

## Neural Networks Based Time-Delay Estimation using DCT Coefficients

Samir J. Shaltaf and Ahmad A. Mohammad

Department of Electronics Engineering, Princess Sumaya University for Technology,  
P.O. Box 1438, Al-Jubaiha 11941 Amman, Jordan

---

**Abstract:** This study dealt with the problem of estimating constant time delay embedded into a received signal that was noisy, delayed and damped image of a known reference signal. The received signal was filtered, normalized with respect to the peak value it achieved and then transformed by the Discrete Cosine Transform (DCT) into DCT coefficients. Those DCT coefficients that were most sensitive to time delay variations were selected and grouped to form the Reduced Discrete Cosine Transform Coefficients set (RDCTC). The time delays embedded in the filtered signals were efficiently encoded into those RDCTC sets. The RDCTC sets were applied to a pre trained multi layer feedforward Neural Network (NN), which computed the time-delay estimates. The network was initially trained with large sets of RDCTC vectors, in which each RDCTC vector corresponded to a signal delayed by a randomly selected constant time-delay. Using the RDCTC as input to the NN instead of the full length incoming signal itself resulted in a major reduction in the NN size. Accurate time delay estimates were obtained through simulation and compared against estimates obtained through classical cross-correlation technique.

**Key words:** Neural networks, time-delay estimation, discrete cosine transform

---

### INTRODUCTION

Time-delay estimation is a problem of many important practical applications. Some of its applications exist in the areas of radar, sonar, seismology, communication systems and biomedicine<sup>[1]</sup>. For example, in active source location applications a reference signal  $s(t)$  is transmitted and a noisy and delayed replica of it is received. The following model represents the received waveform:

$$r(t) = \alpha s(t - d) + w(t) \quad (1)$$

where,  $r(t)$  is the received signal which consists of the reference signal  $s(t)$  after being damped by an unknown attenuation factor  $\alpha$ , delayed by an unknown constant value  $d$  and distorted by an additive white Gaussian noise  $w(t)$ .

A generalized cross-correlation method has been used for estimation of fixed time-delay in which the delay estimate was obtained by the location of the peak of the cross-correlation between the two filtered inputs<sup>[2,3]</sup>. Estimation of constant and time varying delay was considered in<sup>[4,5]</sup>, where Least Mean Square (LMS) adaptive filter was used to correlate the two input data. The resulting delay estimate was obtained as

the location at which the filter obtained its peak value. To obtain the non-integer value of the delay, peak location estimation must be used that involves interpolation. Etter and Stearns<sup>[6]</sup> has used gradient search to adapt the delay estimate by minimizing a mean-square error, which was a function of the difference between the signal and its delayed version<sup>[6]</sup>. Also, So et al. minimized a mean-square error function of the delay, where the interpolating sinc function was explicitly parameterized in terms of the delay estimate<sup>[7]</sup>. The Average Magnitude Difference Function (AMDF) was also exploited for the determination of the correlation peak<sup>[8]</sup>. During in<sup>[8]</sup> has shown that recursive algorithms produced better time-delay estimate than nonrecursive techniques. Conventional prefiltering of incoming signals as used in<sup>[2]</sup> was replaced by filtering one of the incoming signals using wavelet transform<sup>[9]</sup>. Chan *et al.*<sup>[9]</sup> has used the conventional peak detection of the cross-correlation for estimating the delay. Chan et al. reported that their proposed algorithm outperformed the direct cross-correlation method for constant time-delay estimation. In<sup>[10]</sup>, Wang et al. has developed a neural network system that solves a set of unconstrained linear equations using  $L_1$ -norm that optimized the least absolute deviation problem. The time-delay estimation

---

**Corresponding Author:** Samir J. Shaltaf, Department of Electronics Engineering, Princess Sumaya University for Technology, P.O. Box 1438, Al-Jubaiha 11941 Amman, Jordan

problem was converted to a set of linear algebraic equations through the use of higher order cumulants. The unknown set of coefficients represented the parameters of a finite impulse response filter. The parameter index at which the highest parameter value occurred represented the estimated time-delay. The algorithm proposed by Wang et al. was capable of producing delay estimate that was only a multiple integer of sampling interval and did not deal with the case of fractional time-delay. Also, Wang et al. algorithm has utilized the high order spectra, which required heavy computational power.

