Implementation of Time Delay Estimation Using different Weighted Generalized Cross Correlation in Room Acoustic Environments

Kobra Mahdinejad *, Morteza Ziaee Seghaleh **

* Department of Electrical Engineering, Bojnourd Branch, Islamic Azad University, Bojnurd, Iran
Email: k_mahdinejad@yahoo.com
** Department of Electrical Engineering, Asrar Institute of Higher education, Mashhad, Iran
Email: morteza_ziaee@yahoo.com

Time delay estimation (TDE) is one of the most significant topic of research in many fields of study such as radar, sonar, geophysics, seismology, ultrasonic, hands-free communications and any other field each related to positioning a radiating source. TDE is the first step for any application related to identifying, localizing, and tracking radiating sources which used Time Difference of Arrival (TDOA). This paper presents an implementation of TDE using different weighted generalized cross correlation and make a comparison between them. We will discuss the effect of length of observation interval on accuracy and speed of estimation and the influence of distance of microphones and sound source on estimation error based on some experimental results in room acoustic environments with reverberation and noise.

[Kobra Mahdinejad, Morteza Ziaee Seghaleh. **Implementation of Time Delay Estimation Using different Weighted Generalized Cross Correlation in Room Acoustic Environments.** *Life Sci J* 2013;10(6s):846-851] (ISSN:1097-8135). http://www.lifesciencesite.com. 134

Keywords: Time delay estimation (TDE), Time Delay (TD), Cross-Correlation (CC), General Cross-Correlation (GCC), Time Difference of Arrival (TDOA), Maximum Likelihood (ML)

1. INTRODUCTION

Time delay estimation (TDE), the first stage of any detection system which recognize, and locate radiating sources, has a lot of applications in many fields of study such as radar, sonar, geophysics, seismology, ultrasonic, and communications. Noticeable amount of researches have worked on this topic so far.

The estimation of time delay would be an easy task if the two received signals were just a delayed and scaled version of each other. In real world, however, the source of signal is usually covered by noise since natural environment around us is a place where the existence of noise is inevitable. Furthermore, each received signal may contain multiple delayed and attenuated duplicates of signal which emitted from source and that is due to reflections from objects in the environment. This multipath propagation effect introduces echoes and spectral distortions into the observation signal, termed as reverberation, which severely deteriorates the source signal. In addition, the source of the wave front may also move from time to time, resulting in a changing time delay. All these factors make time delay estimation a sophisticated and challenging problem.

A great deal of efforts has been made to improve the robustness of TDE techniques over the past few years. By and large, the improvements are achieved through three different ways. The first way is to incorporate previous knowledge about the distortion sources into the GCC method which lead to improvement of its performance. The second technique is to use more than two sensors. In this way it can be taken the advantages of the redundancy that enhance the accuracy of delay estimation between the two selected sensors. The third way is to consider reverberation in the signal model and exploit the advanced system identification techniques to improve time delay estimation. [1] attempts to summarize these efforts, and reviews all these important techniques, and mentions the recent advances which have significantly improved performance of time delay estimation in adverse environments.

2. TDE Algorithms

Before discussing the TDE algorithms, different mathematical models that can be used to describe an acoustic environment for the TDE problem will introduce. It helps us better understand the problem and form a basic formula for analysis and discussion of various algorithms. Principally, three signal models have been used in the field of time delay estimation. They categorize in three models of the ideal single-path propagation model, the multipath model, and the reverberation model, respectively.

A signal which is emerging from a remote source and monitored at two spatially separated sensors in the presence of noise can be mathematically modeled as**:**

$$x_1(t) = S(t) + n_1(t) \tag{1}$$
$$x_2(t) = \alpha_1\, S(t + D_1) + n_2(t) \tag{2}$$

In this ***ideal model*** that the effect of reverberation has been ignored, $S(t)$ that is received signal and $n_1(t)$ and $n_2(t)$ which are noise, all are real, jointly stationary random processes. Signal $S(t)$ is presumed to be uncorrelated with noise $n_1(t)$ and $n_2(t)$. There are many applications in which it is of interest to estimate the delay D [2]. Some common techniques of estimating TD will mention bellow.

