A Time Delay Estimation Method Based on Wavelet Transform and Speech Envelope for Distributed Microphone Arrays

Zhe CHEN, Shuwen WANG, Fuliang YIN
School of Information and Communication Engineering, Dalian University of Technology, Dalian,116024, China
zhechen@dlut.edu.cn; wsw_1104@yahoo.com.cn; flyin@dlut.edu.cn

***Abstract*—A time delay estimation method based on wavelet transform and speech envelope is proposed for distributed microphone arrays. This method first extracts the speech envelopes of the signals processed with multi-level discrete wavelet transform, and then makes use of the speech envelopes to estimate a coarse time delay. Finally it searches for the accurate time delay near the coarse time delay by the cross-correlation function calculated in time domain. The simulation results illustrate that the proposed method can accurately estimate the time delay between two distributed microphone array signals.**

***Index Terms*—microphone arrays, time of arrival estimation, Wavelet transforms, envelope detectors, speech processing.**

I. INTRODUCTION

In the past decade, the microphone array based speech enhancement, speech separation, sound source localization and speaker tracking techniques have developed rapidly, and have been widely applied to video conference systems, human-machine interface, digital entertainment, robotics and other fields [1-5]. But for the conventional microphone array, the array's geometry information should be known in advance and a regular structure usually is necessary, which results in that changing or moving such an array is difficult. These restrictions lead to great limitations in many applications. To address these problems, the distributed microphone array is developed. Distributed microphone array, also called ad hoc microphone array, is usually a set of microphones which can be arbitrarily placed. According to the recent research work[6,7], a distributed microphone array has the following advantages: the array is easy to establish, with less or even no need for the consideration of its geometry and location information; it has a certain fault tolerance, i.e. the failure of the individual microphone will not lead to the paralysis of the entire array. In recent years, the distributed microphone array gradually becomes a hot topic in the speech processing field.

The speech processing technology based on distributed microphone array is first considered by Aarabi in 2001, and he gives a sound source localization method based on time delay [8,9]. Chen et al present a model of noise and observation error in log domain, and estimates the microphones' and the speaker's positions based on energy-based optimization criterion and the maximum likelihood method [10]. Elahi uses two methods to estimate the sound source location in a distributed microphone array. One method is based on spatial likelihood function and maximum likelihood criteria, and the other method is based on a priori information and the maximum a posteriori criteria [11]. In the applications of microphone array based source localization and speech enhancement, time delay estimation (TDE) is one of the basic problems. The adaptive time delay estimation [12] and the generalized cross-correlation method[13] are two classic time delay estimation methods. For distributed microphone array, most of the current time delay estimation methods are the improved versions of the above two methods. For instance, the method to estimate the time delay in [11,14] is the steered response power with phase transform filter (SRP-PHAT). Ono et al optimize an objective function by minimizing the square errors, in which the iterative update rules are derived by an auxiliary function [15]. Because of the short pitch period of speech signal and the long distance between microphones in a distributed microphone array, the space sampling theorem [16] cannot be met. So the direct application of classic time delay estimation methods will cause the periodic ambiguity [17] in the estimation result.

In order to meet the space sampling theorem and overcome the problem of periodic ambiguity, a low frequency component of the speech signal should be used for time delay estimation. Obviously, the speech envelope is a time-domain signal whose frequency is low enough and it also contains some speech features. So in this paper, the speech envelope is extracted to estimate a coarse time delay first, and then the accurate time delay is searched for within a certain range near the coarse delay. Hilbert transform is often used to extract the signal envelope, but it can only apply to the narrow band random signal [18,19]. For short-term stationary and non-narrow band speech signal, a new solution must be found. The wavelet transform is suitable for non-stationary random signal analysis and without any specific requirement for signal bandwidth, so it can be used for speech envelope estimation. And referring to the envelope definition in Hilbert-Huang transform [20], in this paper the speech envelope is extracted by modulus maxima interpolation approach approximately. Therefore, a method based on wavelet transform and modulus maxima interpolation is used to extract the speech envelope for

This work was supported by the National Natural Science Foundation of China (61172110, 61172107); Specialized Research Fund for the Doctoral Program of Higher Education of China (200801410015); Dalian Municipal Science and Technology Fund Scheme No.2008J23JH025; the Fundamental Research Funds for the Central Universities(DUT13LAB06).

Digital Object Identifier 10.4316/AECE.2013.03007

coarse delay estimation.

The paper is organized as follows. Section II is a brief introduction of wavelet transform and its fast algorithms. The detail of the proposed method is described in Section III. In Section IV, simulation results are discussed. And Section V provides some conclusions.

