International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-04, Issue-06, Sep 2018

Source Localization using SRP-PHAT

[1] Ritu, [2]Priyanka Dalal,

[1,2]Assistant Professor, ECE, GJUS&T, Hisar, India.

Abstract: The Steered Response Power using Phase Transform (SRP PHAT) is a robust technique for source localization however the evaluation of Steered Response Power for large spatial grids makes it computational burdensome which restrain its usage in the real time applications. Hence, this article aims to study the additional parameters which can be used to improve the performance while reducing the spatial grid points. The performance of the ibid algorithm has been studied in a room which is simulated using the image source method.

Keywords – localization, phase transform.

I. INTRODUCTION

Localization of the acoustic signal plays an important role in the modern application such as automatic camera steering, teleconferencing, military purpose, robotics and many more. The required information for localization is extracted from the signal captured by the microphones. The two most popular methods used are Time Delay Estimation (TDE) and Steered Beamforming based [1,5]. The TDE method is being widely used in real time application since it is fast but is less reliable due to its reliance on the first stages time delay estimates which are often not so much accurate in low SNR and high reverberant conditions. The techniques which are rooted from Steered Beamforming are the most precise in highly reverberant and noisy environment. In this technique the beamformer is steered to a particular location and power is calculated for that point. These are called as Steered Response Power (SRP) based techniques. When the SRP techniques is used with PHAse Transform (PHAT) function for pre-filtering or pre-whitening of the cross correlation then it is called SRP-PHAT algorithm. It is easy to implement for multiple cross correlation and outperforms TDE as it uses the entire cross-correlation instead of just the peak. The practical implementation of SRP- PHAT is hindered by its large grid size and hence a lot of efforts are to be made to make it suitable for real time applications [2,3,4,6,7].

II. SYSTEM MODELING

The system has been modeled with the assumptions that the

(1) Source is radiated uniformly in the all directions and the size & shape of the object doesn't affect the propagating acoustic signal.
(2) The microphones used have omnidirectional beampattern.
(3) The medium is homogeneous so that the velocity of the signal remains constant while travelling from source to microphone.
(4) The source is stationary and is emitting radiation continuously.
(5) The signal propagates as the linear wave equation and hence the system is modeled using linear set of equations.

The room of dimensions (5 x 6 x 3.25) simulated using Image source method [8] is used to analyse the performance of the ibid algorithm. To make the modeled system more close to real, the affect of two coherent noise sources has been included.

The signal received at m^{th} microphone at time t is given as (Equation.1)

$$x_m(t) = h_m(t) * s(t-\tau) + n_m(t) \qquad (1)$$

In Equation (1), the m varies from 1 to M where M is the number of microphones. The propagation delay between m^{th} microphone and source s(t) is τ, $h_m(t)$ is the impulse response when signal travels from source to m^{th} microphone, $n_m(t)$ is white noise added to the microphone to include the affect of system noise[1,5].

III. SRP –PHAT

The SRP P at any spatial grid point for any time frame l of length T is calculated by Equation (2)

$$P = \int_{lT}^{(l+1)T} \sum_{m=1}^{M} |w_m x_m(t-\tau)|^2 \mathrm{dt} \qquad (2)$$

Where τ is the delay between the m^{th} microphone and source s(t) and, w_m is the corresponding weight vector.

It has been shown in [10] that the SRP can also be calculated in frequency domain as below

$$R_{XmXn}(\tau) = \int_{-\infty}^{\infty} \psi_{mn}(\omega) X_m(\omega) X_n(\omega)\, e^{j\omega\tau} d\omega \qquad (3)$$

Where τ is the delay, $X_m(\omega)$, $X_n(\omega$ are the Fourier transform of the $x_m(t), x_n(t)$ signals received at m^{th} and n^{th} microphone respectively. The $\psi_{mn}(\omega)$ is the combined weighting function given in the frequency domain. There exist different weighting functions in the literature but the PHAT function has been proved to be very effective in

