Sound Source Localization Based on Time Delay Estimation Using Variable Step Size in LMS Adaptive Filter

Achyutha Ramarao Maddukuri, Fuliang Wang
School of Engineering
Blekinge Institute of Technology
Karlskrona, Sweden
(acma16, fuwa15)@student.bth.se

Wlodek J.Kulesza
School of Engineering
Blekinge Institute of Technology
Karlskrona, Sweden
wka@bth.se

***Abstract*—Time delay estimation is one of the key methods in sound source localization and LMS adaptive filter is one of the popular time delay estimation methods. However the time delay estimation based on LMS adaptive filter has to sacrifice the Real-time performance to improve the positioning accuracy in sound source localization. Compared with the LMS adaptive filter, a variable step size LMS adaptive filter applied on a time delay estimation system is presented in this paper for sound source localization. The results show that time delay estimation based on variable step size LMS adaptive filter has better real-time performance and position accuracy in sound source localization.**

***Index Terms*—least mean square, sound source localization, time delay estimation, variable step size**

I. Introduction

Sound source localization is an indispensable technology in the field of voice communications such as video conferencing, hand-free mobile telephony and so on [1]. Time delay estimation (TDE) is one of the key technologies in sound source localization. The aim of TDE is to estimate the time delays caused by the distance difference between the sound source and different microphones. The sound Source can be estimated by time delays according to the model of microphones.

LMS adaptive filter is one of the popular TDE methods. However, TDE based on LMS adaptive filter has to sacrifice the real-time performance to improve the positioning accuracy.

Instead of that this paper present a TDE method based on a variable step size (VSS) LMS algorithm applied on a three-dimensional microphone model. Simulation results illustrate the new method has better real-time performance and positioning accuracy in sound source localization.

II. Survey of Related Works

The initial widely used TDE methods are generalized cross-correlation (GCC) and LMS adaptive filter. Literature [2] and [3] present the subspace based eigenvalue decomposition (EVD) and acoustic transfer function ratio (ATF-s ratio) methods respectively. Literature [4-5] described the features of different methods by comparison of them. LMS adaptive filter is still popular because it does not have to know the statistical characteristics of the input signals.

III. Problem Statement and Main Contribution

Literature [6] presented a VSS algorithm applied in LMS adaptive filter, but it hasn't been used for TDE in sound source localization. The objective of this paper is to find out whether the new method consisting of the VSS improves the real-time and positioning accuracy in the sound source localization.

The TDE system based on VSS adaptive filter and a three-dimensional microphone array system are modeled and implemented by MATLAB. Finally simulation results verify the real-time performance and positioning accuracy respectively.

IV. Problem Solution

A. Modeling

1) Three-dimensional microphone array system: The model of three-dimensional microphone array is shown in Fig 1.

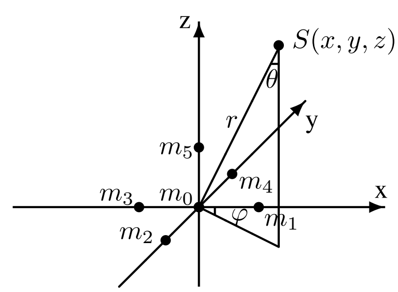


Fig. 1. Hexahydric microphone array model, where $m_0\_m_5$ denote six microphones, r is the distance between sound source $S(x,y,z)$ and reference microphone m_0, θ and φ represent pitch angle and azimuth angle respectively.

Sound source can be estimated by time delays according to the system which is easy to model mathematically.

2) Time delay estimation system: Time delays t_{i0} can be obtained by VSS-LMS adaptive filter as the Fig.2 shown.

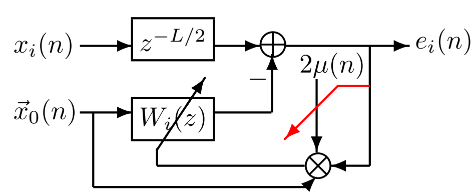


Fig. 2. Time delay estimation based on VSS-LMS adaptive filter, where L is the filter order, $\vec{x}_0(n)$ is the input signal from reference microphone m_0 and $x_i(n), i=1,2,...,5$ are desired signals from remaining microphones.

$$\begin{cases} \vec{w}_i(n+1) = \vec{w}_i(n) - 2\mu(n)e_i(n)\vec{x}_0(n) \\ \mu(n) = \beta\{1 - exp[-\alpha|e_i(n)|^2]\} \end{cases} \tag{1}$$

where $\vec{w}_i(n)$ is filter weight for the desired signal $x_i(n)$, α and β represent two coefficients for updating the step size $\mu(n)$.

