions where $G_{n_1 n_1}(f)$ is large and $\hat{G}_{x_1 x_2}(f)$ is more likely to be in error.

The Smoothed Coherence Transform (SCOT)

Errors in $\hat{G}_{x_1 x_2}(f)$ may be due to frequency bands where $G_{n_2 n_2}(f)$ is large, as well as bands where $G_{n_1 n_1}(f)$ is large. One is, therefore, uncertain whether to form $\psi_R(f) = 1/G_{x_1 x_1}(f)$ or $\psi_R(f) = 1/G_{x_2 x_2}(f)$; hence, the SCOT [11] selects

$$\psi_s(f) = 1/\sqrt{G_{x_1 x_1}(f) G_{x_2 x_2}(f)}. \tag{16}$$

This weighting gives the SCOT

$$\hat{R}^{(s)}_{y_1 y_2}(\tau) = \int_{-\infty}^{\infty} \hat{\gamma}_{x_1 x_2}(f) \, e^{j2\pi f\tau} \, df, \tag{17}$$

where the coherence estimate[2]

$$\hat{\gamma}_{x_1 x_2}(f) \triangleq \frac{\hat{G}_{x_1 x_2}(f)}{\sqrt{G_{x_1 x_1}(f) \, G_{x_2 x_2}(f)}}. \tag{18}$$

For $H_1(f) = 1/\sqrt{G_{x_1 x_1}(f)}$ and $H_2(f) = 1/\sqrt{G_{x_2 x_2}(f)}$, the SCOT can be interpreted through Fig. 1 as prewhitening filters followed by a cross correlation. When $G_{x_1 x_1}(f) = G_{x_2 x_2}(f)$, the SCOT is equivalent to the Roth processor. If $n_1(t) \neq 0$ and $n_2(t) \neq 0$, the SCOT exhibits the same spreading as the Roth processor. This broadening persists because of an apparent inability to adequately prewhiten the cross power spectrum.

The Phase Transform (PHAT)

To avoid the spreading evident above, the PHAT uses the weighting [16]

$$\psi_P(f) = \frac{1}{|G_{x_1 x_2}(f)|}, \tag{19}$$

which yields

$$\hat{R}^{(P)}_{y_1 y_2}(\tau) = \int_{-\infty}^{\infty} \frac{\hat{G}_{x_1 x_2}(f)}{|G_{x_1 x_2}(f)|} e^{j2\pi f\tau} \, df. \tag{20}$$

For the model (1) with uncorrelated noise (i.e., $G_{n_1 n_2}(f) = 0$),

$$|G_{x_1 x_2}(f)| = \alpha G_{s_1 s_1}(f). \tag{21}$$

Ideally, when $\hat{G}_{x_1 x_2}(f) = G_{x_1 x_2}(f)$,

$$\frac{\hat{G}_{x_1 x_2}(f)}{|G_{x_1 x_2}(f)|} = e^{j\theta(f)} = e^{j2\pi fD} \tag{22}$$

has unit magnitude and

$$R^{(P)}_{y_1 y_2}(\tau) = \delta(t - D). \tag{23}$$

The PHAT was developed purely as an ad hoc technique. Notice that, for models of the form of (1) with uncorrelated noises, the PHAT (20), ideally, does not suffer the spreading that other processors do. In practice, however, when $\hat{G}_{x_1 x_2}(f) \neq G_{x_1 x_2}(f)$, $\theta(f) \neq 2\pi fD$ and the estimate of $R^{(P)}_{y_1 y_2}(\tau)$ will not be a delta function. Another apparent defect of the PHAT is that it weights $\hat{G}_{x_1 x_2}(f)$ as the inverse of $G_{s_1 s_1}(f)$. Thus, errors are accentuated where signal power is smallest. In particular, if $G_{x_1 x_2}(f) = 0$ in some frequency band, then the phase $\theta(f)$ is undefined in that band and the estimate of the phase is erratic, being uniformly distributed in the interval $[-\pi, \pi]$ rad. For models of the form of (1), this behavior suggests that $\psi_P(f)$ be additionally weighted to compensate for the presence or absence of signal power. The SCOT is one method of assigning weight according to signal and noise characteristics. Two remaining processors also assign weights or filtering according to S/N ratio: the Eckart filter [5] and the ML estimator or Hannan-Thomson (HT) processor [12].

