A New Statistical Measure of Signal Similarity

CONFERENCE PAPER · MARCH 2007

DOI: 10.1109/IDC.2007.374535 · Source: IEEE Xplore

1 AUTHOR:

Hugh Lachlan Kennedy
University of South Australia

A New Statistical Measure of Signal Similarity

Hugh L. Kennedy [1#]
[1]*Pacific Noise and Vibration (PNV) Pty. Ltd.,*
GPO Box 429, Canberra ACT 2601, Australia, hkennedy@pnv.com.au

Abstract

A new measure of similarity, suitable for both signal processing and image processing applications, is presented. The measure of similarity (Z_M) is an F-distributed test statistic that quantifies the degree of alignment, correlation or coincidence between two or more (time- or space-dependent) waveforms. The test statistic may be used to automate detection, classification, localisation, association and registration operations. When used to estimate time-of- arrival differences, which is typically done using cross correlation or beam formation, the technique provides an efficient means of: maintaining a low and constant false-alarm rate, detecting weak signals in noise and reducing angular errors. The method is applied to the problem of acoustic source detection and localisation. Real data are used to compare the method with cross correlation and normalised cross correlation. The new method yields prominent and narrow peaks, at the true source locations, with false peaks due to misalignment, misassociation, and background noise, suppressed. When processing continuous data streams, the test statistic is computed efficiently in the time domain.

1. Introduction

Measures of similarity are required in a wide range of radar sonar, communications, remote sensing, artificial intelligence and medical applications, where one signal or image is compared with another. Many basic processing operations, such as matched filtering, cross correlation, and beam formation, may be interpreted as being based on measures of similarity. These related operations typically form the foundation of the detection, classification, localisation, association and registration algorithms employed in semi-autonomous sensor systems. Matched filtering involves the projection of an unknown waveform onto a known template; cross correlation is used to determine the translation (in time or space) required to bring one waveform (signal or image) into alignment with another; while beam formation brings multiple signals into phase so that they constructively interfere. A new measure of similarity [1] is derived in this paper and applied to the problem of acoustic source localisation.

Beam-formation [2] and cross-correlation processing techniques [3] are used to compute Time-Of-Arrival-Differences (TOADs) or Time Delay Estimates (TDE) in distributed networks of acoustic sensors [4&5]. Networks may be composed of single-sensor nodes, where low-resolution digitised data are shared over the network [6], or of multi-sensor nodes, where high-resolution digitised data are processed at each node, with only direction vectors [7] (and possibly range [8]) shared over the network. Both processing techniques and both network topologies employ detection and localisation processing stages. Cross correlation and beam formation processing are typically implemented using detect-then-localise and localise-then-detect architectures, respectively. Detection prior to localisation is efficient because a detection filter is applied to only one or more channels of digitised sensor data and the localiser is only applied to short data segments containing transient features of interest, as identified by the detector. Localisation prior to detection, using time-domain delay-and-sum beam-formation, is computationally more demanding, because the detector is applied to the outputs of a large parallel bank of localising filters, however, this architecture allows both continuous and transient signals to be processed.

Localisation by beam formation can be applied to any number of sensors, while localisation by cross correlation is restricted to sensor pairs, with the pairwise results combined to form a consolidated angular estimate during a subsequent (fusion) processing stage. Both cross-correlation and beam-formation processes output a localisation function, computed in either the time or frequency domain. Maxima (peaks) in this function are used to identify the most likely lag (in the case of cross correlation) or beam (in the case of beam formation). The wrong lag or beam may be selected in the presence of: reverberation, multi-path propagation, multiple sources, and background noise. Angular errors due to the selection of the wrong peak are much greater than the errors in the location of the true peak. Unfortunately, when sampling rates are low, and when sound sources are far from broadside, even small TOAD errors result in large Angle-Of-Arrival (AOA) errors. A variety of alternatives to cross correlation, such as normalised cross correlation, normalised covariance and sum-of-squared differences have been proposed and compared in the literature [9], in an effort to mitigate the problems described above. These two-channel methods will be collectively referred to here as correlation, or lag-estimation techniques.

