A comparision of Generalized cross correlation methods for time delay estimation

A Comparison of Generalized Cross-Correlation Methods for Time Delay Estimation

Ritu and Sanjeev Kumar Dhull***

Time Delay Estimation (TDE) between the signals received at the spatial sensor pair has been a topic of significant importance because this estimated delay can be used as input in plenty of applications for further processing of the signal. The estimation of the delay should be accurate enough in order to avoid or minimize the propagation of the error to the subsequent stages. The generalized cross-correlation method is the most popular and widely used technique for TDE due to accuracy and low computational complexity. The paper presents a comparison of weighting functions used in generalized cross-correlation along with their simulation results. It has been shown how each weight affects the peak sharpening in the correlation function. Simulation has been carried out with different types of sound and noise. The effect of varying Signal-to-Noise Ratio (SNR) on the estimated delay has also been studied.

Keywords: Phase Transform (PHAT), Smoothed Coherence Transform (SCOT), ROTH, Maximum Likelihood (ML), Hassan and Boucher (HB), Cross Correlation (CC), Generalized Cross Correlation (GCC), Time Delay Estimation (TDE)

Introduction

Assume a sound signal originates from a source at certain time; it arrives at different microphones at different intervals depending upon their locations. This estimated time delay is processed further to extract the desired information and therefore the accuracy of the signal being processed depends on this preliminary step. It is widely used in applications such as passive source detection, identification, localization, speech enhancement, speech recognition and many more. This task of estimating delay is very challenging because of the non-stationary nature of speech, coherent noise and of room acoustic reverberation (Patel and Shiva, 2015). Peak in the correlation function gets weakened with low Signal-to-Noise Ratio (SNR) and strong reverberation. Not only this, there occurs multiple peaks in correlation function due to strong echoes and hence the desired estimation could not be extracted. The

* Assistant Professor, Department of Electronics and Communication Engineering, GJUS&T, Hisar, Haryana, India; and is the corresponding author. E-mail: rituboora@gmail.com

** Associate Professor, Department of Electronics and Communication Engineering, GJUS&T, Hisar, Haryana, India. E-mail: sanjeevdhull@yahoo.com

problem of Time Delay Estimation (TDE) has large field of applications, making estimation an important research issue. An estimate of time delay between two windowed sensor signals is obtained by computing a similarity measure between the signals and locating the delay that maximizes (or minimizes) the measure.

This paper is organized as follows: first it describes the signal model; then it discusses the time delay estimation techniques and the results, and finally ends with the conclusion.

Signal Model for TDE

If $r_1(t)$ and $r_2(t)$ are the signals received at microphones 1 and 2, respectively, due to source $S(t)$ in any environment, then the received signal can be modeled as Equations (1) and (2):

$$r_1(t) = h_1(t) * S(t) + n_1(t) \qquad ...(1)$$

$$r_2(t) = h_2(t) * S(t) + n_2(t) \qquad ...(2)$$

where $h_1(t)$ *and* $h_2(t)$ are impulse responses of the channels; $n_1(t)$ and $n_2(t)$ are noises added to the source signal while travelling to microphone array. The noise can be Gaussian or non-Gaussian (usually considered Gaussian for simplification) and uncorrelated with sound signal, and the noises are assumed to be self-uncorrelated.

In case of non-reverberant channel, Equations (1) and (2) get simplified and rewritten as Equations (3) and (4).

$$r_1(t) = S(t) + n_1(t) \qquad ...(3)$$

$$r_2(t) = \alpha S(t - \tau) + n_2(t) \qquad ...(4)$$

where α is the attenuation factor, $0 \le t \le T$, T is the observation interval, $c\tau$ is the extra distance the wave has to travel to reach microphone 2 with respect to microphone 1, c is the speed of sound in the medium, and τ is relative delay of $r_2(t)$ with respect to $r_1(t)$.

TDE Algorithms

Cross-Correlation (CC)

CC is the most straightforward and fastest method of TDE. It simply correlates the microphone outputs to give correlation function $R_{r1r2}(\tau)$ (Marinescu *et al.*, 2013) as:

$$R_{r1r2}(\tau) = E[r_1(t)r_2(t - \tau)] \qquad ...(5)$$

where E is the expectation operator. Usually the speech signals are very long and hence measured for fixed time called "observation interval" (Knapp and Carter, 1976) and therefore Equation (5) is rewritten as:

$$R_{r_1r_2}(\tau) = \frac{1}{T}\int_0^{T-\tau} r_1(t)r_2(t-\tau)dt \qquad ...(6)$$

Here T is the time duration of the signal being observed and is the normalization factor. This limited observation interval puts constraint on the maximum delay which can be estimated. The time argument that maximizes this correlation function is the estimated delay (Knapp and Carter, 1976), and can be represented mathematically as Equation (7):

$$D = argmax\ (R_{r1r2}(\tau)) \qquad ...(7)$$

It is usually convenient to use the correlation function in normalized form, defined as Equation (8):

$$\rho_{r1r2}(\tau) = \frac{R_{r1r2}(\tau)}{\sqrt{R_{r1r1}(0)R_{r2r2}(0)}} \qquad ...(8)$$

where $R_{r1r1}(0)$ and $R_{r2r2}(0)$ are the values of the auto-correlation functions at $\tau = 0$. Further, it may be pointed out that it is usually preferred to find the correlation in frequency domain as it eliminates the need for time shifting and multiplication operations needed for correlation in time domain. It is calculated in frequency domain (Knapp and Carter, 1976) as:

$$R_{r1r2}(\tau) = \int_{-\infty}^{\infty} G_{r1r2}(f)e^{j2\pi f\tau}df \qquad ...(9)$$

where

$$G_{r1r2}(f) = r_1(f)r_2^*(f) = [s(f) + n_1(f)][\alpha s(f)e^{-jw\tau} + n_2(f)]^* \qquad ...(10)$$

is the cross-power spectral density (CSD) of the received signals, $r_1(f)$ and $r_2^*(f)$ are the FFTs of $r_1(t)$ and $r_2(t)$, respectively, and * means complex conjugate. If it is assumed that signal and noise are uncorrelated, then Equation (10) gets simplified to Equation (11), as other terms reduce to zero.