The Eckart Filter

The Eckart filter derives its name from work in this area done in [5]. Derivations in [7], [8], [17], and [18] are outlined here briefly for completeness. The Eckart filter maximizes the deflection criterion, i.e., the ratio of the change in mean correlator output due to signal present to the standard deviation of correlator output due to noise alone. For

[1] As discussed earlier, $\psi(f)$ may have to be estimated for this processor and those which follow, because of a lack of *a priori* information. In this case, (11) must be modified by replacing $G_{x_1 x_1}(f)$ with $\hat{G}_{x_1 x_1}(f)$.

[2] A more standard coherence estimate is formed when the auto spectra must also be estimated, as is usually the case.

long averaging time T, the deflection has been shown [8] to be

$$d^2 = \frac{L\left[\int_{-\infty}^{\infty} H_1(f) H_2^*(f)\, G_{s_1s_2}(f)\, df\right]^2}{\int_{-\infty}^{\infty} |H_1(f)|^2\, |H_2(f)|^2\, G_{n_1n_1}(f)\, G_{n_2n_2}(f)\, df}, \tag{24}$$

where L is a constant proportional to T, and $G_{s_1s_2}(f)$ is the cross power spectrum between $s_1(t)$ and $s_2(t)$. For the model (1), $G_{s_1s_2}(f) = \alpha G_{s_1s_1}(f)$ exp $(j2\pi fD)$. Application of Schwartz's inequality to (24) indicates that

$$H_1(f) H_2^*(f) = \psi_E(f)\, e^{+j2\pi fD} \tag{25}$$

maximizes d^2 where

$$\psi_E(f) = \frac{\alpha G_{s_1s_1}(f)}{G_{n_1n_1}(f)\, G_{n_2n_2}(f)}. \tag{26}$$

Notice that the weighting (26), referred to as the Eckart filter, possesses some of the qualities of the SCOT. In particular, it acts to suppress frequency bands of high noise, as does the SCOT. Also note that the Eckart filter unlike the PHAT attaches zero weight to bands where $G_{s_1s_1}(f) = 0$. In practice, the Eckart filter requires knowledge or estimation of the signal and noise spectra. For (1), when $\alpha = 1$ this can be accomplished by letting

$$\psi_E(f) = |\hat{G}_{x_1x_2}(f)|\, \{\hat{G}_{x_1x_1}(f) - |\hat{G}_{x_1x_2}(f)|] \cdot [\hat{G}_{x_2x_2}(f) - |\hat{G}_{x_1x_2}(f)|]\}. \tag{27}$$

The first five processors in Table I can be justified on the basis of reasonable performance criteria, whether heuristic or mathematical.

In the next section, the ML estimator of the parameter D is derived. It is shown to be identical to that proposed by Hannan and Thomson [12].

The HT Processor

To make the model (1) mathematically tractable, it is necessary to assume that $s_1(t)$, $n_1(t)$, and $n_2(t)$ are Gaussian. Denote the Fourier coefficients of $x_i(t)$ as in [3, eq. (3.8)] by

$$X_i(k) = \frac{1}{T}\int_{-T/2}^{T/2} x_i(t)\, e^{-jkt\omega_\Delta}\, dt, \tag{28a}$$

where

$$\omega_\Delta = \frac{2\pi}{T}.$$

Note that the linear transformation $X_i(k)$ is Gaussian since $x_i(t)$ is Gaussian. Further, from [3, eq. (3.13)], as $T \to \infty$ and $K \to \infty$ such that $K\omega_\Delta = \omega$ is constant

TABLE I
CANDIDATE PROCESSORS

Processor Name	Weight $\psi(f) = H_1(f) H_2^*(f)$	Text Reference
Cross Correlation	1	[2]-[4], [19]
Roth Impulse Response	$1/G_{x_1x_1}(f)$	[9]
SCOT	$1/\sqrt{G_{x_1x_1}(f)\, G_{x_2x_2}(f)}$	[11], [16]
PHAT	$1/\lvert G_{x_1x_2}(f)\rvert$	[16]
Eckart	$G_{s_1s_1}(f)/[G_{n_1n_1}(f)\, G_{n_2n_2}(f)]$	[5], [7], [8] [17], [18]
ML or HT	$\dfrac{\lvert\gamma_{12}(f)\rvert^2}{\lvert G_{x_1x_2}(f)\rvert[1 - \lvert\gamma_{12}(f)\rvert^2]}$	[12]