II. WAVELET TRANSFORM

The wavelet transform decomposes signals over dilated and translated wavelets to analyze signal structures of very different sizes [21]. It is a very flexible and common tool for the local time-frequency analysis.

If $\psi(t)$ is the basic wavelet, the continuous wavelet is

$$\psi_{a,b}(t)=\frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}),\quad a,b\in R,a>0 \tag{1}$$

where a is the dilation factor, and b is the translation factor.

Let $\psi^*(t)$ be the complex conjugate of $\psi(t)$, then the continuous wavelet transform of the signal $s(t)$ can be defined as

$$(W_\psi s)(a,b)=\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty}s(t)\psi^*(\frac{t-b}{a})\mathrm{d}t \tag{2}$$

Restricting parameters a and b of continuous wavelet transform to only discrete values, and taking $a=a_0^j$, $b=kb_0a_0^j$ (a_0>1, b_0>1), then discrete wavelet is obtained as

$$\psi_{j,k}(t)=a_0^{-j/2}\psi(a_0^{-j}t-kb_0) \tag{3}$$

So the discrete wavelet transform is defined as

$$d_{j,k}=\int_{-\infty}^{+\infty}s(t)\psi_{j,k}^*(t)\mathrm{d}t \tag{4}$$

Usually take a_0=2, b_0=1, then the discrete wavelet transform is called dyadic wavelet transform. The dyadic wavelet is expressed as

$$\psi_{j,k}(t)=2^{-j/2}\psi(2^{-j}t-k) \tag{5}$$

For dyadic wavelet transform, Mallat algorithm [22] can be used for fast calculation, and the expression is

$$\begin{cases}c_j(k)=\sum\limits_n h(n-2k)c_{j-1}(n)\\ d_j(k)=\sum\limits_n g(n-2k)c_{j-1}(n)\end{cases},\quad j=1,2,\cdots,J \tag{6}$$

where $c_j(k)$ is the smooth signal of the original signal, $d_j(k)$ is the detail signal of the original signal, $g(n)$ is the unit impulse response of a bandpass filter related to the wavelet function, and $h(n)$ is the unit impulse response of a lowpass filter related to the scale function.

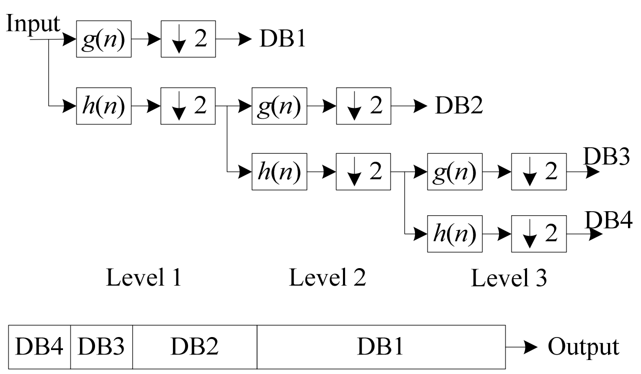

Figure.1 The 3-level wavelet transform using Mallat algorithm

It is noteworthy that in Mallat algorithm, the output of the filter should be twice extracted, and then the next level wavelet transform is conducted to the remaining data, as shown in Fig.1 (taking 3-level wavelet transform as an example). So according to the data form of the output, the signal sequences will have the same number of points before and after wavelet transform.

If the extraction is not conducted in the calculation process, the algorithm converts into à trous wavelet transform [23], as shown in Fig.2. According to à trous algorithm, to ensure the correct result in multi-level wavelet transform without extraction, some adjustments for the filter coefficients are needed before each level wavelet transform. The two sets of filter coefficients for the j-level wavelet transform are

$$h_j(n)=\begin{cases}h_{j-1}(\frac{n}{2}), & n=\text{even}\\ 0, & n=\text{odd}\end{cases} \tag{7}$$

$$g_j(n)=\begin{cases}g_{j-1}(\frac{n}{2}), & n=\text{even}\\ 0, & n=\text{odd}\end{cases} \tag{8}$$

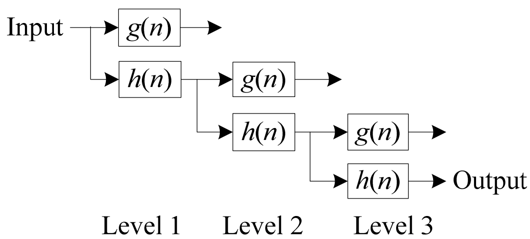

Figure.2 The 3-level wavelet transform using à trous algorithm

In addition, it also ensures that the output signal without extraction has the consistent power with the original signal, thus all the filter coefficients are required to be divided by $\sqrt{2}$ before the first level wavelet transform.