As distributed arrays become larger, relative to the frequencies and the pulse durations of the sounds that they are processing, the issue of TOAD ambiguities becomes more significant. Reducing the quality of the sensor data (sampling rates, number of channels and quantization bits), as required in capacity constrained and disadvantaged networks, only makes matters worse. To meet these challenges, the new measure of similarity described in this paper is used as the basis of a robust and efficient TOAD estimation algorithm.

The TOAD estimation algorithm operates on two or more channels of sensor data and outputs an F-distributed test statistic that is ideal for joint localisation and detection. It has the potential to enhance the performance of distributed acoustic networks, when processing both impulsive transients and continuous waveforms. Computer simulation and extensive live testing using gun-shot data (processed in real time) by PNV have demonstrated improvements in the following areas relative to conventional cross correlation and beam formation techniques: increased angular resolution, reduced (and theoretically constant) false-alarm rates in the presence of background noise, and an increased probability of detecting weak signals. The test statistic may also be used as a measure of localisation quality (or reliability) as misaligned or misassociated waveforms, in the presence of signal distortion or background interference, are likely to yield a low measure of similarity. Filtering out AOA estimates of low quality reduces the average angular error and the false detection rate of the system, however, the probability of detection is also reduced. A measure of AOA reliability has also been derived in [10]; however, it is applied at the array level where multiple pairwise TOADs are combined to form a consolidated AOA vector, not at the signal level, during TOAD estimation.

2. DERIVATION

A statistical treatment of a delay-and-sum beam-former will now be described and used to derive the new measure of signal similarity. The derivation is based on a few standard statistical relationships [11]. A hypothesis test is performed with the null hypothesis being that there is no signal present and that the waveforms entering the beam former contain only zero-mean Gaussian-distributed noise. It is assumed that any Direct Current (DC) offset in the data (e.g. sensor bias) or frequencies that are of no interest (e.g. wind or self noise) have been removed by a pre-whitening stage. If the null hypothesis is rejected then it is assumed that a localisable signal is present. The test statistic for all possible lag combinations corresponding to all physically measurable angles is computed. The most likely direction of the source is set equal to the angular coordinate for which the null hypothesis is least likely, i.e. the test statistic is maximised.

The delay-and-sum beam-former is applied as

$$y(n) = \sum_{m=0}^{M-1} x_m(n), \tag{1}$$

where $x_m(n)$ is the nth sample output from the mth delay channel and $y(n)$ is the beam-formed output. In Eq. (1) it is assumed that the appropriate delays have been applied to steer a beam in a desired direction. The noise statistics of every sample from all sensors are assumed to be identical, so the nth sample in each delay channel is assumed to be an independent observation of the random variable X_n. Analysing the digitised waveforms (in x) over a window of length N, gives a total of N different random variables, with M observations of each variable. Under the null hypothesis the variables have a Gaussian (Normal) distribution

$$X_n \sim \mathrm{N}\left(\mu_n, \sigma_n^2\right). \tag{2}$$

At a given n, using the data from all M channels, the Maximum Likelihood Estimates (MLEs) $\hat{\mu}_n$ and $\hat{\sigma}_n^2$, of the (true) mean and variance μ_n and σ_n^2, are computed using

$$\hat{\mu}_n = \frac{y(n)}{M} \tag{3}$$

and

$$\hat{\sigma}_n^2 = \frac{1}{M}\left\{\sum_{m=0}^{M-1} x_m(n)^2 - \frac{y(n)^2}{M}\right\}. \tag{4}$$

Under the null hypothesis the following relationships hold:

$$\text{If } Z_a = M\frac{\left(\hat{\mu}_n - \mu_n\right)^2}{\sigma_n^2} \text{ then } Z_a \sim \chi^2(1); \tag{5}$$

$$\text{If } Z_b = M\frac{\hat{\sigma}_n^2}{\sigma_n^2} \text{ then } Z_b \sim \chi^2(M-1). \tag{6}$$

