A Fast and Accurate Sound Source Localization Method Using the Optimal Combination of SRP and TDOA Methodologies

Mohammad Ranjkesh
Department of Electrical Engineering, University Of Guilan, Rasht, Iran
m.ranjkesh@khi.ac.ir
Reza Hasanzadeh*
Department of Electrical Engineering, University Of Guilan, Rasht, Iran
hasanzadehpak@guilan.ac.ir

Abstract

This paper presents an automatic sound source localization approach based on a combination of the basic time delay estimation sub-methods namely, Time Difference of Arrival (TDOA), and Steered Response Power (SRP) methods. The TDOA method is a fast but vulnerable approach for finding the sound source location in long distances and reverberant environments and is so sensitive in noisy situations. On the other hand, the conventional SRP method is time consuming, but a successful approach to accurately find sound source location in noisy and reverberant environments. Also, another SRP-based method, SRP Phase Transform (SRP-PHAT), has been suggested for the better noise robustness and more accuracy of sound source localization. In this paper, based on the combination of TDOA and SRP based methods, two approaches were proposed for sound source localization. In the first proposed approach called Classical TDOA-SRP, the TDOA method is used to find the approximate sound source direction and then SRP based methods were used to find the accurate location of sound source in the Field of View (FOV) which is obtained by the TDOA method. In the second proposed approach which called Optimal TDOA-SRP, for more reduction of computational processing time of SRP-based methods and better noise robustness, a new criterion has been proposed for finding the effective FOV which is obtained through the TDOA method. Experiments were carried out under different conditions confirming the validity of the purposed approaches.

Keywords: Steered Response Power; Time Delay Estimation; Steered Response Power Phase Transform; Sound Source Localization; Time Difference Of Arrival; Field Of View.

1. Introduction

Distributed microphone systems have been considered for various applications including human computer/machine interfaces, talker tracking, Robotic domain and beam-forming for signal-to-noise ratio (SNR) enhancements [1,2]. Many of these applications require detecting and localizing the sound sources. Therefore, proposed methods for sound source localization problems with distributed microphone arrays are usually very important. In some practical sound source localization (SSL) applications, the source should be automatically detected for computer driven analyses of the auditory scene [1]. SSL algorithms can be broadly divided into indirect and direct schemes [3]. Indirect algorithms usually follow a two-step procedure. In the first step, the time delay of arrival between each microphone pairs is computed and in the second step, they estimate sound source position based on the estimated delay and the geometry of arrays. The direct algorithm performs time delay of arrival and sound source location estimations in one single step by scanning a set of candidate source locations and selecting the most likely position as an estimated sound source location [4,5]. There are several algorithms for SSL applications categorized in a similar manner. But, two most successful and recently proposed methods which are well-known as Steered Response Power (SRP) and Time Difference of Arrival (TDOA) have been considered in the recent years for direct and indirect approaches, respectively [5,6]. The basic principle of SRP methodology is based on the filter-and-sum (delay-and-sum) beam-forming operation, which leads to noise power reduction proportional to the number of uncorrelated microphone channels used in the operation [6,7]. Although SRP methods have been used properly for applications such as intrusion detection and gunfire location, this kind of SSL method is time consuming which makes real time applications inappropriate [8]. On the other hand, TDOA is another popular SSL method, which is more appropriate for practical and real time applications [9]. This method is nonlinear in its nature, but it has significant computational advantages over any other SSL methods. However, this kind of SSL methods is only able to estimate the direction of the corresponding sound source location in long distances. Therefore, this problem makes TDOA method inappropriate for applications in which the precise detection of the SSL is necessary [8,10].

In this paper, a combination approach has been proposed to estimate the sound source direction using a

---

basic TDOA method and then SRP method has been used to find the final sound source location in the estimated direction.

The experimental results in this paper show that because of the pre-estimation of the sound source direction in the proposed methods, we have a valuable reduction of the computational time and more noise robustness relative to the conventional SSL methods.
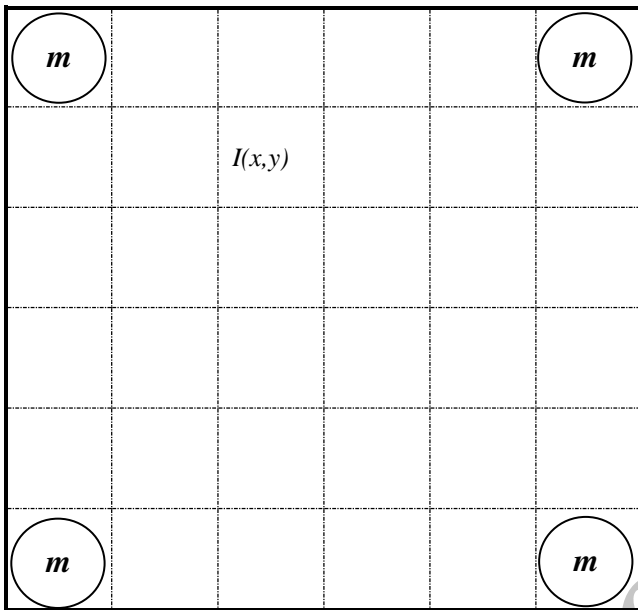

Fig .1. FOV for SRP setup with four microphones. *I(x, y)* is a typical grid point