$$\tilde{X}_i(\omega) = \lim_{T\to\infty} TX_i(k) \tag{28b}$$

$$= \int_{-\infty}^{\infty} x_i(t) e^{-j\omega t}\, dt, \tag{28c}$$

where $\tilde{X}_i$ is the Fourier transform of $x_i(t)$. A more complete discussion on Fourier transforms and their convergence is given in [3, p. 381], [4, pp. 23-25], [19, ch. 1], and [20, p. 11]. From [21], it follows for T large compared to $|D|$ plus the correlation time of $R_{s_1s_1}(\tau)$, that

$$E[X_1(k) X_2^*(l)] \cong \begin{cases} \frac{1}{T} G_{x_1x_2}(k\omega_\Delta), & k = l \\ 0, & k \neq l. \end{cases} \tag{29}$$

Now let the vector

$$X(k) = [X_1(k), X_2(k)]', \tag{30}$$

where $'$ denotes transpose. Define the power spectral density matrix Q such that

$$E[X(k) X^{*\prime}(k)] = E\begin{bmatrix} X_1(k) X_1^*(k) & X_1(k) X_2^*(k) \\ X_2(k) X_1^*(k) & X_2(k) X_2^*(k) \end{bmatrix} \tag{31a}$$

$$= \frac{1}{T}\begin{bmatrix} G_{x_1x_1}(k\omega_\Delta) & G_{x_1x_2}(k\omega_\Delta) \\ G_{x_1x_2}^*(k\omega_\Delta) & G_{x_2x_2}(k\omega_\Delta) \end{bmatrix} \tag{31b}$$

$$\triangleq \frac{1}{T} Q_x(k\omega_\Delta). \tag{31c}$$

Properties of $Q_x(k\omega_\Delta)$ can be used to prove

$$0 \leqslant |\gamma_{x_1x_2}(k\omega_\Delta)|^2 \leqslant 1, \forall\, k\omega_\Delta$$

[4, p. 467]. The vectors $X(k)$, $k = -N, -N+1, \cdots, N$ are, as a consequence of (29), uncorrelated Gaussian (hence, independent) random variables. More explicitly, the probability for $X \equiv X(-N), X(-N+1), \cdots, X(N)$, given the power spectral density matrix Q (or the delay, attenuation, and spectral characteristics of the signal and noises necessary to determine Q) is

$$p(X|Q) = p(X|\alpha, G_{s_1s_1}, G_{n_1n_1}, G_{n_2n_2}, G_{n_1n_2}, D) = c \exp\left(-\tfrac{1}{2} J_1\right) \tag{32}$$

where

$$J_1 = \sum_{k=-N}^{N} X^{*\prime}(k)\, Q_x^{-1}(k\omega_\Delta)\, X(k)\, T \tag{33}$$

and c is a function of $|Q_x(k\omega_\Delta)|$ [15, p. 185]. Replacing $X(k)$ by $\frac{1}{T}\tilde{X}(k\omega_\Delta)$ from (28),

$$J_1 = \sum_{k=-N}^{N} \tilde{X}^{*\prime}(k\omega_\Delta)\, Q_x^{-1}(k\omega_\Delta)\, \tilde{X}(k\omega_\Delta)\frac{1}{T}. \tag{34}$$

For ML estimation (see, for example, [4] or [15]), it is desired to choose D to maximize $p(X|Q,D)$.

In general, the parameter D affects both c and J_1 in (32). However, under certain simplifying assumptions, c is constant or is only weakly related to the delay. Specifically, from (1) and (31), suppressing the frequency argument $k\omega_\Delta$,

$$\begin{aligned}|Q_x| = {} & (G_{s_1s_1} + G_{n_1n_1})(\alpha^2 G_{s_1s_1} + G_{n_2n_2})\\ & -(G_{n_1n_2} + \alpha G_{s_1s_1} e^{-j2\pi fD})\\ & \cdot (G^*_{n_1n_2} + \alpha G_{s_1s_1} e^{+j2\pi fD}),\end{aligned} \tag{35}$$

which is independent of D if $G_{n_1n_2} = 0$ (i.e., the noises are uncorrelated).

