

Localization of underwater moving sound source based on time delay estimation using hydrophone array

S. A. Rahman¹, D. Arifianto¹, T. Dhanardono¹, Wirawan²

¹ Engineering Physics Department, Faculty of Industrial Technology, Institut Teknologi Sepuluh Nopember

² Electrical Engineering Department, Faculty of Industrial Technology, Institut Teknologi Sepuluh Nopember Jl. Arief Rahman Hakim, Surabaya, 0111

E-mail: saptianrahman@gmail.com

Abstract. Signal and noise of an underwater moving sound source is used to track the azimuth of a target. Uniform linear array with four hydrophones is used to detect azimuth of target by obtain the time delay information to get azimuth information. Success rate of time delay estimation influenced by characteristics of sound propagation like reflection, reverberation, etc. Experiment in real environment was done to analyze performance of the cross correlation (CC) and generalized cross correlation with the phase transform (PHAT) weighting to estimate time delay between two signal. The simulation done by convolute two signal that has been given time delay and impulse response of the medium test. Then the time delay of two signal estimated by CC and PHAT algorithm in Matlab in the various SNR. Then the algorithm tested in a pool to detect stationary and moving position of sound source. Result of the simulation and experiment in real environment shown that PHAT better than CC. The best azimuth tracking achieved by using PHAT algorithm with error of 0 - 9.48 degree in stationary position. In moving sound experiments, tracking the bearing and azimuth of the mini vessel (sound source) can be done by time delay estimation using PHAT.

1. Introduction

The existence of an object in the water can be detected by its acoustic signal emission. The resulting emission flows in the form of fluid particles pressure on a medium so can be detected by a hydrophone in variable frequency and gain [1]. There are various methods to detect position of the of the sound source, one of them is the estimated time delay. Time delay estimation has advantages compared with other methods to detect the position of the sound source such as Capon 's method or MUSIC. The advantage is more suitable to estimate time delay of broadband signals type, such as vessels signal. Azimuth tracking of underwater sound sources can be obtained based on the estimated time delay using multiple sensors distributed spatially, so that able to capture coming of sound signal [2]. Then value of time delay obtained and converted into azimuth of sound source relative to the hydrophone array. The most basic TDE methods is the cross-correlation (CC). The CC method assumes an ideal sound propagation model which makes it suitable for low-noise and reverberant-free outdoor environments. The improved version of the CC method is the generalized cross correlation (GCC) method [3].



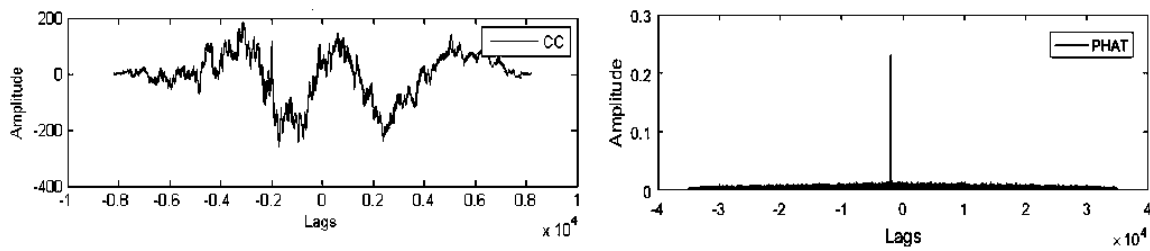


Figure 1. Comparison peak detection between CC and PHAT algorithm

The GCC is algorithm which applies various weight functions to the received signal (pre-filters) in order to improve time delay estimation in noisy or reverberant environment. But, GCC method more complex computation than CC method [3].

Success rate of time delay estimation influenced by characteristics of sound propagation like reflection, reverberation, etc. By using PHAT method, accuracy of time delay estimation able to improve on all kinds of signals (narrowband and broadband). As the study conducted by [4] that PHAT method can be used to track a moving sound source on the high seas with very well. In that research, the medium used is very broad so that can reduce effect of reflection, reverberation, and changes of medium volume due to the movement of a sound source. Hence, needs to be done research on tracking position of the moving sound source in the pool so that the effect of reflections or noise on the environment to estimate time delay can be known. CC and PHAT method chosen to estimate the delay time in this study.