2. Steered Response Power (SRP) Method

The SRP methods use sound's power and create a SRP image to show the sound source location. The SRP method can be affected by different type of uncorrelated and correlated noises [6]. The uncorrelated noise typically results from the independent noise on each microphone channel and the correlated noise, on the other hand, results from coherent noise sources such as sources outside the Field of View (FOV), multiple targets and reverberations [6]. In the SRP method, the correlated noise creates greater challenges for beam-forming compared to the uncorrelated noise [6] and will be used in the experimental results of this paper. In order to reduce the impact of noise on the sound source location estimation, several filters for the SRP method have been proposed for improving performance, such as Maximum Likelihood (ML)[11], Smooth Coherence Transforms (SCOT)[12], Phase Transform (PHAT)[13] and the Roth Processor[14]. The experimental results show that PHAT has a better performance than others in noisy and reverberant environment [15].

2.1 Mathematical Methodology of SRP

Fig. 1 shows a simple fundamental structure of SRP methods in 2-dimentional case such that Sound source *(I)* and microphones *(m)* are at the same *Z* coordinate. In the

SRP method, a microphone array is used to make the beam-form for each point in the FOV [16]. For each grid point of interest, the SRP delays each microphone signal to result in a coherent addition for a sound source traveling from the point of interest. For each point in the region of interest, the received signals are delayed accordingly and summed together coherently, and finally the power of each point in the region of interest is computed, respectively. The detection and location of the sound source is based on value of the estimated power at each point. Also, the power estimation maybe corrupted by noise sources, reverberation and the finite distributions of microphones [15].

As shown in fig. 1, for finding the location of sound source, it can be assumed that the FOV is formed as grid points i.e. *I(x,y)*. By defining a 2-dimensional FOV (assuming that the sound source and microphones are in the same horizontal place e.g. *xy* plane) and *N* microphones and also considering the output from *q*'th microphone is $m_q(t)$, the SRP at the spatial point *X=[x,y]* for a time frame *n* of length *L* can be defined as

$$P_n(X) = \int_{nL}^{(n+1)L} |\sum_{q=1}^{N} m_q(t - \tau(X,q))|^2 dt \qquad (1)$$

In this equation, $\tau(X,q)$ is the direct time of travel from location *X* to microphone *q*. In [17], it is shown that the SRP can be computed by summing the General Cross-Correlation (GCC) for all possible pairs of the set of microphones. The GCC for a microphone pair *(k,l)* is computed as

$$R_{m_k m_l}(\tau) = \int_{-\infty}^{+\infty} M_k(w) M_l^*(w) e^{jw\tau} dw \qquad (2)$$

where τ is the time lag, $*$ denotes complex conjugation, and $M_l(w)$, $M_k(w)$ are the Fourier transform of the microphone signals $m_l(t)$, $m_k(t)$, respectively.

Taking into account the symmetries involved in the computation of (1) and removing some fixed energy term, the part of $P_n(X)$ that changes with *X* is isolated as[5]

$$P_n'(X) = \sum_{k=1}^{N} \sum_{l=k+1}^{N} R_{m_k m_l}(\tau_{kl}(X)) \qquad (3)$$

where $\tau_{kl}(X)$ is the microphone time delay function of each pair is given by,

$$\tau_{kl}(X) \equiv \frac{\| X - X_k \| - \| X - X_l \|}{c} \qquad (4)$$

where X_k, X_l are the microphone locations and *c* is the speed of sound which is calculated by[21],

$$c = 20.05\sqrt{273.15 + T} \qquad (5)$$

In this equation, *c* is sound's propagation speed in *(m/s)* and *T* is environmental temperature (°C).

In the SRP method, $P_n'(X)$ is evaluated on the FOV to find the sound source location, X_s which provides the maximum value [5,7,17].

$$X_s = \arg\max P_n'(X), X \in FOV \qquad (6)$$

2.2 SRP Phase Transform (SRP-PHAT) Method

The basic principle of SRP-PHAT is similar to SRP method, but in this method, a weighting function has been used to increase the accuracy of finding the sound source delays beside the advantage of its simplicity in the implementation [5]. In this method, the weighting function works as a normalizing factor which relates to the phase spectrum information of sound source. Regarding this term, equation (1) can be formalized as follows [5],

$$P_n(X) = \int_{nT}^{(n+1)T} | \sum_{q=1}^{N} w_q m_q(t - \tau(X,q)) |^2 dt \qquad (7)$$

where w_q is weighting factor and $\tau(X,q)$ is the direct time of travel from location X to microphone q.

Therefore, SRP can be computed by summing the GCCs for all possible pairs of the set of microphones [5]. The GCC for a pair *(k,l)* is computed as

$$R_{m_k m_l}(\tau) = \int_{-\infty}^{+\infty} \varphi_{kl}(w) M_k(w) M_l^{*}(w) e^{jw\tau} dw \qquad (8)$$

where τ is the time lag,* denotes complex conjugation, $M_l(w)$ is the Fourier transform of the microphone signal $m_l(t)$ and $\varphi_{kl}(w)$ is a combined weighting function in the frequency domain [5]. In the SRP-PHAT, the weighting function for a reverberant environment is defined as [5],

$$\varphi_{kl}(w) \equiv \frac{1}{| M_k(w) M_l^{*}(w) |} \qquad (9)$$

In SRP-PHAT, GCC is computed using (8) instead of (2) to obtain $P_n'(X)$ which is mentioned in (3). Finally, the sound source location can be evaluated by finding the point source location X_s that provides the maximum value in (6) [5].