 DOI : 10.18231/2454-9150.2018.0712

reverberant environment. The PHAT weighting is defined as Equation (4)

$$\psi_{mn}(\omega) = \frac{1}{|X_m(\omega)X_n(\omega)|} \qquad (4)$$

This weighting function removes the magnitude spectrum from cross-correlation and hence also called as pre-whitening filter. The calculated cross correlation has flat magnitude spectrum and unfiltered phase spectrum. The SRP algorithm, based on GCC-PHAT (Generalised Cross Correlation using Phase Transform) [1,5,9], between the microphone pairs m and n can be re-written as Equation (5)

$$R_{XmXn}(\tau) = \int_{-\infty}^{\infty} \frac{X_m(\omega)X_n(\omega)}{|X_m(\omega)X_n^*(\omega)|} e^{j\omega\tau} d\omega \qquad (5)$$

The SRP [1,5,9] at any point p(x,y,z) in the FOV (field of view) is given as

$$P(p) = \sum_{m=1}^{M} \sum_{n=m+1}^{M} R_{XmXn}(\tau_{mn}(p)) \qquad (6)$$

$\tau_{mn}(p)$ is the inter -microphone time delay of arrival for the microphone m and n from the position p. The entire search space is divided into the spatial grid G by taking the resolution r and, the GCC is evaluated only for these discrete points called candidate locations for M(M-1)/2 microphone pairs. The assumption is that the peak power occurs for true source location as Equation(7)

$$p_s = \underset{p}{argmax}\,(P) \qquad (7)$$

The probability of missing the global peak increases with the coarser grid but increasing the grid size to minimize the effects of local maxima is not a feasible solution in real time applications [2,4].

IV. Performance Analysis

The experiments have been carried out with impulse type acoustic signal shown in Figure. 1, plotted in time and frequency domain.

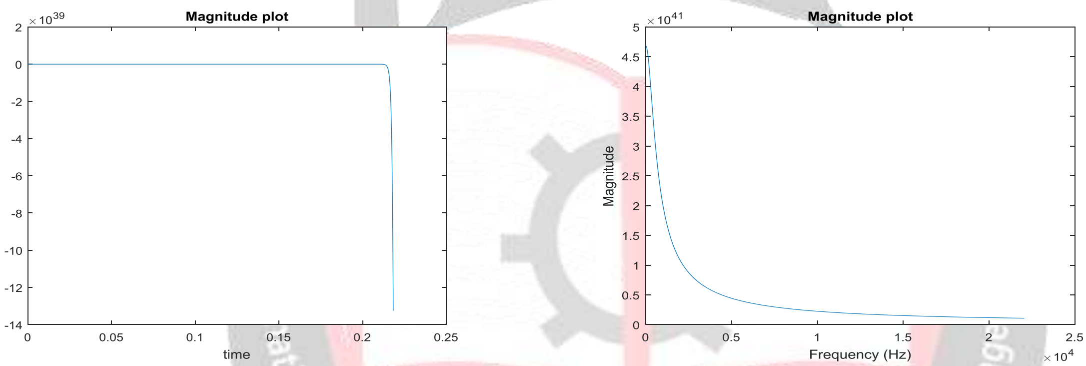
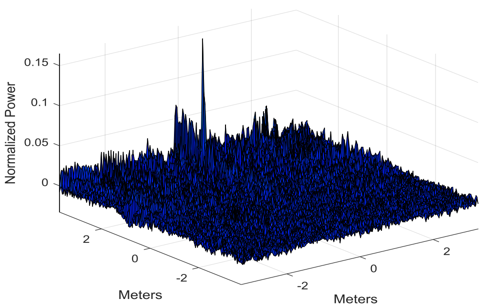
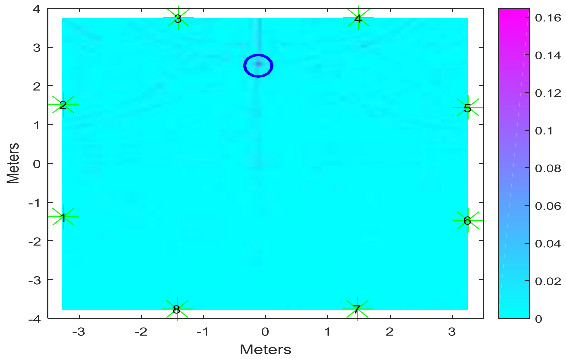