2. Time Delay Estimation

When a sound source is emitting sound in a field, two signals $x_1(t)$ and $x_2(t)$, received at two spatially separated hydrophones, 1 and 2, can mathematically be modeled as:

$$x_1(t) = s_1(t) + n_1(t) \quad (1)$$

$$x_2(t) = \alpha s_1(t + d) + n_2(t) \quad (2)$$

where, $x_1(t)$ are the signals picked up by the microphones m_1 , m_2 respectively, $s_1(t)$ is the sound signal, α is the attenuation factor, $s(t)$ is the signal (direct path) from the source, $n_1(t)$ and $n_2(t)$ are assumed to be mutually uncorrelated noises and d is time delay. This model does not consider multipath or Doppler effects. When there are two microphones are used, it can be estimated based on the delay time difference from propagation of sound waves.

2.1 Cross Correlation

One of the methods commonly used to estimate the time delay d is to calculate the cross-correlation function between the signals received at two microphones. Then look for the maximum peak on the output which is the estimated time delay [5]. Cross power spectrum density function is calculated by the following equation.

$$R_{12}(\tau) = E[x_1(t)x_2(t - \tau)] \quad (3)$$

The time delay can be estimated from $R_{12}(\tau)$:

$$\tau_{12} = \arg \max_{\tau_{ED}} R_{12}(\tau) \quad (4)$$

2.2 Generalized Cross Correlation Phase Transform

Based on the relationship between the cross correlation and cross power spectrum, can be obtained by the following equation:

$$R_{12}(\tau) = \int_0^\pi G_{12}(\omega) e^{j\omega\tau} d\omega \quad (5)$$

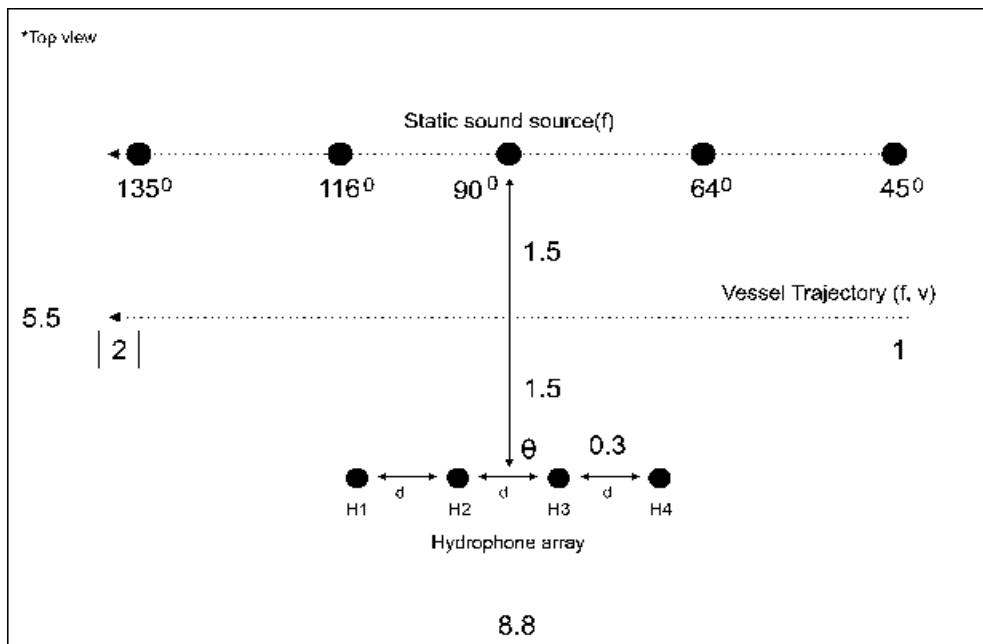


Figure. 2 The experiment setup: stationary and moving sound source straight trajectory