The ideal propagation model takes only into account the direct-path signal. In many situations, however, each sensor receives multiple delayed and attenuated copies of the source signal due to reflections of the wave front from boundaries and objects in addition to the direct-path signal. This model that is known as ***multipath model*** has been intensively studied in [9, 11, 19, 21]. In this case, the received signals are often described mathematically as:

$$x_n[k] = \sum_{m=1}^{M} \alpha_{nm} S[k - t - \tau_{nm}] + n_n[k] \quad n = 0,1,\ldots,N-1 \qquad (3)$$

In (3) α_{nm} shows the attenuation factor related to the unknown source which received the nth sensor via the mth path, t is the propagation time from the source to sensor 0 via direct path, τ_{nm} is the relative delay between sensor n and sensor 0 for path m with $\tau_{01} = 0$, M is the number of different paths, and $n_n[k]$ is stationary Gaussian noise and assumed to be uncorrelated with both the source signal and the noise signals observed at other sensors.

The multipath model is valid for some but not all environments [8]. In addition, if there are many different paths, that is, M is large, it is difficult to estimate all τ_{nm}'s in (3). This model that is known as ***reverberation model*** has been discussed in [1,13, 16, 20, 14].

In this model, the received signals are expressed as:

$$x_n[k] = h_n[k] * S[k] + n_n[k] \qquad (4)$$

where $*$ denotes convolution, $h_n[k]$ is the channel impulse response between the source and the nth sensor. It has to be mentioned that $S[k]$ is broadband and $n_n[k]$ is uncorrelated with $S[k]$ and the noise signals at other sensors.

In this method, time delay estimation is often achieved in two steps. The first step is to estimate the N channel impulse responses from the source to the N receivers. Once the channel impulse responses are measured, the TDOA information between any two receivers is obtained by identifying the two direct paths. Knowing different model of sound propagation, several common methods of TDE will mention below.

2.1. Cross-Correlation (CC): Cross correlation technique is one of the most commonly used methods to estimate TD. It is also the most straightforward and the earliest developed TDE algorithm, which is formulated based on the single-path propagation model. The cross-correlation between two signals define as:

$$R_{x_1x_2}(\tau) = E[x_1(t)x_2(t-\tau)\,] \qquad (5)$$

where E denotes expectation. The argument τ that maximizes (5) provides an estimate of delay [2].

To estimate TDOA, the autocorrelation of $x_1(t)$ and cross-correlation of $x_1(t)$ and $x_2(t)$ has been calculated. The difference between peaks of these two function considered as TDOA which is so accurate in none reverberation environment. Figure (1) shows computing of TDOA using CC.

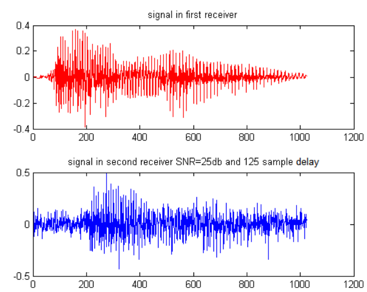

Fig (1.a). Audio siglas recieved by two microphone considering noise and delay.

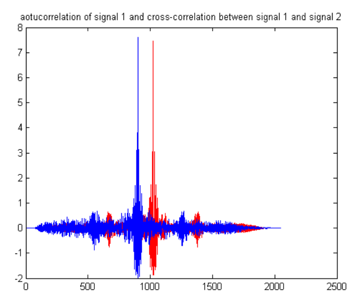

Fig (1.b). Autocorrelation of $x_1(t)$ (left signal), and cross-correlation of $x_1(t)$ and $x_2(t)$ (right signal).

2.2. Generalized cross-correlation method: The generalized cross-correlation (GCC) algorithm can be mentioned as an improved version of the CC method. In this technique some weighting functions (sometimes called a prefilter) are used to give some elements more "weight" or influence on the result than other elements in the same set.