Under the null hypothesis it is also assumed that the noise statistics of the sensor outputs are zero mean and time invariant so the parameters of each distribution are the same:

$$\mu_1 = \mu_2 =,\ldots,= \mu_N = \mu = 0 \qquad (7)$$

and

$$\sigma_1^2 = \sigma_2^2 =,\ldots,= \sigma_N^2 = \sigma^2 . \qquad (8)$$

Using the reproductive property of χ^2 variables, the following aggregate test statistics can be formed and analysed:

$$\text{If } Z_c = \frac{M}{\sigma^2}\sum_{n=0}^{N-1}\hat{\mu}_n^2 \text{ then } Z_c \sim \chi^2\left(N\right); \qquad (9)$$

$$\text{If } Z_d = \frac{M}{\sigma^2}\sum_{n=0}^{N-1}\hat{\sigma}_n^2 \text{ then } Z_d \sim \chi^2\left(N\left(M-1\right)\right) . \qquad (10)$$

So far it has been assumed that the true variance (σ^2) of the (white) noise is known. This is an inconvenient and unnecessary assumption. It can be eliminated by dividing (9) by (10); furthermore, if the numerator and the denominator are scaled by the inverse of their respective degrees of freedom, i.e.

$$Z_M = \frac{Z_c/N}{Z_d/\left(N\left(M-1\right)\right)}, \qquad (11)$$

then a variable distributed according to Snedecor's F distribution results; that is, after substituting (9) and (10) into (11):

$$Z_M = \left(M-1\right)\frac{\sum_{n=0}^{N-1}\hat{\mu}_n^2}{\sum_{n=0}^{N-1}\hat{\sigma}_n^2}, \qquad (12)$$

with

$$Z_M \sim \mathrm{F}\left(N, N\left(M-1\right)\right). \qquad (13)$$

Substituting the expressions for $\hat{\mu}_n$ and $\hat{\sigma}_n^2$, given in (3) and (4) respectively, into (12) yields

$$Z_M = \left(M-1\right)\frac{\frac{1}{M}\sum_{n=0}^{N-1} y\left(n\right)^2}{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} x_m\left(n\right)^2 - \frac{1}{M}\sum_{n=0}^{N-1} y\left(n\right)^2}. \qquad (14)$$

3. IMPLEMENTATION

Note that the above expression (14), contains three sum-of-squares terms that are equivalent to power measurements computed in the time domain. Alternatively, (14) may be written in terms of moments:

$$Z_M = \left(M-1\right)\frac{\sum_{n=0}^{N-1} E\left[x\left(n\right)\right]^2}{\sum_{n=0}^{N-1} E\left[x\left(n\right)^2\right] - \sum_{n=0}^{N-1} E\left[x\left(n\right)\right]^2} \qquad (15)$$

using

$$E\left[x(n)\right]=\frac{1}{M}\sum_{m=0}^{M-1}x_m(n)\,, \tag{16}$$

$$E\left[x(n)^2\right]=\frac{1}{M}\sum_{m=0}^{M-1}x_m(n)^2 \tag{17}$$

and

$$y(n)=\sum_{m=0}^{M-1}x_m(n)\,. \tag{18}$$

The two sum-of-squares terms (16&17) are computed efficiently as moving averages, using sliding windows, implemented using circular buffers. In this way, the Z_M statistic, for each beam, is calculated using only the samples that enter and exit two sliding windows, each of length *N*; as a result, the computational effort does not depend on *N*.

4. INTERPRETATION

The Z_M test statistic is the ratio of two sum-of-squares quantities (12). If the square of the estimated mean (numerator) is regarded as the (delay-and-sum) signal power, and the variance (denominator) the noise power, then it may be convenient to convert Z_M into a Signal-to-Noise Ratio (SNR) in dB. Images may then be formed using many closely-spaced beams, and presented to an operator for visual inspection.

Knowing that under the null hypothesis the Z_M statistic is F distributed, allows a detection threshold to be computed to give the desired probability of falsely rejecting the null hypothesis when it is indeed true (a false alarm). If the computed Z_M value exceeds the threshold then a localisable signal is instead assumed to be present. The necessary threshold is determined using the inverse Cumulative Density Function (CDF) of the F distribution. The two parameters (degrees of freedom) of the function automatically adjust the threshold (increase it) to compensate for the higher variability of the test statistic when low channel counts (*M*) are used and when the data window length (*N*) is small.

In practice, the null hypothesis is rarely entirely true, and false alarms due to nuisance sources are common, so a larger detection threshold is usually appropriate, giving a negligible theoretical false-alarm probability (the size of the test), an acceptable practical false-alarm probability and a reasonable probability of detection (the power of the test).