3. Time Difference of Arrival (TDOA) Method

The TDOA is one of the time delay estimation (TDE) sub-methods that is used in low noise or noise free environments, which leads to a considerable reduction of computational complexity. In this method, at least two microphones should be used to find sound source direction *(θ)*. For finding θ, we need to calculate the time delay between received signals of each microphone, respectively. An approach to estimate the time delay between the received signals at two microphones is cross-correlation [18]. The computed cross-correlation values give the point at which the two signals from separate microphones have their maximum correlations. The cross-correlation of sound signals s_i and s_j received in microphones i and j respectively is given by [8],

$$R_{lk}(i) = E\{s_l^{*}[j]s_k[i + j]\} \qquad (10)$$

where E denotes the expectation operator, i is discrete time shift, j points samples of each sound source signal and *denotes complex conjugate operation. As shown in (11), the discrete time delay between received signals, τ_{kl}, can be obtained by finding argument of the maximum value of cross-correlation, where the signals are best aligned[8],

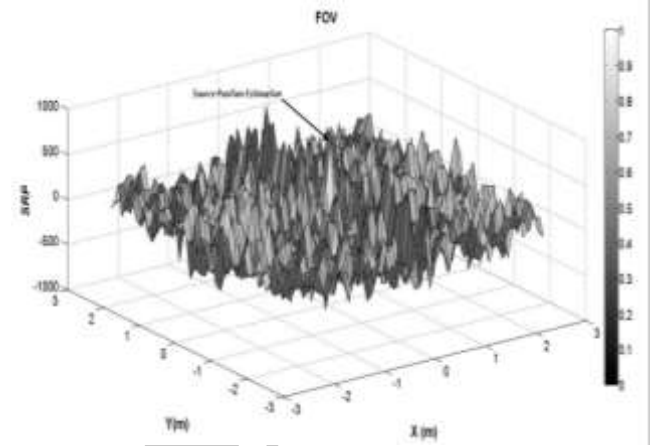$$\tau_{kl} = \arg \max_i (R_{kl}(i)) \qquad (11)$$

Fig. 2. SRP image for 5×5 m² FOV in the presence of noise with *SNR=10dB*.

The time delay between two typical microphones is also given by [8],

$$t_{kl} = \frac{\tau_{kl}}{f_s} \qquad (12)$$

Where f_s is sampling frequency rate of sound source. Therefore, the sound source direction, θ can be given by,

$$\theta = sin^{-1} \left(\frac{c \times t_{kl}}{d}\right) \qquad (13)$$

Where d and c are distances between two microphones and sound's propagation speed, respectively. In this approach. It is assumed that d should not be larger than sound wave length [18].

Fig. 3 shows a typical setup of TDOA method. As shown in Fig. 3, two candidates of θ can be mentioned for sound source direction [18-20]. For solving this problem, two pairs of microphones can be used to find the accurate sound source direction [21,22].
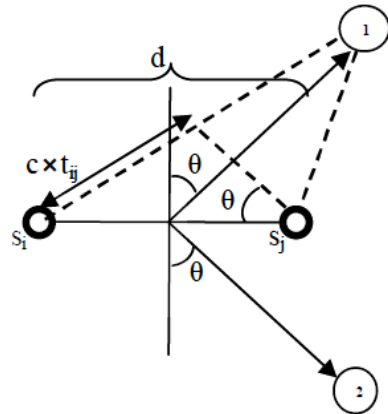

Fig. 3. Calculating the angle of sound source

One of the suggestions for setup of microphone pairs which can be aligned together is shown in fig. 4 and used in this paper.
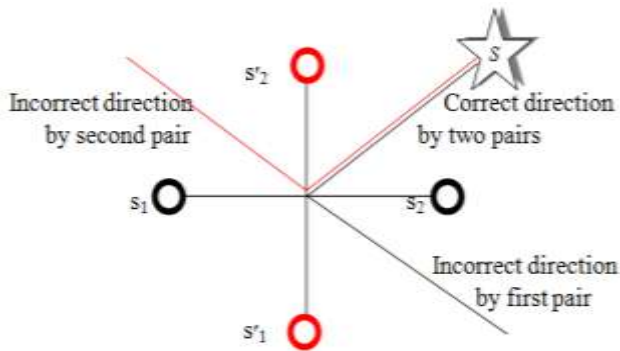
Fig.4. A typical TDOA microphone array (s_1 and s_2 are the first microphone pairs, s'_1, s'_2 are second microphone pairs, S is the sound source position)

4. Combination of SRP/SRP-PHAT and TDOA

As mentioned in section 1, although the SRP method can find sound source location, it is time consuming. On the other hand, although the TDOA is a low computational time method, it is noise effective. A suggestion can be derived using a combination of these methods to decrease the computational time as well as more robustness in the presence of noise. As shown in fig.5, two TDOA setups such as fig. 4 are used at the center of FOV [8] and three additional microphones are also utilized for each quarter [2,5,6,8]. For each quarter, these additional microphones with the central microphone can be used for SRP methodology. Therefore, in this paper, 13 microphones have been used in order to have a symmetric structure.