There are some algorithms in the GCC family depending on how the weighting function is selected. Commonly used weighting functions include the constant weighting, the smoothed coherence transform (SCOT) [17], the Roth processor [18], the

Echart filter [5], the phase transform (PHAT), the maximum-likelihood (ML) processor [3], the Hassab-Boucher transform [7], and so forth. Combination of some of these functions is also reported in use [14].

2.3. LMS [1] -type adaptive TDE algorithm: This method, also based on the ideal propagation model with two sensors [6]. It has been intensively investigated in [4, 12, 10, 5]. Different from the cross-correlation-based approaches, this algorithm achieves time delay by minimizing the mean-square error between signal received by reference receiver and a filtered (FIR filter) version of signal received by reference receiver, and the delay estimate is obtained as the lag time associated with the largest component of the FIR filter.

There are another method of estimating TDE which discussed in details in some references [1].

3. Implementation

For Implementation of TDOA estimation and positioning, an almost ideal echoic environment is needed. Therefore this project has been done in a 2×3 room that the walls covered by accusative and the floor covered by a carpet to minimize reverberation. The hardware of this project which will describe latter putted on a scaled table. Almost more than 100 positions of source and microphones have been tested, different algorithms of TDOA estimation and effect of some parameters on accuracy of estimation are considered. Using the results of these tests, the best way of TDOA estimation has been chosen and based on this method time delay between microphones in an array with 4 microphones has been calculated.

However, algorithms based on sequence repetition do better in reverberant places, but they have low speed and sometimes they do not converge. Therefore in on time implementation they cannot be used and in these project correlation algorithms is so preferred.

For surveying the algorithms and effect of parameters, 17 different arrangements considered for four microphones and for each array, the source of sound has been located in 9 different positions. Totally more than 100 sounds were recorded and different algorithms were tested on them. Maximum sampling frequency which has been used in sound cards is 44100 samples each second. Time interval of recording sound is 10 seconds. First 3 seconds are used for calibrating and synchronizing the sound cards and last 7 seconds are used for TDOA estimating.

In hardware implementation, sound cards are used for receiving sound. Below are four reasons why sound cards are used: first of all it can be recognized by Mathlab software. Secondly receiving sound has high quality. Thirdly sampling frequency is high in sound cards and last of all they connect to computer easily.

Easy installation, appropriate cost and compatibility with sound chipset of motherboard are the reason why C-Media 8738 sound card [15] are used in this project, however in [15] Mr. James Scott and Mr. Boris Dragovic did their implementation in Linux but here the project is done with Windows operating system.

In the beginning of sound cards installation, conflict error was happened. When a peripheral wants to access a resource, it sends an interrupt request to the processor in order to get its attention. The peripherals have an interrupt number that is called an IRQ (Interruption Request). Numbers of IRQ lines in CPU are limited, therefore it cannot respond to a limited number of requests in the same time. For solving this problem, some peripheral devices have to disable in BIOS setting and then IRQ lines will assign to PCI slots which the sound cards installs there.

Figure 2: The whole hardware that are used in TDE estimating

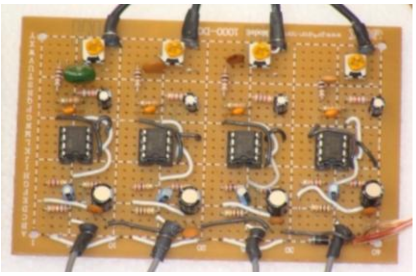

Figure 3: Preamplifiers of capacitor microphones

Hardware used in this project is shown in figure 2. As it can be seen capacitor microphone are used because they are small, so they increase the accuracy of estimation. In addition this kind of microphone is so sensitive and cheap. The only problem with this kind of microphones is that their output signal is so weak. For detecting sound by sound cards, the signal must be at least 23dB, hence after microphones there is a preamplifier for each of them, working based on

1 Least mean squares

low noise operational amplifier (Figure 3). At the end of each preamplifier there is a potentiometer to adjust the amplitude of sound signals. After amplifying the received sound it has been send to the switch.