Where $G_{12}(\omega)$ is a cross - spectrum $x_1(t)$, $x_2(t)$ signal captured by the sensor m_1 and m_2 . By looking at the cross- power spectrum between the two signals and then provide weighting corresponding to the frequency domain, the method of whitening the GCC process and array signal thereby increasing the frequency components with higher SNR and inhibiting the effects of the array. Recently performed inverse transform to the time domain to obtain the GCC function of two signals [5]:

$$R_{12}^g(\tau) = \int_0^\pi \psi_{12}(\omega) G_{12}(\omega) e^{j\omega\tau} d\omega \quad (6)$$

Where $\Psi_{12}(\omega)$ is a weighting function GCC. Selection of weighting function depends on two aspects: the array and the condition of the reflectance of the sound source. Shown at Figure 1, the purpose of the use of weighting functions is to make $R_{12}(\tau)$ has a relatively sharp peaks so as to facilitate estimation of the time delay. R_{12} peak (τ) is the time delay between the two sensors. Weighting is very important to obtain a high SNR value it based on field conditions, where the presence of factors such as echo or reverberation time coming from various sides can reduce the value of SNR [5]. Cross correlation function will calculate the peak of the main signal and the reflected signal to estimate the time delay would be difficult to do at a low SNR values. Therefore, needed a weighting $\Psi_{12}(\omega)$ to suppress the effects of the array and the reverberation time. GCC PHAT types (Phase Transform) is a type of weighting is most commonly used with the following equation:

$$\Psi_{12}(\omega) = \frac{1}{|G_{12}(\omega)|} \quad (7)$$

Then do the inverse transform to the time domain to obtain the GCC PHAT function of two signals, the time delay can be estimated from $R_{12}(\tau)$:

$$\tau_{12} = \arg \max_{\tau_{ED}} R_{12}(\tau) \quad (8)$$

3. Experimental Setup

The measurement site done at pool, its bottom in the form of sedimentary soil a depth of 0.5 meters and the wall in the form of concrete.

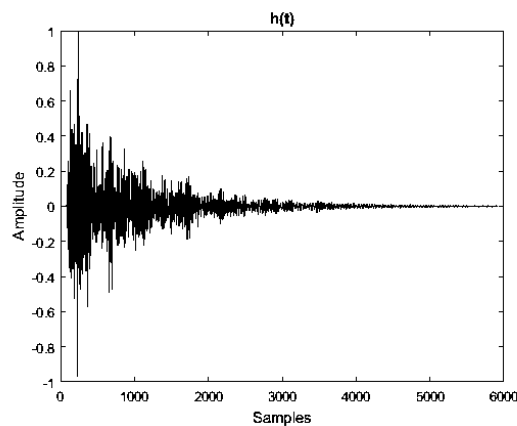


Figure 3. Impulse response of experimental pool

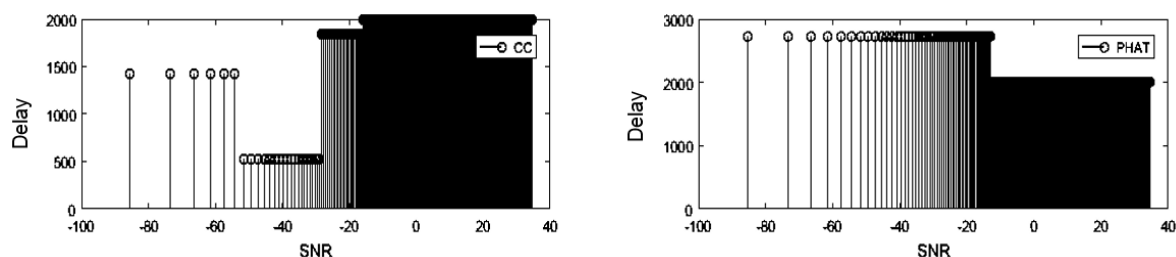


Figure 4. Influence of SNR concerning time delay estimation

Experimental scheme shown at Figure 2. Four hydrophones formed uniform linear array, $d = 30$ cm. The hydrophones were omnidirectional and placed at 0.25 cm below the water surface. A sampling frequency of 44.1 kHz with a receiver dynamic range of 16 bits used. Increasing the sampling rate improves the TDE resolution which in turn lead to a higher DOA resolution. There are two sound source in this experiment, underwater speaker as stationary and remote control vessel as moving sound source. In the figure, the unit in meters. To represent sound of the vessel that were dominant at low frequencies, generated four frequency variation: 200, 500, 800, 1000 Hz in the stationary experiment. In the moving experiment, used remote control vessel type Sea Wolf Submarine with dimension 35 x 9 x 9 cm and rotational speed of the propeller between 40 – 50 Hz. The vessel move from point 2 to point 1 with angular speed as big as 0.2 m/s and at the time the sound of the vessel recorded.

