PERFORMANCE COMPARISON BETWEEN STEERED RESPONSE POWER AND GENERALIZED CROSS CORRELATION IN MICROPHONE ARRAYS FOR SOUND SOURCE LOCALIZATION

Muhammad Fadhil Zuandi[1], Mareska Pratiwi Maharani[1] and Wansu Lim[2]
[1]Telkom University, Bandung, Indonesia
[2]Department of IT Convergence, Kumoh National Institute of Technology, Gumi, Gyeongbuk, South Korea
E-Mail: wansu.lim@kumoh.ac.kr

ABSTRACT

Acoustic-based source localization is being widely developed in target localization due to its advantages, compared to visual-based localization. There are several techniques for acoustic source localization, including time difference of arrival and beamforming. Methods related to those two techniques are GCC-PHAT and SRP-PHAT. GCC-PHAT is related to time difference of arrival, while SRP-PHAT is related to beamforming. In this paper, a comparison of GCC-PHAT and SRP-PHAT methods for acoustic source localization is introduced in order to determine the accuracy and response in speed for acoustic source localization applications. The results obtained from several experiments shows the performance comparison in terms of accuracies and computational times between the two approaches.

Keywords: localization, steered response power, generalized cross correlation.

1. INTRODUCTION

One of the basic aspects for mobile robot navigation is target localization. Vision navigation technology has been widely used for target recognition of mobile robots [1, 2]. However, visual navigation is limited by dark surroundings and its field of view to detect the target. Sound-based navigation system is the answer to those problems.

The use of acoustic or sound source techniques to localize the target has several advantages, given that it provides a better time resolution and it is not limited by visibility. Development of auditory sensors using sound localization advantages can make up the restrictions of other sensors and increase the types of environment information acquired by the robot [3]. The aim of sound source localization is to locate the source of the acoustic signal and to evaluate the angle and distance between the sources of the measuring system [4]. The ability to accurately localize sound sources contributes to perception, decision making and task performance. There are several techniques for sound source localization, such as fingerprinting, binaural based, time difference of arrival (TDOA), and beamforming [5, 6, 7, 8]. Each technique has different computational time, accuracy and effectiveness in sound localization. Comparison between them is necessary to determine characteristics and choosing better methods for localize targets on an environment using sound. Therefore, comparison between TDOA-based method and beamforming-based method for sound source localization is presented in this paper.

TDOA technique in sound localization needs low time consumption, which is suitable for single sound source localization. TDOA technique relies on different propagation distance from the same source signal to each microphone installed that brings about time difference of arrival [9]. The TDOA-based sound source localization is a two-step algorithm and its accuracy depends on the performance of time delay estimation (TDE) [6].

Generalized Cross Correlation (GCC) is a frequently used method constructed under TDOA [10, 11]. Based on [10]; GCC method is adequate for real time processing compared to other methods due to its simple algorithm. GCC is a simple correlation, implemented in the frequency-domain for an efficiency of locating the target, combined with a frequency bin weighing for robustness. GCC method starts by obtaining the time delay estimation. Then, the GCC function for a given time-lag is calculated as the inverse Fourier transform of the received signal cross spectrum scaled by a weighting function. In order to make robust GCC estimations, the correlations should be filtered, e.g., with the phase amplitude transform (PHAT). PHAse Transform (PHAT) is usually used due its ability to avoid peak spreading in the result and its anti-jamming ability while calculating cross correlation to reduce the uncertainty in TDOA. PHAT alters the correlation function such that it has unity gain in the frequency-domain while the phase information is preserved [11]. Generalized cross correlation with phase transform is often referred to as Generalized Cross Correlation - Phase Transformation (GCC-PHAT).

Beamforming is another technique used in sound localization, which is a one-step process technique. Beamforming aims to optimize a spatial filter that isolates a target signal's energy. It scans predefined location and gets reinforcement of the signal from particular direction, which is called Steered Response Power (SRP). Steered Response Power is defined as the output power of a filter-and-sum beamformer steered to a given spatial location. In SRP method, the environment is partitioned into small rectangles referred to as pixels, a delay and sum beamforming is performed to aim the microphone array towards the selected pixel [12]. Hence a power map is produced and the pixel that is associated with the maximum power is interpreted as the source location. Phase transform as a filter and weighing function is then applied to SRP (i.e. SRP-PHAT). In this paper, Stochastic Region Contraction (SRC) is used to optimize SRP-PHAT algorithm. For simplicity, the term SRP-PHAT with SRC will be referred as SRP-PHAT.

These two methods (GCC-PHAT and SRP-PHAT) have different ways in sound source localization using microphones. This paper presents a comparison between the GCC-PHAT and SRP-PHAT method. It also discussed their characteristics and which one is better for sound source localization. The methods were analyzed and compared based on performance factors such as computational time, time to reach preferred destination and results accuracy. A training device was used as the test subject to determine its result. The results can underlie further application of these methods for sound source localization based on their characteristics.