Fig 1: Sound Signal (Time and Frequency Domain)

The performance analysis of SRP PHAT algorithm has been done in the above modeled room with equal reflections for all the four walls. The floor and ceiling have been taken with slightly smaller reflection then the four walls. The spatial grid points where the SRP has to be calculated is defined with step size of 0.04 in the rectangular room. The microphones have been placed on the perimeter of room as shown in Image plot in Figure 2 with green asterisk. We observe a sharp peak in 3D plot for the true location in Figure 2, and the true location has been encircled in the 2D plot in Figure 2. Now if the number of grid points is decreased, the SRP- PHAT gives a degraded performance in same environment as above. This is explained with the help of Figure 3 where the grid step size is increased from the 0.04 to 0.08 which has ultimately decrease the number of grid points . Figure 3a shows that it has become tough to distinguish the peak as many local maxima has aroused and image plot in no more smooth like Figure 2 Image Plot.



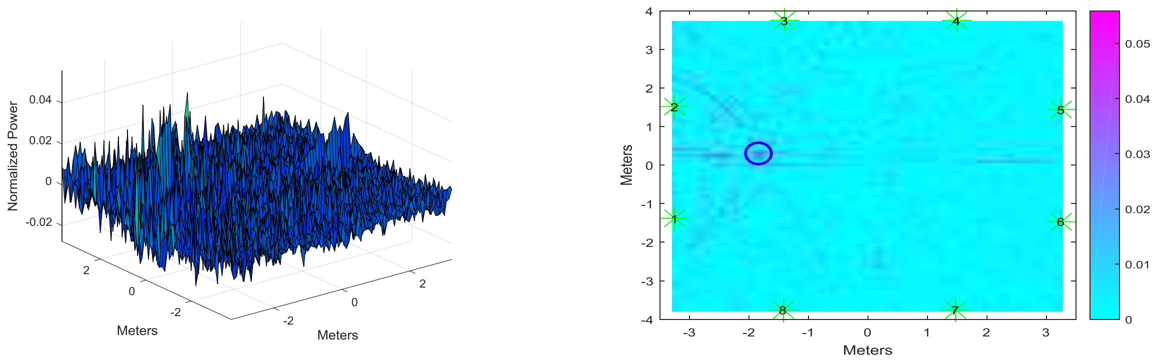
SRP 3-D Plot



SRP 2-D Image Plot

Figure 2: SRP 3D and 2D Plot

DOI : 10.18231/2454-9150.2018.0712



SRP 3D Plot SRP 2D Image Plot

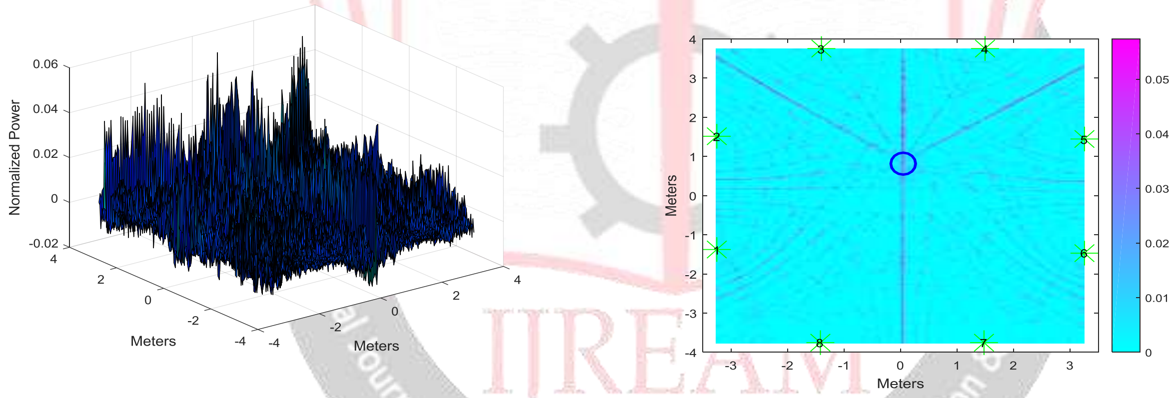
Figure 3. SRP Plot with grid step size =0.08