This method is able to suppress peaks on background noise and produce a single sharp peak on localisable signals because the Z_M statistic for poorly aligned signals is low. The null hypothesis is rejected when the estimated means at each sample in the analysis window ($\hat{\mu}_n$) are far from zero and when the dispersion around each of the means is small (as indicated by small values of $\hat{\sigma}_n^2$). Applying the appropriate delays to steer the beam in the direction of the source results in large $\hat{\mu}_n$ values and small $\hat{\sigma}_n^2$ values because the waveforms in each delay channel are nearly identical (except for disturbances caused by additive white noise). In the presence of non-localisable signals (e.g. flow noise or other diffuse sound sources), or when the incorrect delays are applied to steer the beam in the wrong direction, the $\hat{\mu}_n$ values remain large but so are the $\hat{\sigma}_n^2$ values (due to waveform misalignment), which ensures that Z_M remains small.

The Z_M method is able to detect and localise weak signals in noise because a small variance term in the denominator of (14) yields a large Z_M test statistic, even when the power in the numerator is small.

5. APPLICATIONS

As this technique outputs large values when the delayed waveforms are large in magnitude and aligned with each other (in time or space), it may be used in any application where measures of similarity are required. One of the advantages of this technique relative to other measures of similarity (like cross-correlation) is that it can be applied to an arbitrary number of channels.

When the Z_M test statistic is applied to two data segments (M=2) of finite length, extracted from different channels by a detector it may be used to estimate TOADs, as an alternative to cross correlation. When applied to the problem of lag estimation, using two finite data segments of equal length, one data segment is shifted relative to the other and the summations are performed only over the overlapping portion (of length N) of the data segments. A lag-dependent threshold (as determined by the inverse of the F-distribution CDF) provides an alternative way of managing the problem of high variance at large lags, due to reduced overlap, without biasing results (towards zero) with a triangular window function [12]. The lag-dependent detection thresholds are computed once only on system initialization. In this role, the Z_M method may be used to correlate transmit and receive signals for range estimation in active sensor systems, or to correlate multiple pairs of received signals for AOA estimation in passive systems.

The method may also be used to compare an unknown transient waveform (in a continuous data stream or a finite data segment) with a set of pre-specified signal templates (each a finite data segment), for the purpose of joint detection and classification. In this case, the channel count is also equal to two (M=2). One of the channels contains the unknown waveform, the other contains the template used to represent a given class. The Z_M test statistic is computed for all possible classes. A detection is declared when any of the test statistics exceeds a specified threshold. The class, for which the test statistic is maximised, is assigned to the detection.

The above operations of correlation, detection and classification are also required when analysing imagery data. When the Z_M test statistic is applied to frames of imagery, the summations over time are replaced by summations over all (spatial and spectral) dimensions of the image space, within the region of image overlap. Two or more images may be aligned (correlated or registered) by finding the set of spatial transforms for which the test statistic is maximised.

6. ILLUSTRATION

Using real acoustic data, a single two-channel lag-estimation problem will be used to illustrate the operation, and demonstrate the performance, of the new measure of similarity described in this paper (Z_M), relative to cross-correlation (R_{12}) and normalised cross-correlation (ρ_{12}). A pair of ¼" pre-polarised free-field array-microphones (± 1 dB from 100 to 10000 Hz) were used, each simultaneously digitised at a rate of 50,000 samples per second per channel (F_s), with 24bit quantization, using a data acquisition device (DAQ) with in-built anti-aliasing filters. The microphones, separated by a distance of 0.4m (D), were placed in the corner of a 3m x 5m room, 1m from each wall. A speed of sound of 335m/s (v_c) was assumed. The microphones were exposed to a transient signal, generated using a mechanical clicker (approx. pulse frequency = 2000Hz-2500Hz, approx. pulse duration = 7ms). The emitter was placed 2.5m from the inter-sensor mid-point, at 90° to the inter-sensor axis (broadside, where l = 0). Wave-front curvature for this sensor-source geometry is not insignificant, it was however ignored in this study. The digitised sensor signals were pre-filtered using a DC notch filter (-2dB at 25Hz) implemented in software. Upon detection of the signal, data segments from both channels were extracted and processed. The segment from channel 1 was correlated with the segment from channel 2 at all physically feasible lags ($-L \le l \le +L$), where $L = 60$ (see Eq. 18). The segments were extracted so that the wave-front arrival point, as determined automatically by the detector, was at the centre of both segments. The segment from the first channel was of length N; the segment from the second channel was extended (padded) by L samples at either end, to ensure that the analysis window length (N) is constant for all lags (for uniform variance), giving a total segment length of N+2L. Two analysis window lengths were used to process the same detection ($N = 201$ and $N = 101$) and their results compared. The results are plotted in Figs. 1 to 4.