Since the derivation of the phase function $\phi_i(w)$ of the filter represent the delay of samples, the time delays can be estimated as (2) shown.

$$t_{i0} = \frac{1}{f_s}(L/2 + \frac{d\phi_i(w)}{dw}) \tag{2}$$

where f_s is the sampling frequency and t_{i0} denotes time delay between $x_i(n)$ and $x_0(n)$.

B. Implementation

The implementation is done by MATLAB script as below.

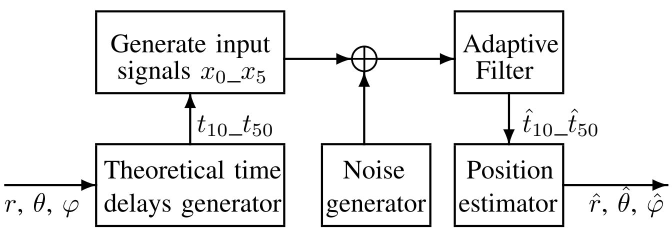


Fig. 3. Block diagram of the whole system

It can be divided into four steps as the following procedure.
1) Input the parameters of actual position r,θ,φ and generate theoretical time delays $t_{10}\_t_{50}$.
2) Generate theoretical received signals $x_0\_x_5$ using theoretical time delays $t_{10}\_t_{50}$.
3) Apply the received signals into adaptive filter using both LMS algorithm and VSS-LMS algorithm after adding noise to get the estimated time delays $\hat{t}_{10}\_\hat{t}_{50}$.
4) Using estimated time delays $\hat{t}_{10}\_\hat{t}_{50}$ to obtain the estimated sound source position $\hat{r}$, $\hat{\theta}$, $\hat{\varphi}$.

V. Simulation Results

A. Real Time Performance

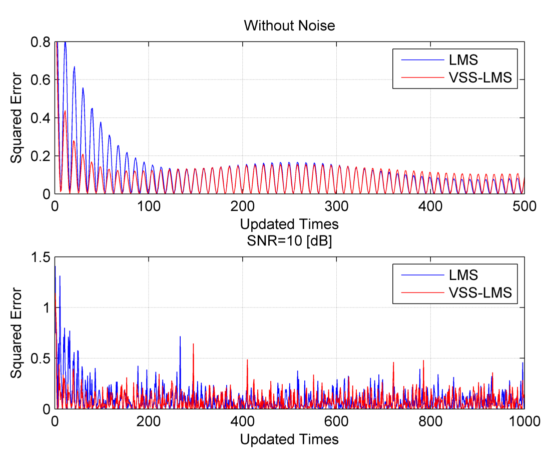


Fig. 4. Convergence speeds of both methods, where fixed step size for LMS $\mu = 0.0002$, variable step size coefficients $\alpha = 0.75$ and $\beta = 0.0025$, input position $P(r, \theta, \varphi) = (50m, 60°, 120°)$, filter length $L = 320$ and sampling frequency $f_s = 10000Hz$.

The real-time performance is observed by the convergence speed of both methods as shown in Fig 4.

B. Positioning Accuracy

Input three positions $P1$, $P2$ and $P3$, the estimation of positions by both methods is shown in Fig 5.

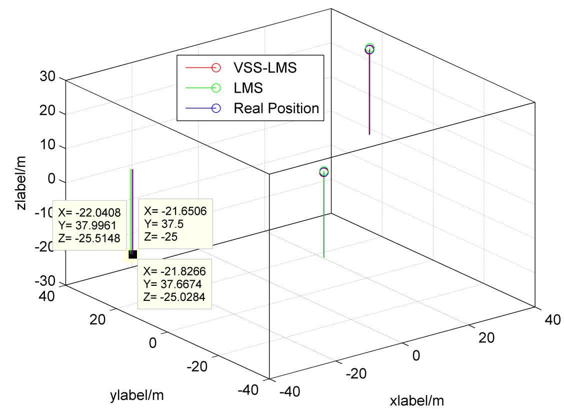


Fig. 5. Position estimation of both methods, where $P1 = (50m, 60°, 30°)$, $P2 = (50m, 60°, -120°)$, $P3 = (50m, 120°, 120°)$, SNR=10 dB and other parameters are the same as Fig 4

VI. Conclusion

As Fig 4 shown, TDE based on VSS-LMS algorithm converges faster than that based on LMS algorithm with and without noise. That means the new method has better real-time performance in sound source localization. On the other hand, positions estimated by the new method consisting of VSS-LMS algorithm has better accuracy comparing to LMS algorithm according to Fig 5.