Other circuits which are used in hardware are driver, multivibrator, switch and sound generator which is used to calculate time delay between cards (Figure 4). When sound cards start receiving the sound, pin 21 of IC (sound card) produce a square wave. This wave enters latch circuit and produce down trigger that runs monostable multivibrator. By entering down trigger, monostable multivibrator produces a pulse with definite time interval. During this time the switch has been activated and sound cards received sound from sound generator (UM66 chipset), therefore the time delay between cards can be calculated. After this time, the circuit receives sound of microphones.

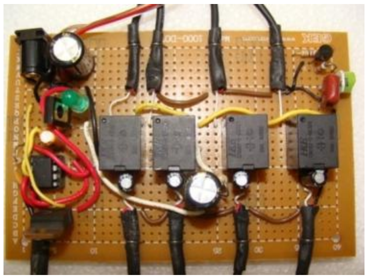

Figure 4: The circuit of multi vibrator and switches

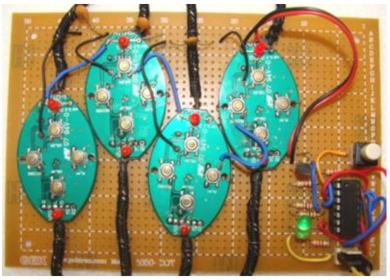

Figure 5: Sound cards using in hardware

Because of small different frequency between crystals of sound cards, in long time interval the delay occurs between cards. In a practical test after receiving 1200000 samples, at most 2 to 10 samples delay have seen between cards. For solving this problem, one side of crystals attached together to equalize frequency of sound cards pulse. The sound cards and the hardware of synchronization are shown in Figure 5.

4. Experimental Results

4.1. Comparison between different weighted generalized cross correlation: Different weighted generalized cross correlation algorithms have been tested in acoustic room mentioned before. Calculating peak of Cross-Correlation (CC) is the basic algorithm using in time delay estimation. For improving this algorithm, it has been weighted by different function and the result is called General Cross-Correlation (GCC). The most important weighted function that are used in time delay estimation are: Maximum Likelihood (ML) with methods which are introduced by Hanon and Tamson, PHAT and SCOT.

Surveying and comparing these algorithms shows PHAT weighted function do better than the other algorithms in room environment. After GCC-PHAT, Maximum Likelihood (ML) weighted function do the best (Figure 6).

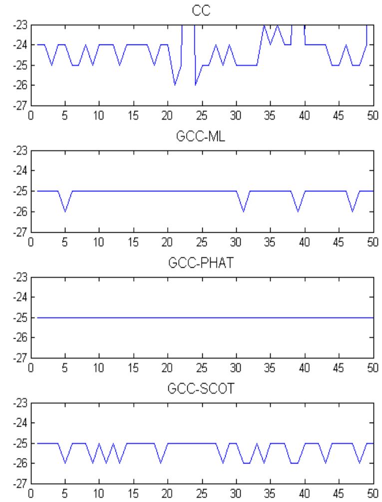

Figure 6: Comparison of TDE using CC and three weighted GCC (ML, PHAT and SCOT)

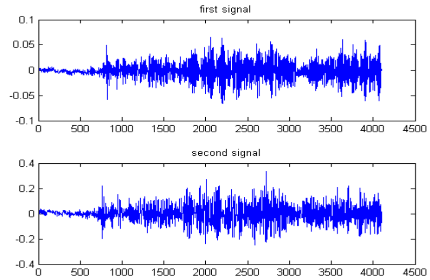

Figure 7-a: A piece of signal received by two microphones

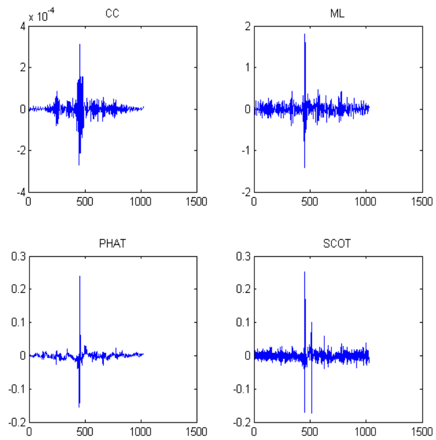

Figure 7-b: Cross-correlation of two signals using different weighted GCC

A test has done on 4096 samples of signals received from two microphones to show why GCC-PHAT do better than the rest of weighted function. In Figure 7 the result of different GCC weighting function is shown.