The author of this article had introduced neural networks for the first time into the direct estimation of constant time-delay<sup>[11]</sup>. Shaltaf<sup>[11]</sup> has trained NN with large set of data representing the reference and the received signals with embedded constant time-delays. The NN were tested with a new set of data and produced accurate estimates of time-delay. The signals were introduced to the NN in vector form, in which the signal vector size was either 128 or 256 samples. This has resulted into using NN with large number of inputs, which in turn resulted in large NN systems.

Many compression techniques exist for reducing data size, one of which was the DCT. The DCT was established as a major compression technique and has been known for its excellent energy compaction property and is widely used in signal and image compression<sup>[12]</sup>. This study utilized the compression property of the DCT in order to compress the signal into a small set of RDCTC. Only those few DCT coefficients which possessed high sensitivity to time-delay variations were chosen amongst the RDCTC set. The resulting set of RDCTC was then used as input to the NN, hence, resulting into using a small size NN. In<sup>[11]</sup>, the size of the NN was large due to the direct use of the reference and received signals as input to the NN. The input layer size was as large as 128 and 256, which was equal to the signal vector size.

As a start, a damped, delayed and noisy image of a reference signal was received as modeled in Eq. 1. The received signal was sampled, filtered by a band pass digital filter, normalized with respect to the peak value it achieved and then compressed by the DCT into a small set of RDCTC. Those DCT coefficients that possessed high sensitivity to time delay variations were selected as members of the RDCTC set. The time-delay embedded in the filtered signal was then indirectly encoded into the RDCTC set. A large collection of RDCTC sets were applied to a pre trained multi layer feedforward NN which computed the time-delay estimates. About one thousand training sets of RDCTC with its corresponding time-delays were used to train

the NN. Accurate and fast estimates of time-delay resulted from performing one pass of the RDCTC through the NN. This estimation process was fast when compared with the classical techniques for time-delay estimation. Classical techniques first rely on generating the computationally demanding cross-correlation between the two signals. Then it utilizes a peak detection algorithm to find the time at which the peak exists.

### TIME DELAY ESTIMATION ALGORITHM

The reference signal  $s(t)$  was assumed to be sinusoidal signal with frequency  $\Omega_0$  rad sec<sup>-1</sup>. and sampled with a sampling period  $T$  seconds. The resulting discrete reference signal was:

$$s(n) = \sin(\omega_0 n) \quad (2)$$

where,  $\omega_0 = \Omega_0 T$  was the frequency of the sampled reference signal. Assuming the received signal in (1) has been filtered by an anti-aliasing filter and sampled, then its discrete form is:

$$r(n) = \alpha s(n-D) + w(n) \quad (3)$$

where,  $s(n-D)$  was the delayed reference signal,  $D$  was an unknown constant delay measured in sampling intervals and related to the time-delay  $d$  by the relation  $D = d/T$  and  $w(n)$  was assumed a zero-mean Gaussian noise with variance  $\sigma_w^2$  uncorrelated with the signal  $s(n)$ . Since the received signal  $r(n)$  was noisy, it was best that it got filtered in order to obtain better estimates for the unknown time-delay.

A fourth order type II Chebyshev band pass digital filter was used to filter the received signal. The filter was designed to have a narrow pass bandwidth equal to 0.02 radian with 2 dB attenuation for the pass frequencies and a stop band bandwidth of 0.32 radian with 40 dB attenuation for the stop frequencies. The center frequency  $\omega_c$  of the band pass filter was set equal to the reference signal frequency  $\omega_0 = 0.3$  and was set to be exactly equal to the geometric mean of the pass and stop frequencies. The resulting band pass digital filter order that satisfied the above conditions was found to be 4. The strict narrow bandwidth condition resulted in a band pass filter that was capable of reducing the input noise power to about 1.6% of its value at the filter output. This meant a noise reduction factor equal to 62.5, which implied a signal to noise ratio improvement by 18dB. This improvement on the signal to noise ratio resulted in improving the accuracy of the time-delay estimates.