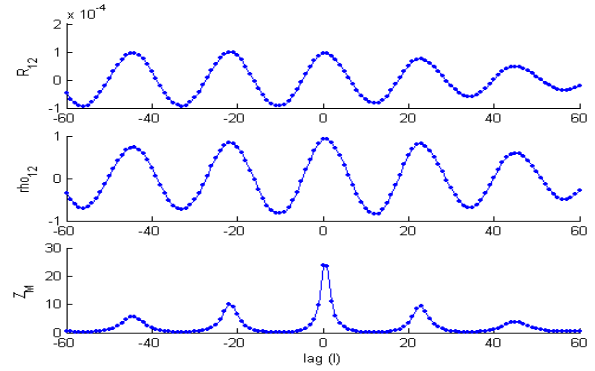

Fig. 1: Localisation functions (with *N*=201).

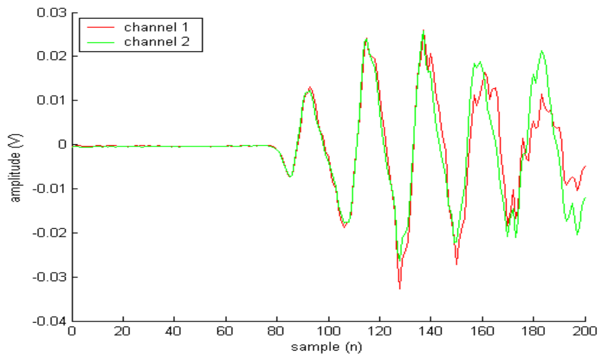

Fig. 2: Digitised waveforms over the analysis window, aligned using the most likely lag (as determined using Z_M with *N*=201).

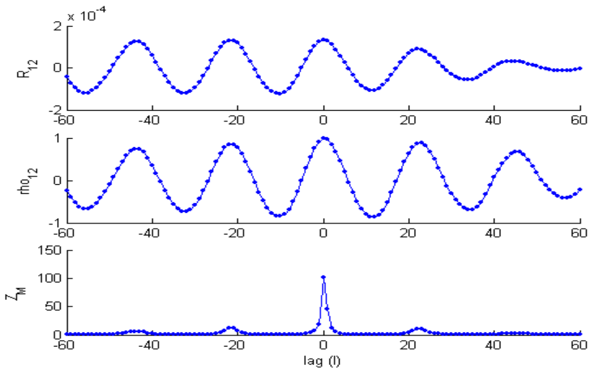

Fig. 3: Localisation functions (with *N*=101).

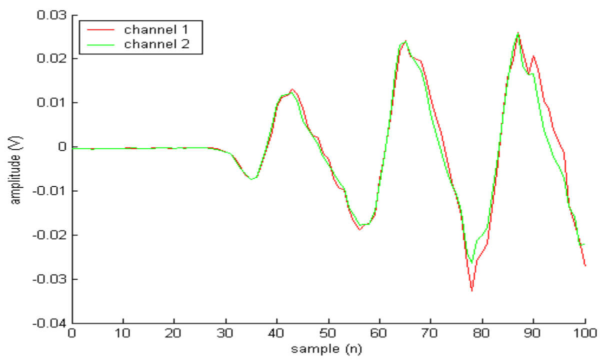

Fig. 4: Digitised waveforms over the analysis window, aligned using the most likely lag (as determined using Z_M with *N*=101).