The peak of GCC is really important in time delay estimation. As indicated in the pictures sharp peaks can be obtained by GCC-PHAT and then GCC-ML. Sharp peaks can be seen in GCC-SCOT as well, but at it has been indicated in the picture there are another peaks except the main one bring about noticeable error in some cases. Totally the sharper main peak as well as the lower other part of GCC means the best noise elimination.

4.2. Effect of length of observation interval on accuracy and speed of estimation: For considering length of observation interval, two 204800 samples of sound that were received by two microphones, has been tested. The received sounds were divided in 400 interval of 512 samples, 200 interval of 1024 samples,100 interval of 2048 samples, 50 interval of 4096 samples and 25 interval of 8192 samples respectively. In each case the error of estimation has been noticed.

As GCC-PHAT has been showed the best result, therefore the effect of observation interval length was considered on this algorithm. The other algorithms showed the same results as well.

The outcomes indicated for observation interval less than 4096 samples, the system has not been working well as in 512 samples interval no definite result was obtained for time delay estimation. Generally it was proved that the least appropriate interval is 4096 samples and the error will decrease by increasing observation interval but, it has to be noticed enlarging the interval bring about reducing in speed of system consequently, some problem will appear for online implementation.

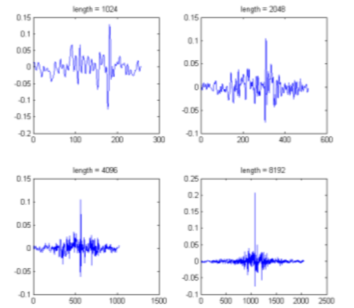

Figure 8: GCC-PHAT for different length of observation interval

With regard to GCC-PHAT in different observation interval, it can be seen that increasing time interval causes sharper peak of cross-correlation function and as cross-correlation is a Statistical function it is obvious enlarging samples can increase accuracy of estimation (Figure 8).

In addition, limited sampling rate to 44100 samples causes 22.68μs accuracy in distinguishing, considering speed of sound, this time is equal to distance of 7.846mm. Very small errors in TDE can cause noticeable errors in positioning. Increasing sampling rate is not possible due to hardware limitation; however, using interpolation has increased estimation accuracy. In this method except the peak value, the previous and latter samples have been considered as well.

4.3. Distance of microphones and sound source and estimation error: In the tests which have done, sound source has located in different location compared to microphones. The results have been showed when sound source is near to microphones (less than 90cm), a noticeable growth in error will happen, and going far from source cause a remarkable decrease in estimation error so that distance of 150cm showed the best result.

4. Conclusion

In this paper different algorithms of TDE have been implemented and investigated. Results show that among weighted GCC, GCC-PATH and ML do the best respectively. Furthermore, the research shows for observation interval less than 4096 samples, system does not work well. On the other hand increasing observation interval which enhances accuracy of estimation is not practical in real time implementation because it reduces the speed of system. To achieve better result with 44100 samples in second interpolation has been used. Based on tests which have done on distance between source and receivers, increasing this distance, leads to more accurate results in which sound propagate as a surface wave.

Acknowledgements

The authors would like to thank the Ista Ertebat Kavoshgar company (Mashhad, Iran), and the Amn Gostare Mooud (AGM) company (Bojnourd, Iran). Part of this work was supported by these two companies.

References:

[1] Jingdong Chen, Jacob Benesty, and Yiteng (Arden) Huang “Time Delay Estimation in Room Acoustic Environments: An Overview”, EURASIP Journal on Applied Signal Processing,Volume 2006, Article ID 26503, Pages 1–19.