$$G_{r1r2}(f) = [\alpha G_{ss}(f)e^{-jw\tau} + G_{n1n2}(f)] \qquad ...(11)$$

If the distance between sensors recording the two signals is greater than the spatial de-correlation distance of the noise process, then $n_1(t)$ and $n_2(t)$ are uncorrelated, therefore $G_{n1n2}(f)$ becomes negligible (Knapp and Carter, 1976) and Equation (11) reduces to:

$$G_{r1r2}(f) = \alpha G_{ss}(f) e^{-jw\tau} \quad ...(12)$$

Usually the signal and noise are correlated and hence contribute to CSD, which simply means that the other two terms in Equation (10) will not reduce to zero. The measured signal may vary with time due to noise and attenuation and hence the correlation may vary. As the SNR value decreases, the correlation peak gets spread and flattens up, as shown in Figures 1d and 2c for the CC, and thus it is tedious to estimate the correct delay (Aguilar, 2012). The discussed frequency domain method can only be used to estimate short-time delays, as we use finite length FFT and limited observation interval. If FFT length is increased further for large delays, then computational complexity is increased dramatically.

The spreading of peak can also be explained by expanding the correlation term Equation (6) as:

$$R_{r1r2} = \alpha R_{ss}(\tau) * \delta(t-\tau) + R_{sn2}(\tau) + R_{n1s}(\tau) + R_{n1n2}(\tau) \quad ...(13)$$

The sample cross-correlation converges to the first term and the other three terms constitute zero mean disturbance affecting the peak resolution of the first term only if signal and noise are uncorrelated and noises are self-uncorrelated. But due to finite observation interval and other factors, the last three do not reduce to zero and thus contribute noise and spreading to peak in the correlation function (Knapp and Carter, 1976). This suggests the use of pre-filtering or pre-whitening the signal before correlating them, which has been explained in Generalized Cross-Correlation (GCC).

GCC Methods

GCC was proposed by Knapp and Carter (1976) as an improvement over cross-correlation, and it is the most widely used technique for TDE due to its accuracy and moderate computational complexity. This approach to the TDE problem is derived under single-path conditions and produces theoretical optimal results only under ideal noise conditions. Frequency domain weights are applied to the cross-power spectral density $G_{r1r2}(f)$ of the signals and then Wiener-Khinchin theorem is used to calculate cross-correlation as:

$$R_{r1r2}(\tau) = \int_{-\infty}^{\infty} \psi(f) G_{r1r2}(f) e^{j2\pi f\tau} df \quad ...(14)$$

where $\psi(f)$ is the weighting function, and the delay can be estimated in Equation (15) as:

$$D = \arg max[R_{r1r2}(\tau)] \quad ...(15)$$

For p microphones Equation (14) can be written as:

$$R_{r1r2}(\tau) = \sum_{i=1}^{p} \sum_{k=1}^{p} \int_{-\infty}^{\infty} \psi_{ik}(f) r_1(f) r_2^*(f) e^{j2\pi f(\tau_i - \tau_k})df \quad \text{...(16)}$$

The objective behind using these whitening filters is to sharpen the peak in cross-correlation function, which gets flattened with noise, by de-emphasizing the signal spectrum affected by it. GCC aims solely to calculate the largest peak corresponding to the delay between signals and not the correct estimation of cross-correlation, and hence the use of the weighting function does not affect our results. It also has certain limitations as performance degrades considerably in the presence of harmonic sound or generally pseudo-periodic sounds. Different weights used in GCC are discussed below:

Phase Transform (PHAT)

It is the traditional and most widely used weighting function which places equal importance on each frequency by dividing cross-spectrum by its magnitude, due to which it is sub-optimal under ideal conditions but tends to be less vulnerable to inconsistent conditions, particularly reverberation (Marinescu *et al.*, 2013). It is defined as:

$$\psi_{PHAT}(f) = 1/\,|\,G_{r1r2}(f)\,| \quad \text{...(17)}$$

Substitute Equation (17) in Equation (14), and we get:

$$R_{r1r2}(\tau) = \int_{-\infty}^{\infty} \frac{G_{r1r2}(f)}{G_{r2r2}(f)} e^{j2\pi f \tau} df \quad \text{...(18)}$$

In this case, output will be a delta function at time D if $G_{r1r2}(f) = |G_{r1r2}(f)|$, which means the cross-correlation becomes independent of source signal and depends only on the channel impulse response and therefore results in a sharp narrow spreading peak (Aguilar, 2012). This also shows that output delta function is independent of frequency of signal and thus of the filter. Denominator tends to be zero under small signal energy, so it will increase the error. In some improved methods, it is considered to add a fixed constant in the denominator. PHAT can be considered as special case of Maximum Likelihood (ML) algorithm when environment noise effect is negligible, which explains its good performance under low noise environments even under relatively heavy reverberation. PHAT is based on heuristic approach hence explanation of its efficiency has not fully been explored yet (Zhang *et al.*, 2008). The optimality of PHAT is independent of reverberation which proves its robustness to reverberation. It has certain disadvantages as it does not consider the noises in the signals $r_1(t)$ and $r_2(t)$; and as it is inversely proportional to $|G_{r1r2}(f)|$, it accentuate the errors, especially when the SNR is low. Impulsive noise and high reverberation degrade the performance of PHAT and therefore an improved method based on order statistics and via signal detection has been proposed (Patel and Shiva, 2015).

Modified Cross Spectral Phase (M-CSP)

M-CSP tries to improve the performance of PHAT in certain environments. This has been done with a target to control the whitening of signal and limit the amount of degradation from the independent noise (Marinescu *et al*., 2013). Mathematically, the weighting function can be expressed as:

$$\psi_{M-csp}(f) = 1/\,|\,G^{m}_{r1r2}(f)\,| \qquad \text{...(19)}$$

Maximum Likelihood

The ML weighting (Marinescu *et al*., 2013) function aims at giving the likelihood solution for TDE as:

$$\psi_{ML}(f) = \frac{1}{|\,G_{r1r2}(f)\,|}\frac{|\,\gamma_{r1r2}(f)\,|^2}{1-|\,\gamma_{r1r2}(f)\,|^2} \qquad \text{...(20)}$$

where

$$|\,\gamma_{r1r2}(f)\,|^2 = \frac{|\,G_{r1r2}(f)\,|^2}{G_{r1r1}(f)G_{r2r2}(f)} \qquad 0 < |\,\gamma_{r1r2}(f)\,|^2 < 1 \qquad \text{...(21)}$$

is the magnitude-squared coherence function which enables us to identify significant frequency-domain correlation between the two time series and therefore CC becomes:

$$R_{r1r2}(\tau) = \int_{-\infty}^{\infty} \frac{G_{r1r2}(f)}{|\,G_{r1r2}(f)\,|}\frac{|\,\gamma_{r1r2}(f)\,|^2}{1-|\,\gamma_{r1r2}(f)\,|^2}e^{j2\pi f\tau}df \qquad \text{...(22)}$$