When $N = 201$ (see Fig. 1), the maxima in the Z_M and ρ_{12} localisation functions are at the correct lag ($l = 0$). After quadratic interpolation around the Z_M maximum, a non-integer lag of 0.5 is obtained, yielding an AOA (θ) of 89.5°, using (for far-field sources):

$$\theta = \cos^{-1}\left(\frac{l}{L}\right) \text{ with } L = \mathrm{ceil}\left(\frac{F_s D}{v_c}\right). \qquad (18)$$

In the R_{12} localisation function, one of the peaks at an incorrect lag ($l = -22$) is slightly larger than the peak at the correct lag (see Fig. 1). Use of this incorrect (integer) lag yields an AOA of 111.5°.

Reducing the length of the analysis window from $N = 201$ to $N = 101$, excludes the newer data from the computations, where the waveforms begin to lose coherence (compare Fig. 2 and Fig. 4). All methods now yield maxima at the correct lags (see Fig. 3). The value of the Z_M localisation function at the true lag ($l = 0$) is now much greater than values at the adjacent lags ($l = \pm 1$), and much greater than maxima of adjacent peaks at the wrong lags ($l = \pm 22$). The interpolated lag value of 0.1 yields an improved AOA of 89.9°.

In general, the ambiguous peaks in the R_{12} localisation function tend to be similar in height; as a result, random signal perturbations may cause the height of a false peak to be greater than that of the true peak. In comparison to the ρ_{12} and R_{12} methods, the Z_M method yields true peaks that are tall and narrow, with flattened false peaks. Further investigations (not shown here) with different sound sources, inter-sensor separations, angles of arrival, and noise environments, have revealed that the ρ_{12} and the Z_M methods are much less likely to fail (i.e. to identify the incorrect lag) than the R_{12} method; however, they are not perfect and do fail in some cases. Even though the ρ_{12} false peaks are still tall, relative to the true peak, it was noted that the ρ_{12} and Z_M methods tend to succeed and fail in the same cases; therefore, the subtle suppression of the ρ_{12} false peaks is sufficient in most cases. The more obvious suppression of the Z_M false peaks is still a desirable characteristic when the localisation function is converted into an image because the image is sharper and visually less ambiguous.

7. DISCUSSION

According to Eq. (18), the theoretical angular resolution is increased by increasing the inter-sensor separation, by increasing the sampling rate, or both (assuming the speed of sound is constant). However, when the source frequency is above the array frequency, as is the case in the above example, the cyclic ambiguity gives rise to multiple peaks during TOAD estimation, which may result in misalignment errors, yielding very large AOA errors. By reducing the incidence of misalignment errors (spatial aliasing), the Z_M measure of similarity allows the use of large arrays, with greater angular resolution, to be used to localise transient sounds, over a greater range of frequencies, with a greater degree of accuracy.

In the preceding example, the Z_M measure of similarity resolves the cyclic ambiguity through the effective utilization of envelope information. If the detector is late in determining the onset of the signal, so that the rise of the envelope is not captured within the analysis window, preferably near its centre, then the height difference between the true peak and the ambiguous peaks, in the localisation function, is reduced. The performance of the detector and the length of the analysis window (relative to the signal wavelength, the duration of the modulating envelope, and the extent of signal coherence) are critical design parameters of all correlation methods, when processing transients using a detect-then-localise architecture.

The Z_M measure of similarity reduces the incidence of misalignment errors when processing data collected from a compact array. It is also able to reduce the incidence of misassociation errors when processing data collected from widely dispersed arrays, where the sensors may be separated by hundreds of meters. Association is used here to describe the process whereby all signals, due to the same source, measured using different sensors, are identified. The Z_M test statistic will be small, if a signal from sensor 1, due to source A, is correlated with a signal from sensor 2, due to source B, if the waveforms of A and B are dissimilar.

A disadvantage of the Z_M measure of similarity is that it requires the signal power levels in each channel to be approximately equal. Signal power levels are determined by the received Sound Pressure Level (SPL), the sensor/preamplifier sensitivity, and the signal conditioner gain. Even if the digitised sensor signals due to a common source have been correctly aligned (in time), a high variance (in amplitude) will result if the power of the signals in each channel are different. This effect is minimised by sensor calibration or pre-normalization of the data, in each channel, over the analysis window. Normalised cross correlation is unaffected by signal power level differences, however, like the Z_M method, it is sensitive to DC offsets. This shortfall motivated the development of the normalised covariance method.