[2] Charles H. KNAPP, G. Clifford Carter "The Generalized Correlation Method for Estomation of Time Delay", IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-24, NO. 4, AUGUST 1976.

[3] C. H. Knapp and G. C. Carter, “The generalized correlation method for estimation of time delay,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 24, no. 4, pp. 320–327, 1976.

[4] H. C. So, P. C. Ching, and Y. T. Chan, “A new algorithm for explicit adaptation of time delay,” IEEE Transactions on Signal Processing, vol. 42, no. 7, pp. 1816–1820, 1994.

[5] P. L. Feintuch, N. J. Bershad, and F. A. Reed, “Time delay estimation using the LMS adaptive filter—dynamic behavior,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 29, no. 3, pp. 571–576, 1981.

[6] F. A. Reed, P. L. Feintuch, and N. J. Bershad, “Time delay estimation using the LMS adaptive filter—static behavior,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 29, no. 3, pp. 561–571, 1981.

[7] J. C. Hassab and R. E. Boucher, “Performance of the generalized cross correlator in the presence of a strong spectral peak in the signal,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 29, no. 3, pp. 549–555, 1981.

[8] C. S. Clay and H.Medwin, Acoustical Oceanography, John Wiley& Sons, New York, NY, USA, 1977.

[9] T. G. Manickam, R. J. Vaccaro, and D. W. Tufts, “A leastsquares algorithm for multipath time-delay estimation,” IEEE Transactions on Signal Processing, vol. 42, no. 11, pp. 3229–3233, 1994.

[10] P. C. Ching and Y. T. Chan, “Adaptive time delay estimation with constraints,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 36, no. 4, pp. 599–602, 1988.

[11] J.-J. Fuchs, “Multipath time-delay detection and estimation,” IEEE Transactions on Signal Processing, vol. 47, no. 1, pp. 237–243, 1999.

[12] D. H. Youn, N. Ahmed, and G. C. Carter, “On using the LMS algorithm for time delay estimation,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 30, no. 5, pp. 798–801, 1982.

[13] J. Benesty, “Adaptive eigenvalue decomposition algorithm for passive acoustic source localization,” Journal of the Acoustical Society of America, vol. 107, no. 1, pp. 384–391, 2000.

[14] Y. (Arden) Huang and J. Benesty, “Adaptive multichannel time delay estimation based on blind system identification foracoustic source localization,” in Adaptive Signal Processing—Applications to Real-World Problems, J. Benesty and Y. (Arden) Huang, Eds., chapter 8, pp. 227–248, Springer, Berlin, Germany,2003.

[15] James Scott and Boris Dragovic, “Audio Location: Accurate Low-Cost Location Sensing” Intel Research Cambridge.2006.

[16] S. Doclo and M.Moonen, “Robust adaptive time delay estimation for speaker localization in noisy and reverberant acoustic environments,” EURASIP Journal on Applied Signal Processing,vol. 2003, no. 11, pp. 1110–1124, 2003.

[17] G. C. Carter, A. H. Nuttall, and P. G. Cable, “The smoothed coherence transform,” Proceedings of the IEEE, vol. 61, no. 10, pp. 1497–1498, 1973.

[18] P. R. Roth, “Effective measurements using digital signal analysis,” IEEE Spectrum, vol. 8, no. 4, pp. 62–70, 1971.

[19] P. P. Moghaddam, H. Amindavar, and R. L. Kirlin, “A new time-delay estimation inmultipath,” IEEE Transactions on Signal Processing, vol. 51, no. 5, pp. 1129–1142, 2003.

[20] T. G. Dvorkind and S. Gannot, “Time difference of arrival estimation of speech source in a noisy and reverberant environment,” Signal Processing, vol. 85, no. 1, pp. 177–204, 2005.

[21] J. P. Ianniello, “Large and small error performance limits for multipath time delay estimation,” IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 34, no. 2, pp. 245–251, 1986.

4/13/2013