4. Result and Discussion

4.1 Simulation Results

Two algorithms have been simulated in Matlab, which are Cross Correlation (CC) and Phase Transform (PHAT). In order to simulate the real situation, a pure tone 1000 Hz has been chosen as a sound source signal. Then the sound source convolute by impulse response of the medium so as if generated in the real medium. An additional noise with various signal to noise ratio (SNR) was used to evaluate the methods. A white noise/ Gaussian noise was added to each sound signal. The two noise signals are assumed to be uncorrelated and to have a zero mean value. One of the two sound signals is delayed 2000 samples to simulate the time that the signal needs to reach the further microphone in the real environment. Figure 4 shows the simulation results, it can be seen that all methods was able to estimate the time delay until negative SNR. Figure 4 shows that, as the SNR becomes larger, the localization error becomes smaller. The result of simulation is given in Figure 4.

4.2 Experimental Results

In this subsection, we present the results of time delay based localization experiment. We convert

estimated time delay into azimuth information. Consider a setup for DOA estimation consisting of an array of four hydrophones separated by a small distance d and a source that lies in the far field of the array. The azimuth θ is related to the time delay, which for 1D microphone array is expressed as.

Table 1. Tracking result of stationary sound source

Actual Azimuth	Measured CC	Measured PHAT	Deviation CC	Deviation PHAT	Error CC	Error PHAT
45	70.11	40.14	28.73	0.00	25.11	4.86
64	73.10	61.66	0.00	0.00	9.10	2.34
90	91.50	93.33	17.53	0.00	1.50	3.33
116	136.95	125.48	28.80	1.01	20.95	9.48
135	118.45	135.93	24.90	13.81	16.55	0.93

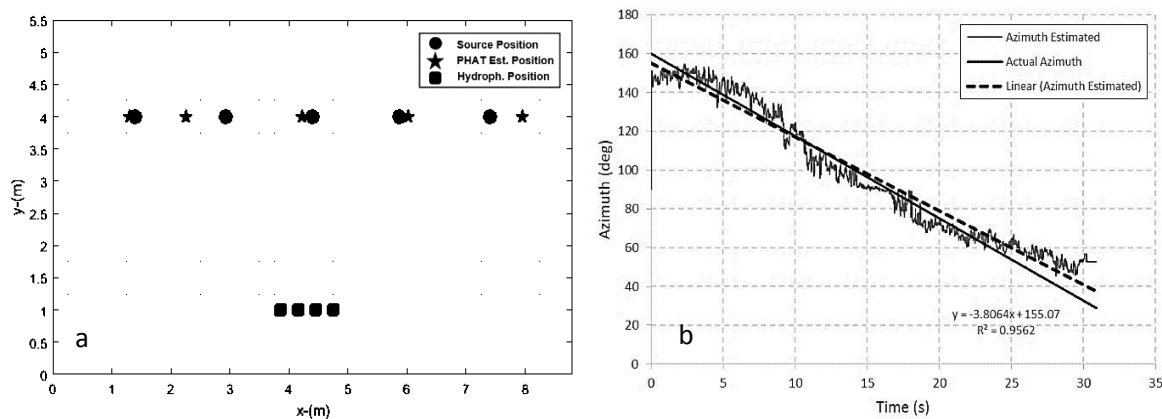


Figure 5. Tracking result of (a) stationary sound source (b) moving sound source

$$\theta = \sin^{-1} \frac{R}{D} = \cos^{-1} \frac{c \cdot \Delta t}{D} \quad (9)$$

Where c is the sound propagation velocity in the water at 1492.1 m/s. Results of several localization experiments for two different algorithms are presented. In this experiment, was used underwater speaker and remote control vessel as the target. For stationary measurement, underwater speaker placed one by one in line position with angle variation (45, 64, 90, 116, 135 degree). The result of the localizing of azimuth are shown below in the degrees unit.