ML has two prefilters Equation (22); the first one $\frac{1}{|\,G_{r1r2}(f)\,|}$ is similar to PHAT weighting Equation (17) and performs peak sharpening; the fundamental term $\frac{|\,\gamma_{r1r2}(f)\,|^2}{1-|\,\gamma_{r1r2}(f)\,|^2}$ weights the cross-spectral phase according to the variance of phase such that its big weights are used when coherence is near unity (high) and de-emphasizes the frequencies with near zero coherence (Aguilar, 2012). This leads to minimum variance of TDE. But here errors will be enlarged when energy of the signal is small. ML is equivalent to Smoothed Coherence Transform (SCOT) at low SNR. Similarly, at certain conditions, it is equivalent to Hannon Thomson and MacDonald and Schultheiss (Knapp and Carter, 1976). This method is also efficient for ideal conditions such as low reverberation, large observation interval, uncorrelated stationary Gaussian signal and noises, and in addition spectra of the noise is known a priori. Usually coherence function is not available beforehand and therefore estimated from the given data by temporal averaging technique which becomes problematic for any non-stationary signal (Roth, 1971). When reverberations are introduced, the cross-spectrum phase noise terms may be significantly biased, rendering ML weighting inappropriate and producing significant error in the corresponding estimators (Hero and Schwartz, 1985). The modification of the weighting by use of the PHAT dissipates some of these effects which is considered as a special case of ML. Dranka and Coelho (2015) provided the robust ML method in the case of non-Gaussian and uncorrelated noise which can severely degrade the performance. Champagne *et al.* (1996) studied the effect of reverberation on time delay estimation using maximum likelihood method. The effect of reverberation has been developed using the image source method.

Roth Filter

With a ROTH weighting, cross-correlation can be given as (Roth 1971; and Marinescu *et al.*, 2013):

$$R_{r1r2}(\tau) = \int_{-\infty}^{\infty} \frac{G_{r1r2}(f)}{G_{r1r1}(f)} e^{j2\pi f\tau} df \qquad ...(23)$$

where

$$\psi_{ROTH} = \frac{1}{G_{r1r1}(f)} \qquad ...(24)$$

The cross-spectrum is divided by auto spectrum of $r_1(t)$ signal and this division has justification only when $r_1(t)$ is the input to a linear system. When there is no such physical interpretation, then it leads to confusion as to divide by auto spectra of $r_1(t)$ or $r_2(t)$ as error in $G_{r1r2}(f)$ may be due to $G_{n1n1}(f)$ or $G_{n2n2}(f)$ (Roth, 1971). This problem led to the development of SCOT filter. In ROTH filter, both the amplitude and the phase information are saved in contrast to the SCOT, where the amplitude information is lost. Both the argument and the amplitude are required to calculate an impulse response to obtain a perfect peak in CC, which is an advantage over SCOT filter. It is equivalent to the Wiener filtering, which can effectively suppress the big-noise band, but will broaden the peak of correlation function.

If $G_{n1n1} \neq 0$ then $G_{r1r1}(f) = G_{ss}(f) + G_{n1n1}$ and correlation becomes (Knapp and Carter, 1976).

$$R_{r1r2}(\tau) = \delta(\tau - \tau_{cc}) * \int_{-\infty}^{\infty} \frac{G_{r1r2}(f)}{\{G_{ss}(f) + G_{n1n1}\}} e^{j2\pi f \tau} df \qquad ...(25)$$

ROTH has the desirable effect of suppressing the frequency regions in which power spectral noise density, $G_{n1n1}(f)$ or $G_{n2n2}(f)$ is large (Roth, 1971) and the estimate of the cross-power spectral signal density, $G_{x1x2}(f)$ is likely to be in error. However, this filter does not minimize the spreading effect of the delta function whenever the power spectral noise density is not equal to some constant (including zero) times the power spectral density of the signal, $G_{s1s1}(f)$ (Knapp and Carter, 1976).

Smoothed Coherence Transform (SCOT)

A variation of the ROTH weight is the Smoothed Coherence Filter (SCOT) (Carter *et al.*, 1972). It also acts upon the SNR-based weighting concept, but allows the signals being compared to have a different spectral noise density. Whitening of the signal with SCOT is lesser as compared to PHAT. It is based on heuristic technique and was basically designed to deal with correlated noise corrupted broadband signals. This filter being an intuitive correlation type, is not much sensitive to deviation from the assumed model, whereas Eckart (EK), Hannon Thomson (HT) and Hassab Boucher (HB) are the optimum filters (they maximize some performance criteria) and therefore very sensitive to deviation from the assumed characteristics of signal and noise (Hero and Schwartz, 1985). A SCOT filter is defined as Fourier transform of the weighted coherence function.

$$R_{r1r2}(\tau) = \int_{-\infty}^{\infty} \frac{G_{r1r2}(f)}{\sqrt{G_{r1r1}(f) G_{r2r2}(f)}} e^{j2\pi f \tau} df \qquad ...(26a)$$

$$R_{r1r2}(\tau) = \int_{-\infty}^{\infty} \rho_{r1r2}(f) e^{j2\pi f \tau} df \qquad ...(26b)$$

where ρ is the coherence estimate. When $G_{r1r1}(f) = G_{r2r2}(f)$, it becomes a ROTH filter which can be considered its special case (Roth, 1971).

The SCOT estimator can be interpreted similar to ML as two filtering processes as:

$$\psi_{scot} = \gamma_{r1r2}(f) / |G_{r1r2}(f)| \qquad ...(27)$$

Here numerator does the attenuation in low SNR and denominator does the pre-filtering or pre-whitening similar to PHAT. If the power of uncorrelated noise components is greater than the signal, true peak may be obscured. These noise components may not be received by filtering since signal may contain both broadband and narrowband components.