For large T, (34) becomes

$$J_1 \cong \int_{-\infty}^{\infty} \tilde{X}^{*\prime}(f)\, Q_x^{-1}(f)\, \tilde{X}(f)\, df. \tag{36}$$

From (31),

$$Q_x^{-1}(f) = \frac{\begin{bmatrix} G_{x_2x_2}(f) & -G_{x_1x_2}(f)\\ -G^*_{x_1x_2}(f) & G_{x_1x_1}(f)\end{bmatrix}}{G_{x_1x_1}(f)\, G_{x_2x_2}(f) - |G_{x_1x_2}(f)|^2} \tag{37a}$$

$$= \frac{1}{[1 - |\gamma_{12}(f)|^2]} \cdot \begin{bmatrix} 1/G_{x_1x_1}(f), & -G_{x_1x_2}(f)/\{G_{x_1x_1}(f)\cdot G_{x_2x_2}(f)\}\\ -G^*_{x_1x_2}(f)/\{G_{x_1x_1}(f)\, G_{x_2x_2}(f)\}, & 1/G_{x_2x_2}(f)\end{bmatrix}, \tag{37b}$$

which will exist provided $|\gamma_{12}(f)|^2 \neq 1$; i.e., $x_1(t)$ and $x_2(t)$ cannot be obtained perfectly from one another by linear filtering [14], or equivalently for the model (1) that observation noise *is* present.

When $G_{n_1n_2}(f) = 0$,

$$G_{x_1x_1}(f) = G_{s_1s_1}(f) + G_{n_1n_1}(f), \tag{38}$$

$$G_{x_2x_2}(f) = \alpha^2 G_{s_1s_1}(f) + G_{n_2n_2}(f), \tag{39}$$

$$G_{x_1x_2}(f) = \alpha G_{s_1s_1}(f)\, e^{-j2\pi fD}, \tag{40}$$

and it follows that

$$J_1 = \int_{-\infty}^{\infty} \tilde{X}^{*\prime}(f)\, Q_x^{-1}(f)\, \tilde{X}(f)\, df = J_2 + J_3, \tag{41}$$

where

$$J_2 = \int_{-\infty}^{\infty} \left[\frac{|\tilde{X}_1(f)|^2}{G_{x_1x_1}(f)} + \frac{|\tilde{X}_2(f)|^2}{G_{x_2x_2}(f)}\right] \cdot \frac{1}{[1 - |\gamma_{12}(f)|^2]}\, df, \tag{42a}$$

$$-J_3 = \int_{-\infty}^{\infty} A(f) + A^*(f)\, df, \tag{42b}$$

$$A(f) = \tilde{X}_1(f)\, \tilde{X}_2^*(f) \cdot \frac{\alpha G_{s_1s_1}(f)\, e^{j2\pi fD}}{G_{x_1x_1}(f)\, G_{x_2x_2}(f)\, [1 - |\gamma_{12}(f)|^2]}. \tag{42c}$$

In order to relate these results to [12] and interpret how to implement the ML estimation technique, note that for $x_1(t)$ and $x_2(t)$ real, $A^*(f) = A(-f)$. Then (42b) can be rewritten as

$$-J_3 = \int_{-\infty}^{\infty} A(f)\, df + \int_{-\infty}^{\infty} A(-f)\, df = 2\int_{-\infty}^{\infty} A(f)\, df. \tag{43}$$

Letting $T\hat{G}_{x_1x_2}(f) \triangleq \tilde{X}_1(f)\, \tilde{X}_2^*(f)$, (43) and (42c) can be rewritten as

$$-J_3 = 2T\int_{-\infty}^{\infty} \hat{G}_{x_1x_2}(f) \frac{1}{|G_{x_1x_2}(f)|} \frac{|\gamma_{12}(f)|^2}{[1 - |\gamma_{12}(f)|^2]} \cdot e^{j2\pi fD}\, df. \tag{44}$$

Notice that the ML estimator for D will minimize $J_1 = J_2 + J_3$, but the selection of D has no effect on J_2. Thus, D should maximize $-J_3$. Equivalently, when $\tilde{X}_1(f)\, \tilde{X}_2^*(f)$ is viewed as T times the estimated cross power spectrum, $T\hat{G}_{x_1x_2}(f)$, the ML estimator selects as the estimate of delay the value of τ at which

$$R^{(\mathrm{HT})}_{y_1y_2}(\tau) = \int_{-\infty}^{\infty} \hat{G}_{x_1x_2}(f) \frac{1}{|G_{x_1x_2}(f)|} \cdot \frac{|\gamma_{12}(f)|^2}{[1 - |\gamma_{12}(f)|^2]}\, e^{j2\pi f\tau}\, df \tag{45a}$$

achieves a peak.