In this experiment, it was obtained good results of azimuth sound localization. Using four kinds of frequency variation, was obtained azimuth value from each frequency at a certain angle of sound source. The azimuth from each frequency averaged to determine the deviation readings thus the effect of kinds frequency can be known. Shown at Table 1, that PHAT algorithm achieved better accurate reading than CC with smallest error 2.34 degrees and biggest error 9.48 degrees. Meanwhile, CC algorithm achieved smallest error 1.5 degrees and biggest error 25.1 degrees. From the result, also show that PHAT algorithm achieved better precision reading than CC with smallest deviation 0 degrees and biggest error 13.81 degrees. Meanwhile, CC algorithm achieved smallest deviation 0 degrees and biggest error 28.8 degrees. Overall, PHAT has better performance than CC because more accurate and precise in tracking a sound source azimuth. Result of measurement different than simulation, where CC and PHAT cannot obtain actual time delay in SNR 25 dB in the real environment. Inaccurate result be expected because there is some correlation among the signals from the same sound source, we can estimate TDOA value by calculating the

correlation function between the signals received by different hydrophones. However, in the real environment, due to the impact of noise and reverberation, the maximum peak of the correlation function will be weakened and sometimes there are multiple peaks, which have caused difficulties in the actual peak detection [5]. The GCC method is to weight signals in the power spectral domain to highlight the relevant signal components and suppress the components disturbed by noise so as to make the peak of relevant function at the time delay more prominent. However, due to the presence of reverberation, there are multiple echo components in the signals, so the cross-correlation function calculated will include the peaks formed by the direct wave and reflected wave that will make the time-delay estimation detection difficult [5].

After that, was done localization of moving sound source by using remote control vessel. From the stationary experiment result, achieved that PHAT algorithm better than CC to localize azimuth of sound source at reverberant, scattering, and reflecting medium. Thus, chosen PHAT algorithm to localize azimuth of moving sound source.

Combination between pair H1 - H2, H2 - H3, H3 - H4 used in this experiment then tracked azimuth in each pair averaged so achieved azimuth value from hydrophone array. Result of the tracking shown in Figure 5(b) where estimated azimuth is in the form of oscillating data. The oscillation occurs because the source of the move resulted in the medium through which it passes also change in volume of the medium. This changes influence direction of propagation so the time delay also changed. But, from the oscillated result can attracted a linear regression so achieved linear result from the tracking. The linear regression is azimuth estimated from hydrophone array. Figure. 5(b) show that PHAT algorithm achieved good tracking, with smallest error 0 degrees and biggest error 6.8 degrees. In this experiment, not occur Doppler effect because ratio between source speed and sound speed in the water is too high so frequency shifting too small.

5. Conclusion

The two algorithms presented in this paper worked well in the simulated environment, where there was added the noise signals were assumed to be uncorrelated and to have a zero mean value. However, PHAT method achieved the best results in the experimental part and CC method had the worst results in experimental part. Our results show also that there are some different between stationary and moving sound source experiment. Where, in stationary experiment achieved that the tracking more accurate and precision than measurement in moving sound condition. The study indicates that hydrophone arrays with CC and PHAT can be used for moving target positioning with satisfying results.

6. References

- [1] Urlick RJ. 1983. "Principles of Underwater Sound 3rd Edition" New York: Mc Graw Hill.
- [2] Khaddour H 2011. "A Comparison of Algorithms of Sound Source Localization Based on Time Delay Estimation" *Elektrorevue*, ISSN 1213-1539, **2**, p. 1, April
- [3] Knapp CH dan Carter GC 1976. *IEEE Transactions on Acoustics* **24**, p 4
- [4] Zetterberg V, Pettersson MI and Claesson I 2005. *Proceedings of OCEANS 2005 MTS/IEEE*
- [5] Lin C, Yongchun L, Fancheng K. 2011. *Procedia Engineering*, **15**, pp. 4912-4919, 2011.