2. GENERALIZED CROSS CORRELATION-PHASE TRANSFORM (GCC-PHAT)

Time delay estimation (TDE) is one of the basic tools of statistical acoustic signal processing. Generalized cross correlation function is widely used due to its simplicity and short time decision for real time processing [10]. Cross-correlation functions describe the degree of correlation of two-time series at different periods in time. The signals received at each sensor, in this case a pair of microphones, are from the same source and there is a certain correlation, therefore by calculating the correlation function of the different sensors to receive the same signal, the time difference between different sensors can be estimated.

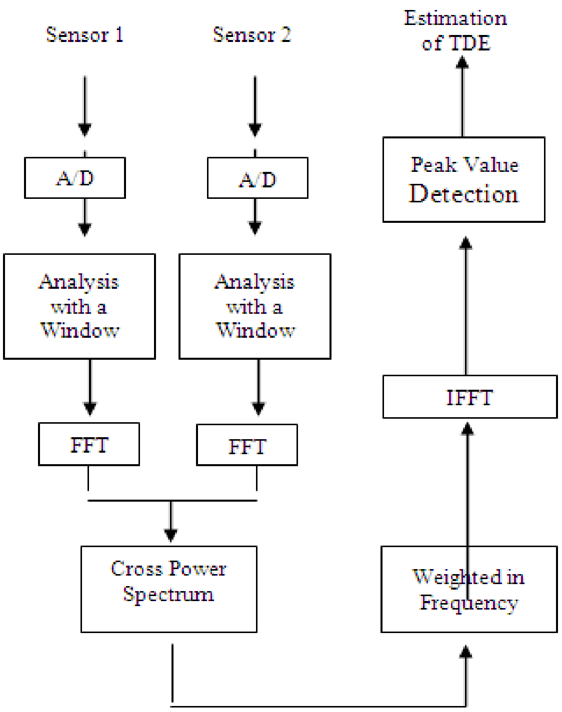

Figure-1. GCC process algorithm used in this paper with case of two microphones.

Generalized cross correlation time delay estimation method computes the two-related cross power spectrum signals in the spectrum domain by weighted processing, and inhibition of noise and reverberation part, so that the peak of the correlation function is more prominent, then inverse transformation to the time domain. Finally calculate the value of the delay [13]. Because of that, GCC is also known as Cross-Power Spectrum Phase (CSP). In this paper, generalized cross correlation algorithm is simulated in circular microphones array case. The algorithm process for generalized cross correlation simulation used is explained in Figure-1.

The signals received by two microphones used in the array, $x_1(t)$ and $x_2(t)$, can be described with:

$$x_1(t) = h_1(t) * s(t) + v_1(t) \qquad 0 \le t \le T$$
$$x_2(t) = h_2(t) * s(t-\tau) + v_2(t) \# \qquad (1)$$

where τ is the relative signal delay of interest, $h_1(t)$ and $h_2(t)$ represent the impulse responses of the reverberant channels, $s(t)$ is the speech signal, $v_1(t)$ and $v_2(t)$ correspond to uncorrelated noise added by microphone, and * denotes linear convolution [14].

The Generalized Cross Correlation (GCC) method then may be expressed as:

$$R_{x_1x_2}^{GCC}(\tau) = \int_{-\infty}^{\infty} \psi(w) X_1(w) X_2^*(w)\, e^{jwt}\, dw \#(2)\#$$

While the delay time between signals $x_1(\tau)$ and $x_2(\tau)$ can be obtained by the following expression

$$\hat{\tau}^{GCC} = \arg max\, R_{x_1x_2}^{GCC}(\tau)\#(3)$$

$X_1(w)$ and $X_2(w)$ denote the signals recorded by microphones 1 and 2 respectively, w is the angular frequency, [.]* stands for the conjugate transpose operation, the spectrum $X_2^*(w)$ is also the Fourier transform of the microphone signals $X_2(w)$ but with complex conjugation processing, and $j = \sqrt{-1}$. The weighting function $\psi(w)$ is designed to optimize a given performance criteria, which is reciprocal of $X_1(w)X_2^*(w)$ [15] Weighting approach can highlight the peak value of GCC spectrum to reduce the influence of reverberation and noise in order to make more robust GCC results. Thus,$\psi(w)$was used as a frequency weighing function also called as Phase Transform (PHAT)with equation as follows:

$$\psi(w) = \frac{1}{\left| X_{1(w)} X_{2(w)}^* \right|} \#(4)$$

By substituting (4) to (2), then

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

$$R_{x_1x_2}^{GCC}(\tau) = \int_{-\infty}^{\infty} \frac{X_1(w)X_2^*(w)}{\left|X_1(w)X_2^*(w)\right|} e^{jw\tau} dw \#(5)$$

Under ideal conditions, PHAT weighting function delivers an ideal GCC:

$$R_{x_1x_2}^{PHAT}(\tau) = \int_{-\infty}^{\infty} e^{j2\pi f(p-\tau)} df = \begin{cases} \infty, p = \tau \\ 0, otherwise \end{cases} \#(6)$$

There are several weighting functions can be used in GCC, as shown in table below:

Table-1. List of weighting function in GCC [16].