V. RESULTS AND DISCUSSIONS

Effect of Parameters Variations

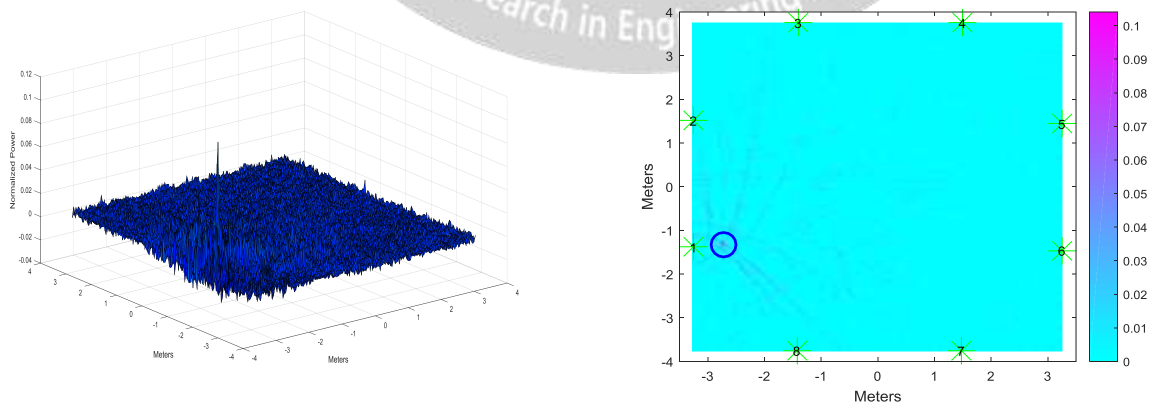
In order to improve the performance of the SRP- PHAT while reducing the grid size, there is utmost requirement to see the parameters which affect its performance.

- **Sampling Frequency**: Once the sampling frequency (Fs) is increased the performance improves considerably. This is clearly explained with help of Figure 4 and Figure 5. Figure 4 shows the calculated SRP in region of interest with Fs=16000Hz while Figure 5 displays SRP with Fs=24000Hz. The Figure 5 (3 D plot) has sharp peak corresponding to true location when compared to Figure 4. The Image plot in Figure 4 is very rough while a fine smooth Image plot has been obtained in Figure 5. However the fast processors are needed to handle the large data aroused due to increase of Fs.



