

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
27 September 2007 (27.09.2007)

PCT

(10) International Publication Number  
**WO 2007/106950 A1**

- (51) International Patent Classification:  
*G06F 17/15* (2006.01)
- (21) International Application Number:  
PCT/AU2007/000364
- (22) International Filing Date: 23 March 2007 (23.03.2007)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
2006901479 23 March 2006 (23.03.2006) AU
- (71) Applicant (for all designated States except US): **RDA PTY. LIMITED** [AU/AU]; Unit 12/71 Constitution Avenue, Canberra, ACT 2612 (AU).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

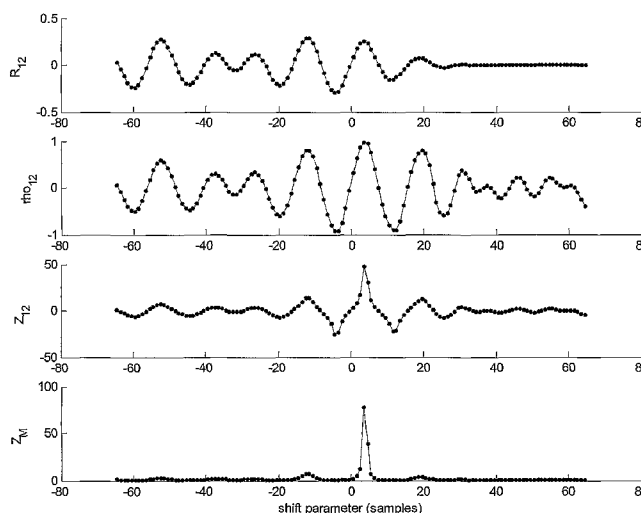
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

- (72) Inventor; and
- (75) Inventor/Applicant (for US only): **KENNEDY, Hugh** [AU/AU]; Unit 12/71 Constitution Avenue, Canberra, ACT 2612 (AU).
- (74) Agent: **MADDERNS**; Level 1, 64 Hindmarsh Square, Adelaide, S.A. 5000 (AU).

Published:  
— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SIGNAL ANALYSIS METHODS



(57) Abstract: The invention relates to a method of comparing two or more series of signal analysis data samples, to produce a measure of their similarity. Corresponding data samples of each series are compared, and measures of the magnitude of the signal and the variation between corresponding data samples are calculated. The magnitude measure may simply be the average power of corresponding data samples. The produced measure of the similarity of the series is a test statistic having Snedecor's F distribution. When compared to typical cross correlation methods, a test statistic of the present invention provides a greater distinction between series displaying a high degree of similarity and series with a low degree of similarity. Accordingly, the present invention has applicability to the detection, classification and angular localisation of signals from acoustic or electromagnetic sources.

WO 2007/106950 A1

## SIGNAL ANALYSIS METHODS

### FIELD OF THE INVENTION

The present invention relates to the field of digital signal processing. In particular, the present invention relates to assessing the similarity of two or more series of signal data samples, and applications of the present invention include (although are not necessarily limited to) the detection, classification and localisation of signals.

### 10 BACKGROUND OF THE INVENTION

Sensors can be used to detect signal emissions - for example, acoustic or electromagnetic waves (e.g. radar). However, in a noisy environment, such as may be expected on a battlefield, the established methods of detection via amplitude discrepancies in a relatively steady background of noise have been found to result in a relatively high False Alarm Rate (FAR) as any intruding noise of high enough level will trigger the detector.

Accordingly, Constant False Alarm Rate (CFAR) detectors have been developed which provide a specified (low) FAR. CFAR characteristics may be achieved by many different mechanisms. In addition to the current measurement they typically incorporate an estimate of the environmental state. The 'environmental' estimate may be used to set an adaptive threshold, or used to scale the current measurement which is then compared against a constant threshold.

25

Once a signal is detected, the received waveform is typically classified by cross correlation with a template waveform stored in a library. Each template waveform corresponds to a particular class of signal source - for example, the class of weapon fired.

30

Furthermore, if multiple sensors are used, the direction of the source can also be determined. This can be accomplished by a process of lag estimation – the time delay or lag between receiving the signal at each sensor is calculated and then used to estimate the angle of arrival of the signal. In the lag estimation process, sensors are typically partitioned into pairs, and lag estimates are calculated for each pair. Series of data samples taken from each sensor are cross correlated at different lags, and the lag with the highest correlation value is selected and used to determine the angle of arrival.

- 10 There are a number of known cross correlation techniques which can be used for classification or lag estimation purposes. The standard cross correlation of two signals  $x_1(n)$ ,  $x_2(n)$  of N samples, where  $x_2$  is shifted by  $l$  samples relative to  $x_1$ , can be written as:

$$r_{12} = \frac{1}{N} \sum_{n=0}^{N-1} x_1(n)x_2(n) \quad (1)$$

15

Standard cross correlation gives a positive figure for positive correlation, but a negative figure for negative correlation; however, the figures mean little in an absolute sense. Accordingly, normalised cross correlation is an improvement on standard cross correlation, giving a value that always lies between -1 and +1. +1 means 100% correlation, whereas -1 means 100% correlation in the opposing sense (e.g. signals in antiphase). The normalised cross correlation of two signals  $x_1(n)$ ,  $x_2(n)$  of N samples can be written as:

20

$$rho_{12} = \frac{r_{12}}{\frac{1}{N} \left[ \sum_{n=0}^{N-1} x_1^2(n)x_2^2(n) \right]^{1/2}} \quad (2)$$

- 25 In preference to standard or normalised cross correlation, the generalized correlation method is also used (see *The Generalized Correlation Method for Estimation for Time Delay*, Knapp and Carter, August 1976, IEE Transactions on

Acoustics, Speech, and Signal Processing, Vol. ASSP-24, No. 4, 320). However, each of these techniques requires a preceding detection step. Furthermore each technique can sometimes result in either misclassification of a signal or an incorrect lag estimation - and accordingly, significant angular errors.

5

### SUMMARY OF THE INVENTION

In a first aspect of the present invention, there is provided a method of assessing the similarity of two or more series of signal data samples comprising:

10 forming a test statistic which is inversely proportional to a variation measure, the variation measure indicating the variation between the series, whereby the value of the test statistic therefore provides a measure of the similarity of the series.

15 Of course, the formation of the test statistic may only be an intermediate step of the assessment process, and other calculations may be performed in addition to the formation of the test statistic when assessing the similarity of the series.

20 The test statistic may be proportional to a magnitude measure, the magnitude measure indicating the magnitude of the signal.

Preferably, the method of assessing the similarity of two or more series of signal data samples may further comprise:

25 calculating, for each data sample, a sample variation measure indicating the variation between corresponding data samples of each series, and calculating the variation measure from the sample variation measures.

In some embodiments, the variation measure is  $\sum_{n=0}^{N-1} \hat{\sigma}_n^2$ ,

where N is the number of samples in each series, and

where  $\hat{\sigma}_n$  is the observed standard deviation of the set of the *n*th samples of each series, at a specific lag.

5 Further, the test statistic may be proportional to  $\frac{\sum_{n=0}^{N-1} \hat{\mu}_n^2}{\sum_{n=0}^{N-1} \hat{\sigma}_n^2}$ ,

where  $\hat{\mu}_n$  is the observed mean of the *n*th samples of each series, at a specific lag.

Further, the test statistic may be equal to  $(M - 1) \frac{\sum_{n=0}^{N-1} \hat{\mu}_n^2}{\sum_{n=0}^{N-1} \hat{\sigma}_n^2}$ ,

where M is the number of series of data samples.

10

Like other methods, this test statistic can be used to assess the similarity of two series. However, unlike other methods, the test statistic described above may also be used to assess the similarity of three or more series of data samples.

15

Preferably, the test statistic has a known distribution, which for the test

statistic equal to  $(M - 1) \frac{\sum_{n=0}^{N-1} \hat{\mu}_n^2}{\sum_{n=0}^{N-1} \hat{\sigma}_n^2}$  will be Snedecor's F distribution.

20 In a second aspect of the present invention, there is provided a test statistic for assessing the similarity of two or more series of signal data samples, wherein

the test statistic is inversely proportional to a variation measure, the variation measure indicating the variation between the series.

One application of the method and test statistic of the present invention is in angle of arrival estimation, wherein series of signal data samples are compared at a variety of different lags. At different lags, a data sample in one series will correspond to different data samples in other series, depending on the lag. For three or more series of signal data samples, lag combinations may be selected to correspond to possible angles of arrival.

Accordingly, in a third aspect of the present invention, there is provided a method of aligning two or more series of signal data samples comprising:

calculating a test statistic at a first lag combination, wherein the test statistic is inversely proportional to a variation measure, the variation

measure indicating the variation between the series;

calculating the test statistic at one or more second lag combinations;

and

selecting the lag combination for which the test statistic is greatest, thereby aligning the series.

As will be understood, where there are only two series to be aligned, each lag combination will consist of only one lag.

Of course, there are other applications for the present invention. Classification of a signal of interest is one other obvious example of where it may be desirable to compare two series of signal data samples (for instance, comparing a signal of interest to a template representing a particular class of signal). However, another problem to which the present invention may be

applied is the detection of a signal of interest, which conventionally does not assess the similarity between different series of signal data samples.

Accordingly, in a fourth aspect of the present invention, there is provided a method of detecting a signal of interest (SOI) by:

obtaining a first series of signal data samples from a first sensor;

obtaining a second series of signal data samples from a second sensor;

calculating, from the first and second series, a test statistic which

provides a measure of the similarity between the first and second series at a

first lag of zero or more data samples;

calculating the test statistic at one or more second lags of zero or more data samples; and

declaring a detection if the largest calculated test statistic is more extreme than a threshold value.

15

Where the test statistic has a known distribution, this means that the threshold value can be set to provide a desired CFAR.

In further aspects, the present invention provides computer readable media encoded with data representing computer programs, that can be used to direct a programmable device to perform the above methods.

20

In yet further aspects of the present invention, computer program elements are provided comprising computer program code means to make a programmable device execute the above methods.

25

In yet further aspects of the present invention, there are provided computer processing devices configured to perform the above methods.

A detailed description of one or more embodiments of the invention is provided below, along with accompanying figures that illustrate by way of example the principles of the invention. While the invention is described in connection with such embodiments, it should be understood that the invention is not limited to any embodiment. On the contrary, the scope of the invention encompasses numerous alternatives, modifications and equivalents. For the purpose of example, specific details are set forth in the following description in order to provide a thorough understanding of the present invention.

10

For the purpose of clarity, technical material that is known in the technical fields related to the invention has not been described in detail so that the present invention is not unnecessarily obscured.

## 15 BRIEF DESCRIPTION OF THE DRAWINGS

An illustrative embodiment of the present invention will be discussed with reference to the accompanying drawings wherein:

20

FIGURE 1 depicts an incoming signal to be received by a linear array of sensors;

FIGURE 2 is a flow diagram of one embodiment of the present invention;

25

FIGURE 3 is a plot of two series of signal data samples for a signal of interest, aligned at the most likely lag;

FIGURE 3A is a plot of test statistics calculated for the data in Figure 2, plotted as a function of lag for four candidate lag-estimation methods;

FIGURES 4, 5, 6, 7, 8 and 9 are each a plot of two series of signal data samples for different signals of interest, aligned at the most likely lag; and

FIGURES 4A, 5A, 6A, 7A, 8A and 9A are each a plot of test statistics

5 calculated for the data in Figures 4, 5, 6, 7, 8 and 9 respectively, plotted as a function of lag for four candidate lag-estimation methods.

### DETAILED DESCRIPTION

10 This embodiment of the present invention is best described in relation to the problem of angular localisation, although the present invention does have broader application to other circumstances where it is desirable to assess the similarity of two or more series of signal data samples.