Name	Value
CC	$\psi_{cc}(f) = 1$
PHAT	$\psi_{PHAT}(f) = \frac{1}{\lvert X_{r1r2}(f)\rvert}$
ML	$\psi_{ML}(f) = \frac{1}{\lvert X_{r1r2}(f)\rvert} \frac{\lvert\gamma_{r1r2}(f)\rvert^2}{1-\lvert\gamma_{r1r2}(f)\rvert^2}$ where $\lvert\gamma_{r1r2}(f)\rvert^2 = \frac{\lvert X_{r1r2}(f)\rvert^2}{X_{r1r1}(f)X_{r2r2}(f)}$
ROTH	$\psi_{Roth}(f) = \frac{1}{X_{r1r1}(f)} or \frac{1}{X_{r2r2}(f)}$
SCOT	$\psi_{SCOT}(f) = \frac{1}{X_{r1r1}(f)X_{r2r2}(f)}$

Among them, PHAT has been found to have an outstanding performance for acoustic localization in reverberant environments [16], leading to the GCC-PHAT method when used with GCC. PHAT can be seen as a filter which discards the amplitude and preserves the phase of the signal. One of the advantages is that no assumptions are made about the signal or room conditions, which are typically unknown. Based on [17], GCC-PHAT method performs in general better than the CC and SCOT methods for TDOA estimation with respect to a speech sound source. This procedure has received considerable attention due to its simplicity and robustness, compared to other techniques.

3. STEERED RESPONSE POWER PHASE TRANSFORM (SRP-PHAT) USING STOCHASTIC REGION CONTRACTION (SRC)

Steered Response Power (SRP) is a method based on beamforming for acoustic source localization. Beamforming based methods are single stage, given that localization here is one step process, unlike TDE. Instead of working on pair wise time delay, it exploits the multitude of microphones to overcome the limitation given by early decision and reverberation. Beamforming processes converge or focus on the specific location and any specific direction in terms of the volume of signal generation on microphones or sensors. The purpose of this sound source localization is to estimate the position of sound sources in a search space using the audio observation or search scan. When the focus direction matches its actual source location in the systems, the Steered Response Power (SRP) will create a peak. The idea behind localizing an acoustic source using a steered beamformer is based on the assumption that the location of the signal source radiates more energy than all other locations [18], but the analysis of this fact is complicated and deferred for a future publication.

SRP addresses its work by covering the entire of Field of View (FOV) in the test environment. This method, partition the FOV environment in to small rectangles referred to as pixels, a delay and sum beamforming is performed to aim the microphone array towards the selected pixel. Hence a power map is produced and the pixel that is associated with the maximum power is interpreted as the source location [12].

According to [19], the SRP-PHAT algorithm is used to calculate the real-valued function for the 3-D spatial vector $\vec{x}$, which is obtained by steering a delay-and-sum beamformer (SRP), named $P_n(\vec{x})$. Those areweak irrelevant peaks, and by using that, the global maximum can be found, i.e. the sound source location$\hat{x}_s$.

$$\hat{x}_s = \arg\max P_n(\vec{x}) \#(7)$$

Say $\vec{x}$is the spatial position vector of a sound source. Given an n microphones system for example, S(t) is the source signal, then we can get the low frequency sound signal of sound source after we set S(t) filter by a low pass filter. So, we suppose $m_i(t)$ is the signal received by microphone i, and then the SRP is defined as [20]

$$P_n(\vec{x}) = \int_{nT}^{(n+1)T} \left| \sum_{i=1}^{N} w_i m_i\left(t - \tau(\vec{x}, i)\right) \right|^2 dt \quad \#(8)$$

where T is the length of data frames, w_i is a weight and $\tau(\vec{x}, i)$ is the direct time of travel from sound source location $\vec{x}$ to microphone i. Expanding Equation by going to the frequency domain using more general, frequency-dependent weights $W_l^*(\omega)$and considering Parseval's theorem can obtain equation's Fourier transform,

$$P_n(\vec{x}) = \sum_{k=1}^{M}\sum_{l=1}^{M}\int_{-\infty}^{+\infty} W_k(\omega)W_l^*(\omega) \times \int_{-\infty}^{+\infty} M_k(\omega)M_l^*(\omega)e^{jw(\tau(\vec{x},l)-\tau(\vec{x},k)}d\omega \quad (9)$$