However, the new method applied on sound source localization hasn't been experimented on an actual device. Furthermore, this paper doesn't consider aliasing noise and the influence of the microphone array model. These problems discussed above can be researched in the future.

References

[1] J. G. Ryan and R. A. Goubran, Application of near-field optimum microphone arrays to hands-free mobile telephony, IEEE Transactions on Vehicular Technology, vol. 52, no. 2, pp. 390400, Mar. 2003.

[2] S. Doclo and M. Moonen, Robust time-delay estimation in highly adverse acoustic environments, in Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the, 2001, pp. 5962.

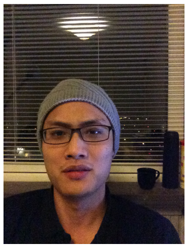
[3] T. G. Dvorkind and S. Gannot, Time difference of arrival estimation of speech source in a noisy and reverberant environment, Signal Processing, vol. 85, no. 1, pp. 177204, Jan. 2005.

[4] C. Weiwei, C. Zhigang, and W. Jianqiang, Time delay estimation techniques in source location, Journal of Data Acquisition Processing, vol. 22, no. 1, pp. 9099, 2007.

[5] S. Dhull, S. Arya, and O. P. Sahu, Comparison of time-delay estimation techniques in acoustic environment, International Journal of Computer Applications, vol. 8, no. 9, pp. 2931, 2010.

[6] Y. Qian, A new variable step size algorithm applied in LMS adaptive signal processing, in 2016 Chinese Control and Decision Conference (CCDC), 2016, pp. 43264329.

Biographies



Fuliang Wang was born in Hangzhou City, Zhejiang province, China. He received the B.Sc in electrical engineering and automation from Zhejiang University of Technology, China. He is pursuing the M.Sc degree in electrical engineering with emphasis on signal processing at Blekinge Institute of Technology, Sweden.

His current research interests include automatic control, speech enhancement and pattern recognition.

Mr.Wang has is the member of Know About Business which is the project instituted by international labor organization.

Maddukuri Achyutha Ramarao was born in Pedanadipalli, Andhra Pradesh, 1995. He received B.Sc. in Electrical and communication engineering from Jawaharlal Nehru Technological University, Kakinada, Andhra Pradesh, India, in 2015.

He is currently pursuing the M.sc degree in electrical engineering at Blekinge Institute of Technology, Karlskrona

He was awarded as certification of merit from Srishti a workshop on mobile making organized by Horizon held at Visakhapatnam, Andhra Pradesh, India, in December 2012. He was the member of IETE workshop held at JNTUK, Kakinada in 2014.

Wlodek J. Kulesza received the M.Sc. and the Ph.D. degrees from Lodz University of Technology, Poland, and a docent degree from Linkping University, Sweden. In 2001 he became Professor in Measurement Science at the University of Kalmar, Sweden. Since 2005 he has held a Professor position at the Blekinge Institute of Technology, Sweden.

His current research interests are multi-sensor systems and wireless sensor networks.

Prof. Kulesza has been IEEE member since 1995. iversity, Sweden.

Appendix

A. The Mathematical Model of Microphone array system

$$\begin{cases} x^2+y^2+z^2=r^2 \\ (x-l)^2+y^2+z^2=(r+t_{10}c)^2 \\ x^2+(y-l)^2+z^2=(r+t_{20}c)^2 \\ (x+l)^2+y^2+z^2=(r+t_{30}c)^2 \\ x^2+(y+l)^2+z^2=(r+t_{40}c)^2 \\ x^2+y^2+(z-l)^2=(r+t_{50}c)^2 \end{cases} \qquad \text{(A-1)}$$

where r is the distance between sound source $P(x, y, z)$ and the reference microphone m_0, and l is the distance between other microphones m_i and sersor m_0, t_{i0} represents the time delay between m_i and m_0, and r is the sound speed.

B. The Table of Positioning Estimation

TABLE A-I

POSITIONING ESTIMATION

$P(r, \theta, \varphi)$	Estimated using LMS	Estimated using VSS-LMS
$(50, 60, 30)$	$(50.40, 59.81, 30.13)$	$(49.86, 60.11, 30.11)$
$(50, 60, -120°)$	$(50.62, 59.78, -120.09)$	$(49.81, 60.14, -120.09)$
$(50, 120, 120)$	$(50.22, 120.16$, $120.07)$	$(50.35, 119.97, 120.06$