The band pass filter was used to filter the received noisy signal and improve the signal to noise ratio. To eliminate the effect of the presence of the damping factor  $\alpha$ , the filtered signal was normalized with respect to the peak value it achieved. The filtered signal was then normalized in order to prevent the NN from being affected by the different amplitude variations of the signal due to the different and unknown damping factors. The filtered and normalized signal was DCT transformed and the most sensitive DCT coefficients were used to form the RDCTC set. The resulting RDCTC set was then applied as input to the NN. By applying the RDCTC of the filtered and normalized signal to the NN, the NN should then be capable of producing accurate time-delay estimates.

The resulting band pass digital filter  $H(z)$  turned to assume the following transfer function form:

$$H(z) = \frac{0.0102 - 0.037z^{-1} + 0.0537z^{-2} - 0.037z^{-3} + 0.0102z^{-4}}{1 - 3.7585z^{-1} + 5.4682z^{-2} - 3.6397z^{-3} + 0.9378z^{-4}} \quad (4)$$

Let  $h(n)$  be the impulse response of the digital filter  $H(z)$  in Eq. 4, by applying the received signal  $r(n)$  as input to the filter, the corresponding output was:

$$r_f(n) = \sum_{k=0}^n h(k)r(n-k), \quad n = 0, 1, \dots, N-1. \quad (5)$$

Let  $r = [r_f(0), r_f(1), \dots, r_f(N-1)]$  be a vector consisting of  $N$  samples of the received and filtered signal. The vector  $r$  was DCT transformed and only  $M$  of the most sensitive DCT coefficients were selected to form the RDCTC which was used as input to the NN. The constant time-delay was used as the NN output in the training phase.

### DCT TRANSFORM AND SELECTION OF SENSITIVE DCT COEFFICIENTS

In this research the received and filtered signal  $r_f(n)$  was DCT transformed and the most  $M$  sensitive DCT coefficients were used to form the RDCTC set. The time-delay embedded in this signal was then indirectly encoded into those RDCTC coefficients. The NN was supposed to decode the RDCTC into an accurate time-delay estimate.

The following was the DCT transformation used for transforming the filtered signal  $r_f(n)$  which resulted in  $R_m$  as the DCT coefficients:

$$R_m = \begin{cases} \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} r_f(n) \cos\left(\frac{(2n+1)m\pi}{2N}\right), & m = 1, 2, \dots, N-1 \\ \sqrt{\frac{1}{N}}, & m = 0 \end{cases} \quad (6)$$

Out of the  $N$  DCT coefficients, the best  $M$  DCT coefficients which possessed the highest sensitivity to time-delay variations were concatenated to form the RDCTC vector  $R$ .

To show the capability of the DCT in encoding time-delay values into the RDCTC coefficients, a sinusoidal reference signal  $s(n) = \sin(0.3n)$  was delayed with several delay values in the range  $[0.0-10.0]$  in increments of 0.25 sampling interval. The signal length was chosen to be equal to  $N = 128$  samples. The signal was filtered by the band pass filter and only the first 64 DCT coefficients curves  $[R_0, R_1, \dots, R_{63}]$  were plotted against the time-delay in Fig. 1a. The rest of the DCT coefficients had negligible changes versus time delay variations. A clear and direct relationship that shows how the DCT coefficients change against time-delay variations is clearly shown in Fig. 1a.