15 With an array of sensors, a signal of interest (SOI) arriving at an angle as shown in Figure 1 will be received by each sensor 12 at a different time. In Figure 1, the right-most sensor receives the SOI first, the middle sensor receives the SOI after a time delay, and the left-most sensor only receives the SOI after another time delay. The angle of arrival (AOA) of the SOI can be determined by the duration of these time delays. For a three dimensional  
20 array (at least four sensors arranged in a non-coplanar configuration) the determined AOA can be uniquely determined in three dimensional space.

As depicted in Figure 2, series 22 of signal data samples can be obtained 20 from each sensor 12, and may be measured in volts. To obtain 20 the series 22, 25 a continuous data segment can be extracted from the digitised data samples taken by each sensor 12, for a given analysis window. The data within this analysis window will provide a series 22 of signal data samples, which can be used in the detection, classification or localisation of a signal of interest.

According to the present invention, a statistical treatment of the series of data samples is performed with the null hypothesis being that there is no SOI and that the series contain only zero-mean Gaussian distributed noise. It is herein assumed that any DC offset in the data (e.g. sensor or DAQ bias) or  
 5 frequencies that are of no interest (e.g. wind or self noise) have been removed by a pre-whitening stage. The test statistic for all possible lag combinations corresponding to all physically measurable angles is calculated 30, and the largest calculated test statistic 32 is identified 40. If the largest calculated test statistic is greater than a threshold value 42 (preferably set to give a specified  
 10 CFAR) then the null hypothesis is rejected, and a detection is declared 50. Furthermore, in this way the series 22 can be aligned, and the most likely direction of the source can be set equal to the angular coordinate for which the null hypothesis is least likely, i.e. the angle of arrival is estimated 60 by selecting the angle which corresponds to the lag combination for which the  
 15 test statistic is maximised.

The test statistic is a measure of similarity, applied to two or more series of signal data samples. In this embodiment, the signals are delayed acoustic sensor signals.

20

The delay-and-sum (beam-formed) total of the two or more series is applied as:

$$y(n) = \sum_{m=0}^{M-1} x_m(n) \quad (3)$$

where  $x_m(n)$  is the  $n$ th sample of the  $m$ th sensor at a specified lag of data  
 25 samples, and  $y(n)$  is the delay-and-sum total. It is assumed that the data in the channel  $x_m$  has already been delayed to steer the beam in the desired direction.

The noise statistics of each sensor are assumed to be the same, so the  $n$ th sample in each delay channel is assumed to be an independent observation of the random variable  $X_n$ , giving a total of  $N$  different random variables, with  $M$  observations for each variable. Under the null hypothesis the variables

5 have a Gaussian (or normal) distribution:

$$X_n \sim N(\mu_n, \sigma_n^2) \quad (4)$$

The Maximum Likelihood Estimates (MLE)  $\hat{\mu}_n$  and  $\hat{\sigma}_n^2$ , of the mean and variance are computed using:

$$\hat{\mu}_n = \frac{y(n)}{M} \quad (5)$$

10 and

$$\hat{\sigma}_n^2 = \frac{1}{M} \left\{ \sum_{m=0}^{M-1} x_m(n)^2 - \frac{y(n)^2}{M} \right\} \quad (6)$$

Under the null hypothesis, the following relationships hold:

a. If  $Z_a = M \frac{(\hat{\mu}_n - \mu_n)^2}{\sigma_n^2}$  then  $Z_a \sim \chi^2(1)$  (7)

15

b. If  $Z_b = M \frac{\hat{\sigma}_n^2}{\sigma_n^2}$  then  $Z_b \sim \chi^2(M-1)$  (8)

That is, the quantities  $Z_a$  and  $Z_b$  are distributed as  $\chi^2$  (chi-squared) variables with 1 and  $M-1$  degrees of freedom respectively.

20 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:

$$\mu_1 = \mu_2 = \dots = \mu_N = \mu = 0 \quad (9)$$

and

$$\sigma_1^2 = \sigma_2^2 = \dots = \sigma_N^2 = \sigma^2 \tag{10}$$

Using the fact that the sum of  $\chi^2$  variables is a  $\chi^2$  variable, the following aggregate test statistics can be formed and analysed:

5

c. If  $Z_c = M \frac{\sum_{n=0}^{N-1} \hat{\mu}_n^2}{\sigma^2}$  then  $Z_c \sim \chi^2(N)$  (11)

d. If  $Z_d = M \frac{\sum_{n=0}^{N-1} \hat{\sigma}_n^2}{\sigma^2}$  then  $Z_d \sim \chi^2(N(M-1))$  (12)

10 So far it has been assumed that the true variance ( $\sigma^2$ ) of the white noise is known. This is an inconvenient and unnecessary assumption. It can be eliminated by dividing (11) by (12). Furthermore, if the numerator and the denominator are scaled by the inverse of their respective degrees of freedom then a variable distributed according to Snedecors' F distribution results:

15 
$$Z_M = \frac{Z_c / N}{Z_d / (N(M-1))} \tag{13}$$

$$Z_M = (M-1) \frac{\sum_{n=0}^{N-1} \hat{\mu}_n^2}{\sum_{n=0}^{N-1} \hat{\sigma}_n^2} \tag{14}$$

$$Z_M \sim F(n, n(m-1)) \tag{15}$$

The  $Z_M$  value obtained can be used as a test statistic according to this embodiment of the present invention. The term  $\sum_{n=0}^N \hat{\sigma}_n^2$  accordingly provides a variation measure which indicates the variation between the series. Substituting the expressions for  $\hat{\mu}_n$  and  $\hat{\sigma}_n$  into Eq. (14) yields:

$$5 \quad Z_M = (M-1) \frac{\frac{1}{M} \sum_{n=0}^{N-1} y(n)^2}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} x_m(n)^2 - \frac{1}{M} \sum_{n=0}^{N-1} y(n)^2} \quad (16)$$

Note that Eq. (16) appears to contain sum-of-squares terms that are equivalent to power measurements computed in the time domain. Alternatively, Eq. (16) may be written in terms of moments:

$$Z_M = (M-1) \frac{\sum_{n=0}^{N-1} E[x(n)]^2}{\sum_{n=0}^{N-1} E[x(n)^2] - \sum_{n=0}^{N-1} E[x(n)]^2} \quad (17)$$

10 using

$$E[x(n)] = \frac{1}{M} \sum_{m=0}^{M-1} x_m(n) \quad (18)$$

$$E[x(n)^2] = \frac{1}{M} \sum_{m=0}^{M-1} x_m(n)^2 \quad (19)$$

and

$$y(n) = \sum_{m=0}^{M-1} x_m(n) \quad (20)$$

15

When determining the AOA of a SOI, the  $Z_M$  test statistic can be computed efficiently using sliding windows, implemented in software as circular buffers, to estimate the moments using the delay line inputs  $x_m(n)$ .

Furthermore, the final test statistic is the ratio of two squared quantities. If the square of the estimated mean is regarded as the signal power, and the variance the noise power, then  $Z_M$  may be converted into a signal to noise ratio (SNR) in dB using:

5

$$\text{SNR} = 10 \log_{10} (Z_M) \quad (21)$$

Knowing that under the null hypothesis the  $Z_M$  test statistic is F distributed, allows a detection threshold 42 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.

In practice the null hypothesis is rarely entirely true, and false alarms due to nuisance detections are common. Accordingly, a larger detection threshold 42 is usually appropriate, giving a negligible theoretical false alarm probability and an acceptable practical false alarm probability and probability of detection.

20 The null hypothesis is rejected when the estimated means at each point ( $\hat{\mu}_n$ ) are a long way from zero and when the dispersion around the mean is small (as indicated by a small  $\hat{\sigma}_n$ ). When the null hypothesis is indeed false, applying the appropriate lags results in large  $\hat{\mu}_n$  values and small  $\hat{\sigma}_n$  s values because the waveforms in each delay channel are nearly identical (except for  
25 disturbances caused by additive white noise).

In the presence of non-localisable signals (e.g. wind noise), or when the incorrect delays are applied, the  $\hat{\mu}_n$  values are large but so are the  $\hat{\sigma}_n$  values,

which ensures that  $Z_M$  remains small. It is this characteristic that makes this method an effective detector and localiser, with prominent peaks marking the directions to true sources set against a relatively flat and featureless background.

5

The  $Z_M$  test statistic may be used, for example, for beam formation from multiple sensors. The test statistic may be formed for all angles, and the most likely lag used to declare a detection of a SOI and to calculate its AOA. It may be used on pairs of series of signal data samples obtained from respective sensor pairs, or alternatively it could be used on multiple sensors at once.

10

With most planar or three dimensional sensor arrays, and with the limits of current processing technology, it is too computationally expensive, using software implementations, to continuously form beams in bearing and elevation, to cover the surveillance surface (hundreds of beams must be formed to give the desired angular resolution). Accordingly, a predetection stage may be performed such that the  $Z_M$  test statistics are only formed when processing data likely to contain a signal of interest. The predetector may, for example, look for a large and/or sudden increase in received signal power.

15

The predetector could then select data segments of potential interest for further processing by a method according to the present invention.

20

An alternative embodiment encompassed within the scope of the present invention will now be described, wherein a test statistic (in this case  $Z_{12}$ ) is again inversely proportional to a measure of the variation between the series.  $Z_{12}$  is a measure of similarity, between (exactly) two signals. This method will be described with reference to classification of signals, although it can also be applied to the problem of angular localisation by lag estimation. Similarly,

25

methods utilising the  $Z_M$  test statistic may be applied to the problem of classification.

For the  $Z_{12}$  test statistic, it is assumed that the features being classified are  
 5 impulsive or transient, i.e. they have a finite duration, typically with a rapid onset and a more gradual decay. A template is used to represent a given class, for example the type of weapon fired. Separate templates may, for example, be used for the 'crack' (the supersonic shockwave emitted by the passing bullet) and 'thump' (the muzzle blast) of a particular weapon.

10

A sensor is used to obtain a series of signal data samples of a waveform for a signal of interest, which is then compared against the template. The comparison is done using a least-squares fit. Large fitting coefficients suggest that the received waveform may belong to the class represented by the  
 15 template. Optionally, the significance of the fitting coefficient is assessed using the residual of the fit. The comparison may be done by identifying the start of the feature in the waveform being classified, or if this cannot be done reliably then the fit may be implemented for all possible displacements using a matched filter.

20

This process may be formalised as follows. Let the template be the series  $x_1(n)$ , and the waveform containing the feature to be classified  $x_2(n)$ . It is assumed that the onset of the signature in the template and the waveform is at  $n = 0$  and that both  $x_1(n)$  and  $x_2(n)$  are of equal length. The null hypothesis is that  
 25 there is no correlation between the two series. The signature in the waveform has been corrupted by additive white noise. The model is:

$$x_2(n) = \beta_{12}x_1(n) + \varepsilon \quad (22)$$

We are assuming that the template is a perfect representation of the class, i.e. that it does not contain noise. This representation may be constructed from a theoretical model, from filtered or enhanced measurements, or simply set equal to raw measurements, which are assumed to be noise free. The most appropriate method depends on how well the signature can be measured and how well the feature can be modelled. It is also assumed here that the onset point of the signature in  $x_1(n)$ , can be identified, as misalignment will cause errors. When misalignment is unavoidable, the calculation described below may be repeated over a range of displacements, with the value of the best displacement used for decision-making purposes.