III. A SPEECH ENVELOPE BASED TIME DELAY ESTIMATION METHOD

In order to meet the space sampling theorem and overcome the periodic ambiguity problem in time delay estimation, the speech envelope, a low frequency signal, is used to calculate the cross-correlation function. So to estimate an accurate time delay of speech signals in a distributed microphone array, the following steps are required. First, the speech signals are processed with multi-level discrete wavelet transform, then the modulus maxima connecting lines of the transformed signals are extracted to calculate a coarse time delay, and at last the accurate delay is searched for within a speech signal period near the coarse delay.

In the proposed method, only the low frequency signal after wavelet transform is required for speech envelope estimation, and there is no need for data extraction after each level transform, so the à trous wavelet transform method is adopted. By equations (6), (7) and (8) the J-level discrete wavelet transform can be conducted to the speech signal $s(n)$, and a low frequency signal $s_l(n)$ with mild amplitude variations can be obtained. If the bandwidth of

$s(n)$ is f_{BW}, the bandwidth of $s_J(n)$ is $\frac{f_{BW}}{2^J}$.

After J-level discrete wavelet transform, the transformed signal is still a real signal, so the modulus value of each sampling point is also its absolute value. Search for the modulus maxima of $s_J(n)$ in the time-amplitude plane, then two connecting lines constituted respectively by the maxima and the minima of the signal can be obtained by interpolation method, which can be regard as an envelope approximation of $s_J(n)$, as shown in Fig.3, in which the solid line represents the speech signal and the dotted lines represent signal envelopes.

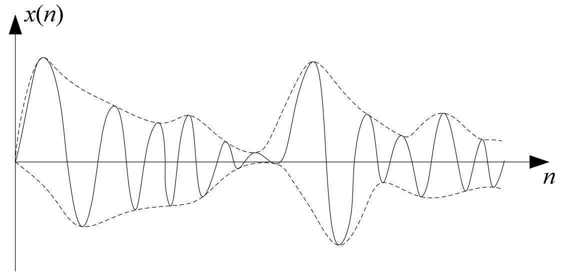

Figure.3 Schematic diagrams of speech signal and its envelopes

Though there are many interpolation methods, in this paper, linear interpolation [24] is adopted for its less computational complexity. For example, when the upper envelope $p(n)$ of signal $s_J(n)$ is extracted, if two adjacent maxima appear respectively in the n_1-th and the n_2-th sampling point, with the corresponding maximum value $p(n_1)$ and $p(n_2)$, and the integer n meets the condition of $n_1<n<n_2$, representing a sampling point between n_1 and n_2, then the value of envelope $p(n)$ corresponding to the n-th sampling point is obtained approximately as

$$p(n) = p(n_2)\frac{n-n_1}{n_2-n_1} + p(n_1)\frac{n_2-n}{n_2-n_1} \tag{9}$$

In the next step, the cross-correlation function of the upper envelopes (or the lower envelopes) of the two speech signals is calculated, and a coarse delay of signals can be obtained according to the positions of the maximum value. Define the upper envelopes of the two speech signals $s_1(n)$ and $s_2(n)$ are $p_1(n)$ and $p_2(n)$ respectively, and their corresponding discrete Fourier transform (DFT) are $P_1(k)$ and $P_2(k)$, i.e.

$$\begin{cases} P_1(k) = \sum_{n=0}^{N-1} p_1(n)\mathrm{e}^{-j\frac{2\pi}{N}nk} \\ P_2(k) = \sum_{n=0}^{N-1} p_2(n)\mathrm{e}^{-j\frac{2\pi}{N}nk} \end{cases} \tag{10}$$

where N is the length of the envelope, also the length of the speech signal.

Thus, the cross-correlation function $r_{p_1p_2}(m)$ of $p_1(n)$ and $p_2(n)$ can be calculated by their cross-power spectrum [25], i.e.

$$r_{p_1p_2}(m) = \frac{1}{N}\sum_{k=0}^{N-1} P_1(k)P_2^*(k)\mathrm{e}^{j\frac{2\pi}{N}mk} \tag{11}$$

where $m = 0, 1, \cdots, N-1$.