In image or signal compression the best DCT compression coefficients are chosen based on a constant or adaptive thresholding technique. The DCT coefficient with a value that is larger than the specified threshold level is selected, otherwise it is ignored. In

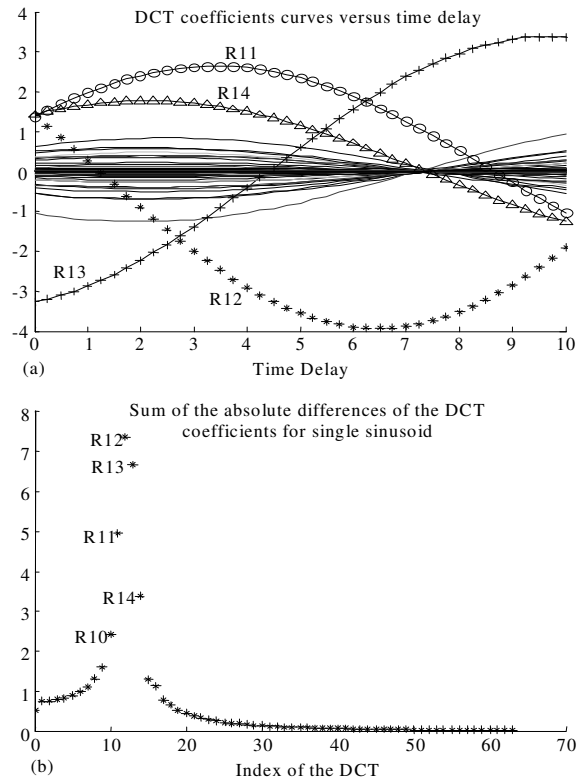


Fig. 1: (a): DCT coefficients curves versus time delay for single sinusoid signal. (b): SAD values of the DCT coefficients

Table 1: Sum of the absolute difference (SAD) values of the most sensitive DCT coefficients for the single sinusoid signal

DCT Coef	R <sub>12</sub>	R <sub>13</sub>	R <sub>11</sub>	R <sub>14</sub>	R <sub>10</sub>	R <sub>9</sub>	R <sub>8</sub>	R <sub>15</sub>	R <sub>16</sub>	R <sub>7</sub>
SAD	7.35	6.65	4.95	3.38	2.42	1.61	1.33	1.31	1.14	1.11

this research, the selection criterion of the DCT coefficients was based on the DCT coefficient sensitivity with respect to time-delay variations. Only M DCT coefficients that possessed the highest sensitivity were selected as the appropriate RDCTC set. The time-delay embedded in the filtered signal was clearly observed to be encoded into those RDCTC coefficients. The NN was supposed to learn this relationship and encode it into its internal structure parameters in the training phase.

A simple technique was devised to obtain the most sensitive DCT coefficients to form the RDCTC set. The gradient of each DCT coefficient curve with respect to time-delay was approximated by its difference. The Sum of the Absolute Difference (SAD) was performed for each DCT coefficient curve seen in Fig. 1. Each DCT coefficient curve which resulted in a high SAD value was chosen as one of the most sensitive RDCTC coefficients. Table 1 shows the SAD values of the first ten most sensitive RDCTC values in decreasing order. It is observed that R<sub>12</sub> was the most sensitive of all coefficients. Then it is followed by R<sub>13</sub>, R<sub>11</sub> and R<sub>14</sub>. This result can be inspected visually by looking at Fig. 1a-b. Figure 1a shows the DCT coefficients curves versus time delay, where it is observed that the curve of the DCT coefficient R<sub>12</sub> possessed the highest gradient of all RDCTC coefficients. Figure 1b shows the SAD values of each DCT coefficients where it was observed that [R<sub>12</sub>, R<sub>13</sub>, R<sub>11</sub>, R<sub>14</sub> R<sub>10</sub>] were the most sensitive DCT coefficients in decreasing order. It must be mentioned that although R<sub>12</sub> has the highest SAD value, it must be mentioned that although it is most sensitive in the intervals [0.0-6.0] and [7.0-10.0], it has the lowest sensitivity within the interval [6.0-7.0] as can be seen in Fig. 1a. This implied that R<sub>12</sub> would contribute to the accuracy of the time-delay estimate within the intervals in which it is most sensitive and its contribution to the accuracy of the time-delay estimate would be in its lowest within the interval in which it is most insensitive. Luckily, it was observed from looking at Fig. 1a that in the interval [6.0-7.0] in which R<sub>12</sub> was most insensitive, the DCT coefficients R<sub>11</sub>, R<sub>13</sub> and R<sub>14</sub> were most sensitive. Therefore when one DCT coefficient was least sensitive in one interval, it was observed that other DCT coefficients were most sensitive within that interval and hence they contributed to the accuracy of the time-delay when that coefficient