Weiner

It is a linear filter (Hero and Schwartz, 1985) that removes additive noise of given variance from a noisy signal. Based on the channel's linearity, it gives the optimal linear estimate of the original signal from the observation $r_1(t)$ and channel output signal from $r_2(t)$, by minimizing the linear mean square errors (Marinescu *et al.*, 2013). In this way, given the channel characteristics, the solution results in Wiener filters, which yield the Wiener weighting function.

$$\psi_{Weiner} = |\gamma_{r1r2}(f)|^2 \quad \text{...(28)}$$

The filter needed to minimize the MSE are defined as Equation 29a & 29b, where c(ω) stands for the channel characteristics

$$H_1(\omega) = \frac{1}{c(\omega)} * \frac{G_{12}(\omega)}{G_{11}(\omega)} \quad \text{...(29a)}$$

$$H_2(\omega) = \frac{1}{c(\omega)} * \frac{G_{12}(\omega)}{G_{22}(\omega)} \quad \text{...(29b)}$$

Eckart

Finally, the Eckart filter (Eckart, 1952) maximizes the deflection criterion, i.e., the ratio of the change in mean correlation output due to signal present compared to the standard deviation of correlation output due to noise alone (Marinescu *et al.*, 2013). In practice, the Eckart filter requires knowledge of the signal and noise spectra. Similar to the SCOT, it suppresses frequency bands of high noise and assigns zero weight to bands, where $G_{s1s1}(f) = 0$. For low frequency regions, it is similar to HT function (Carter *et al.*, 1972; and Knapp and Carter, 1976).

$$\psi_{Eckart} = \frac{G_{r1r2}(f)}{[G_{r1r1}(f) - |G_{r1r2}(f)|][G_{r2r2}(f) - |G_{r1r2}(f)|]} = \frac{G_{r1r2}(f)}{G_{n1}(f)G_{n2}(f)} \quad \text{...(30a)}$$

$$\psi_{Eckart} = \frac{s(w)s(w)}{n_1(w)n_1^*(w)n_2(w)n_2^*(w)} \quad \text{...(30b)}$$

Hassab and Boucher

This was presented by Hassab and Boucher (HB). HB varies from the SCOT filter as it suppresses the CSD of the regions of high SNR so that the strong tonal can be diminished from the observed signal along with low SNR regions (Marinescu *et al.*, 2013).

$$\psi_{HB} = \frac{|G_{r1r2}(f)|}{G_{r1r1}(f)G_{r2r2}(f)} \quad ...(31)$$

Results and Discussion

These GCC methods have been shown to be effective in reducing the peak spreading by minimizing the effect of noise and reverberation up to some extent. This paper presented the GCC for acoustic signals in frequency as it eliminates the need for selection of cut-off frequencies required for the fine-tuning of filters in time domain, but at the expense of a small loss of accuracy. It can be concluded that the maximum value of time delay measured using cross-correlation in frequency domain is limited by the length of FFT. Increase in FFT length adds computational burden, which increases the calculation time requirement and hence puts constraint on its real-time use. PHAT gives a delta function located at the exact time delay but does not consider the effect of background noise, whereas SCOT and ML estimators do so. ML estimator has the effect of overemphasizing and underemphasizing at certain frequencies and therefore not preferred over SCOT. These GCC methods have given satisfactory results only if noise and interference are not temporally and spectrally coincident with $s(t)$. GCC methods require the difference in TDOA for each signal to be greater than the width of the correlation function, else it cannot resolve the two overlapping signals. As these methods are not signal selective and therefore produce peak for interference as well when it is spectrally coincident with signal.

In Figure 1c, CC has found using signal in 1a and noise in 1b. From this, we can see that the correlation peak, which got spread with conventional technique, has been sharpened with GCC weighting. Regions with low SNR have also got diminished as compared to CC by using these weighting functions. Difference of using M-CSP is sudden fall in amplitude near peak. We can notice the difference that ROTH has a sharper peak than SCOT but suppression for delay other than true is less. Most of the weighting function has worked out satisfactorily for this high SNR. Figure 1d shows that how correlation values for the signal and noise in Figure 1a and Figure 1b degrades for the low SNR values.

It can be easily observed that due to increase in noise, correlation values have increased for delay other than true and it becomes tedious to estimate the delay. We can see that SCOT, HB and PHAT give accepectable results, whereas ML and ROTH

Figure 1a: Signal

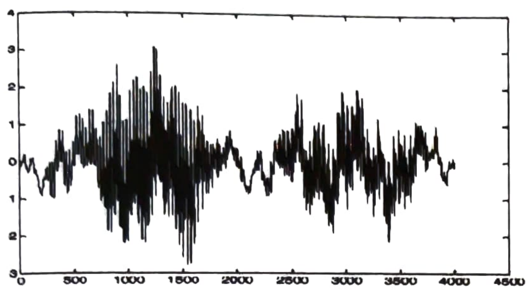


Figure 1b: Noise Signal

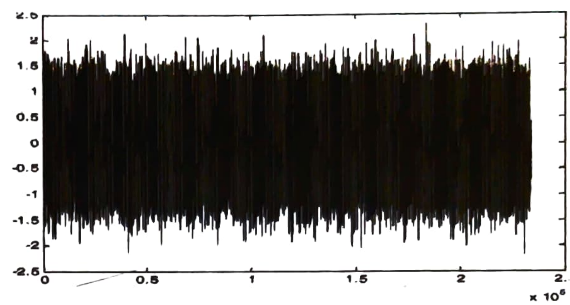


Figure 1c: GCC with Different Weights for Signal 1a and Noise 1b with SNR = 16.4558 dB