The sensitivity of the Z_M method is an advantage in most cases; however, detections on weak but well correlated signals due to nuisance or low-interest sources (acoustic clutter) or internal noise (e.g. electromagnetic interference) may also be generated. Normalised cross correlation is similarly affected, while cross correlation is relatively unaffected. In practice, when a localise-then-detect architecture is adopted, and the Z_M statistic computed, a constant minimum variance term is added to the denominator in Eq. (14), to reduce the sensitivity and compensate for the finite measurement resolution of the DAQ.

One of the undesirable characteristics of cross correlation is that detection thresholds must be arbitrarily set. This is not a problem if each data segment is known to contain one localisable signal; in this case, the lag of the maximum in the cross correlation vector is selected, regardless of its level. Normalised (covariance and cross correlation) methods are more favourable in this respect because perfectly aligned waveforms have a coefficient of unity. On this basis, an appropriate detection threshold (0.90 or 0.99, for example) may be chosen. When the Z_M method is used, the detection threshold is chosen to give the desired theoretical false-alarm rate, when the null hypothesis is assumed to be true, using the inverse of the F-distribution CDF.

8. Conclusion

The method described in this paper has the potential to improve a wide range of signal processing algorithms where a measure of signal similarity is required. The method is particularly well-suited to the problem of source localisation, where it has been shown to improve angular accuracy, and reduce false-alarm rates. The method was used in this paper to correlate two channels of acoustic data containing transient signals; however, it may also be used to estimate TOADs in more than two channels.

Acknowledgement

This work was (in part) conducted under the sponsorship of a Capability and Technology Demonstrator (CTD) program (Land 148), provided by the Commonwealth of Australia (CoA) and administered by the Land Engineering Agency (LEA). Initial technical guidance was provided by the Defence Science and Technology Organization (DSTO), Maritime Operations Division.

References

[1] Patent Pending, Provisional Application No. 2006901479, "Signal Analysis Methods", Mar. 2006.

[2] B.D. Van Veen and K.M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Mag*., vol. 5, pp. 4-24, Apr. 1988.

[3] C.H. Knapp and G.C. Carter, "The Generalized Correlation Method for Estimation of Time Delay," *IEEE Trans. Acoust., Speech, and Sig. Proc.*, vol. 24, no. 4, pp. 320-327, Aug. 1976.

[4] J.C. Chen, Kung Yao and R.E. Hudson, "Source localization and beamforming", *Sig. Proc. Mag. IEEE*, vol. 19, no. 2, pp. 30–39, Mar. 2002.

[5] Qing Zhao, A. Swami, Lang Tong, "The interplay between signal processing and networking in sensor networks", *Sig. Proc. Mag. IEEE*, vol. 23, no. 4, pp. 84–93, Jul. 2006.

[6] P. W. Boettcher and G. A. Shaw, "A Distributed Time-Difference of Arrival Algorithm for Acoustic Bearing Estimation", in *Proc. 2001 Int. Conf. on Information Fusion*, vol. 1, Montreal, Aug. 2001.

[7] B.G. Ferguson, L.G. Criswick and K.W. Lo, "Locating far-field impulsive sound sources in air by triangulation", *J. Acoust. Soc. Am.*, vol. 111, no. 1, pt. 1, pp. 104-16, Jan 2002.

[8] B.G. Ferguson and K.W. Lo, "Passive ranging errors due to multipath distortion of deterministic transient signals with application to the localization of small arms fire", *J. Acoust. Soc. Am.*, vol. 111, no. 1, pp. 117-128, Jan. 2002.

[9] F. Viola and W.F. Walker, "A comparison of the performance of time-delay estimators in medical ultrasound", *Ultra., Ferro. Freq. Contr., IEEE Trans. on*, vol 50, no. 4, pp. 392–401, Apr. 2003.

[10] T. Pirinen, P. Pertila, A. Visa, "Toward intelligent sensors - reliability for time delay based direction of arrival estimates", in *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*, 2003, vol. 5, pp. 197-200.

[11] L.J. Bain and M. Engelhardt, *Introduction to Probability and Mathematical Statistics*, 2nd Ed., California, Duxbury Press, 1992.

[12] J.G. Proakis and D.G. Manolakis, *Digital Signal Processing. Principles, Algorithms and Applications*, 3rd. Ed. New Jersey, Prentice-Hall International, 1996.