contribution was the least. This observation told that at least two RDCTC coefficients must be used as inputs to the NN in order to obtain an accurate time delay estimates over the range [0.0-10.0] as can be shown in Fig. 1a.

## NEURAL NETWORK SYSTEMS

The neural network systems used in this study were two layers feedforward networks. Hyperbolic tangent nonlinearity was used for the hidden layer neurons, while linear transfer function was used for the output neuron. Improved version of the backpropagation training algorithm was used for training the network. It was called resilient backpropagation<sup>[13]</sup>. Riedmiller *et al.*<sup>[13]</sup> had noticed that the nonlinear transfer functions; the hyperbolic tangent and the log sigmoid; of the neurons have very small gradient values for large input values. The small gradient values resulted in slow convergence for the NN in the training phase because the backpropagation was gradient based learning algorithm. In order to overcome this problem, the sign of the gradient was used instead of its small value to update the NN parameters. This has resulted in a major improvement on the speed of convergence of the NN.

Multiple hidden layers NN were not considered in this research since it was shown in<sup>[11]</sup> that it did not add significantly to the improvement of the time delay estimates.

## SIMULATION RESULTS

In this research, a search for the optimal NN structure that would result in the smallest delay estimation error variance was performed. Then the optimal NN structure found was used for estimating the unknown time delay.

**Search for optimal neural network:** To find out the optimal NN system that would result in the smallest delay estimation error variance, different NN structures were trained with 1000 noise free data sets with time delay values chosen randomly within the range [0.0-10.0]. The trained NN were then tested with a new 1000 signals that were delayed with delay values starting from 0.0 and ending with 10.0 samples in increments of 0.01 samples. The variance between the exact time delay values and the estimated delay values was then obtained and tabulated in Table 2 for the case of the single sinusoidal reference signal. Table 2 shows the estimation error variance for NN systems with a single hidden layer. The hidden layer consists either of 5, 10,

Table 2: Delay estimation error Variance for a single sinusoid using most sensitive DCT coefficients with noise free data

No. of RDCTC Coef.	No. of Hidden units					
	5	10	15	20	25	30
2	9.63E-4	7.23E-4	11.62E-4	10.47E-4	11.30E-4	14.09E-4
4	6.35E-4	4.63E-4	6.88E-4	7.77E-4	7.75E-4	4.95E-4
6	4.70E-4	7.74E-4	7.03E-4	5.05E-4	4.40E-4	8.52E-4
8	5.20E-4	3.89E-4	4.70E-4	5.70E-4	6.94E-4	7.63E-4
10	10.10E-4	4.13E-4	4.99E-4	5.81E-4	5.11E-4	6.24E-4
12	5.64E-4	4.04E-4	3.51E-4	5.25E-4	5.79E-4	5.33E-4
14	5.34E-4	4.62E-4	6.13E-4	4.99E-4	4.70E-4	4.45E-4
16	5.35E-4	5.53E-4	4.78E-4	5.46E-4	4.11E-4	4.19E-4

15, 20, 25, 30 neurons. Each one of these NN systems was tested with 2, 4, 6, 8, 10, 12, 14, 16 RDCTC input coefficients. Looking into Table 2, it was observed that the NN with 15 hidden neurons and 12 RDCTC coefficient inputs was the best of all NN, where it was observed that it resulted in the smallest delay estimation error variance of 3.51E-4. Hence, the NN with 15 hidden neurons and 12 RDCTC coefficient inputs was chosen as the optimal NN.