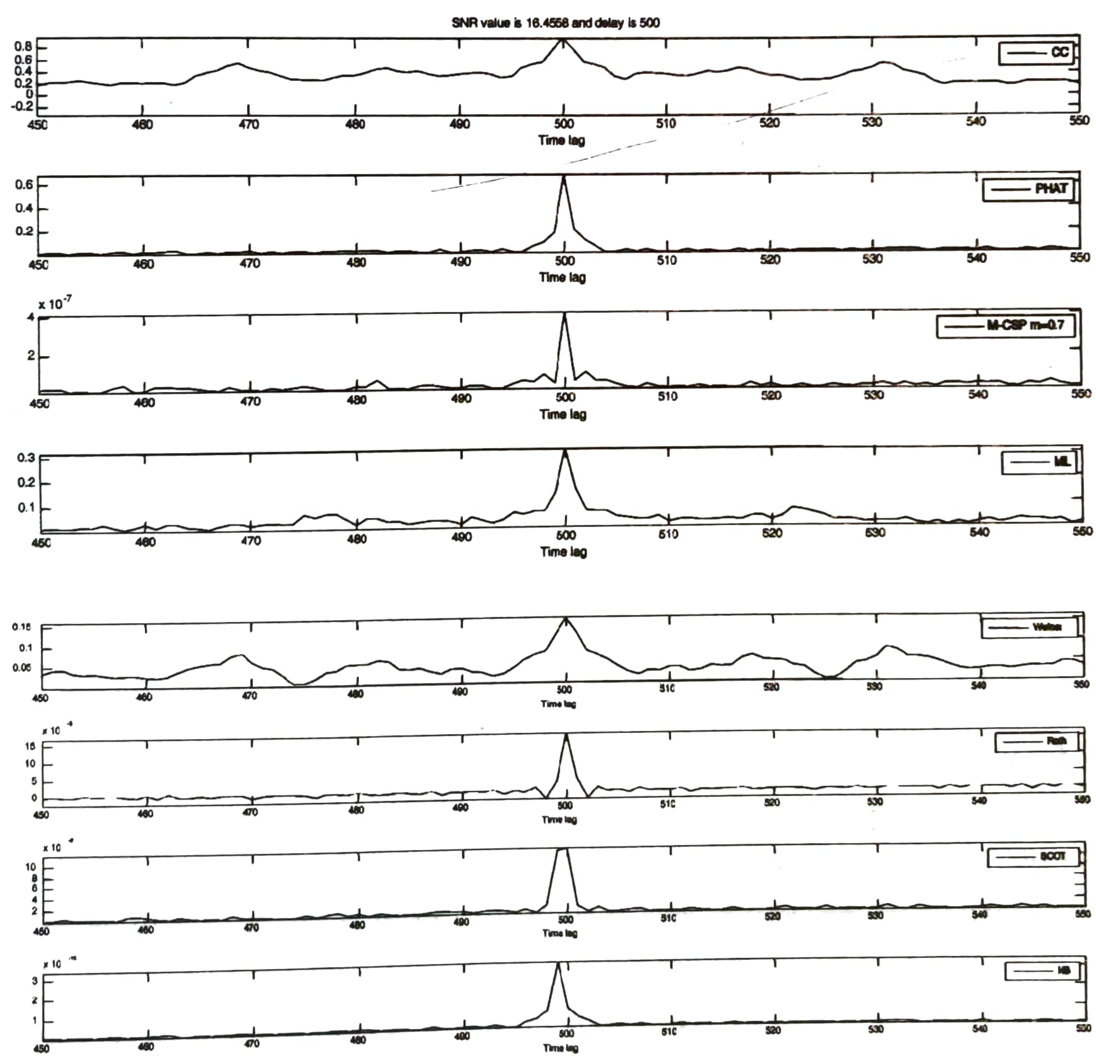
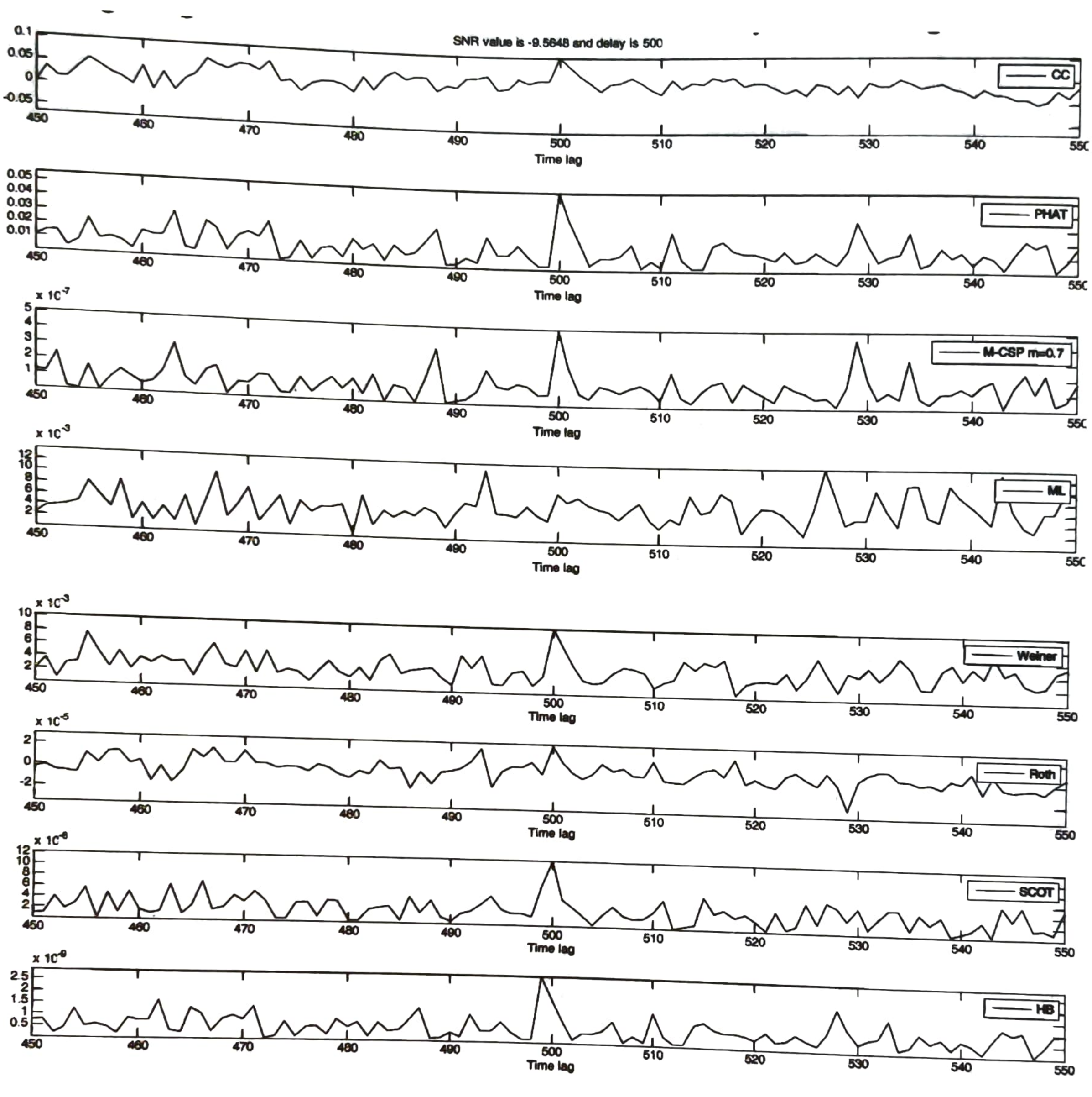