Then a combined weighting function is defined,

$$\psi_{kl}(\omega) = W_k(\omega)W_l^*(\omega) \#(10)$$

VOL. 13, NO. 9, MAY 2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

The integral is seen to be the cross-power spectrum for microphones k and l with the direct waves in alignment. Noting the elements summing to $P_n(\vec{x})$ form a symmetric matrix with fixed energy terms on the diagonal, the part of $P_n(\vec{x})$ that changes with $\vec{x}$ is defined as $P_n'(\vec{x})$, i.e.,

$$P_n'(\vec{x}) \equiv \sum_{k=l}^{M} \sum_{l=k+1}^{M} \int_{-\infty}^{\infty} \Psi_{kl}(\omega)$$

$$\times \int_{\infty}^{\infty} M_k(\omega) M_l^*(\omega) e^{jw(\tau(\vec{x},l)-\tau(\vec{x},k))} d\omega \qquad (11)$$

The phase transform (PHAT) is an effective weight for using GCC to obtain a time delay of arrival (TDOA) in highly reverberant environment, and it is an inverse of the magnitudes of the frequency components. Thus, it will be used in the test with SRP which leads to SRP-PHAT method. PHAT itself can be defined as:

$$\psi_{lk(\omega)} = \frac{1}{|X_l(\omega) X_k^*(\omega)|} \#(12)$$

Thus, the process is to explore $P_n'(\vec{x})$over the whole focal volume and ultimately find the set of one or more distinct maxima $\hat{x}_s$ (k). The calculation of any particular point of $P_n'(\vec{x})$will be called a functional evaluation (FE). For the SRP-PHAT functional, it is necessary to determine a point-source location in the room that gives the maximum value of $P_n'(\vec{x})$.

Improvements have been made to make more reliable SRP-PHAT algorithm for acoustic localization. On [19], [20]Stochastic Region Contraction (SRC) is used to realize real-time localization SRP-PHAT in terms of search for the global maximum. Hoang, *et al.* [19] used stochastic region contraction (SRC) to find the global maximum instead of a grid-search, which requires FE on a fine grid throughout the room. It is also proven in [21], that the use of SRC can optimize SRP-PHAT algorithm by reducing its calculation process.

The basic idea of the SRC algorithm is, given an initial rectangular search volume which containing the desired global optimum and perhaps many local maxima or minima. Then gradually, in an iterative process, contract the original volume until a sufficiently small sub volume is reached in which the global optimum is trapped (the uncertainty voxel (volume V_u). Each contraction operation on iteration i is based on a stochastic exploration of the $P_n'(\vec{x})$ functional in the current sub volume. [w] While $P_n'(\vec{x})$ satisfies equation (14), completes once search

$$P_n(\vec{x}) > \lambda \cdot max P_n(\vec{x}) \#(13)$$

where λ is a threshold factor, and it's in the range of 0~1. Select the SRP values which satisfy equation (14), and then narrow the search area for second search. Repeat the operation above until only the global maximum in the search area, namely the sound source location, as shown in Figure-2. This SRC process is proven to optimize the SRP-PHAT algorithm, thus, SRP-PHAT using SRC algorithm is used in this paper for comparison test.

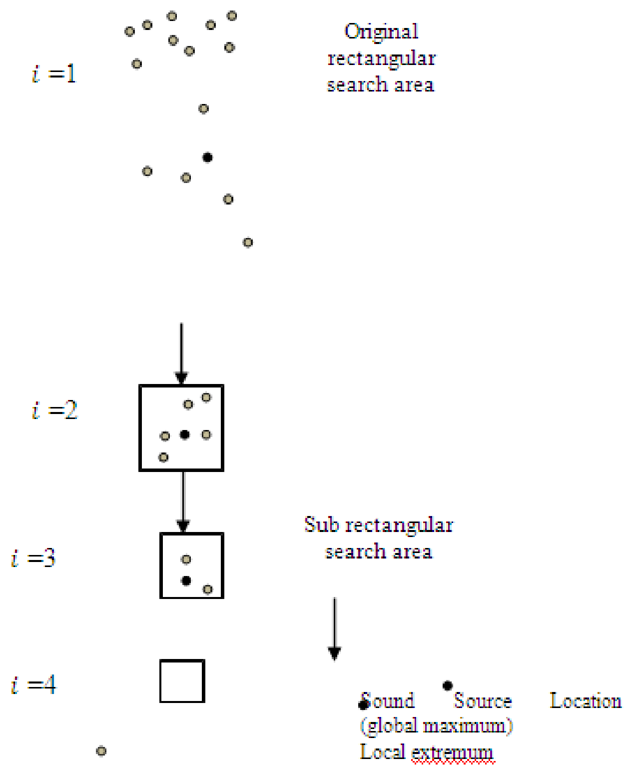

Figure-2. Example of SRC process.