**Simulation using the optimal NN structure:** In the previous section, the optimal NN structure was found to have a total of 12 RDCTC inputs and 15 hidden neurons. This optimal NN was trained with RDCTC sets corresponding to filtered signals that experienced noise with standard deviation levels of 0.1, 0.3 and 0.5. About one thousand training patterns were applied to the NN in patch training mode. The additive Gaussian noise used in this data had a standard deviation value chosen randomly within [0.0-0.5] range. The time-delay values were chosen randomly within [0.0-10.0] sampling interval. The first 12 most sensitive RDCTC coefficients obtained under different noise levels and different time-delay values were applied to the optimal NN structure as inputs and the corresponding time delays as the output.

The trained NN was tested with a set of 1000 RDCTC set corresponding to received signals with imbedded unknown time delays. The damping factor was chosen to have a value of 1.0 in those sets of data. Figure 2 shows the histograms for the delay estimation errors. Figure 2a shows the histogram for the delay estimation error where the used noise had a standard deviation of 0.1. It was observed that all of the delay estimation errors were within the range of  $\pm 0.2$  sampling interval. The delay estimation error variance was found to be 0.0026 which corresponded to a standard deviation of 0.051. It was also observed that the histogram took the Gaussian shape which implied that the delay estimation error assumed the Gaussian density shape. This implied that 99.9% of the delay

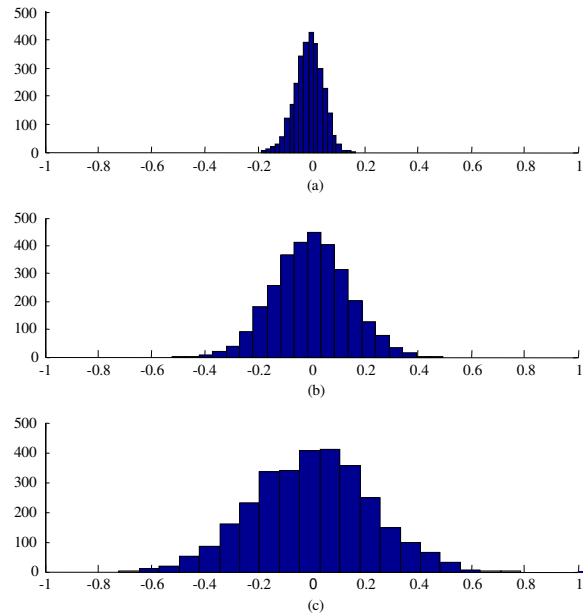


Fig. 2: Delay estimation error histograms. (a): Noise standard deviation is 0.1, (b): Noise standard deviation is 0.3, (c): Noise standard deviation is 0.5

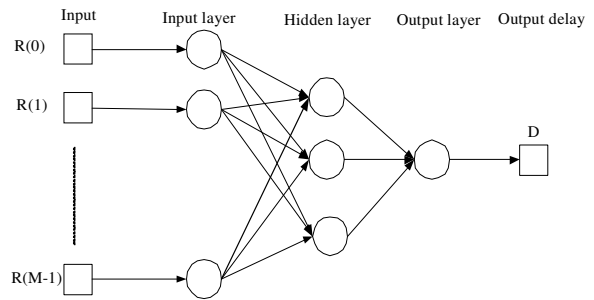


Fig. 3: Two layer feed forward neural network utilizing the first M RDCTC vector R as its input

Table 3: Time-delay estimation based on the optimal NN

Standard deviation of data noise	0.1	0.3	0.5
Delay estimation error variance	0.0214	0.0186	0.0508

estimation errors laid within  $\pm 3$  standard deviations which correspond to  $\pm 0.153$  sampling interval. Figure 2b and c correspond to noise with standard deviation of 0.3 and 0.5 respectively. It was obvious from Fig. 2 that as the noise level increased the delay estimation error increased as well. The delay estimation error laid within  $\pm 0.5$  for Fig. 2b and within  $\pm 0.8$  for Fig. 2c and Fig. 3.