The weighting in (6),

$$\psi_{\mathrm{HT}}(f) = \frac{1}{|G_{x_1x_2}(f)|} \cdot \frac{|\gamma_{12}(f)|^2}{[1 - |\gamma_{12}(f)|^2]}, \tag{45b}$$

where (as required for Q_x^{-1} to exist) $|\gamma_{12}(f)|^2 \neq 1$, achieves the ML estimator. When $|G_{x_1x_2}(f)|$ and $|\gamma_{12}(f)|^2$ are known, this is exactly the proper weighting. When the terms in (45b) are unknown, they can be estimated via techniques of [22]. Substituting estimated weighting for true weighting is entirely a heuristic procedure whereby the ML estimator can approximately be achieved in practice.

Note that, like the HT processor, the PHAT computes a type of transformation on

$$\frac{\hat{G}_{x_1x_2}(f)}{|G_{x_1x_2}(f)|} = \exp\,[j\hat{\theta}\,(f)].$$

However, the HT processor, like the SCOT, weights the phase according to the strength of the coherence.

From [4, p. 379],

$$\operatorname{var}[\hat{\theta}(f)] \cong \frac{1-|\gamma|^2}{|\gamma|^2}\cdot\frac{1}{L_1}, \tag{46a}$$

where L_1 is a proportionality constant dependent on how the data are processed. Thus,

$$R_{y_1y_2}^{(\mathrm{HT})}(\tau) \cong \frac{1}{L_1}\int_{-\infty}^{\infty} e^{j\hat{\theta}(f)}\cdot\frac{1}{\operatorname{var}[\hat{\theta}(f)]}\, e^{j2\pi f\tau}\, df. \tag{46b}$$

Comparison of (46b) and (20) with (21) reveals that the ML estimator is the PHAT inversely weighted according to the variability of the phase estimates.

In interpreting the similarity of the HT processor to the other processors listed in Table I, it can be shown that if $G_{n_1n_1}(f) = G_{n_2n_2}(f) = G_{nn}(f)$ is equal to a constant times $G_{s_1s_1}(f)$, then the last five processors in Table I are the same except for a constant, but the cross-correlation processor ($\psi(f) = 1$, $\forall f$) is a delta function smeared out by the Fourier transform of the signal (noise) power spectrum.

Interpretation of Low S/N Ratio of ML Estimator

Good delay estimation is most difficult to achieve in the case of low S/N ratios. In order to compare estimates under low S/N ratio conditions, let $\alpha = 1$. Then,

$$\psi_{\mathrm{HT}}(f) = \frac{1}{G_{s_1s_1}(f)} \cdot \frac{G^2_{s_1s_1}(f)}{\{[G_{s_1s_1}(f)+G_{n_1n_1}(f)]\cdot[G_{s_1s_1}(f)+G_{n_2n_2}(f)] - G^2_{s_1s_1}(f)\}} \tag{47a}$$

$$= \frac{\dfrac{G_{s_1s_1}(f)}{G_{n_1n_1}(f)\,G_{n_2n_2}(f)}}{\left[1+\dfrac{G_{s_1s_1}(f)}{G_{n_2n_2}(f)}+\dfrac{G_{s_1s_1}(f)}{G_{n_1n_1}(f)}\right]}, \tag{47b}$$

which agrees with [21, eq. (28)] if in (47b) $G_{n_1n_1}(f) = G_{n_2n_2}(f)$.

For low S/N ratio,

$$\frac{G_{s_1s_1}(f)}{G_{n_1n_1}(f)} << 1 \quad\text{and}\quad \frac{G_{s_1s_1}(f)}{G_{n_2n_2}(f)} << 1,$$

it follows that

$$\psi_{\mathrm{HT}}(f) \cong \frac{G_{s_1s_1}(f)}{G_{n_1n_1}(f)\,G_{n_2n_2}(f)} = \psi_E(f). \tag{48}$$

That is, for $\alpha = 1$ and low S/N ratio, the HT processor is identical to the Eckart filter. Similarly, for low S/N ratio,

$$\psi_s(f) \cong 1/\sqrt{G_{n_1n_1}(f)\,G_{n_2n_2}(f)}. \tag{49}$$

Therefore, if $\alpha = 1$,

$$\psi_{\mathrm{HT}}(f) \cong \frac{G_{s_1s_1}(f)}{\sqrt{G_{n_1n_1}(f)\,G_{n_2n_2}(f)}}\,[\psi_s(f)]. \tag{50a}$$

Furthermore, for $G_{n_1n_1}(f) = G_{n_2n_2}(f) = G_{nn}(f)$,

$$\psi_{\mathrm{HT}}(f) \cong \frac{G_{s_1s_1}(f)}{G_{nn}(f)}\,[\psi_s(f)] = \left[\frac{G_{s_1s_1}(f)}{G_{nn}(f)}\right]^2 \psi_p(f). \tag{50b}$$

Thus, under low S/N ratio approximations with $\alpha = 1$, both the Eckart and HT prefilters can be interpreted either as SCOT prewhitening filters with additional S/N ratio weighting or PHAT prewhitening filters with additional S/N ratio squared weighting.