Then the maximum value of the cross-correlation $r_{p_1p_2}(m)$ is searched for, and the delay sample $\hat{m}$ corresponding to the maximum value is the lag of $p_1(n)$ relative to $p_2(n)$, i.e. the coarse time delay we desire.

$$\hat{m} = \arg\max_m r_{p_1p_2}(m) \tag{12}$$

After the coarse time delay, i.e., the lag of $p_1(n)$ relative to $p_2(n)$ is calculated as equations (11) and (12), the accurate time delay can be obtained by searching the maximum value position in the cross-correlation function $r_{s_1s_2}(m)$ of the original signals $s_1(n)$ and $s_2(n)$ within a speech signal pitch period near the coarse time delay. Because of the less points involved in calculation, the cross-correlation function $r_{s_1s_2}(m)$ can be calculated directly in time domain by its definition, i.e.

$$r_{s_1s_2}(m) = \sum_{n=0}^{N-1} s_1(n)s_2(n+m) \tag{13}$$

$$\hat{m}' = \arg\max_m r_{s_1s_2}(m) \tag{14}$$

where $\hat{m}'$ is the lag of $s_1(n)$ relative to $s_2(n)$, and N is the length of the speech signal.

To sum up, the calculation process of the proposed method can be summarized as shown in Fig.4.

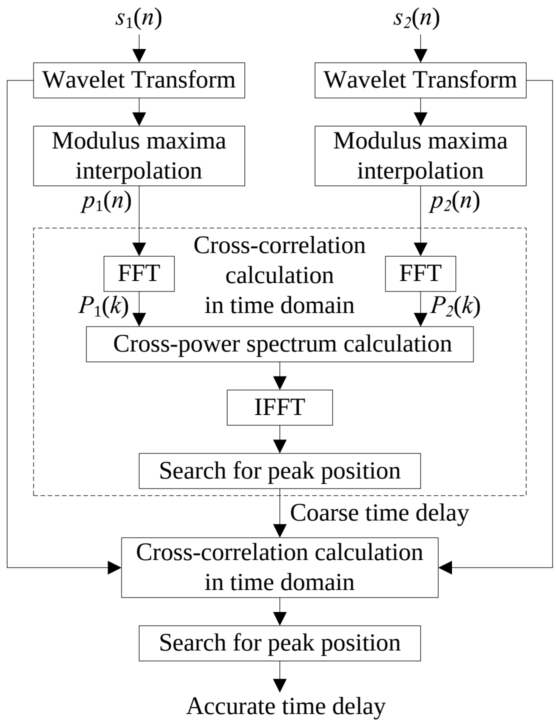

Figure.4 Calculation process of the proposed method

IV. SIMULATIONS AND RESULT DISCUSS

In order to verify the effectiveness of the proposed method, we conduct a set of computer simulations. In these simulations, the third-order Daubechies wavelet is used for signal wavelet transform, and the coefficients of the corresponding lowpass filter and bandpass filter, exported from Matlab, respectively, are

$h(n)$ = [0.03522629188210, -0.08544127388224, -0.13501102001039, 0.45987750211933, 0.806891509313 34, 0.33267055295096]

$g(n)$ = [-0.33267055295096, 0.80689150931334,

-0.45987750211933, -0.13501102001039, 0.085441273882 24, 0.03522629188210]

Three simulations with the same approach and different input signals are conducted under the sampling rate of 48KHz. In these simulations, the 7-level discrete wavelet transform is adopted to filter out the high components of the speech signal and obtain a signal with low frequency. The obtained signal is equivalent to a signal getting through a system formed by seven cascaded lowpass filters, and the total amplitude-frequency response of the cascaded lowpass filters is shown in Fig.5.

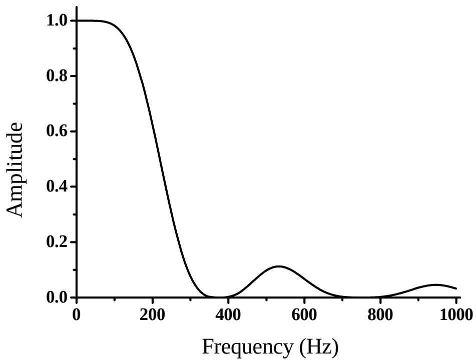

Figure.5 Total amplitude-frequency response at low frequency of 7-level wavelet transforms