The MLE estimate of  $\beta_{12}$  is computed using:

$$\hat{\beta}_{12} = \frac{\sum_{n=0}^{N-1} x_1(n) x_2(n)}{\sum_{n=0}^{N-1} x_1(n) x_1(n)} \quad (23)$$

or

$$\hat{\beta}_{12} = (\mathbf{x}' \cdot \mathbf{x})^{-1} \mathbf{x}' \cdot \mathbf{y} \quad (24)$$

where  $\mathbf{x} = x_1(n)$  and  $\mathbf{y} = x_2(n)$

If the template has been pre-normalised before it is used, so that  $\mathbf{x}' \cdot \mathbf{x} = 1$ , then:

$$\hat{\beta}_{12} = \sum_{n=0}^{N-1} x_1(n) x_2(n) \quad (25)$$

or

$$\hat{\beta}_{12} = \mathbf{x}' \cdot \mathbf{y} \quad (26)$$

and

$$\hat{\sigma}_{12} = \sqrt{\frac{1}{N-1} (\mathbf{y} - \mathbf{x} \hat{\beta}_{12})' (\mathbf{y} - \mathbf{x} \hat{\beta}_{12})} \quad (27)$$

The following test statistic can be formed:

$$Z_{12} = \frac{\hat{\beta}_{12} - \beta_{12}}{\sqrt{\tilde{\sigma}_{12}^2}} \quad (28)$$

with

$$5 \quad Z_{12} \sim t(N-1) \quad (29)$$

That is,  $Z_{12}$  is distributed as a Student's t variable. It is also inversely proportional to the variation measure  $\sqrt{\tilde{\sigma}_{12}^2}$  which provides a measure of the variation between the series.

10

Classification can therefore be conducted by calculating the  $Z_{12}$  test statistic for each template, and selecting the template with the best fit.

15 Alternatively,  $Z_{12}$  can be used for detection/localisation purposes wherein the null hypothesis is that the data contains only noise. A confidence level or false alarm rate can be set according to the magnitude of the  $Z_{12}$  test statistic.

20 When processing impulsive transients due to gunshots, that the duration (the time interval containing approximately 90% of the signals energy) of a 'crack' is approximately 0.001 seconds, or 20 samples at 20,000 samples per second. The duration of a 'thump' is approximately 0.01 seconds, or 200 samples. Ideally, the entire feature should be used for lag estimation and for classification. However, it has been found that coherence decreases, and variability increases, as the time after the initial wavefront onset time  
 25 increases, primarily due to multi-path propagation and reverberation. It has also been found that it is useful to include sections of 'quiet' data in the series analysed, both before and after the feature of interest. The series analysed

may therefore include an additional 25%-50% of the duration of the feature in quiet data before the feature, and an additional 50%-75% after the feature.

### EXAMPLES

5 By way of example, with respect to lag estimation and angular localisation, Figure 2 depicts a plot of two series of acoustic signal data samples for a signal of interest, obtained from two respective acoustic sensors. The data have been aligned as for the most likely lag – that is, the lag for which the test statistics according to the present invention are maximised. Figures 3, 4, 5, 6,  
10 7, and 8 simply represent data series obtained for different signals of interest.

Figure 2A is a plot of test statistics computed using four candidate lag estimation methods plotted as a function of the lag (or time-shift parameter), for the series shown in Figure 2. Figures 3A, 4A, 5A, 6A, 7A and 8A  
15 correspond to the series shown in the other respective figures. The four candidate methods are standard cross correlation ( $r_{12}/R_{12}$ ), normalised cross correlation ( $\rho_{12}$ ),  $Z_M$  test statistics according to the present invention, and  $Z_{12}$  test statistics according to the present invention. As can be seen from the figures, the test statistics according to the present invention provide much  
20 sharper and more prominent peak at the most likely lag, with the  $Z_M$  test statistic providing the sharpest peaks of all.

In Figures 2A, 3A and 4A, this makes little quantitative difference to the final selection of the best lag estimate. However, for Figures 5A, 6A, 7A and 8A, the  
25 highest peak for the  $r_{12}$  method is at the incorrect lag. This results in large angular errors being calculated for the incoming signal of interest.

In addition to improved accuracy in the estimation of the lag between series, the  $Z_M$  and  $Z_{12}$  variables have a known distribution, which allows a

confidence level to be placed on the lag estimate. Poor lag estimates, for which the calculated test statistics are beneath a threshold value, can therefore be considered unreliable, and ignored. This improves accuracy by eliminating bad localisation results (or allowing for unknown classes).

5

Whilst both methods according to the present invention give a greater 'dynamic range' than normalised cross correlation (giving greater difference between good and bad lag estimates), the  $Z_M$  method requires far less computational power than the  $Z_{12}$  method. It can be calculated efficiently using sliding windows and, relative to normalised cross correlation, the  $Z_M$  method requires only a few more floating point operations per sample pair. On the other hand, the  $Z_{12}$  method requires two loops through the data vector – the first to determine the MLE fitting coefficient, and the second to determine the variation measure.

10

Yet another advantage of the  $Z_M$  method is that it can be extended to give the similarity of any number of series of signal data samples, whilst the other methods are only useful for two series at a time.

15

Although an embodiment of the present invention has been described in the foregoing detailed description, it will be understood that the invention is not limited to the embodiment disclosed, but is capable of numerous rearrangements, modifications and substitutions without departing from the scope of the invention. Modifications and variations such as would be apparent to a skilled addressee are deemed within the scope of the present invention. For example, although particular reference has been made to acoustic signal analysis, the present invention has more broad application in the analysis of other types of signal, e.g. radar. Furthermore, although the  $Z_M$  and  $Z_{12}$  methods were described in relation to localisation and classification

20

25

respectively, both could be used for either application or for any situation where it is desirable to assess the similarity of series of signal data samples – i.e. the present invention may replace operations that require cross correlation, matched filtering or template matching giving greater expected performance (higher probability of detection and lower false alarm rate). Such operations are typically performed in radar and (active) sonar range processing, feature detection in images, classification of signals and passive direction finding of (intercepted) communication signals.

10 It should be also appreciated that the present invention can be implemented in numerous ways, including as a process, an apparatus, a system, or a computer readable medium such as a computer readable storage medium or a computer network wherein program instructions are sent over optical or electronic communication links. Furthermore, the methods of the present  
15 invention could be implemented in hardware (as a computer processing device), on dedicated co-processing chips, possibly using Field Programmable Gate Arrays (FPGAs). This may enable real-time beam formation with 3-dimensional sensor arrays. It should be noted that the order of the steps of disclosed processes may be altered within the scope of the invention.

20

Throughout this specification and the claims that follow unless the context requires otherwise, the words 'comprise' and 'include' and variations such as 'comprising' and 'including' will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer  
25 or group of integers.

The reference to any prior art in this specification is not, and should not be taken as, an acknowledgment or any form of suggestion that such prior art forms part of the common general knowledge.

**THE CLAIMS:**

1. A method of assessing the similarity of two or more series of signal data samples comprising:

forming a test statistic which is inversely proportional to a variation  
 5 measure, the variation measure indicating the variation between the series,  
 whereby the value of the test statistic therefore provides a measure of  
 the similarity of the series.

2. A method according to claim 1, wherein the test statistic is proportional  
 10 to a magnitude measure, the magnitude measure indicating the magnitude of  
 the signal.

3. A method according to claim 1 or 2, further comprising:  
 calculating, for each data sample, a sample variation measure  
 15 indicating the variation between corresponding data samples of each series,  
 and  
 calculating the variation measure from the sample variation measures.

4. A method according to any one of the preceding claims, wherein the  
 20 variation measure is  $\sum_{n=0}^{N-1} \hat{\sigma}_n^2$ ,  
 where N is the number of samples in each series, and  
 where  $\hat{\sigma}_n$  is the observed standard deviation of the set of the  $n$ th  
 samples of each series, at a specific lag.

25 5. A method according to claim 4, wherein the test statistic is proportional

to  $\frac{\sum_{n=0}^{N-1} \hat{\mu}_n^2}{\sum_{n=0}^{N-1} \hat{\sigma}_n^2}$ ,

where  $\hat{\mu}_n$  is the observed mean of the  $n$ th samples of each series, at a specific lag.

6. A method according to claim 5, wherein the test statistic is equal to

$$5 \quad (M-1) \frac{\sum_{n=0}^{N-1} \hat{\mu}_n^2}{\sum_{n=0}^{N-1} \hat{\sigma}_n^2},$$

where M is the number of series of data samples.

7. A method according to any one of the preceding claims, wherein the number of series of data samples is more than two.

10

8. A method according to any one of the preceding claims, wherein the test statistic has a known distribution.

9. A method according to claim 8, wherein the known distribution is

15 Snedecor's F distribution.

10. A method according to claim 1, further comprising:

obtaining at least one of the series of signal data samples from a sensor.

20 11. A test statistic for assessing the similarity of two or more series of signal data samples, wherein the test statistic is inversely proportional to a variation measure, the variation measure indicating the variation between the series.

25 12. A test statistic according to claim 1, wherein the test statistic is proportional to a magnitude measure, the magnitude measure indicating the magnitude of the signal.

13. A test statistic according to claim 11 or 12, wherein the variation

measure is  $\sum_{n=0}^{N-1} \hat{\sigma}_n^2$ ,

where N is the number of samples in each series, and

5 where  $\hat{\sigma}_n$  is the observed standard deviation of the set of the  $n$ th samples of each series, at a specific lag.

14. A test statistic according to claim 13, wherein the test statistic is

proportional to  $\frac{\sum_{n=0}^{N-1} \hat{\mu}_n^2}{\sum_{n=0}^{N-1} \hat{\sigma}_n^2}$ ,

10 where  $\hat{\mu}_n$  is the observed mean of the  $n$ th samples of each series, at a specific lag.

15. A method according to claim 14, wherein the test statistic is equal to

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

15 where M is the number of series of data samples.

16. A method of aligning two or more series of signal data samples comprising:

calculating a test statistic at a first lag combination, wherein the test  
 20 statistic is inversely proportional to a variation measure, the variation measure indicating the variation between the series;

calculating the test statistic at one or more second lag combinations;

and

selecting the lag combination for which the test statistic is greatest,  
thereby aligning the series.

17. A method according to claim 16, wherein there are three or more series  
5 of signal data samples.

18. A method according to claim 17, wherein each series is obtained from a  
different sensor, and the method is used for determining the angle of arrival  
of a signal relative to the three or more sensors, and wherein each lag  
10 combination corresponds to a possible angle of arrival of the signal.

19. A method according to any one of claims 16 to 18, wherein the test  
statistic is proportional to a magnitude measure, the magnitude measure  
indicating the magnitude of the signal.  
15

20. A method according to any one of claims 16 to 19, wherein the  
calculation of the test statistic comprises:

calculating, for each data sample, a sample variation measure  
indicating the variation between corresponding data samples of each series,  
20 and

forming the variation measure from the sample variation measures.

21. A method according to any one of claims 16 to 20, wherein the

variation measure is  $\sum_{n=0}^{N-1} \hat{\sigma}_n^2$ ,

25 where N is the number of samples in each series, and

where  $\hat{\sigma}_n$  is the observed standard deviation of the set of the  $n$ th  
samples of each series, at a specific lag.

22. A method according to claim 21, wherein the test statistic is

proportional to 
$$\frac{\sum_{n=0}^{N-1} \hat{\mu}_n^2}{\sum_{n=0}^{N-1} \hat{\sigma}_n^2},$$

where  $\hat{\mu}_n$  is the observed mean of the  $n$ th samples of each series, at a specific lag.

5

23. A method according to claim 22, wherein the test statistic is equal to

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

where M is the number of series of data samples.

10 24. A method of detecting a signal of interest by:

obtaining a first series of signal data samples from a first sensor;

obtaining a second series of signal data samples from a second sensor;

calculating, from the first and second series, a test statistic which

provides a measure of the similarity between the first and second series at a

15 first lag of zero or more data samples;

calculating the test statistic at one or more further lags of zero or more data samples; and

declaring a detection if the largest calculated test statistic is more extreme than a threshold value.