Table 3 shows the results of testing the optimal NN with noisy input data with random delay values within

Table 4: Time-delay estimation based on the cross-correlation technique

Standard deviation of data noise	0.1	0.3	0.5
CROSS1 delay estimation error variance	0.0021	0.0163	0.0413
CROSS2 delay estimation error variance	0.0151	0.0309	0.0655

[0.0-10.0] sampling intervals. The input data was corrupted with three noise levels of 0.1, 0.3 and 0.5. The delay estimation error variances for the optimal NN were presented. It was noted that as the noise level increased the estimation error variance increased.

Table 4 shows the time delay estimation error variance using the cross correlation technique. The noise free reference signal was correlated with the noisy received signal and the delay estimation error variance was presented in the second row of table 4. This cross-correlation was denoted as CROSS1. Another cross-correlation technique was performed by correlating the filtered noise free signal with the filtered received noisy signal and was denoted by CROSS2. In all simulations, a set of 1000 noisy and delayed signals were used in which the noise standard deviation levels were 0.1, 0.3 and 0.5. It was shown that CROSS1 resulted in delay error variance less than that obtained by the optimal NN for all noise levels. But the optimal NN performed better than CROSS2 for noise levels larger than 0.1.

### CONCLUSION

A new time-delay estimation scheme was developed through the use of neural networks and RDCTC. The reduced set of RDCTC coefficients was used as the NN input instead of the filtered signal. This has resulted in a major reduction in the NN size. It was shown in this study that a small size NN with a reduced set of RDCTC coefficients was capable of producing accurate time-delay estimates that were comparable to those obtained by the cross-correlation technique.

### REFERENCES

1. Special Issue On Time-delay Estimation, IEEE Transaction on Acoustics, 198. Speech and Signal Processing, vol. ASSP 29(3),.

2. Knapp, C.H. and G.C. Carter, 1976. The generalized correlation method for estimation of time delay. *IEEE Trans. Acoust. Speech Signal Process.*, 24: 320-327.
3. Carter, G.C., 1987. Coherence and time-delay estimation. *Proc. IEEE*, 75: 236-255.
4. Feintuch, P.L., N.J. Bershard and F.A. Reed, 1981. Time-delay estimation using the LMS adaptive filter-dynamic behavior. *IEEE Trans. Acoust. Speech Signal Process.*, 29: 571-576.
5. Chan, Y.T., J.M.F. Riley and J.B. 1981. Plant, modeling of time delay and its application to estimation of non-stationary delays. *IEEE Trans. Acoust. Speech Signal Process.*, 29: 577-581.
6. Etter, D.M. and S.D. Stearns, 1981. Adaptive estimation of time delays in sampled data systems. *IEEE Trans. Acoust. Speech Signal Process.*, 29: 582-587.
7. So, H.C., P.C. Ching and Y.T. Chan, 1994. A new algorithm for explicit adaptation of time delay. *IEEE Trans. Signal Process.*, 42: 1816-1820.
8. During, C., 1997. Recursive versus nonrecursive correlation for real-time peak detection and tracking. *IEEE Trans. Signal Process.*, 45: 781-785.
9. Chan, Y.T., H.C. So and P.C. 1999. Ching, approximate maximum likelihood delay estimation via orthogonal wavelet transform. *IEEE Trans. Signal Process.*, 47: 1193-1198.
10. Wang, Z., J.Y. Cheung, Y.S. Xia and J.D.Z. Chen, 2000. Neural implementation of unconstrained minimum  $l_1$ -norm optimization-least absolute deviation model and its application to time delay estimation. *IEEE Trans. Cir. Syst. II: Analog Digital Signal Process.*, 47.
11. Shaltaf, S.J., 2004. Neural networks based time delay estimation. *EURASIP J. Applied Signal Process.*, 3: 378-385.
12. Jain, A., 1989. *Fundamentals of Digital Image Processing*, Prentice Hall.
13. Riedmiller, M. and H. Braun, 1993. A direct adaptive method for faster backpropagation learning: The RPROP algorithm. In: *Proceedings of the IEEE International Conference on Neural Networks*.