4. SYSTEM DESIGN

In the simulation, different environment scenarios were created in order to compare two different techniques (GCC and SRP) for measuring the source location. These environments affect the result measurements that were caused by the difference of position between the active source, microphones and conditions such as noise. We implemented the measurement scenarios in the simulation by these conditions,

a) Free Field Environment - In free field scenario, the microphone or the array of microphones is situated over appropriate conditions, because it has no reverberation or multipath propagation conditions. This kind of conditions also mean that the fixed-point source like the microphone, is located in an ideal situation which has no external factor involved in observation, such as reflection or reverberation.
b) Near Field Environment - In this kind of scenario, the microphone is located much closer to the effective source; and the distance between the source and microphones is quite near. In this case, it creates almost lossless condition in the system that can affect the response of microphones or sensors.
c) Far Field Environment- In this kind of scenario, the microphone and the source are situated far away from

ARPN Journal of Engineering and Applied Sciences

each other. The distance is much further than near field environment, making different conditions such losses, take part in measurement.

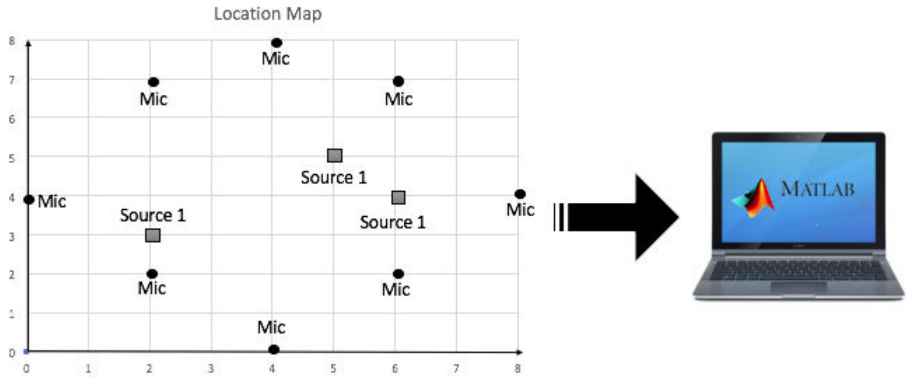

Figure-3. Location map used in the simulation using MATLAB.

The simulation environment was designed with free condition, as described above, with different distances for each source simulated towards the microphones array. The condition in Figure-3 is implemented in the simulation. It illustrates the system created for comparison test measurement of acoustic source localization with GCC-PHAT and SRP-PHAT method, where τ is delay and d is distance between microphones. In the simulation, sources and microphones were located in a room of 8 x 8 x 3 m^3 in size. The system consists of eight microphones, arranged into a circular array with the same distance between them. Those microphones are connected into a computer which runs the simulation. Simulation will be performed in MATLAB. MATLAB Simulink will be used in a computer connected to the system, enabling the implementation of the two methods (GCC-PHAT and SRP-PHAT) alternately, and review the results. The distance from the microphone array to the source varied for each scenario, explained in the table below.

Table-2. Sources Coordinate in the Simulation.

Source	Distance between mic (d)	Coordinate (x,y)
1	30 cm	(5,5)
2	30 cm	(6,4)
3	30 cm	(2,2)

Table-2 shows the coordinate of sources used for simulation. In 8 x 8 x 3 simulated environments, the sources position is varied inside the microphone array. Source 1 is located in middle right inside microphone array. Source 2 is located much closer to the array, beside source 1 location. And source 3 is located in the bottom left inside the array. The varied location of sources can affect the measurement result which can produce various accuracy and precision. The simulations were performed in MATLAB with the two procedures alternatively in order to determine the true location of the simulated sources inside the array. Performance result will be based on the following criteria:

a) Estimation Accuracy - comparison on how accurate the two methods are, for determining the true location of the source, and how big is the error of the estimated position in the simulation.
b) Computational Time - comparison on how fast the two methods are, for determining the true location when the simulation run.

Performance result for each simulation reviewed on a computer by using MATLAB Simulink software.

5. PERFORMANCE EVALUATION

In order to do a comparison and analysis of the two methods; simulations and experiments are conducted using GCC-PHAT and SRP-PHAT as method for sound source localization. The above environment explanation is used in the simulation. In simulation, sound data is simulated by recording a sample and placed in the same coordinate, like Table-2, and Figure-3. Two methods are used in circular microphone array alternatively to determine the simulated sources inside the array. The center of the microphone array is located at (4, 4, and 0.5 m) in the room. Sound sampling rate used is 44100 Hz. Speed of sound in the air is suppose be 340 m/s. The performance of each sound localization methods, is compared with different frequency conditions (800, 1500 and 2000 Hz) to test their suitability and accuracy in low, medium and high frequencies.