3-D shaded surface plot SRP 2D Image Plot

Figure 4. SRP Plot , Fs = 16000 Hz

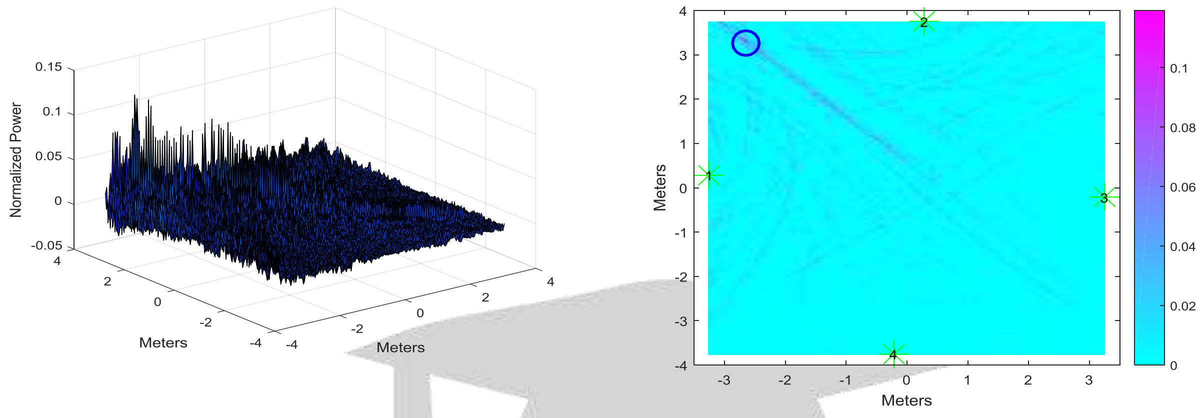


3-D shaded surface plot SRP 2D Image Plot

Figure 5. SRP Plot, Fs=24000 Hz

 DOI : 10.18231/2454-9150.2018.0712

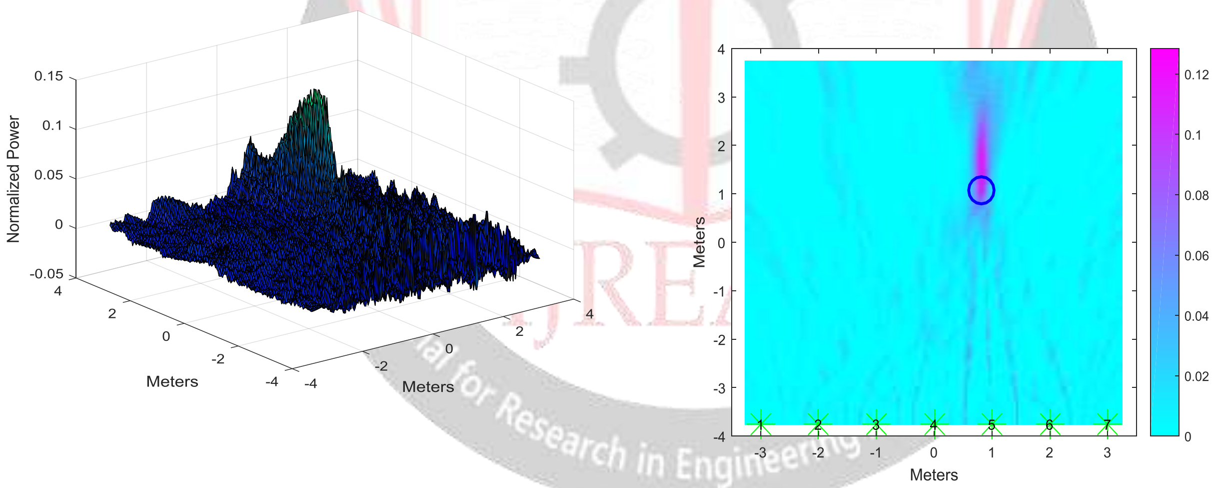
- **Array size:** The array of sensors provides large effective apertures when compared to single large sensor as to capture the effect of acoustic pressure and convert it into electrical signal. This idea has been taken from SONAR and RADAR. Arrays provide redundant information which can be used to improve localization accuracy. However, with increase in the array size there arises the need for additional hardware and computational demands also increase. With decrease in number of microphones from 8 (Figure. 4) to 4 (Figure 6), the power from undesired locations has increased and hence many local maxima can be seen in the Figure 6 below.



SRP 3D plot SRP 2D Image Plot

Figure 6. SRP Plot with reduced array size(M=4)

- **Geometry:** There are different array geometries proposed in literature but the Linear array geometry is most widely used due to its simplicity in implementation. SRP-PHAT algorithm has been simulated with linear array in Figure 7 with M=7. It can be observed that performance has degraded significantly when compared to parametric array and hence the microphone placement/ geometry is another vital factor which should be used to improve the performance.