Figure 1d: GCC with Different Weights (Labeled) for Signal 1a and Noise Signal 1b and SNR = 9.568 dB



have failed for SNR value. In the case of m-CSP, peaks other than for true delay have approximately same vlaue of CC and hence this leads to erroneous results.

Figure 2c shows that different weights have been tested on real world noise (Figure 2b) and signal recorded in anechoic chamber in Figure 2a. Here the wiener and ML filter outputs are totally unacceptable.

In Figure 3c, the correlation again tested on Gunshot sound signal (Figure 3a) and wind noise (Figure 3b). Here we can see that wiener and conventional filter gave a straight line for the correlation values and hence we cannot distinguish the peak.

Figure 2a: Signal Recorded in Anechoic Chamber

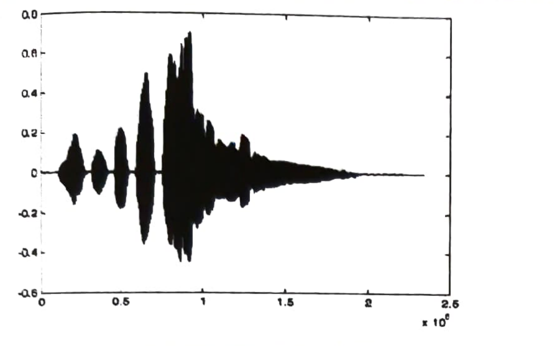


Figure 2b: Wind Noise

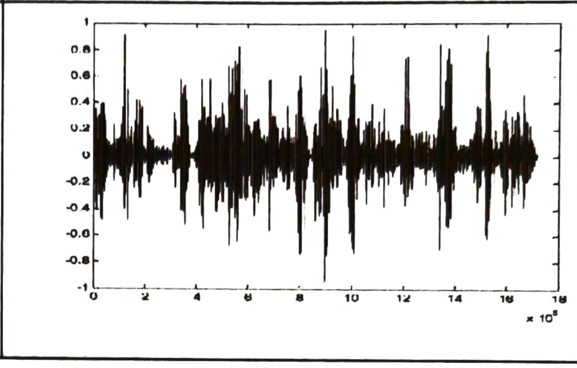


Figure 2c: GCC with Different Weights for the Signal 2a and Wind Noise 2b

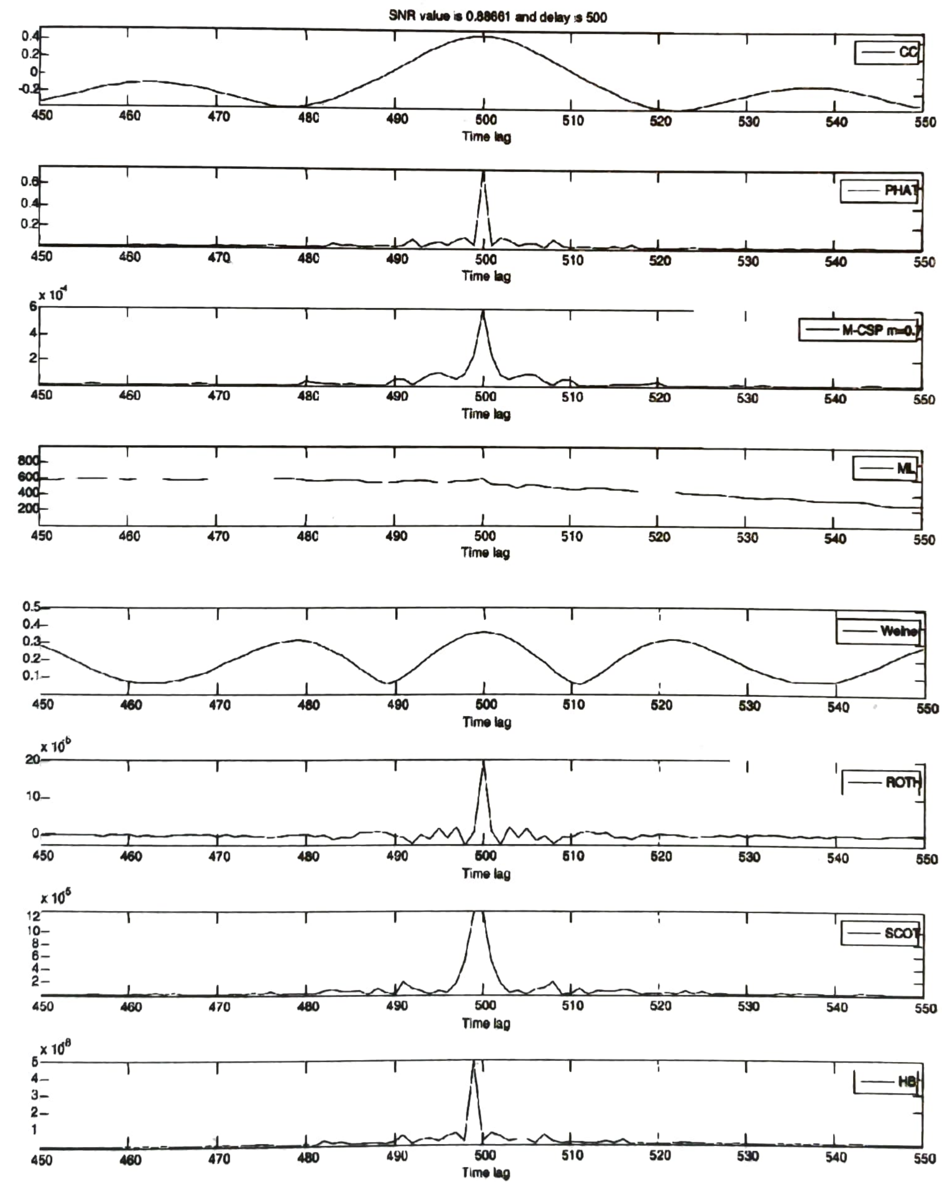


Figure 3a: Signal (Gun Shot)