20

25. A method according to claim 24, wherein the test statistic is inversely proportional to a variation measure, the variation measure indicating the variation between the series.

26. A method according to claim 24 or 25, wherein the test statistic is proportional to a magnitude measure, the magnitude measure indicating the magnitude of the signal.

5 27. A method according to any one of claims 24 to 26, wherein the

variation measure is  $\sum_{n=0}^{N-1} \hat{\sigma}_n^2$ ,

where N is the number of samples in each series, and

where  $\hat{\sigma}_n$  is the observed standard deviation of the set of the *n*th samples of each series, at a specific lag.

10

28. A method according to claim 27, wherein the test statistic is

proportional to  $\frac{\sum_{n=0}^{N-1} \hat{\mu}_n^2}{\sum_{n=0}^{N-1} \hat{\sigma}_n^2}$ ,

where  $\hat{\mu}_n$  is the observed mean of the *n*th samples of each series, at a specific lag.

15

29. A method according to claim 28, wherein the test statistic is equal to

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

where M is the number of series of data samples.

20 30. A method according to any one of claims 24 to 29, further comprising:  
obtaining at least one further series of signal data samples from at least one further sensor,

wherein the test statistic provides a measure of the similarity between the first, second and further series, and the test statistic is calculated at a first

combination of lags of zero or more data samples, and at one or more further combinations of lags of zero or more data samples, and

and wherein each combination of lags relates to an angle of arrival for the signal relative to the first, second and further sensors.

5

31. A method according to claim 30, wherein the angle of arrival of the signal of interest is determined by identifying the combination of lags at which the test statistic is highest.

10 32. A computer readable medium encoded with data representing computer programs, that can be used to direct a programmable device to perform a method as claimed in any one of claims 1 to 10, or 16 to 31.

15 33. A computer processing device, configured to perform the method of any one of claims 1 to 10, or 16 to 31.