A. Time delay estimation of amplitude modulation (AM) signal

Amplitude modulation signal [26] is one of the signals prone to periodic ambiguity. The AM signal selected in this simulation is

$$s(n) = [16384 + 6553.6\sin(4\pi n)]\cos(1600\pi n) \qquad (15)$$

In this simulation, we regard the AM signal $s(n)$ as the received signal of the first microphone (i.e. reference signal), and the delayed signal $s(n\text{-}1440)$ as the received signal of the second microphone, where 1440 samples correspond to 30ms delay or 10.2m distance between two microphones. To exatrct the upper envelope and the lower envelope of each transformed signal, the 7-level discrete wavelet transform is performed on the reference signal and the delayed signal respectively, and then the cross-correlation function of the corresponding envelopes with enough points is calculated to estimate the coarse delay. In general, in order to estimate the delay of the signal correctly, the number of the points required to calculate the cross-correlation function should be at least equivalent to five times of the sampling points of the delay time. For this simulation, taking into account the calculation accuracy of cross-correlation function and that the number of sampling points used for time to frequency transform by fast Fourier transform(FFT) must be an integer power of 2, the data involved in cross-correlation calculation should not be less than 8192 points. After the coarse time delay is obtained, consider the coarse time delay as a center point and search the accurate time delay within a period of carrier wave, i.e. 1.25ms corresponding to the carrier frequency of 800Hz.

Table I shows the estimation results of the coarse delay and the accurate delay of the signals in the cases of different point numbers involved in cross-correlation calculation. It can be seen from Table I that when 4096 or 8192 point data are used to calculate the cross-correlation function, the delay estimation error occur in some frames, but the error frames are fewer when point number is 8192 than that when 4096 points are used. When the point number increases to 16384 or more, time delay of AM signals can be accurately estimated by the proposed method without any periodic ambiguity problem in simulation results.

TABLE I. ESTIMATION RESULTS OF COARSE DELAY AND ACCURATE DELAY OF AM SIGNALS (UNIT: MS)

Number of points	Frame number	Using upper envelope		Using lower envelope	
		Coarse time delay	Accurate time delay	Coarse time delay	Accurate time delay
4096	1	30.00	30.00	30.00	30.00
	2	30.00	30.00	30.00	30.00
	3	30.00	30.00	30.00	30.00
	4	2.50	2.50	0.00	0.00
	5	0.00	0.00	-3.75	-3.75
8192	1	30.00	30.00	30.00	30.00
	2	30.00	30.00	30.00	30.00
	3	-1.25	-1.25	-3.75	-3.75
	4	30.00	30.00	30.00	30.00
	5	30.00	30.00	30.00	30.00
16384	1	30.00	30.00	30.00	30.00
	2	30.00	30.00	30.00	30.00
	3	30.00	30.00	30.00	30.00
	4	30.00	30.00	30.00	30.00
	5	30.00	30.00	30.00	30.00

B. Time delay estimation of clean speech signal

Select a segment of clean male speech as the received signal of the first microphone (i.e. reference signal), and the speech delayed with 30ms as the received signal of the second microphone.

TABLE II. ESTIMATION RESULTS OF COARSE DELAY AND ACCURATE DELAY OF CLEAN SPEECH SIGNALS (UNIT: MS)

Number of points	Frame number	Using upper envelope		Using lower envelope	
		Coarse time delay	Accurate time delay	Coarse time delay	Accurate time delay
8192	1	30.00	30.00	30.00	30.00
	2	22.44	22.52	29.94	30.00
	3	-11.14	-10.17	5.10	6.08
	4	29.77	30.00	29.88	30.00
	5	30.00	30.00	30.00	30.00
16384	1	22.46	22.52	29.94	30.00
	2	29.96	30.00	29.98	30.00
	3	30.00	30.00	29.98	30.00
	4	30.00	30.00	30.00	30.00
	5	29.98	30.00	29.98	30.00
32768	1	30.00	30.00	30.00	30.00
	2	30.00	30.00	30.00	30.00
	3	30.00	30.00	30.00	30.00
	4	30.00	30.00	30.00	30.00
	5	30.00	30.00	30.00	30.00

Since common pitch frequency ranges from 50 to 500Hz, 2ms, equivalent to 96 sampling points, is taken as the search range of accurate delay.

The schematic diagrams of the reference signal and the delayed signal, the corresponding transformed signals of them after 7-level wavelet transform, and the extracted upper and lower envelopes of the two signals in this simulation are shown in Fig.6. Table II gives the estimation results of the coarse delay and the accurate delay of the clean speech signals with different point numbers in cross-correlation calcultion. As shown in Table II, when 16384 point data are used to calculate the cross-correlation function in frequecny domain, there are still some estimation error occuring in some frames, but the total number of the error frames is few. And accurate delay estimation result can be obtained when the 32768 sampling points are used to calculate the cross-correlation function.