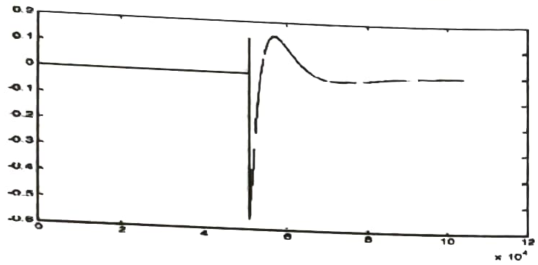


Figure 3b: Wind Noise

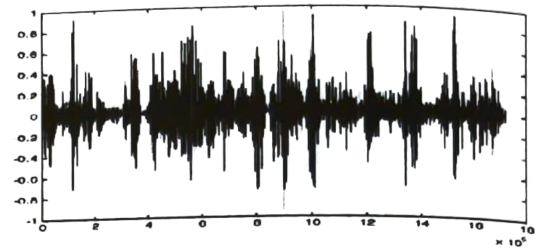
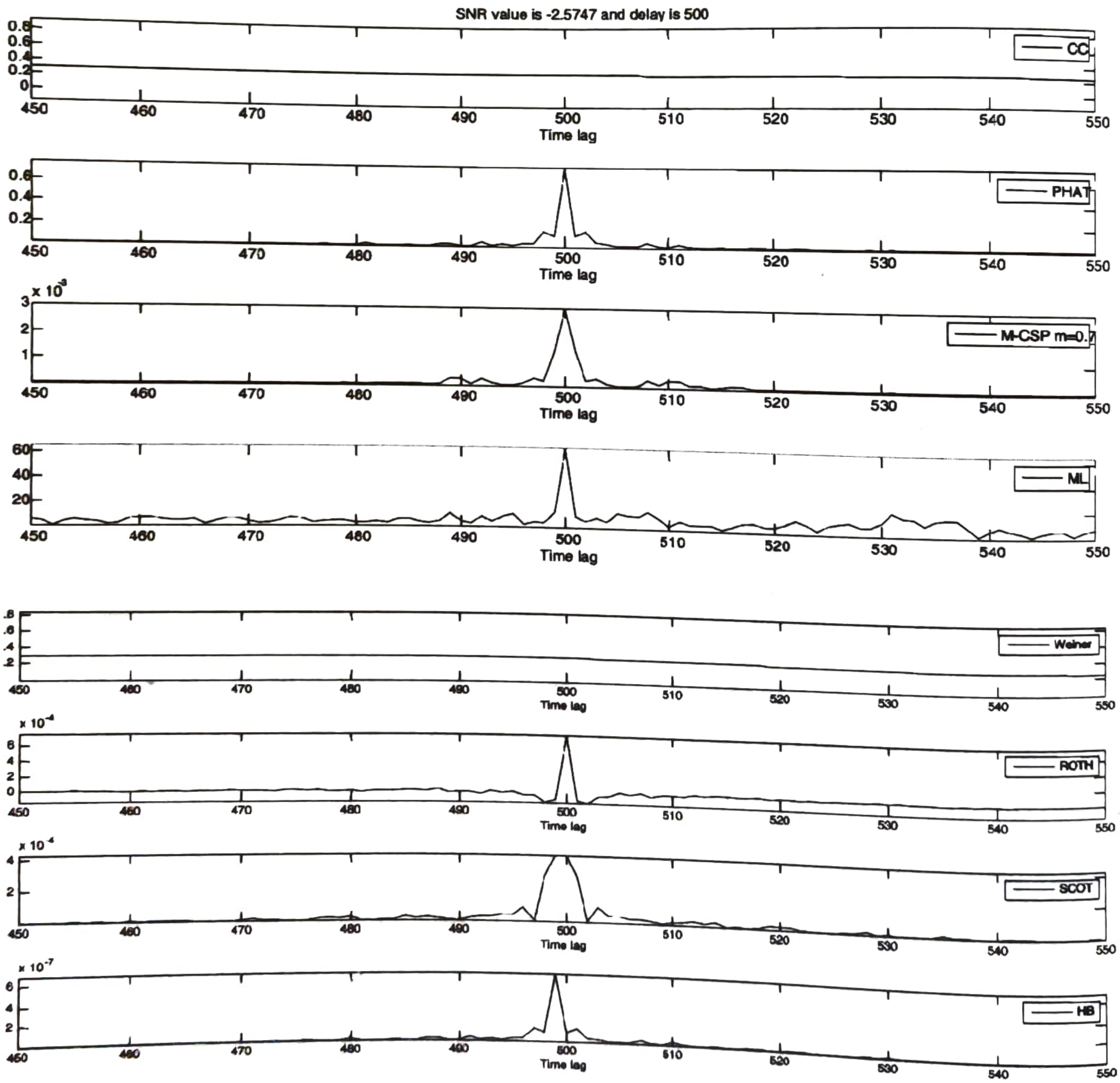