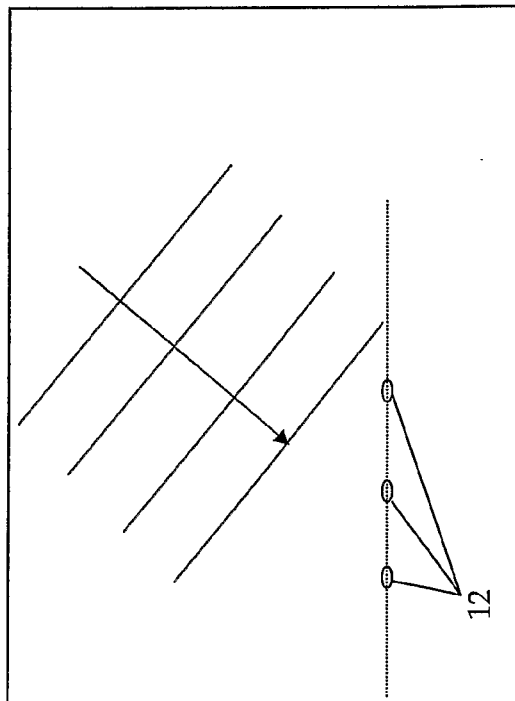


Figure 1

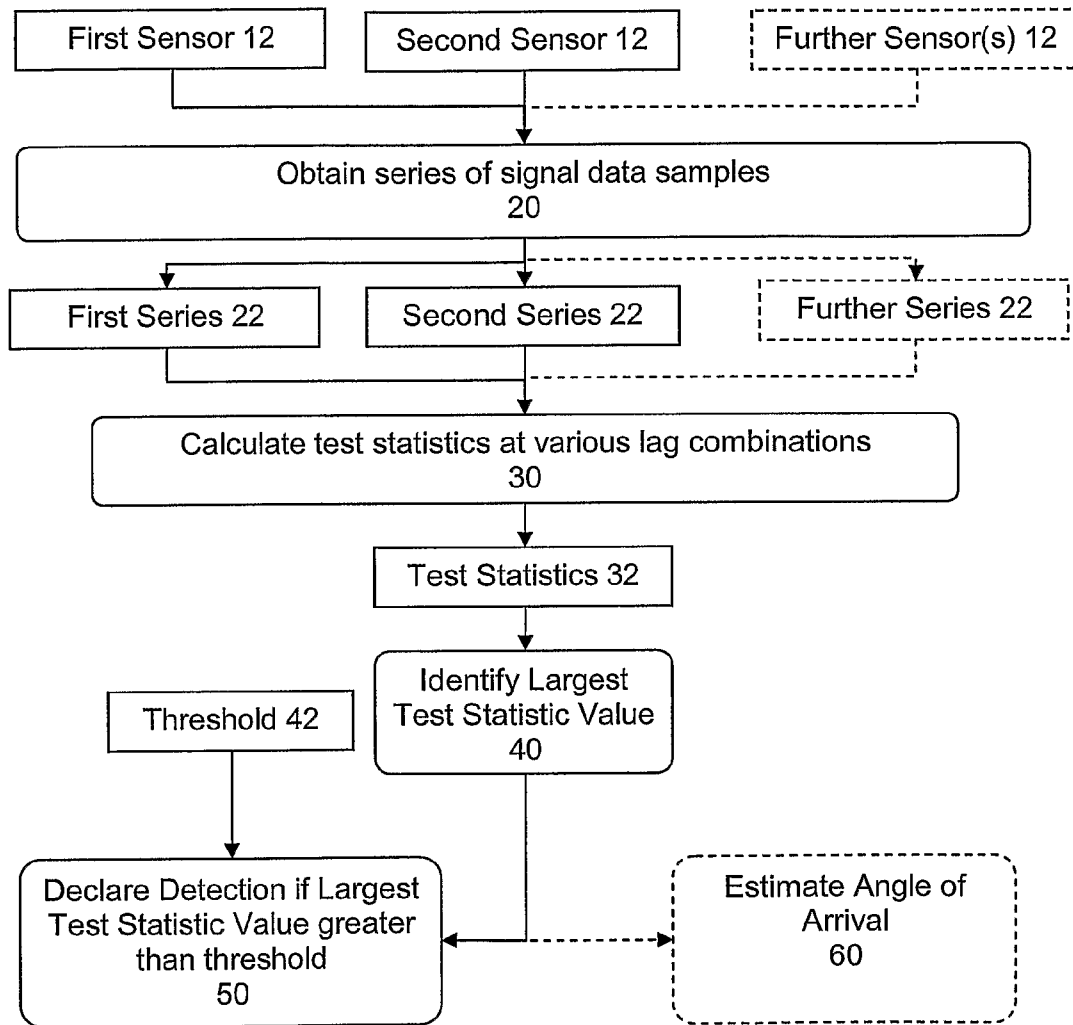


Figure 2

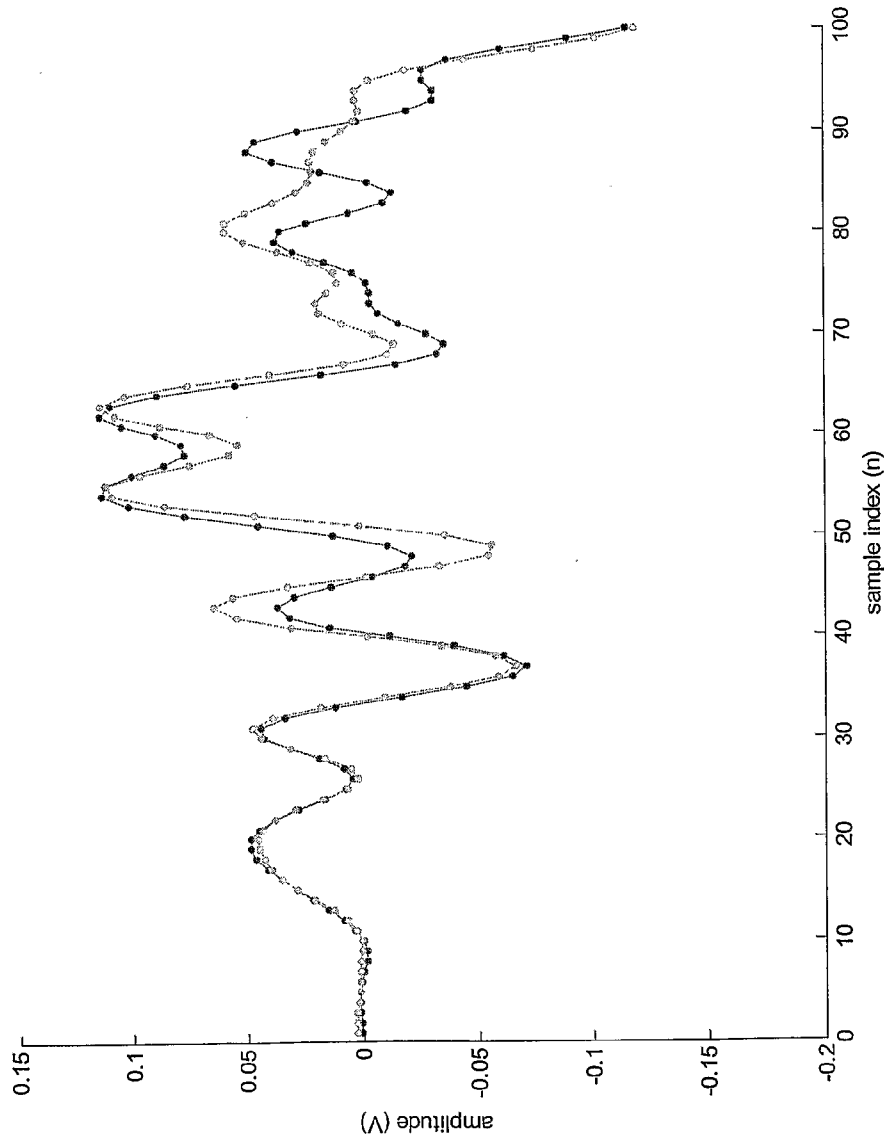


Figure 3

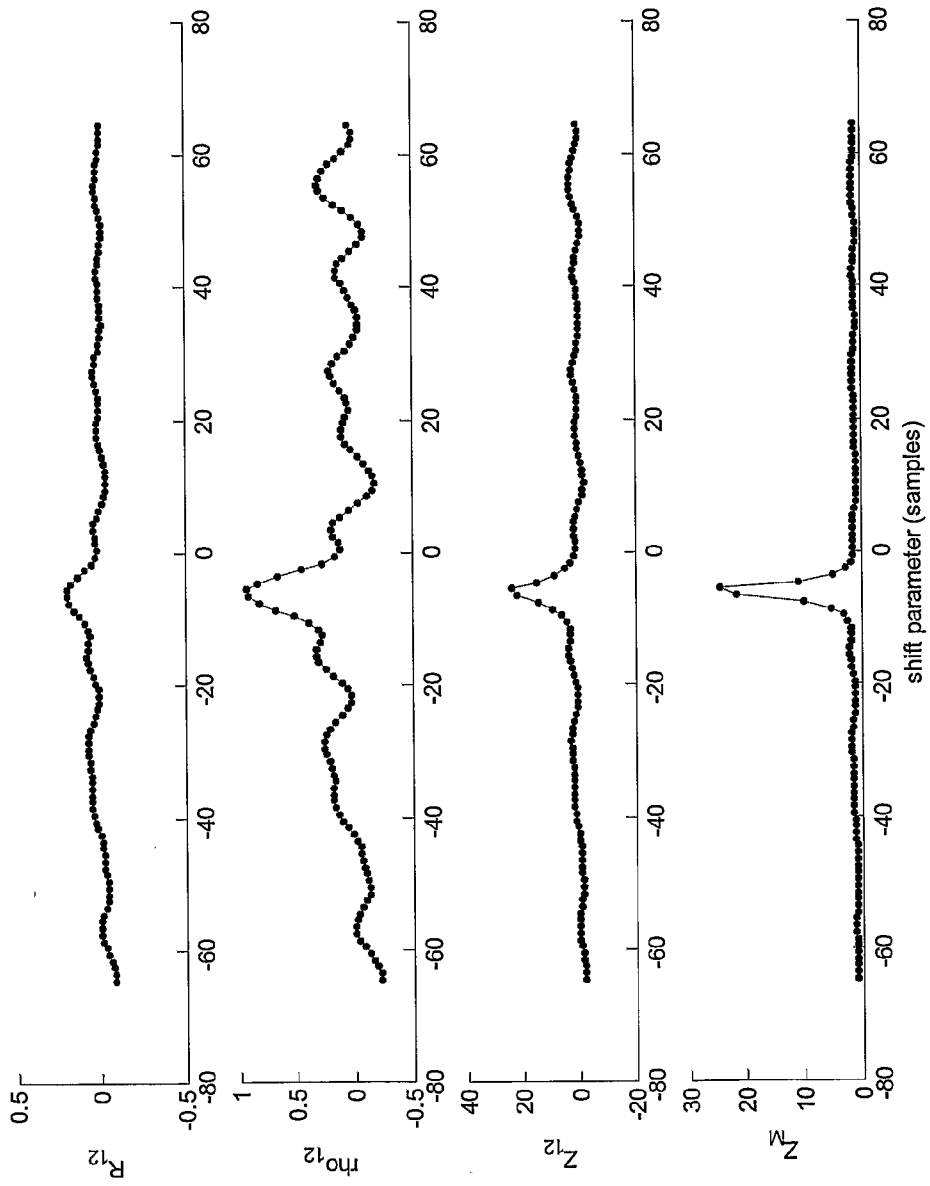


Figure 3A

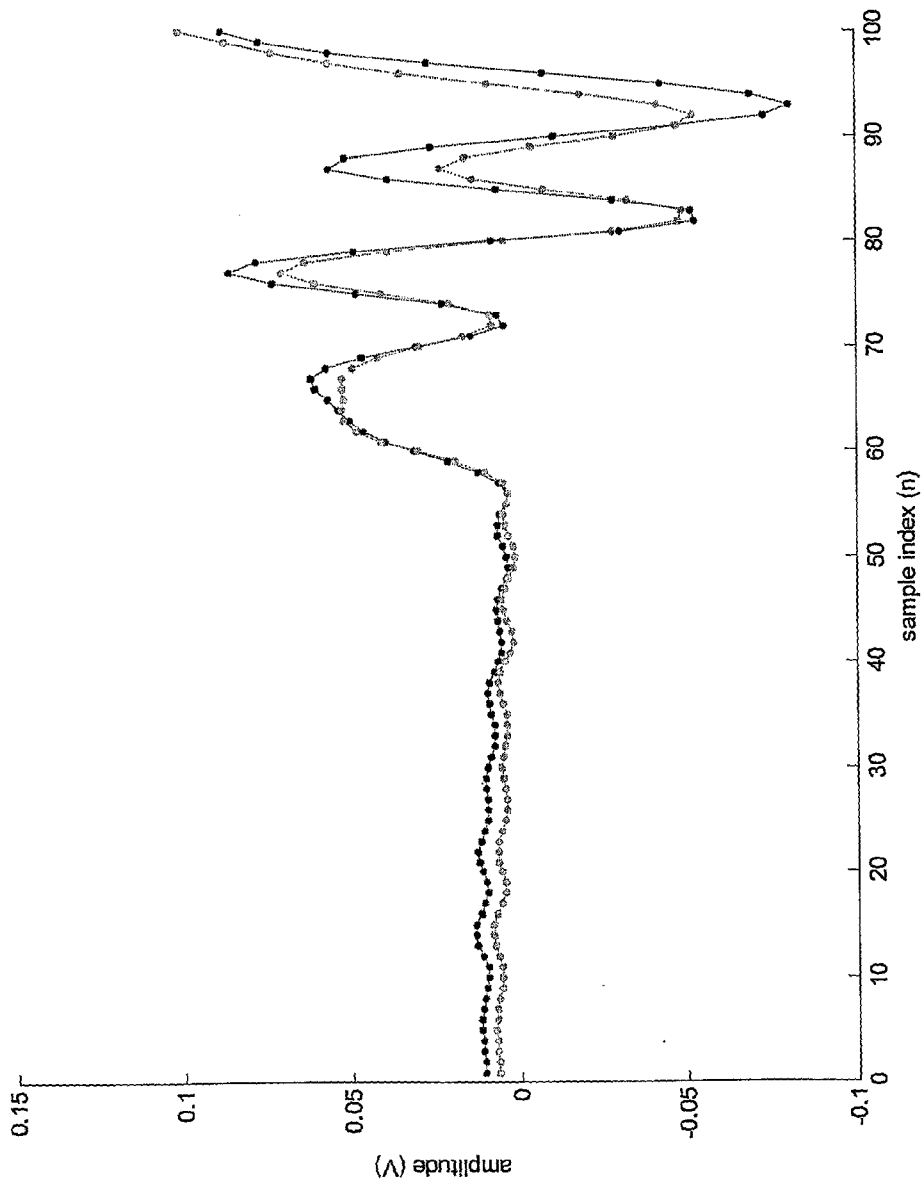


Figure 4

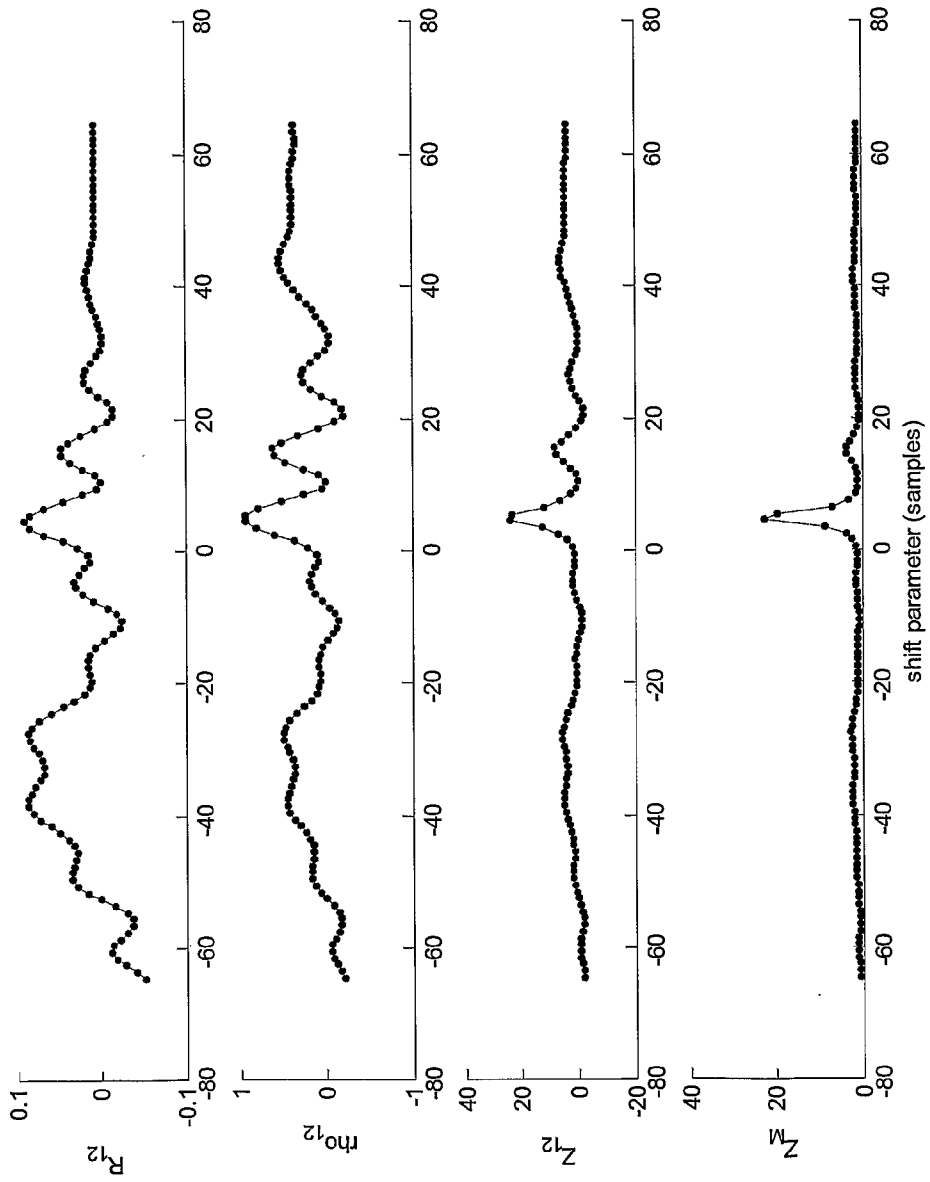


Figure 4A

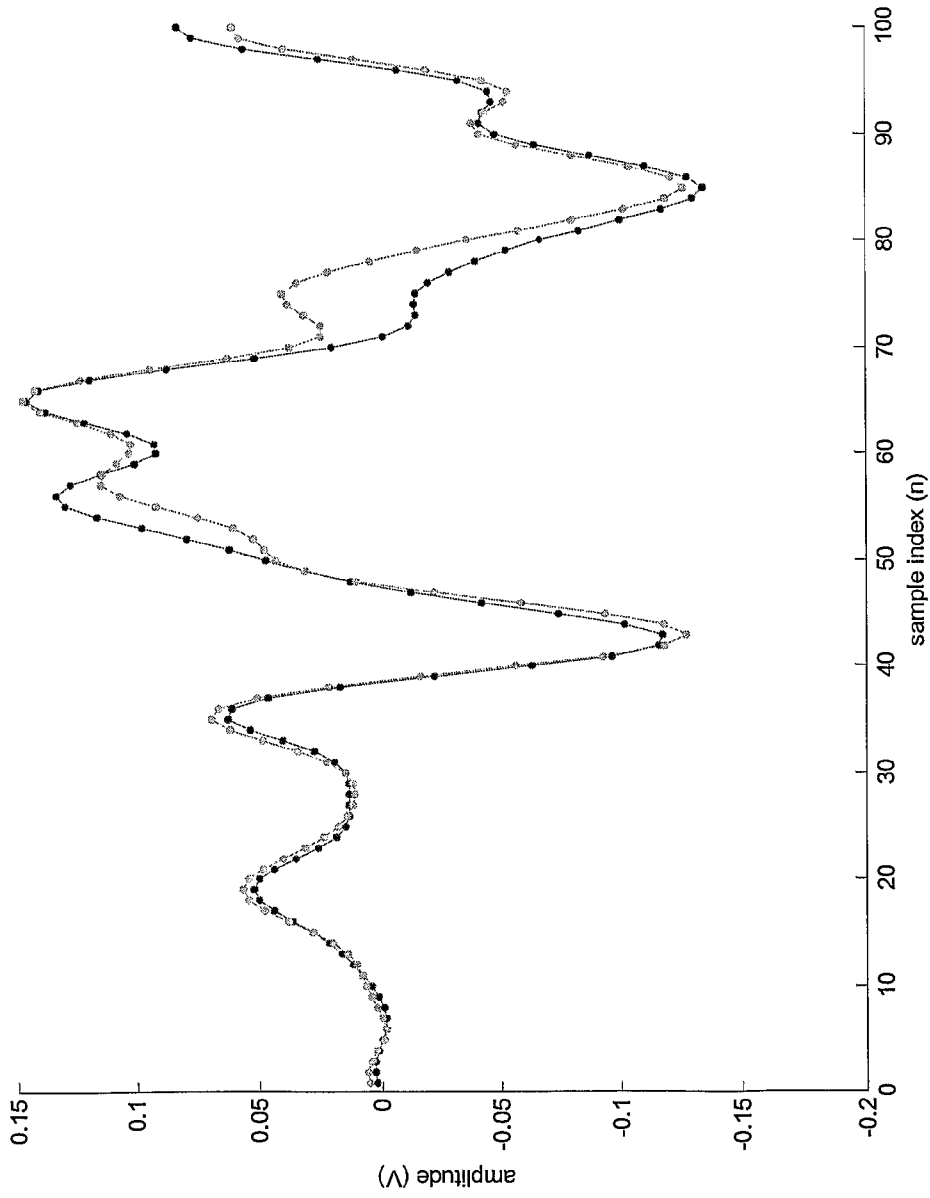


Figure 5

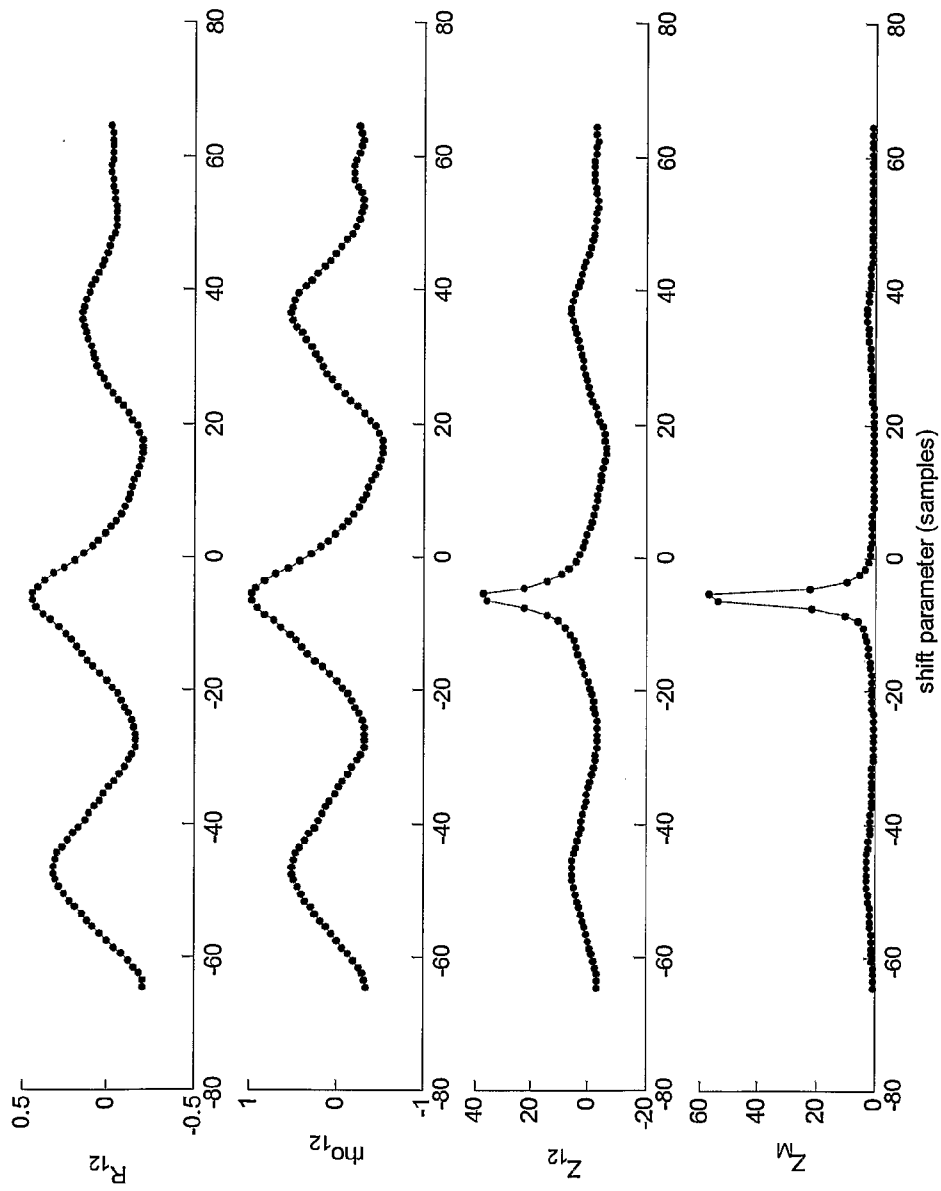


Figure 5A

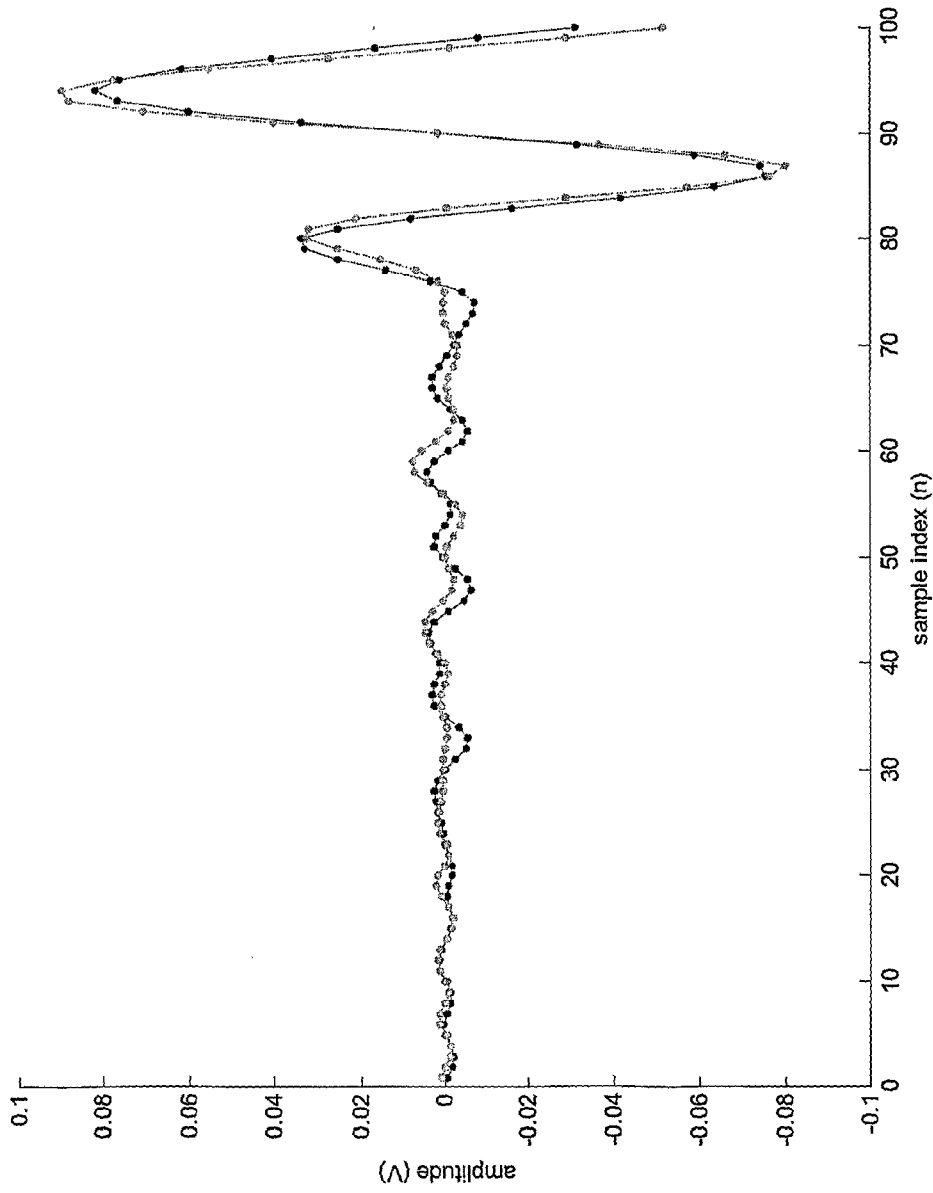


Figure 6

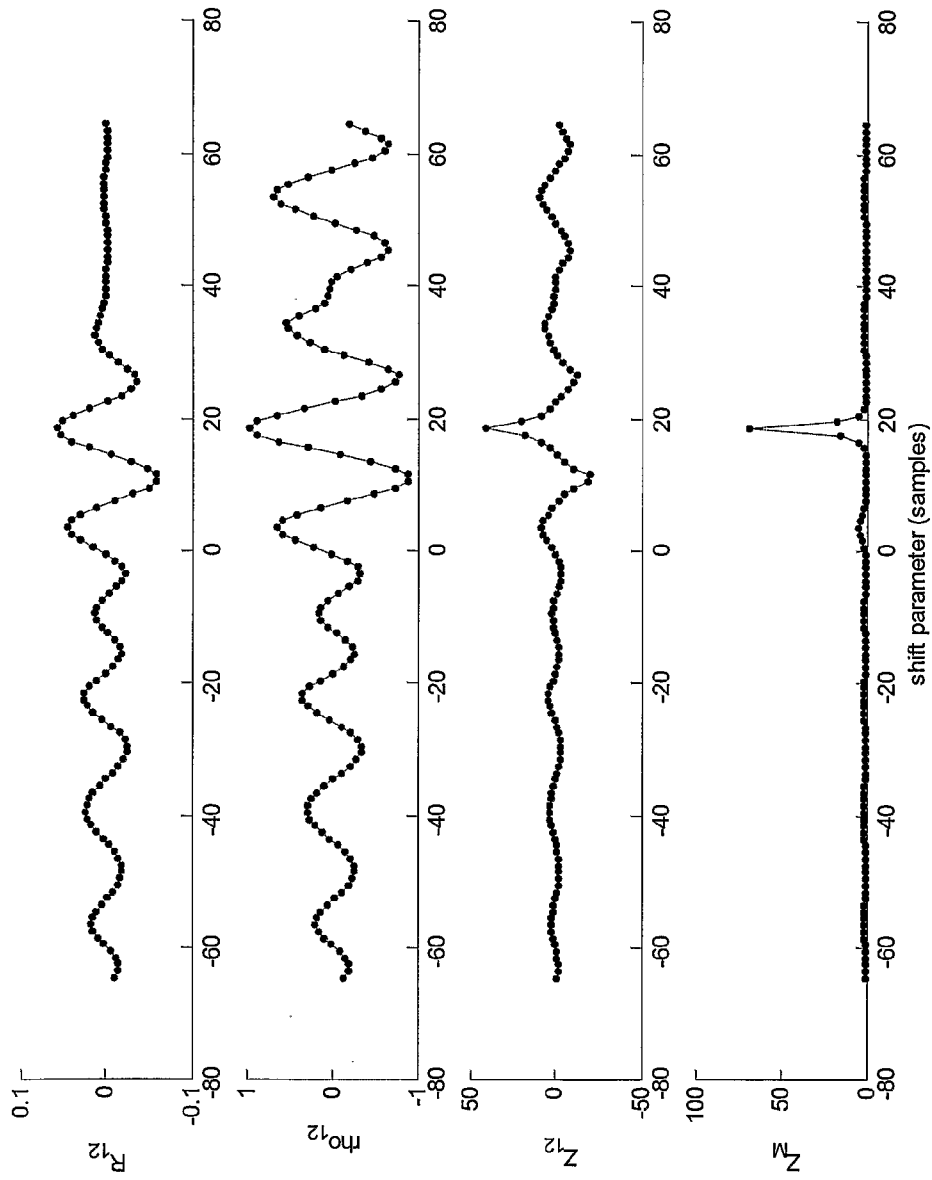


Figure 6A

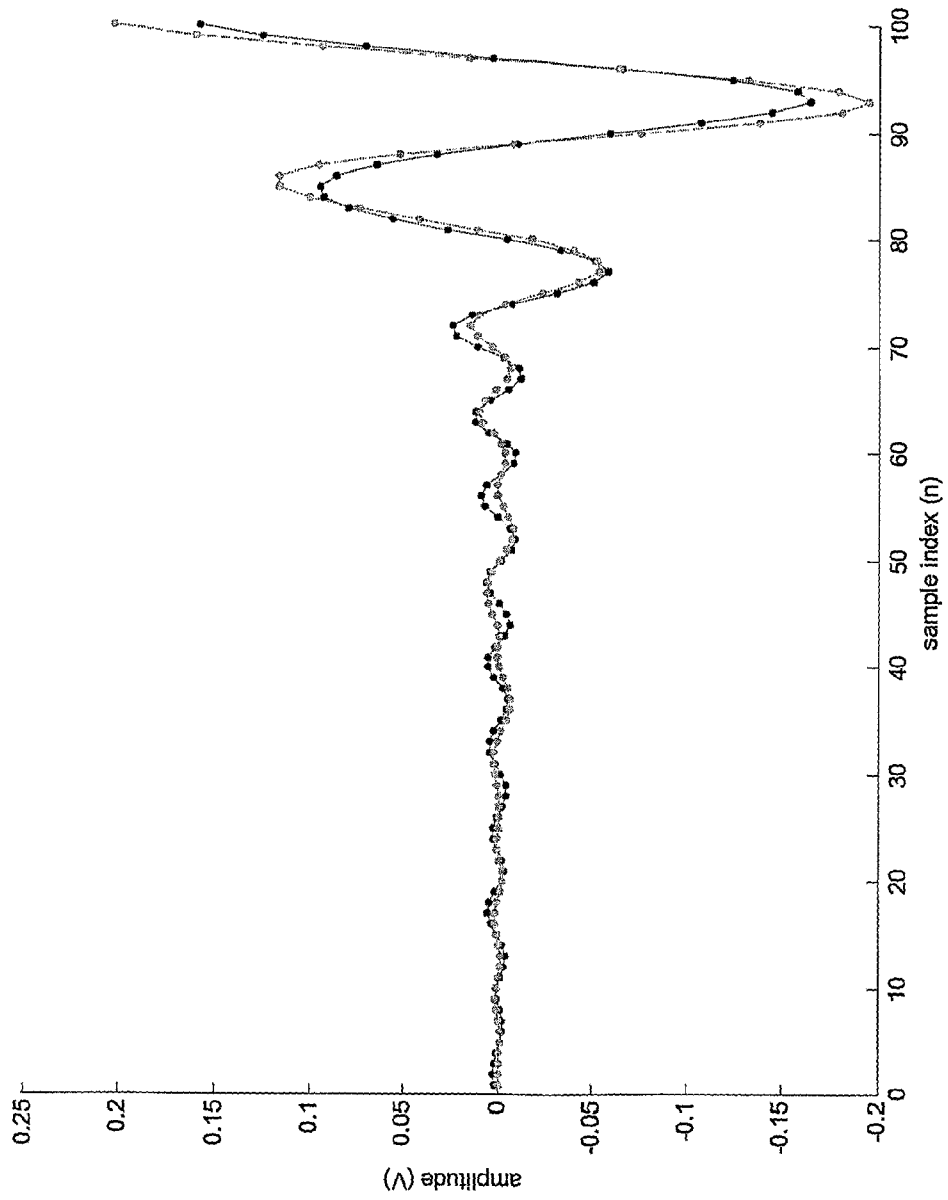


Figure 7

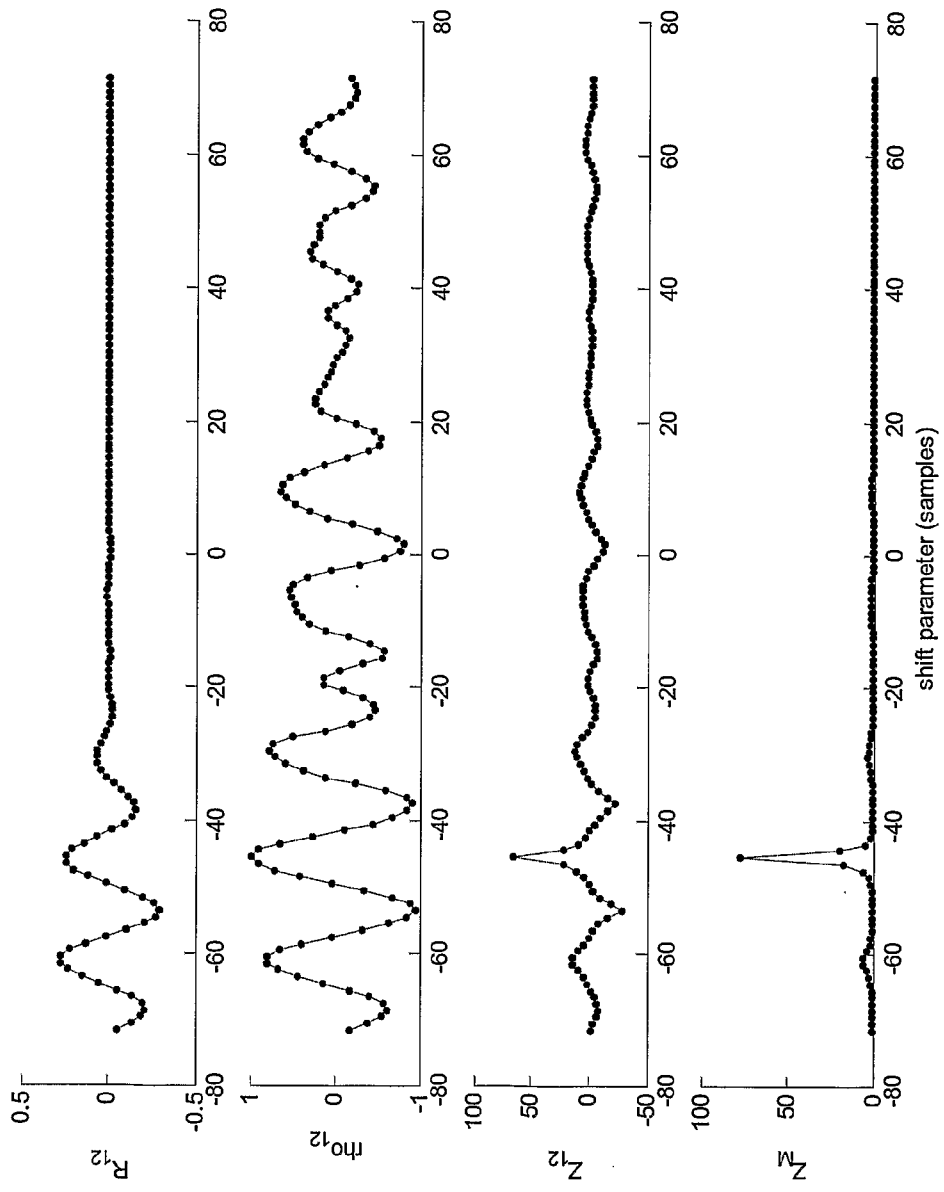


Figure 7A

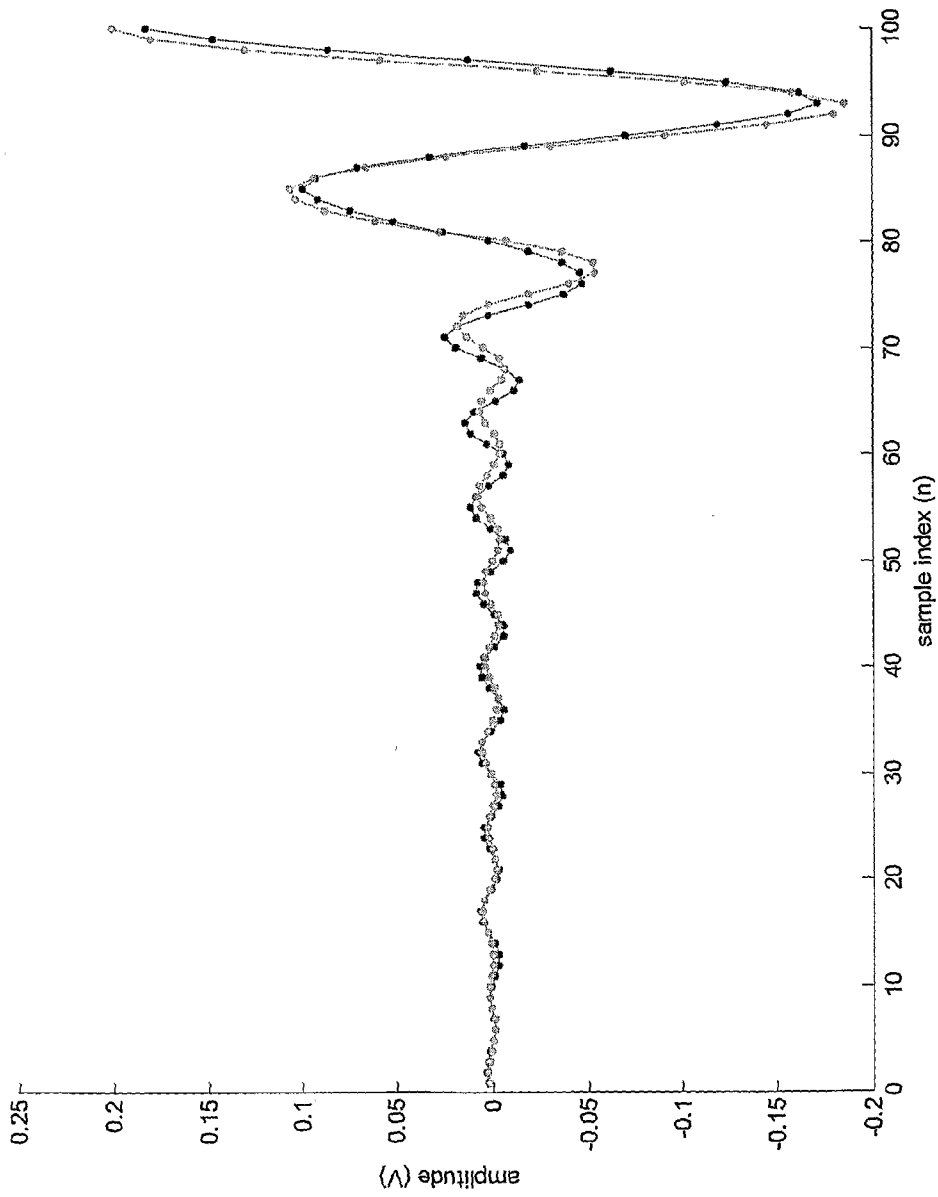


Figure 8

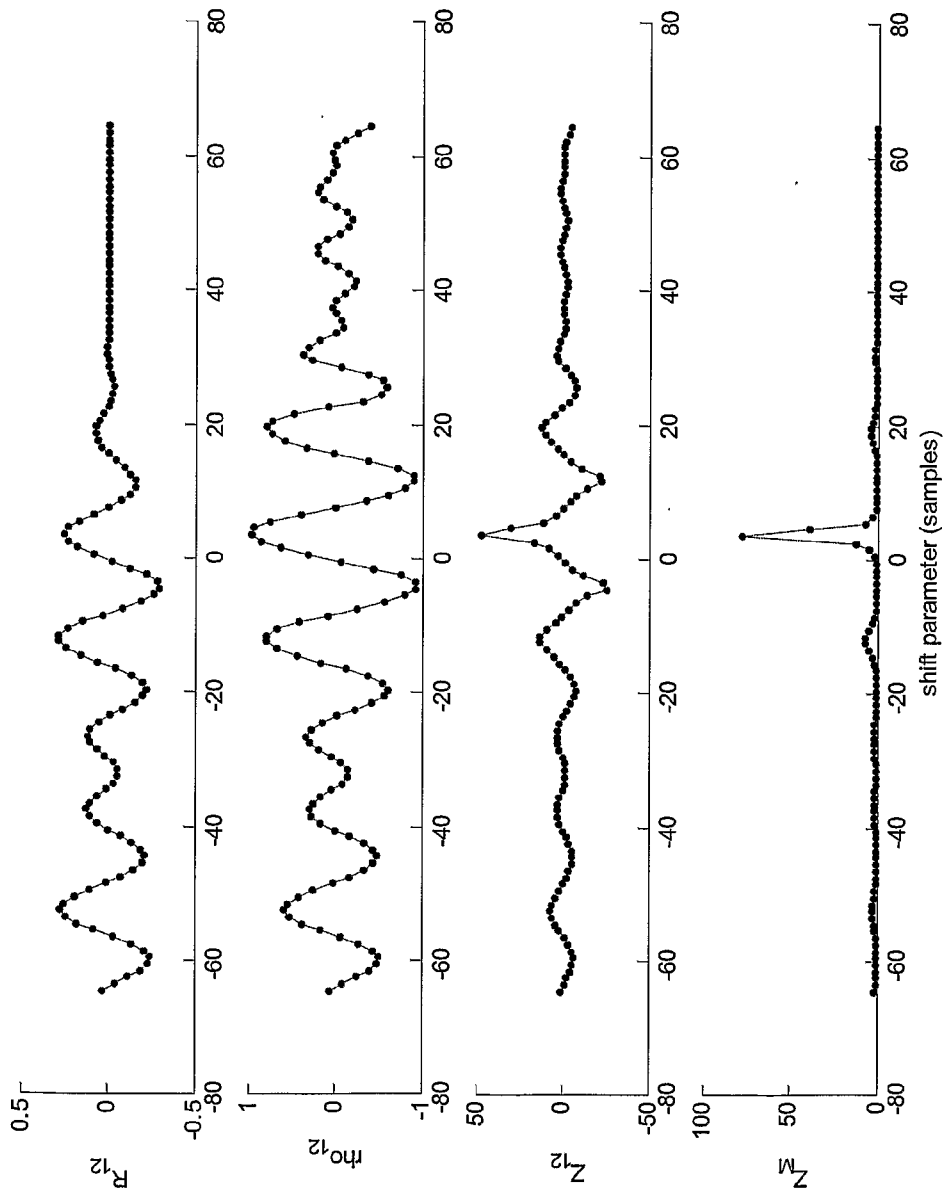