VOL. 13, NO. 9, MAY 2018 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

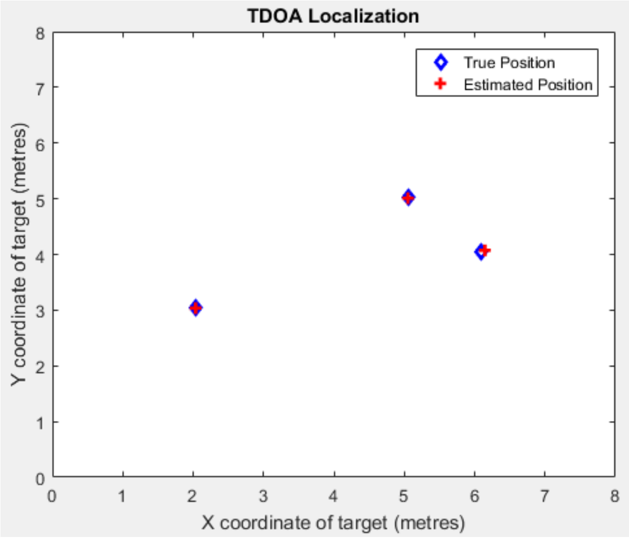

Figure-4. TDOA GCC simulation result.

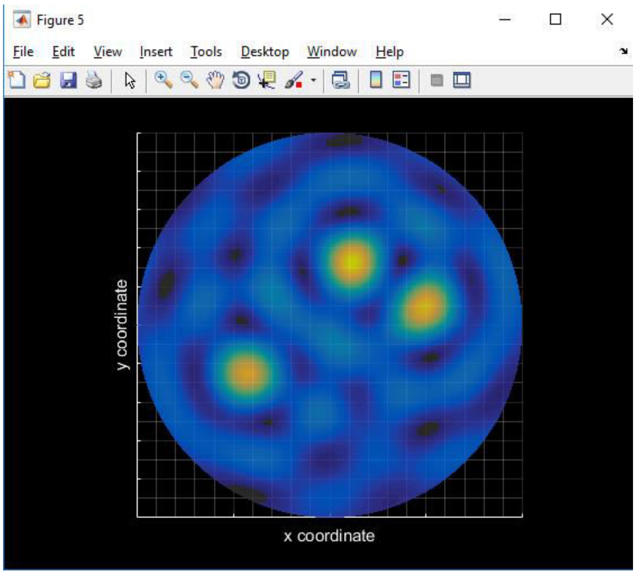

Figure-5. Beamforming SRP simulation result.

The simulation implemented two algorithms in order to localize the recorded sample sound. Figure-4 shows the simulation result of GCC algorithm in determining sound location, while Figure-5 shows the simulation result of SRP algorithm. In Figure-4 the estimated position by GCC algorithm is showed by red plus on the cartesian coordinate. In Figure-5 the estimated position by SRP algorithm is showed by yellow area inside beamforming environment which is represented by blue area. Frequency in SRP can be changed, in this case with 800, 1500 and 2000 Hz, giving the results in Figure-6 and Figure-7.

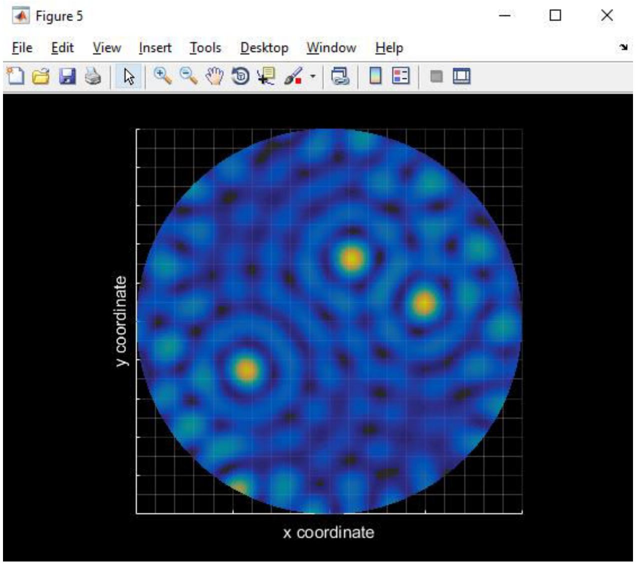

Figure-6. Beamforming SRP with 1500 Hz.