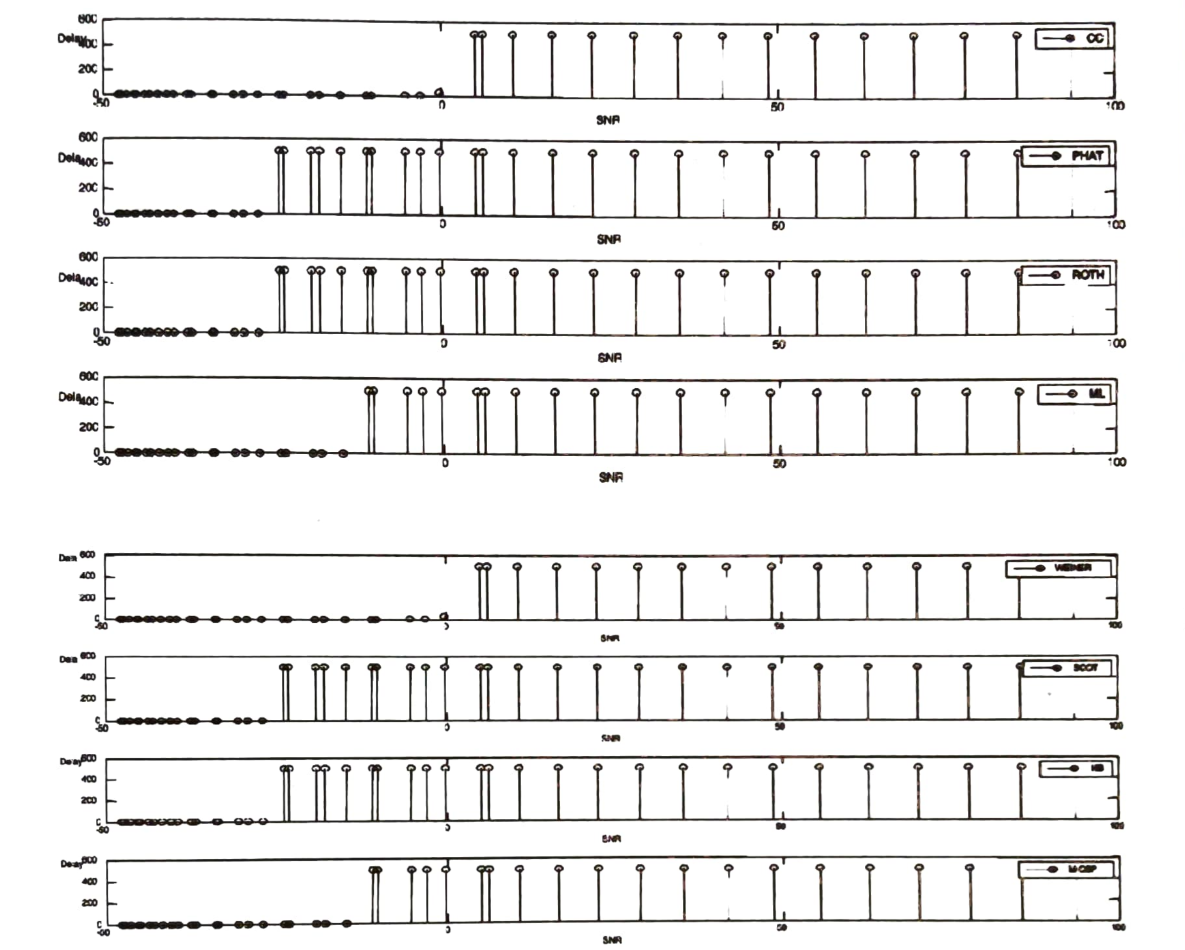
Figure 3c: GCC with Different Weights for the Signal 3b and Wind Noise 3c



Note: The difference can be noted in ROTH and SCOT peak spreading.

Figure 4 shows that different weights work for different lower SNRs. PHAT, HB, SCOT and ROTH have worked effectively for comparatively low SNR as compared to ML.

Figure 4: Delay Estimation for Varying SNR for Signal Figure 1a and Wind Noise Figure 2b



Note: The difference can be noted in ROTH and SCOT peak spreading.

Conclusion

This study has been done with only one additive noise, and no effect of reverberation has been considered. To reach a conclusion about the selection of a weight, we have to see the type of environment, bandwidth of the signal, application area and many more. This field has future scope to reduce the effect of coherent noise, reverberation, improving performance in low SNR and estimating large delays with minimum complexity and time requirement and making them suitable for real-time application.

References

1. Aguilar Cosme Llerena (2012), "A Comparative Study of Time-Delay Estimation Techniques for Convolutive Speech Mixtures", *Advances in Computer Science*,

Proceedings of the 6th WSEAS European Computing Conference (ECC '12), Prague, Czech Republic, September, pp. 24-26.

2. Carter G C, Nuttall A H and Cable P G (1972), "The Smoothed Coherence Transform (SCOT)", *Tech. Rep.*, Naval Underwater Center, New London Lab.
3. Champagne B, Bédard S and Stéphenne A (1996), "'Performance of Time-Delay Estimation in the Presence of Room Reverberation", *IEEE Trans. Speech Audio Process*, Vol. 4, No. 2, pp. 148-152.
4. Dranka E and Coelho R (2015), "Robust Maximum Likelihood Acoustic Energy Based Source Localization in Correlated Noisy Sensing Environments", *IEEE Journal of Selected Topics in Signal Processing*, Vol. 9, No. 2, pp. 259-268.
5. Eckart (1952), "Optimal Rectifier System for Detection of Steady Signals", Univ. California, Scripps Inst. Oceanography, Marine Physical Lab. Rep SIO 12692, SIO Ref 52-11.
6. Hero Alfred and Schwartz Stuart (1985), "A New Generalised Cross Correlator", *IEEE Transction on Acoustic Speech and Signal Processing*, Vol. 33, No. 1.
7. Knapp C and Carter G (1976), "The Generalized Correlation Method for Estimation of Time Delay", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 24, No. 4, pp. 320-327.
8. Marinescu Radu-Sebastian, Buzo Andi, Cucu Horia and Burileanu Corneliu (2013), "Applying the Accumulation of Cross-Power Spectrum Technique for Traditional Generalized Cross-Correlation Time Delay Estimation", *International Journal on Advances in Telecommunications*, Vol. 6, Nos. 3&4.
9. Patel Bipin and Shiva Ram Krishan Vadali (2015), "On Methods to Improve Time Delay Estimation for Underwater Acoustic Source Localization", *International General of Geo Marine Science*, Vol 44, No. 2, pp. 237-244.
10. Roth P R (1971), "Effective Measurements Using Digital Signal Analysis", *IEEE Spectrum*, Vol. 8, No. 4, pp. 62-70.
11. Zhang Cha, Florêncio Dinei and Zhang Zhengyou (2008), "Why Does Phatwork Well in Low Noise, Reverberative Environments?", *ICASSP*, pp. 2565-2569, IEEE.

Reference # 70J-2016-11-03-01