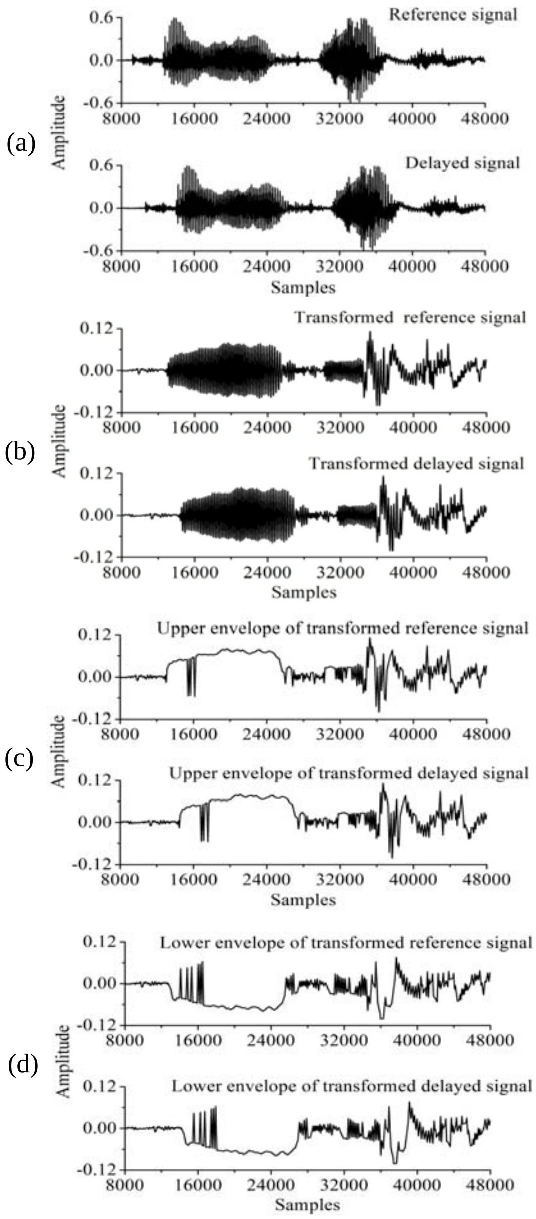

Figure.6 Schematic diagrams of clean speech signals: (a) the reference signal and the delayed signal; (b) the transformed reference signal and the transformed delayed signal after 7-level wavelet transform; (c) the extracted upper envelopes of the two signals; (d) the extracted lower envelopes of the two signals

C. Time delay estimation of noisy speech signal

Add white noises, which are irrelevant to each other and irrelevant to each speech signal, respectively, to the reference signal and the delayed signal used in simulation B, meeting the condition of SNR = 10dB. Then take them as the received signals of the two microphones. Process the two signals with 7-level wavelet transform and estimate the coarse delay using the extracted envelopes. The range for accurate delay searching is also 2ms (equivalent to 96 sampling points).

The schematic diagrams of the reference signal and the delayed signal, the corresponding transformed signals of them after 7-level wavelet transform, and the extracted upper and lower envelopes of the two signals in this simulation are shown in Fig.7. Table III shows the estimation results of coarse delay and accurate delay of the signals in the cases of different point numbers.