4.1 Classical Combination of TDOA and SRP/ SRP-PHAT

First, based on TDOA method mentioned in section 3, the sound source direction can be determined using four microphones placed in the center of FOV.
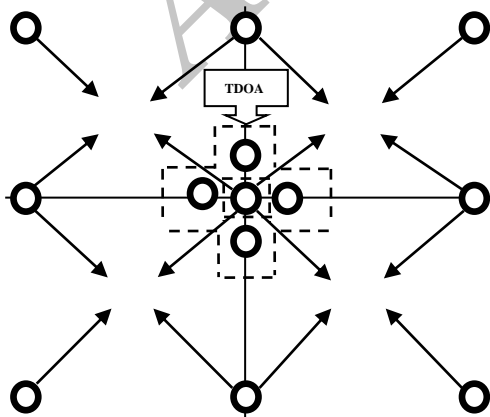

Fig. 5. Combination FOV *(circles are microphone positions)*

The next step is to find which quarter contains the accurate sound source location. Finally, one of the SRP or SRP-PHAT methods can be used to find the actual location of sound source in the selected quarter. For a better discrimination, the two proposed methods are briefly named as TDOA-SRP and TDOA-SRP-PHAT, respectively.
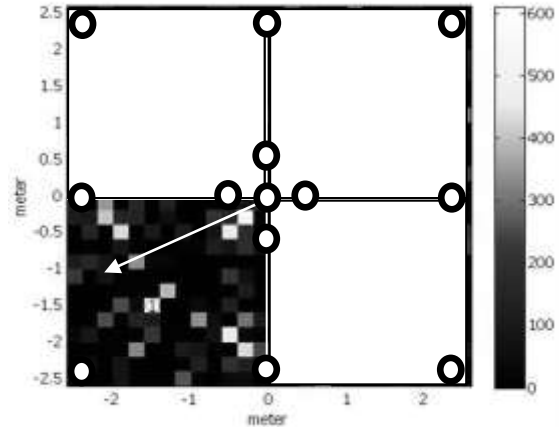

Fig. 6. SRP/SRP-PHAT method computes sound source location in the quarter in the FOV selected using TDOA method. *(Circles are microphone positions)*

Fig. 6 shows a general FOV that TDOA method has been proposed to find sound source direction. Afterward, this direction can be recognized to find the quarter contained sound source location. Each quarter containing the direction arrow is the goal quarter in the first step.

As shown in fig. 7, the selected quarter can be used by SRP or SRP-PHAT grid search methods to find the actual sound source location. In this selected quarter, the SRP or SRP-PHAT search in grid points and find sound source location.

4.2 Optimal Combination of TDOA and SRP/SRP-PHAT

The classical combination of TDOA and SRP/SRP-PHAT mentioned in subsection 4-1 can reduce the search area to a quarter of grid points. But, it should be noted that for noise-free or low-noise environments, the SRP/SRP-PHAT methods just need to span the grid points along the direction which has been estimated by TDOA method. On the other hand, in the heavy noise environment and based on the noise effective nature of TDOA method, the SRP/SRP-PHAT methods should span nearly all of the grid points mentioned in the selected quarter to find the actual sound source location. Considering the time consuming nature of SRP-based methods, it seems for usual environmental noise that it is over qualified to seek all grid points of a quarter to find actual sound source location.

Our experimental results show that for the successful detection of sound source location in the real environment with a different noise level, SRP-based methods can be computed in a region with a deviation, δ, around the direction obtained by TDOA method. Our empirical results indicate that this parameter can be selected proportional to $\delta = \sigma$, where σ is noise standard deviation.
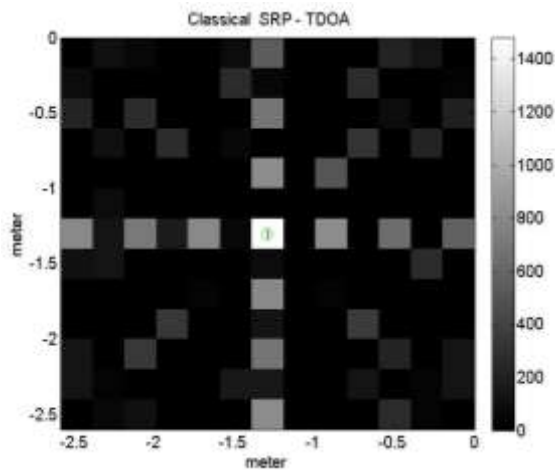
Fig. 7. Sound Source Localization in the selected quarter of Fig.6
(Classical TDOA-SRP/SRP_PHAT)

Fig. 8 shows a typical example of this approach. For better discrimination, two proposed methods optimized through the new structure are briefly named O-TDOA-SRP and O-TDOA-SRP-PHAT, respectively.
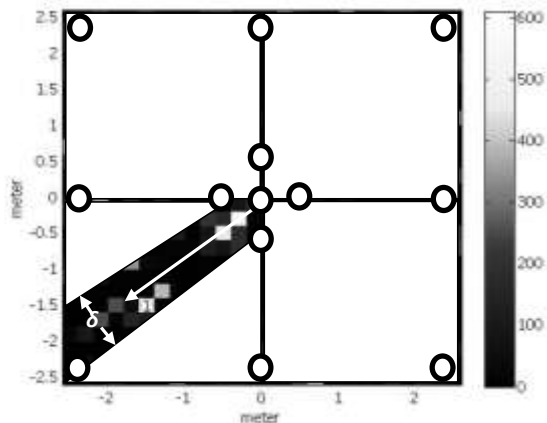

Fig. 8. Optimal TDOA-SRP/SRP_PHAT *(Circles are microphone positions)*

5. Experimental Results

To obtain the experimental results, a PC with the following software and hardware specifications has been used. Software: MATLAB R2013a, Hardware: PC Core(TM)i7-3632QM, CPU 2.20 GHz, RAM 8 GB.