VARIANCE OF DELAY ESTIMATORS

It can be shown, by extending a result from [21], that the variance of the time-delay estimate in the neighborhood of the true delay for general weighting function $\psi(f)$ is given by

$$\operatorname{var}[\hat{D}] = \frac{\displaystyle\int_{-\infty}^{\infty} |\psi(f)|^2(2\pi f)^2 G_{x_1x_1}(f)\,G_{x_2x_2}(f)[1-|\gamma(f)|^2]\,df}{T\left[\displaystyle\int_{-\infty}^{\infty}(2\pi f)^2|G_{x_1x_2}(f)|\,\psi(f)\,df\right]^2}. \tag{51}$$

The variance of the HT processor (substituting from (45b) and using the definition of coherence) is

$$\operatorname{var}^{\mathrm{HT}}[\hat{D}] = \left\{2T\int_0^{\infty}(2\pi f)^2|\gamma(f)|^2/[1-|\gamma(f)|^2]\,df\right\}^{-1}. \tag{52}$$

In particular, the HT processor achieves the Cramér–Rao lower bound (see Appendix). It should be pointed out that (51) and (52) evaluate the local variation of the time-delay estimate and thus do not account for ambiguous peaks which may arise when the averaging time is not large enough for the given signal and noise characteristics. Indeed, when T is not sufficiently large, local variation may be a poor indicator of system performance and the envelope of the ambiguous peaks must be considered [21, p. 40], [23], [24, p. 41]. Further, (51) and (52) predict system performance when signal and noise spectral characteristics are known; for T sufficiently large, these spectra can be estimated accurately. However, in general, (51) and (52) must be modified to account for estimation errors; alternatively, system performance can be evaluated by computer simulation. Empirical verification of expressions for variance has not been undertaken by simulation, because to do so without special purpose correlator hardware would be computationally prohibitive. For example, for a given $G_{s_1 s_1}(f)$, $G_{n_1 n_1}(f)$, $G_{n_2 n_2}(f)$, α, and averaging time T, an estimated generalized cross-correlation function can be computed, from which only one number (the delay estimate) can be extracted. To empirically evaluate the statistics of the delay estimate (which would be valid *only* for these particular signal and noise spectra) many such trials would need to be conducted. We have conducted one such trial (with T large) and verified that useful delay estimates can be obtained by inserting estimates $|\hat{G}_{x_1 x_2}(f)|$ and $|\hat{\gamma}_{12}(f)|^2$ in place of the true values. (This might have been expected since the estimated optimum weighting will converge to the true weighting as $T \rightarrow \infty$. In practice, T may be limited by the stationary properties of the data and (52) may be an overly optimistic prediction of system performance when signal and noise spectra are unknown.)

Conclusions and Discussion

The HT processor has been shown to be an ML estimator for time delay under usual conditions. Under a low S/N ratio restriction, the HT processor is equivalent to Eckart prefiltering and cross correlation. These processors have been compared with four other candidate processors to demonstrate the interrelation of all six estimation techniques. The derivation of the ML delay estimator, together with its relation to various ad hoc techniques of intuitive appeal, suggests the practical significance of HT processing for determination of delay and, thence, bearing. Finally, interpretation of the results leads one to believe that, if the coherence is slowly changing as a function of time, the ML estimation of the source bearing will still be a cross correlator preceded by prefilters that must also vary according to time-varying estimates of coherence.