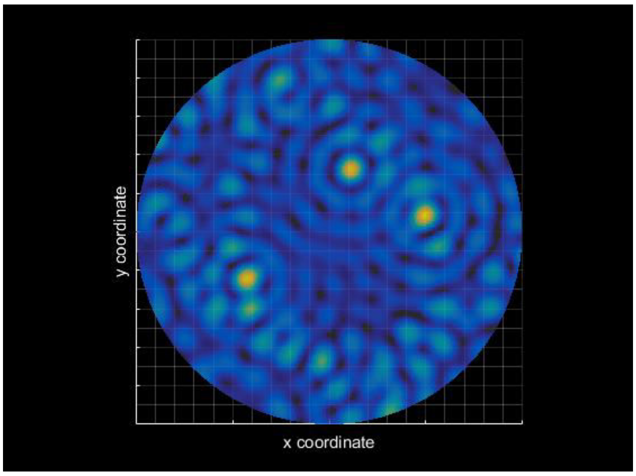

Figure-7. Beamforming SRP with 2000 Hz.

From Figure-6 and Figure-7 it is showed that the higher the frequency, the more highest energy (source) is. In this case is represented with smaller and more focused yellow area.

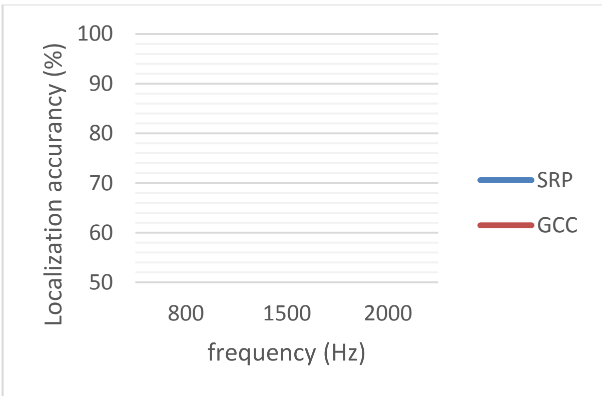

Figure-8. GCC and SRP accuracies compared in two and four microphones cases.

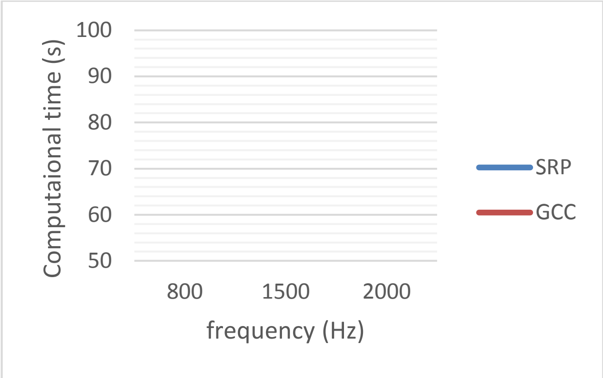

Figure-9. GCC and SRP computational time compared in two and four microphones cases.

The simulation was performed with circular microphone arrays for the two different methods, then analyzed under the performance criteria described before, which are localization accuracy and computational time. Result values were then averaged from several simulations to make the comparison. The result of simulation with two methods compared in terms of localization accuracy and computational time is shown in Figure-7, while the result of simulation with two methods compared in terms of computational and execution time is shown in Figure-8. It is shown in Figure-7 that SRP gives higher accuracy, the higher the frequencies are. Figure-8 shows that the computational time of GCC is faster than SRP, in 3 different frequency conditions.

6. CONCLUSIONS

Generalized Cross Correlation (GCC) and Steered Response Power (SRP) are two methods used in sound source localization that has their own different way in determining sound location. Comparison between two methods mentioned above is required for determining the characteristics of those methods in sound source localization and its environmental conditions suitability. Generalized Cross Correlation depends on time difference of arrival for estimating its results, while Steered Response Power depends on beamforming. Comparison between these methods for determining the performance difference between the two analyzed methods based on its accuracy and its computational time. It is shown in the simulation that the higher frequency is, the higher accuracy will get in beam forming SRP method. After the simulation, it is concluded that SRP is more suited in wider environments because it collects large data from the environment beamforming, which produces more accurate results than depending on delay in large environments. While GCC is suited for near environments where the source is located near the microphone array, which causes TDOA to have minimum error calculation and faster GCC computational time, over is SRP.

ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP; Ministry of Science, ICT & Future Planning) (No. 2017R1C1B5016837).

REFERENCES

[1] W. Yuan, Z. Li and C.-Y. Su. 2016. RGB-D Sensor-based Visual SLAM for Localization and Navigation of Indoor Mobile Robot. 2016 International Conference on Advanced Robotics and Mechatronics (ICARM). pp. 82-87.