In this experiment, the resolution of grid points is assumed to be *100 and 200 mili meters* in 5-1 subsection and is assume as 200 mili meters in other parts. The dimensions of FOV for sound source location are *5×5×2* meters in length, width and height, respectively. The sound source used for this experiment is *Chainsaw* sound in wav format with the time spectrum mentioned in fig. 9. It's number of bits per sample is 16. Maximum frequency of the sound is 21.956 kHz, and the sound source sampling frequency is 44.1 kHz according to Nyquist sampling theorem. The processing was carried out using a sampling rate of 44.1 kHz, with the time windows of 4096 samples of length and 50 % overlap.

The fundamental of the work presented here is based on SRP methodology and TDOA method only has been used to reduce FOV's area by detecting sound source direction (not location). Therefore, experimental results of the proposed methods have been compared with the other SRP base methods which can find both of sound source direction and location.

5.1 Comparison of the Proposed Methods in the Presence of Noise

In this section, the deviation between actual and estimated sound source location of proposed methods is evaluated. In this comparison, for each level of noise, sound sources were degraded h times by noise. Then, the accuracy of the proposed SSL methods have been computed h times using the deviation of estimated and actual SSL methods in terms of Root Mean Square Error (RMSE) [20] as follows,

$$RMSE = \sqrt{\frac{1}{h}\sum_{k=1}^{h}(r_{ref} - r_k)^2} \tag{14}$$

where r_{ref} is the distance between actual sound source location and center of FOV, and r_k is the distance between estimated sound source location and center of FOV. We use h=10, and also several SNR = 40, 25, 10, 0 and -10dB are considered to evaluate the performance of proposed methods.
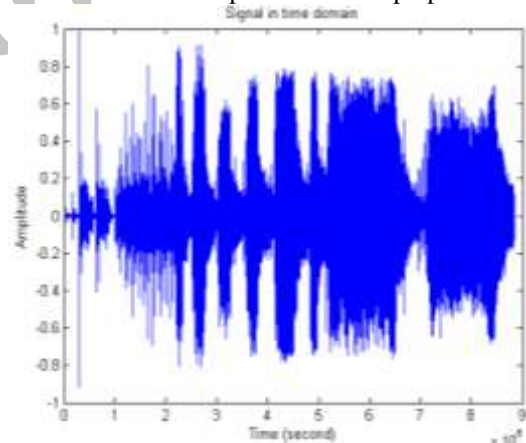

Fig. 9. time spectrum of Chainsaw sound source

Fig. 10 (*a*) and (*b*) show the SSL performance of proposed methods in the presence of different levels of noise for grid resolution 100 and 200 millimeter, respectively. As shown in fig. 10, for the all ranges of SNR, we have a significant difference between SRP-based methods and the proposed methods through classical and optimal combination of TDOA and SRP-based methods. In all methods, by increasing SNR till SNR= 0dB, the RMSE will be reduced and also the classical and optimal TDOA-SRP and TDOA-SRP-PHAT methods have better performances than SRP and SRP-PHAT methods.

On the other hand, due to the reduction of searching region of true sound source location, O-TDOA-SRP and O-TDOA-SRP-PHAT methods can successfully eliminate the similar sound source locations and led to a better performance than C-TDOA-SRP and C-TDOA-SRP-PHAT

methodologies, respectively. An overall evaluation can also show that O-TDOA-SRP-PHAT has the best robustness and accuracy in the presence of different levels of noise.