Appendix

Cramér-Rao Lower Bound on Variance of Delay Estimators

The Cramér-Rao lower bound is given [15, p. 72] by

$$\sigma_{\hat{D}}^2 \geqslant \left. \frac{-1}{E\left\{\dfrac{\partial^2 \ln p(X|Q,\tau)}{\partial \tau^2}\right\}} \right|_{\tau = D}. \tag{A1}$$

The only part of the log density which depends on τ, the hypothesized delay, is J_3 of (44). More explicitly,

$$E\left\{\frac{\partial^2}{\partial \tau^2} \ln p(X|Q,\tau)\right\} = \frac{\partial^2}{\partial \tau^2} E\left(\frac{-1}{2} J_3\right). \tag{A2}$$

If $G_{x_1 x_2}(f) = |G_{x_1 x_2}(f)| e^{-j2\pi f D}$, then since $E[\hat{G}_{x_1 x_2}(f)] = G_{x_1 x_2}(f)$, we have

$$E\left(\frac{-1}{2} J_3\right) = T \int_{-\infty}^{\infty} e^{j2\pi f(\tau - D)} \frac{|\gamma_{12}(f)|^2}{[1 - |\gamma_{12}(f)|^2]}\, df. \tag{A3}$$

Hence the minimum variance is

$$\text{minimum var}\,(\hat{D}) = \left[T \int_{-\infty}^{\infty} (2\pi f)^2 \frac{|\gamma_{12}(f)|^2}{[1 - |\gamma_{12}(f)|^2]}\, df\right]^{-1}. \tag{A4}$$

But this is the variance which the HT processor achieves [see (52)].

References

[1] A. H. Nuttall, G. C. Carter, and E. M. Montavon, "Estimation of the two-dimensional spectrum of the space-time noise field for a sparse line array," *J. Acoust. Soc. Amer.*, vol. 55, pp. 1034–1041, 1974.

[2] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 1965.

[3] W. B. Davenport, Jr., *Probability and Random Processes*. New York: McGraw-Hill, 1970.

[4] G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications*. San Francisco, CA: Holden-Day, 1968.

[5] C. Eckart, "Optimal rectifier systems for the detection of steady signals," Univ. California, Scripps Inst. Oceanography, Marine Physical Lab. Rep SIO 12692, SIO Ref 52-11, 1952.

[6] H. Akaike and Y. Yamanouchi, "On the statistical estimation of frequency response function," *Ann. of Inst. Statist. Math.*, vol. 14, pp. 23–56, 1963.

[7] C. H. Knapp, "Optimum linear filtering for multi-element arrays," Electric Boat Division, Groton, CT, Rep. U417-66-031, Nov. 1966.

[8] A. H. Nuttall and D. W. Hyde, "A unified approach to optimum and suboptimum processing for arrays," Naval Underwater Systems Center, New London Lab., New London, CT, Rep. 992, Apr. 1969.

[9] P. R. Roth, "Effective measurements using digital signal analysis," *IEEE Spectrum*, vol. 8, pp. 62–70, Apr. 1971.

[10] E. J. Hannan and P. J. Thomson, "The estimation of coherence and group delay," *Biometrika*, vol. 58, pp. 469–481, 1971.

[11] G. C. Carter, A. H. Nuttall, and P. G. Cable, "The smoothed coherence transform," *Proc. IEEE* (Lett.), vol. 61, pp. 1497–1498, Oct. 1973.

[12] E. J. Hannan and P. J. Thomson, "Estimating group delay," *Biometrika*, vol. 60, pp. 241–253, 1973.

[13] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[14] G. C. Carter and C. H. Knapp, "Coherence and its estimation via the partitioned modified chirp-Z-transform," *IEEE Trans. Acoust., Speech, Signal Processing* (*Special Issue on 1974 Arden House Workshop on Digital Signal Processing*), vol. ASSP-23, pp. 257–264, June 1975.

[15] H. L. Van Trees, *Detection, Estimation and Modulation Theory, Part I*. New York: Wiley, 1968.

[16] G. C. Carter, A. H. Nuttall, and P. G. Cable, "The smoothed coherence transform (SCOT)," Naval Underwater Systems Center, New London Lab., New London, CT, Tech. Memo TC-159-72, Aug. 8, 1972.

[17] D. W. Hyde and A. H. Nuttall, "Linear pre-filtering to enhance correlator performance," Naval Underwater Systems Center, New London Lab., New London, CT, Tech. Memo 2020-34-69, Feb. 27, 1969.

[18] A. H. Nuttall, "Pre-filtering to enhance clipper correlator performance," Naval Underwater Systems Center, New London Lab., New London, CT, Tech. Memo TC-11-73, July 10, 1973.

[19] L. H. Koopmans, *The Spectral Analysis of Time Series*. New York: Academic, 1974.