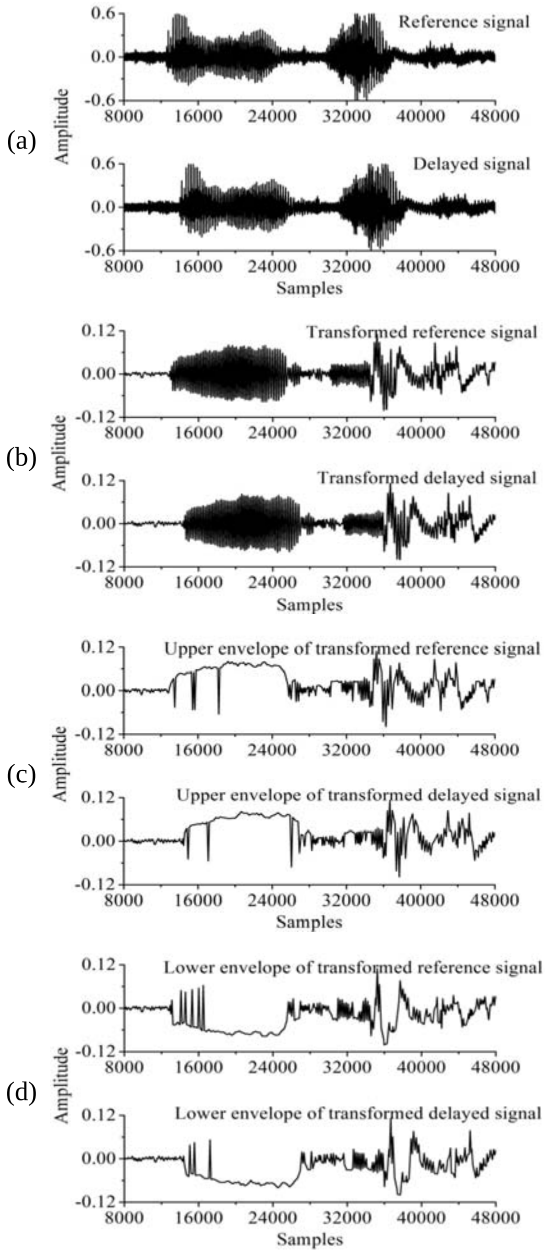

Figure.7 Schematic diagrams of noisy speech signals: (a) the reference signal and the delayed signal; (b) the transformed reference signal and the transformed delayed signal after 7-level wavelet transform; (c) the extracted upper envelopes of the two signals; (d) the extracted lower envelopes of the two signals

The existing noise in speech signal further increases the difficulty of time delay estimation. Data in Table III show that the accurate estimation results could not be obtained until 65536 or more sampling points are used for cross-correlation calculation.

TABLE III. ESTIMATION RESULTS OF COARSE DELAY AND ACCURATE DELAY OF NOISY SPEECH SIGNALS (UNIT: MS)

Number of points	Frame number	Using upper envelope		Using lower envelope	
		Coarse time delay	Accurate time delay	Coarse time delay	Accurate time delay
16384	1	29.85	30.00	19.98	20.02
	2	29.81	30.00	29.88	30.00
	3	29.98	30.00	29.90	30.00
	4	29.94	30.00	29.67	30.00
	5	30.13	30.00	30.17	30.00
32768	1	29.98	30.00	25.25	25.06
	2	29.96	30.00	29.88	30.00
	3	30.02	30.00	30.06	30.00
	4	30.02	30.00	29.96	30.00
	5	29.96	30.00	30.19	30.00
65536	1	29.96	30.00	29.88	30.00
	2	30.02	30.00	30.00	30.00
	3	30.00	30.00	30.13	30.00
	4	30.00	30.00	30.02	30.00
	5	30.02	30.00	30.02	30.00

Comparing Table III with Table II, it can be seen that to achieve substantially the same accuracy, there are more sampling points needed to calculate the cross-correlation function of noisy specch siganls than that as for clean speech signals. Despite the influence of noise, the more sampling points are used, the more accurate result of time delay estimation is obtained.

V. CONCLUSIONS

A time delay estimation method based on wavelet transform and speech envelope for distributed microphone array is proposed in this paper. This method first processes the signals with multi-level discrete wavelet transform and extracts the speech envelopes by modulus maxima interpolation. Next, a coarse time delay is estimated using the speech envelopes, and then the accurate time delay is searched within a certain range near the coarse time delay. The proposed method can avoid the periodic ambiguity problem in time delay estimation for distributed microphone arrays. And with the increase of the points involved in cross-correlation calculation, the estimated coarse time delay is more and more accurate, so that it is easier to search the accurate delay. However, no extraction operation is conducted after wavelet transform in this method, so there are still some redundant data in transformed signal which can be eliminated or utilized in further study.