By comparing diagrams (a) and (b) in fig. 10, it can be concluded that decreasing the grid point resolution from 200 millimeters to 100 millimeters can reduce significantly the value of RMSE and improve all methods performance, respectively.
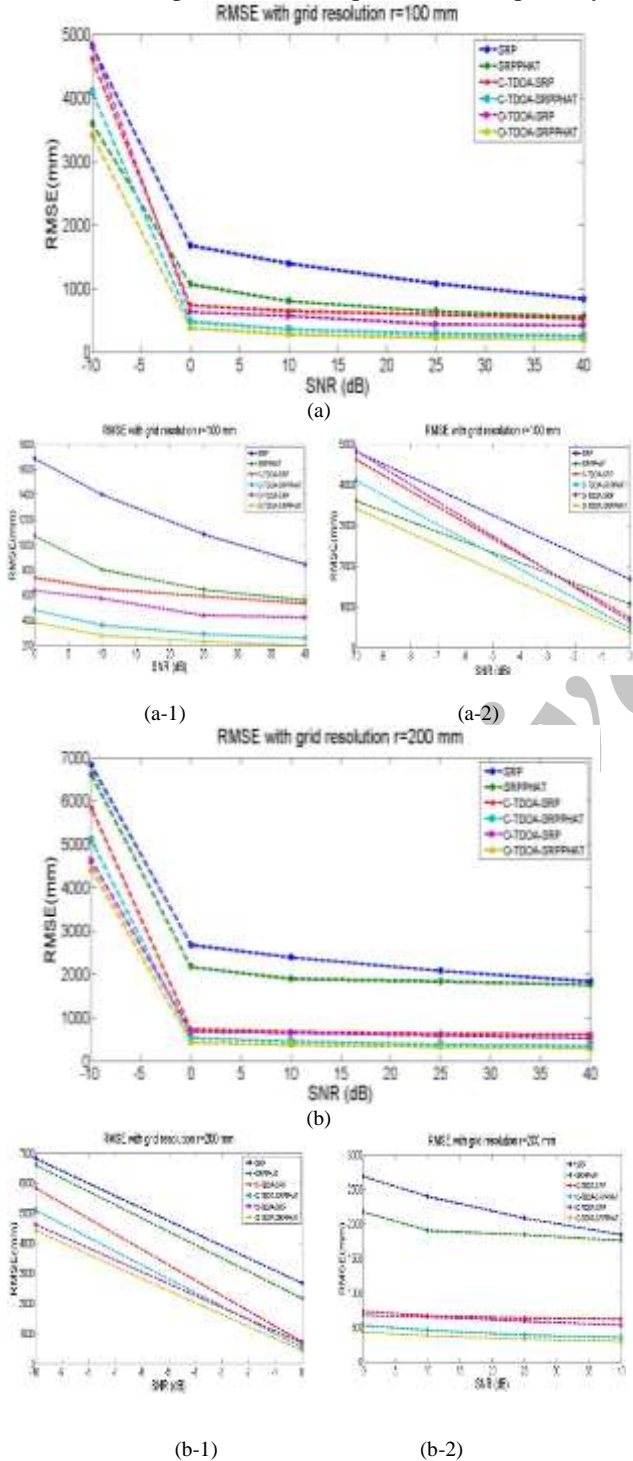

(a)

(a-1)   (a-2)

(b)

(b-1)   (b-2)

Fig. 10. Sound Source localization performance in terms of RMSE for proposed methods when different SNR are applied, (a) Grid resolution r = 100mm (a-1: -10 to 0 dB), (a-2: 0 to 40 dB) and (b) Grid resolution r = 200mm (b-1: -10 to 0 dB), (b-2: 0 to 40 dB).

As shown in fig. 10 (a-1, b-1), for SNR less than 0dB, RMSE is increased abruptly and the performance of methods is reduced effectively. Furthermore, it can be seen in SNR= -10dB that SRP and SRP-PHAT methods have a better performance than their classical combination methods. It is due to this fact that the proposed classical combination methods seek all of the selected quarter to find the sound source location. Therefore, it may lead to several outputs that satisfy the true conditions of real sound source direction. This problem has been solved in the optimized version of proposed methods due to limitation of the seeking area of sound source location.

5.2 Comparison Speed of Proposed Methods

In tables I and II, the computation times of the proposed methods are compared for SNR=10,-10dB in three different dimensions of FOV (these dimensions are in meter). As mentioned, for this entire situation, FOV's height is 2 meters. To calculate these times, each method runs ten times and the mean value of the processing time is reported.

Table 1. Comparison of the Time of Processing for SNR=10dB

Proposed Methods		Dimensions (m2)		
		5×5	10×10	20×20
SRP	Processing Time (Sec)	49	186	728
SRP-PHAT		47	193	781
C-TDOA-SRP		13	50	187
C-TDOA-SRP-PHAT		15	55	194
O-TDOA-SRP		8	24	90
O-TDOA-SRP-PHAT		8	25	100

Table 2. Comparison of the Time of Processing for SNR= -10dB

Proposed Methods		Dimensions (m2)		
		5×5	10×10	20×20
SRP	Processing Time (Sec)	52	189	735
SRP-PHAT		48	193	784
C-TDOA-SRP		11	51	188
C-TDOA-SRP-PHAT		14	54	189
O-TDOA-SRP		7	25	92
O-TDOA-SRP-PHAT		8	26	102

As shown in tables I and II, because of the ability of quarter selection in C-TDOA-SRP and C-TDOA-SRP-PHAT, they have lower computational time than SRP and SRP-PHAT. Also, because of the limitation in the region of process (grid point search), O-TDOA-SRP and O-TDOA-SRP-PHAT have less computational time than any other methods brought in this experiment. Therefore, for a real-time SSL, the optimal combination methods e.g. O-TDOA-SRP and O-TDOA-SRP-PHAT can have better abilities.

The comparison between tables I and II shows that there are no comparable differences between computation time of proposed methods in the presence of high and low SNR (respectively 10 , -10 dB). Therefore, it can be assumed that the computational time of proposed methods can be considered independent relative to SNR variations and have similar performance.