[2] J. Shen and H. Hu. 2006. Visual based Localization for mobile robots with Support Vector Machines. IECON 2006 - 32nd Annual Conference on IEEE Industrial Electronics. pp. 4176-4181.

[3] Z. Linan, Y. Peng, S. Hao and C. Lingling. 2010. Sound Source Target Localization System of Mobile Robot. 2010 11th International Conference on Control Automation Robotics & Vision. pp. 2289-2294.

[4] J. Tuma, P. Janecka, M. Vala and L. Richter. 2012. Sound Source Localization. Proceedings of the 13th International Carpathian Control Conference (ICCC). pp. 740-743.

[5] K. I. Gettapola, R. R. W. M. H. D. Ranaweera, G. M. R. I. Godaliyadda and M. N. F. Imara. 2017. Location Based Fingerprinting Techniques for Indoor Positioning. 2017 6th National Conference on Technology and Management (NCTM). pp. 175-180.

[6] R. F. Armansyah, F. D. Ikram, S. S. Nolika and T. Adiono. 2016. Efficient Sound-Source Localization System using Low Cost TDOA Computation. 2016 International Symposium on Electronics and Smart Devices (ISESD). pp. 315-319.

[7] Z. Lei and K. Yang. 2015. Sound Sources Localization Using Compressive Beamforming with a Spiral Array. 2015 International Conference on Information and Communications Technologies (ICT 2015). pp. 1-5.

[8] X. Zeng and H. Cai. 2014. Binaural sound source localization based on upward zero-crossing point's detection in reverberant fields. 2014 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC). pp. 720-723.

[9] Y. Kan, P. Wang, W. Sheng, F. Zha, M. Li and B. Song. 2016. A New Joint Localization Model Using Multiple Microphone Arrays for Passive Acoustic

Source Localization System. 2016 9th International Symposium on Computational Intelligence and Design (ISCID). 2: 157-160.

[10] S. Lee, Y. Park and Y.-s. Park. 2010. Cleansed PHAT GCC based sound source localization. ICCAS 2010, pp. 2051-2054.

[11] B. V. D. Broeck, A. Bertrand, P. Karsmakers, V. Bart, H. Van hamme and M. Moonen. 2012. Time-domain generalized cross correlation phase transform sound source localization for small microphone arrays. 2012 5th European DSP Education and Research Conference (EDERC). pp. 76-80.

[12] H. R. Zarghi, M. Sharifkhani and I. Gholampour. 2011. Implementation of a cost efficient SSL based on an Angular beamformer SRP-PHAT. 2011 18th IEEE International Conference on Electronics, Circuits, and Systems. pp. 49-52.

[13] D. Ye, J.-Y. Lu, X.-J. Zhu and H. Lin. 2016. Generalized Cross Correlation Time Delay Estimation Based on Improved Wavelet Threshold Function. 2016 Sixth International Conference on Instrumentation & Measurement, Computer, Communication and Control (IMCCC). pp. 629-633.

[14] U. T. K. Nguyen and T. V. Pham. 2011. Performance Assessment of Generalized Cross-Correlation Based Algorithms for Multisource Point-Based Localization and Detection. The 2011 International Conference on Advanced Technologies for Communications (ATC 2011). pp. 303-306.

[15] J. Velasco, M. J. Taghizadeh, A. Asaei, H. Bourlard, C. J. Martín-Arguedas, J. Macias-Guarasa and D. Pizarro. 2015. Novel GCC-PHAT model in diffuse sound field for microphone array pairwise distance based calibration. 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2669-2673.

[16] C. Knapp and G. Carter. 1976. The generalized correlation method for estimation of time delay. IEEE Transactions on Acoustics, Speech, and Signal Processing. 24(4): 320-327.

[17] Y. Yang, X. Gu and X. Wan. 2016. Sound Source Localization Method Based on LDA Classifier. 2016 8th International Conference on Wireless Communications & Signal Processing (WCSP). pp. 1-5..

[18] J. P. Dmochowski, J. Benesty and S. Affes. 2007. A Generalized Steered Response Power Method for Computationally Viable Source Localization. IEEE Transactions on Audio, Speech, and Language Processing. 15(8): 2510-2526.

[19] H. Do, H. F. Silverman and Y. Yu. 2007. A Real-Time SRP-PHAT Source Location Implementation using Stochastic Region Contraction (SRC) on a Large-Aperture Microphone Array. 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07, 1: I-121-I-124.

[20] Y. Z. Zhang and B. Shen. 2015. Sound Source Localization Algorithm Based on Wearable Acoustic Counter-sniper Systems. 2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC). pp. 340-345.

[21] Y. Zhang and S. Meng. 2014. Sound Source Localization Algorithm Based on a Helmet-Mounted Microphone Array. 2014 Sixth International Symposium on Parallel Architectures, Algorithms and Programming. pp. 183-186.