[20] R. K. Otnes and L. Enochson, *Digital Time Series Analysis*, New York: Wiley, 1972.
[21] V. H. MacDonald and P. M. Schultheiss, "Optimum passive bearing estimation," *J. Acoust. Soc. Amer.*, vol. 46, pp. 37–43, 1969.
[22] G. C. Carter, C. H. Knapp, and A. H. Nuttall, "Estimation of the magnitude-squared coherence function via overlapped fast Fourier transfrom processing," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 337–344, Aug. 1973.
[23] D. Signori, J. L. Freeh, and C. Stradling, personal communication.
[24] B. V. Hamon and E. J. Hannan, "Spectral estimation of time delay for dispersive and non-dispersive systems," *J. Royal Stat. Soc. Ser. C (Appl. Statist.)*, vol. 23, pp. 134–142, 1974.

Signal Analysis by Homomorphic Prediction

ALAN V. OPPENHEIM, SENIOR MEMBER, IEEE, GARY E. KOPEC, STUDENT MEMBER, IEEE, AND JOSÉ M. TRIBOLET, STUDENT MEMBER, IEEE

Abstract—**Two commonly used signal analysis techniques are linear prediction and homomorphic filtering. Each has particular advantages and limitations. This paper considers several ways of combining these methods to capitalize on the advantages of both. The resulting techniques, referred to collectively as homomorphic prediction, are potentially useful for pole-zero modeling and inverse filtering of mixed phase signals. Two of these techniques are illustrated by means of synthetic examples.**

Introduction

Two classes of signal processing techniques which have been applied to a variety of problems are homomorphic filtering or cepstral analysis [1], [2] and linear prediction or predictive deconvolution [3]–[5]. Separately, each has particular advantages and limitations. It appears possible, however, to combine them into new methods of analysis which embody the advantages of both. In this paper we discuss several ways of doing this.

Linear prediction is directed primarily at modeling a signal as the response of an all-pole system. Its chief advantage over other identification methods is that for signals well matched to the model it provides an accurate representation with a small number of easily calculated parameters. However, in situations where spectral zeros are important linear prediction is less satisfactory. Furthermore, it assumes that the signal is either minimum phase or maximum phase, but not mixed phase. Thus, for example, linear prediction has been highly successful for speech coding [3], [5], [6] since an all-pole minimum phase representation is often adequate for this purpose. It has also been applied in the analysis of seismic data, although limited by the fact that such data often involve a significant mixed phase component.

Homomorphic filtering was developed as a general method of separating signals which have been nonadditively combined. It has been used in speech analysis to estimate vocal tract transfer characteristics [7]–[9] and is currently being evaluated in seismic data processing as a way of isolating the impulse response of the earth's crust from the source function [10]–[12]. Unlike linear prediction, homomorphic analysis is not a parametric technique and does not presuppose a specific model. Therefore, it is effective on a wide class of signals, including those which are mixed phase and those characterized by both poles and zeros. However, the absence of an underlying model also means that homomorphic analysis does not exploit as much structure in a signal as does linear prediction. Thus, it may be far less efficient than an appropriate parametric technique when dealing with highly structured data.

The basic strategy for combining linear prediction with cepstral analysis is to use homomorphic processing to transform a general signal into one or more other signals whose structures are consistent with the assumptions of linear prediction. In this way the generality of homomorphic analysis is combined with the efficiency of linear prediction. In the next section we briefly review some of the properties of homomorphic analysis that suggest this approach. We then discuss several specific ways of combining the two techniques.

Homomorphic Signal Processing

Homomorphic signal processing is based on the transformation of a signal $x(n)$ as depicted in Fig. 1. Letting $\hat{X}(z)$ and $X(z)$ denote the z transforms of $\hat{x}(n)$ and $x(n)$, the system $D_*[\cdot]$ is defined by the relation

$$\hat{X}(z) = \log X(z) \tag{1}$$

where the complex logarithm of $X(z)$ is appropriately defined

Manuscript received April 3, 1975; revised September 30, 1975 and February 3, 1976. This work was supported in part by the Advanced Research Projects Agency monitored by the ONR under Contract N00014-75-C-0951 and in part by the National Science Foundation under Grant ENG71-02319-A02. The work of G. Kopec was supported by the Fannie and John Hertz Foundation. The work of J. Tribolet was supported by the Instituto de Alta Cultura, Portugal.

The authors are with the Department of Electrical Engineering and Computer Science, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139.