5.3 Stability Comparison

Another comparison can be mentioned through evaluating the stability of the proposed methods in the presence of several level of noise. The standard deviation (SD) between actual and estimated sound

source location can be used as a suitable objective manner to compare the stability of the proposed methods [5]. As shown in fig. 11, the SD of the proposed methods has been computed for three different SNRs (40, 25, 10 dB) and in three different FOV dimension. A lower SD points to the more stability of each method.
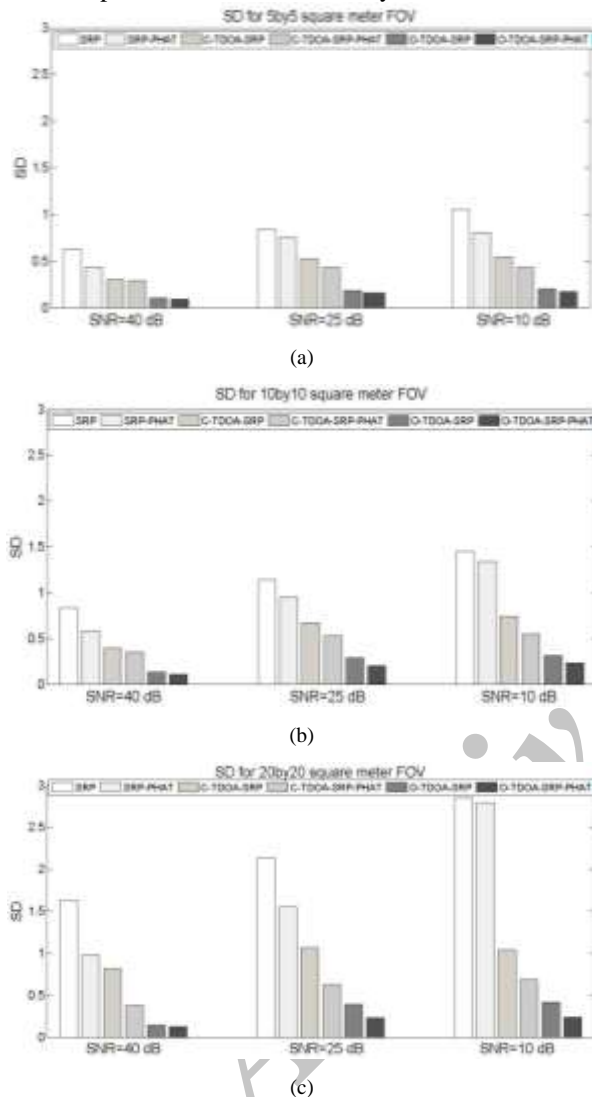

(a)

(b)

(c)

Fig. 11. Stability Comparison with Standard Division examination (each of the six columns is relevant to a SNR)
(a):5*5m^2 FOV; (b):10*10m^2 FOV; (c):20*20m^2 FOV

Based on fig. 11, each sub-fig shows increasing the level of noise can lead to lower stability (more SD) for all proposed methods.

But, in different level of noise (different signal to noise ratio), O-TDOA-SRP-PHAT has the best stability. Furthermore, for different level of SNR, Optimal combination methods (O-TDOA-SRP and O-TODA-SRP-PHAT) have fewer variations in SD level and more stability compared to other methods.

Comparison of figs 11 (a, b and c) show although increasing FOV's dimension reduce stability of methods but, optimal combination methods remain more stable than others.

6. Conclusion and Feature Work

Although SRP-based methods are practical and suitable ways for sound source localization in the noisy and reverberant environment, they need valuable processing time. On the other hand, although TDOA is a low computational approach for sound source localization, this method is very noise effective.

In this paper, experimental results show a combination approach based on TDOA and SRP/SRP-PHAT methodologies optimized and simplified by reducing the initial search region, and can decrease the time of processing as well as the better suppression of noise effect. Results indicate that the proposed sound source localization methods have better robustness and lower computational time relative to the simple SRP method. This reduction has been shown to be sufficient for the development of real-time sound source localization applications. Also, results show that SRP-PHAT method can have a better performance than SRP, even when combined with a basic TDOA. Moreover, the combination of the SRP and SRP-PHAT with the TDOA method increases their stability in different signals to noise ratio level. The limitation of the proposed methods (combined methods) is the number of microphones which may make these approaches inappropriate for some practical applications. The next research challenge for authors is how the number of microphones can be reduced besides keeping the appropriate performances of the proposed methods.

References

[1] K. D. Donohue, K. S. McReynolds, A. Ramamurthy, "Sound Source Detection Threshold Estimation using Negative Coherent Power," Proceeding of the IEEE, Southeast Con., pp. 575-580, April 2008.

[2] K. D. Donohue, S. M. Saghaian Nejad Esfahani and JingjingYu,"Constant False Alarm Rate Sound Source Detection with Distributed Microphones," EURASIP Journal on Advances in Signal Processing, Vol.2011, no.1, pages 12, Article ID 656494,doi:10.1155/2011/656494, Mar. 2011.

[3] N. Madhu and R. Martin, "Acoustic source localization with microphone arrays," in Advances in Digital Speech Transmission. Hoboken, NJ: Wiley, pp. 135–166, 2008.

[4] J. Chen, J. Benesty, and Y. Huang, "Time delay estimation in room acoustic environments: Anoverview," EURASIP J. Appl. SignalProcess., vol. 2006, pp. 1–19, 2006.

[5] M. Cobos, Amparo Marti, Jose J. Lopez, "A modified SRP_PHAT functional for robust real time sound source localization with scalable spatial sampling, " IEEE Signal Processing Letters, vol. 18, no. 1, pp. 71-74, 2011.