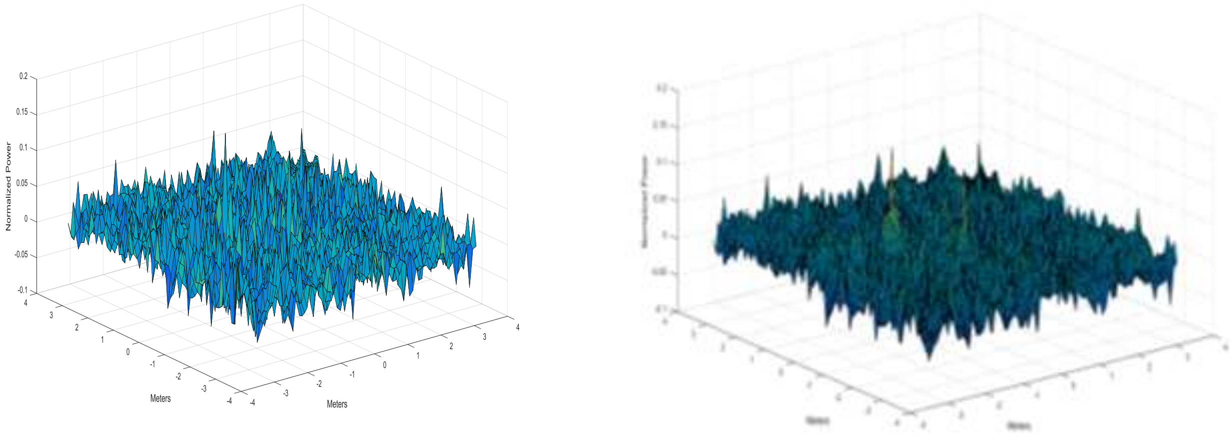
SRP 3D plot SRP 2D Image Plot

Figure 7. SRP Plot with reduced linear array

- **Interpolation:** Targeting to reduce the grid size in order to reduce the computations, the interpolation serves to be a good option to get smoother plots. The power at intermediate points of the spatial grid is evaluated using interpolation to get smooth graphs. On comparing the Figure 8a and Figure 8b, we can see that the plot has become smoother and clearer in Figure 8b with interpolation and hence the readability has improved considerably.

 DOI : 10.18231/2454-9150.2018.0712

ISSN : 2454-9150 Vol-04, Issue-06, Sep 2018



(a) (b)

Figure 8:Effect of Interpolation on SRP Plot(a) without interpolation(b) with Interpolation

VI. CONCLUSION

The SRP-PHAT is a robust method for source localization but the steered response power has to be evaluated at large number of spatial positions to achieve the accuracy. The article has evaluated the other parameters which can be used simultaneously with other spatial grid reduction techniques to improve the performance of SRP PHAT. Parametric array gives better resolution over linear arrays, the local maxima can be minimized further with increase in array size. The sampling frequency of signal should be optimum to obtain the sharp peaks and SRP value can be interpolated to get the power at intermediate points. The interpolation will help to get smooth curve while keeping the grid size smaller.

REFERENCES

[1]. J. H. DiBase, H. F. Silverman, M. S. Brandstein, “Robust Localization in Reverberant Rooms,” in *Microphone Array- Signal Processing Techniques & Applications.*

[2]. J. Dmochowski, J. Benesty, and S. Affes, “A Generalized Steered Response Power Method for Computationally Viable Source Localization,” in IEEE Transactions on Ausio, V.15, p.2510-2526, Nov. 2007.

[3]. H. Do and H. F. Silverman, “A fast microphone array SRP-PHAT source location implementation using coarse-to-fine region contraction (CFRC),” in Proc. IEEE Workshop Appl. Signal Process. Audio Acoust., New York, NY, USA, Oct. 2007, pp. 295–298.

[4]. J. Dmochowski, J. Benesty, and S. Affes, “Fast Steered Response Power Source Localization Using Inverse Mapping of Relative Delays,” in ICASSP, p. 289- 292, 2008.

[5]. J. Benesty, J. Chen, and Y. Huang, Microphone Array Signal Processing.Berlin, Germany Springer, 2010.

[6]. M. Cobos, A. Marti, and J. J. Lopez, “A modified SRP-PHAT functional for robust real-time sound source localization with scalable spatial sampling,” IEEE Signal Process. Lett., vol. 18, no. 1, pp. 71–74, Jan. 2011.

[7]. A. Said, B. Lee, and T. Kalker, “Fast steered response power computation in 3D spatial regions,” HP Labs, Palo Alto, CA, USA, Tech. Rep. HPL-2013-40, Apr. 2013.

[8]. J. B. Allen and D. A. Berkley, “Image method for efficiently simulating small-room acoustics,” J. Acoust. Soc. Amer., vol. 65, no. 3, pp. 943–950, Apr. 1979.

[9]. JOHN Samson Mwela, “ Acoustic source localization in a constrained search space”, Masters Thesis, June 2010.

[10]. J. H. DiBiase, “A high accuracy, low-latency technique for talker localization in reverberant environments using microphone arrays,” Ph.D. dissertation, Brown University, Providence, RI, May 2000.

 DOI : 10.18231/2454-9150.2018.0712