REFERENCES

[1] M. Mukul, R. Prasad, M. Choudhary, et al, “Steering of camera by stepper motor towards active speaker using microphone array,” The Society of Instrument and Control Engineers (SICE) Annual Conference, Tokyo, Japan, 2008, pp.19-24.
[2] H. Buchner, W. Kellermann, “An acoustic human-machine interface with multi-channel sound reproduction,” IEEE Fourth Workshop on Multimedia Signal Processing, Cannes, France, 2001, pp.359-364.
[3] B. Mrazovac, M. Bjelica, I. Papp, et al, “Smart audio/video playback control based on presence detection and user localization in home environment,” The 2nd Eastern European Regional Conference on the Engineering of Computer Based Systems (ECBS-EERC), Bratislava, Slovakia, 2011, pp.44-53.
[4] Y. Sasaki, Y. Tamai, S. Kagami, et al, “2D sound source localization on a mobile robot with a concentric microphone array,” IEEE International Conference on Systems, Man and Cybernetics, Hawaii, USA, 2005, Vol.4, pp. 3528-3533.
[5] Y. Jia, Y. Luo, Y. Lin, “Distributed microphone arrays for digital home and office,” IEEE International Conference on Acoustics, Speech and Signal Processing, Toulouse, France, 2006, pp.V1065-V1068.
[6] H. Noguchi, T. Takagi, K. Kugata, et al, “Low-traffic and low-power data-intensive sound acquisition with perfect aggregation specialized for microphone array networks,” The Fourth International Conference on Sensor Technologies and Applications, Venice, Italy, 2010, pp.157-162.
[7] T. Damarla, L. M. Kaplan, G. T. Whipps, “Sniper localization using acoustic asynchronous sensors,” IEEE Sensors Journal, 2010, Vol. 10, No. 9, pp.1469-1478.
[8] P. Aarabi, The Integration and Localization of Distributed Sensor Arrays, Ph.D. Thesis, Stanford University, USA, May 2001.
[9] P. Aarabi, “Self-localizing dynamic microphone arrays,” IEEE Trans. on System, Man, and Cybernetics, 2002, Vol.32, No.4, pp.474-484.
[10] M. Chen, Z. Liu, L. He, et al, “Energy-based position estimation of microphones and speakers for ad hoc microphone arrays,” IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York, USA, October 2007, pp.22-25.
[11] E. Elahi, “Sound localization and tracking using distributed microphones fusion: maximum likelihood or maximum a-posteriori approach,” IEEE International Conference on Computer, Control and Communication, Karachi, Pakistan, 2009, pp.1-6.
[12] D. H. Youn, N. Ahmed, G C Carter, “On using the LMS algorithm for time delay estimation,” IEEE Trans. on Acoustics, Speech, and Signal Processing, 1982, Vol.30, No.5, pp.798-801.
[13] C. Knapp, G. Carter, “The generalized correlation method for estimation of time delay,” IEEE Trans. on Acoustics, Speech, and Signal Processing, 1976, Vol.ASSP-24, No.4, pp.320-327.
[14] C. Zhang, D. Florêncio, D. E. Ba, et al, “Maximum likelihood sound source localization and beamforming for directional microphone arrays in distributed meetings,” IEEE Trans. on Multimedia, 2008, Vol.10, No.3, pp. 538-548.
[15] N. Ono, H. Kohno, N. Ito, et al, “Blind alignment of asynchronously recorded signals for distributed microphone array,” IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, USA, 2009, pp.161-164.
[16] A. V. Oppenheim, R. W. Schafer, Discrete-Time Signal Processing (3rd Edition). New Jersey, US: Pearson Education, 2009.
[17] J. S. Picard, A. J. Weiss, “Localization based on periodic signals with ambiguity,” IEEE 26-th Convention of Electrical and Electronics Engineers in Israel, Eliat, Israel, 2010, pp.1007-1011.
[18] J. Y. Lee, J. K. Kim, G. Yoon, “A digital envelope detection filter for blood pressure measurement,” The 23rd Annual EMBS International Conference, Istanbul, Turkey, 2001, pp.226-228
[19] M. A. Poletti, “The homomorphic analytic signal,” IEEE Trans. on Signal Processing, 1997, Vol. 45, No. 8, pp.1943-1953.
[20] N. E. Huang, Z. Shen, S. R Long, et al, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” Proceedings of the Royal Society of London, Series A, 1998, Vol.454, No.1971, pp. 903-995.
[21] S. Mallat, A Wavelet Tour of Signal Processing (3rd Edition). San Diego, CA, USA: Academic Press, 2008.
[22] I. Daubechies, Ten Lectures on Wavelets. Philadelphia, USA: SIAM (Society of Industrial and Applied Mathematics), 1992.
[23] M. J. Shensa, “The discrete wavelet transform: wedding the à trous and Mallat algorithms,” IEEE Trans. on Signal Processing, 1992, Vol.40, No.10, pp.2464-2482.
[24] P. S. Hagan, G. West, “Interpolation methods for curve construction,” Applied Mathematical Finance, 2006, Vol.13, No.2, pp.89-129.
[25] N. V. Thakor, J. G. Webster, W. J. Tompkins, “Estimation of QRS complex power spectra for design of a QRS filter,” IEEE Trans. on Biomedical Engineering, 1984, Vol. BME-31, No. 11, pp. 702-706.
[26] T. S. Rappaport, Wireless Communications: Principles and Practice (Second Edition). New Jersey, U.S.A.: Prentice Hall, 2001.