[6] K. D. Donohue, J. Hannemann, H. G. Dietz, "Performance of phase transform for detecting sound sources with microphone arrays in reverberant and noisy environments," Signal Processing, vol.87, no.7, pp.1677-1691, July2007.

[7] Unnikrishnan, Harikrishnan. "Audio Scene Segmentation Using a Microphone Array and Auditory Features," M.s Thesis, 2010.

[8] N. M. Kwok, J. Buchholz. "Sound source localization: microphone array design and evolutionary estimation," IEEE International Conference Industrial Technology (ICIT), pp.281-286, December 2005.

[9] D-H. Kim, Y. Park, "Development of sound source localization system using explicit adaptive time delay estimation," International Conferenceon Control Automation and Systems (ICCAS), Muju Resort, Jeonbuk, Korea, 2002.

[10] I. A. Mc Cowan, "Robust Speech Recognition Using Microphone Arrays," Ph.D. dissertation, Univ. Technology, Queensland, Australia, pp.1-38, April 2001.

[11] J.H. DiBiase, H.F. Silverman, M.S. Brandstein, "Robust localization in reverberant rooms," Microphone Arrays, Signal Processing Techniques and Applications, Springer, Berlin, pp. 157–180, 2001.

[12] C. Knapp, G. Carter, "The generalized correlation method for estimation of time delay," IEEE Trans. Acoustic Speech Signal Process, Vol.24, pp. 320—32. Aug, 1976.

[13] J. Kuhn, "Detection performance of the smooth coherence transform (SCOT)," IEEE International Conference on Acoustics, Speech, and Signal Processing'78(ICASSP '78), Hartford, CT, Vol.3, pp.678-683, Ap., 1978.

[14] T. Qiu, H.Wang. "An Eckert-weighted adaptive time delay estimation method," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vo.44, No.9, pp.2332 – 2335, Sep., 1996.

[15] A. Ramamurty, H.Unnikrishnan, K. D. Donohue, "Experimental Performance Analysis of Sound Source Detection with SRP PHAT β," IEEE Trans. Southeast con. , pp. 422-427. March, 2009.

[16] K. D. Donohue, "Audio systems array processing toolbox," (for MATLAB), Audio Systems Laboratory, Department of Electrical and Computer Engineering, University of Kentucky,(Updated: 27-10-2009), http://www.Engr.uky.Edu /~Donohue/audio/Arrays/MAT Toolbox. Html

[17] J.H. Dibiase "A High-Accuracy, Low-Latency Technique for Talker Localization in Reverberant Environments Using Microphone Arrays,"Ph.D. Thesis, Brown University, Providence, RI, May 2001.

[18] J. C. Murray, H. Erwin and S. Wermter, "Robotic sound-source localization and tracking using interaural time difference and cross-correlation," Proceedings of Neurobotics Workshop, Germany, pp.89-97, September 2004.

[19] M. Brandstein, H. Silverman, "A Practical Methodology for Speech Source Localization with Microphone Arrays," Computer, Speech, and Language, Vol.11, no.2, pp.91-126, April 1997.

[20] A. Brutti, M. Omologo, P. Svaizer, " Comparison between different sound source localization techniques based on a real data collection," in proceedings of the Hands-free Speech Communication and Microphone Arrays (HSCMA), pp.69-72, May 2008 Proceedings of the IEEE, Southeast Con., pp. 575-580, April 2008.

[21] A. Pourmohammad, S. M. Ahadi," TDE-ILD-based 2D half plane real time high accuracy sound source localization using only two microphones and source counting," Electronics and Information Engineering (ICEIE) ,IEEE, vol. 1, pp. 566-572, Aug 2010.

[22] S. Astapov, J. Berdnikova, and J. S. Preden, "A method of initial searchregion reduction for acoustic localization in distributed systems," inProc. 20th Int. Conf. Mixed Design of Integrated Circuits and Systems (MIXDES), pp. 451–456, June 2013.

Mohammad Ranjkesh Eskolaki was born in Rasht. He received the B.Sc degree in Electronic Engineering from the department of electrical engineering, the University of Guilan, Rasht, Iran, in 2013. He is currently pursuing the M.Sc degree in Communication System from the department of communication system, KHI, Mashhad, Iran.
His current research interests include statistical signal processing and applications, digital audio processing, different sound source localization methods, MRI image processing, wireless communication channel estimation, 4G and 5G mobile systems.

Reza Hasanzadeh was born in Rasht, Iran, in 1978. He received the B.Sc degree (Hons) in Electrical Engineering from the University of Guilan, Rasht, Iran, in 2001 and the M.Sc and Ph.D degrees (Hons) in Electrical Engineering from the Amirkabir University of Technology (Tehran Poly-technique), Tehran, Iran, in 2003 and 2008, respectively.
He is currently an Associate Professor in the Department of Electrical Engineering, and has been the Director of the Digital Signal Processing (DSP) Research Laboratory, University of Guilan, Iran, since 2008. From 2005 to 2007, he was the Researcher of the Niroo Research Institute (NRI) in communication group, Tehran, Iran. Since 2008, he has been the Supervisory member of the Guilan Science and Technology Park.
His current research interests include fuzzy logic and signal/image processing techniques for industrial and medical applications.