Figure 8A

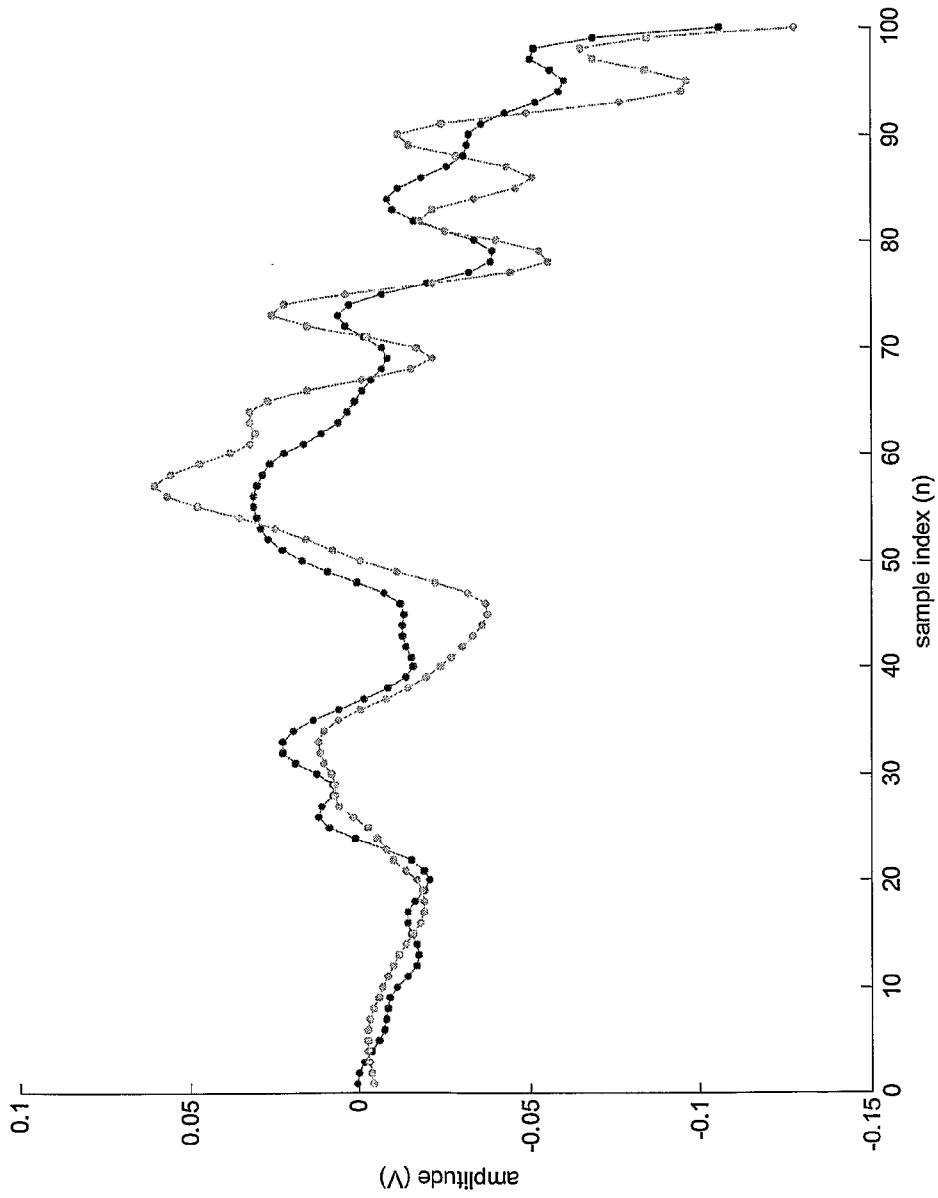


Figure 9

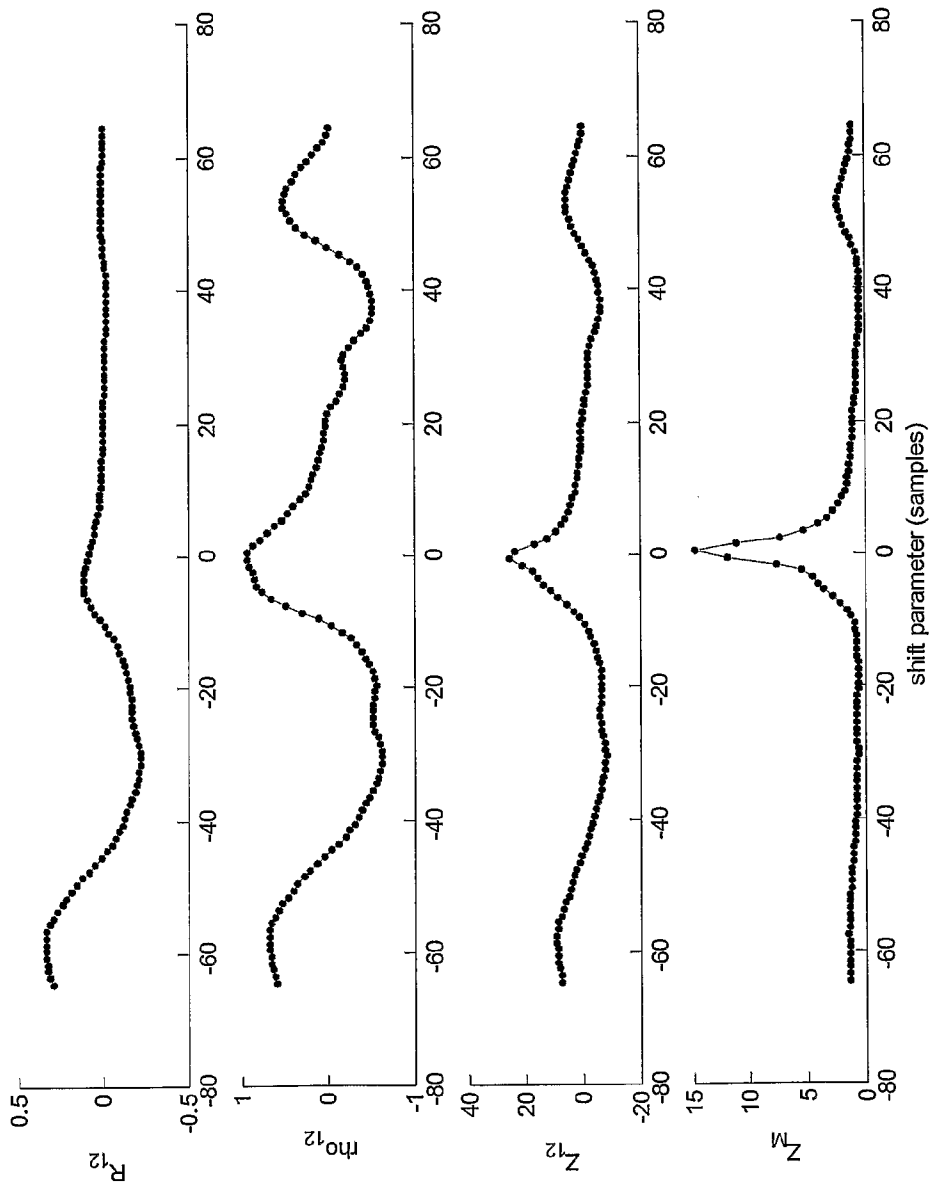


Figure 9A

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/AU2007/000364

A. CLASSIFICATION OF SUBJECT MATTER		
Int. Cl.		
G06F 17/15 (2006.01)		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
WPAT, USPTO, IEEE Keywords: signal data sampl; compar; measur; plural; variat; similar; deviat; stat; standard deviation and like keywords		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6304515 B1 (SPIESBERGER) 16 October 2001 Whole document	1, 11, 16, 24
X	US 5454436 A (JARDINE et al) 3 October 1995 Abstract Col 1 line 42 – Col 2 line 12; Col 2line 50 – Col 4line 4 Figs 1, 2a, 2b	1-3, 11, 12, 16, 17
X	US 6952662 B2 (WEGERICH et al) 4 October 2005 Whole document	11, 24
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C <input checked="" type="checkbox"/> See patent family annex		
* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family	
"O" document referring to an oral disclosure, use, exhibition or other means		
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 03 May 2007	Date of mailing of the international search report 9 MAY 2007	
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaaustralia.gov.au Facsimile No. (02) 6285 3929	Authorized officer ATA MAQBOOL AUSTRALIAN PATENT OFFICE (ISO 9001 Quality Certified Service) Telephone No : (02) 6283 7905	

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/AU2007/000364

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 1143632 A2 (NTT DOCOMO INC, TOKYO 100-6150 (JP)) 10 October 2001 Whole document	24

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

**PCT/AU2007/000364**

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report	Patent Family Member					
US 6304515						
US 5454436	FR 2706942	GB 2279381				
US 6952662	US 2002055826					
EP 1143632	AU 35063/01	CN 1322074	CN 1617460			
	CN 1617461	CN 1617462	CN 1617463			
	JP 2001352275	JP 2006222983	SG 102622			
	US 7050482	US 2001030991	US 2005272460			
<p>Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.</p> <p style="text-align: right;">END OF ANNEX</p>						