Sound Source Localization by a Single Linear Moving Microphone

Meifal Rusli,[a,*] Lovely Son,[a] Mulyadi Bur,[a] and Agus Arisman,[a]

[a)]*Mechanical Engineering Department, Faculty of Engineering, Andalas University, Padang 25163-Indonesia*

*Corresponding author: meifal@ft.unand.ac.id

Paper History

Received: 25-October-2014
Accepted: 15-November-2014

ABSTRACT

This paper discusses about the implementation of a single microphone that is moving in linear track to substitute some channels of linear microphone array and a static microphone as a reference in beamforming method of sound source localization. The single microphone moves at constant velocity from a reference point. All recorded data from the moving and the reference microphone are split into several data represent each discrete microphone position. By this method, the sound localization system is modified from single microphone into artificial linear microphone array. Time delay for each artificial linear microphone is obtained by cross-correlation function between signal from moving microphone and signal from reference microphone. Time Domain beamforming method is performed by the delay-and-sum algorithm for stationary microphone. It is found that the methods can predict the direction of sound source. The shorter track and higher microphone speed can reduce the possibility of aliasing at high frequency sound. However, if the microphone moves closer to the sound source at higher speed, the possibility distortion of angle estimation increase caused by the Doppler effect. Furthermore, implementation of cross-correlation in beamforming can minimize the effect of random noise to predict the sound source direction.

KEY WORDS: *Sound Source Localization; Time Domain Beamforming; Microphone Array; Moving Microphone*
.

NOMENCLATURE

B	Number of beam
M	Number of Microphone
$P(t)$	Microphone Signal
$P_m(t)$	Beamforming Output Signal
Δ_m	Time Delay of Microphone Signal
ψ_b	Direction angle
w_m	Weighting Function

1.0 INTRODUCTION

Sound source localization and visualization are complex task that most acoustics engineers face today. Some standard and highly functional methods based on microphone arrays are used throughout many industries to find and to analyze the noise source. In general, the methods fall into three categories: near-field acoustic holography, beamforming, and inverse methods. The methods were developed using some microphones and data acquisition channels. Most of them use 20 channel microphones or more.[1].

Beamforming is a process to perform spatial filtering, where the response of a sensors array is made sensitive to signals coming from a specific direction while signals from other directions are attenuated. Various methods of beamforming have been developed in time and frequency domain, depending on the processing speed and the signal types. Beamforming is typically referred to in SONAR [2] and RADAR [3]. Recently, its applications extend into seismic, medical ultrasonic imaging, and various other applications [4].

The beamforming technique was first developed for submarines and environmental applications. In the far field, sound waves hitting the microphone array are planar waves. Under these conditions, it is possible to propagate the measured sound field directly to the test object. All microphone signals measured by the beamforming array are added together, taking into account the delay corresponding to the propagation distance. The pressure can be calculated at any point in front of the array, allowing propagation to any kind of surface. Beamforming is sometimes called “sum and delay,” since it considers the relative delay of sound waves reaching different microphone positions. Beamforming requires that all data be measured simultaneously. It is typically done with a measurement system of 40 channels or more [1].

Recently, sound source localization by beamforming has been addressed to the problem of tracking moving sources [5-6]. This paper will discuss the implementation of a single microphone that is moving in linear track to substitute some channels of linear microphone array. The signal from moving microphone is compared with a static microphone as reference signal. Beamforming is performed using cross-correlation function between signal from moving microphone and reference signal from static microphone.

2.0 TIME DOMAIN BEAMFORMING ON MOVING MICROPHONE

The visualization of sound sources on arbitrary objects can be of great utility to locate, compare and demonstrate the causes of noise. To do so, an object can be modeled by placing a grid of hypothetical point monopoles on its surface. Focusing each of these sources by running a beamforming algorithm on measurements done with a microphone array, the influence of a source to the overall noise can be determined. To replace microphone array, a moving microphone and a static reference microphone is possibly developed. The idea is inspired by implementation of a moving subject like robot to identify or localize sound source objects. As preliminary research, we focus on locating the sound sources by adapting a classic time-domain beamforming method.

2.1. Beamforming On Static Microphone Array

The approach used in this paper is based on delay-and-sum beamforming that assuming point monopoles and spherical wave propagation. The beamformers output signal $p(t)$ is obtained by [6]

$$p(t) = \frac{1}{M}\sum_{m=1}^{M} w_m p_m\left(t+\Delta_m\right) \tag{1}$$

That is summation of M microphone signals $p_m(t)$ after each of them is delayed by the time Δ_m and weighted by the factor w_m. The delay times, are usually calculated with

$$\Delta_m = \Delta t_{nm} - \Delta t_{n0} \tag{2}$$

where Δt_{nm} is the sound propagation time between source and microphone and Δt_{n0} is a reference time between the source and the array center. Because Δt_{n0} is equal for all microphone signals it only changes the absolute time frame of the output signal. This is irrelevant for most evaluations that can be run on the output signal, so this offset can be neglected. The weighting factors, that assure that all microphones have the same impact on the result no matter how far away they are from the source, are calculated with

$$w_m = \frac{r_{nm}}{r_{n0}} \tag{3}$$

Im many cases, spatial weighting function w_m is defined as a window function to control main-lobe width and side-lobe magnitude.

To form a beam in direction ψ_b, the discrete delay required to bring the output of the M sensors into time alignment is equal to q_{mb}, which are integer multiplies of the sample spacing t_s, if the beam index b, is an integer (b=0, ±1, ±2, ±3, …) then the beam directions that we can realize are those for which

$$\Delta t_{nm} \sin\psi_b = b \tag{4}$$

and effectively quantizes the beam steering directions ψ_b, such that

$$\psi_b = \sin^{-1}\left(b\cdot\frac{1}{\Delta t_{nm}}\right) \tag{5}$$

2.2. Beamforming On Moving Microphone

When a fixed source is measured with a moving microphone over a period of time, the distance between source and microphone is no longer constant, but as a function of time. Figure 1 shows the consequences this has on phase and amplitude of the recorded signal by taking two events that take place when moving microphone is approaching source at the times t_1 and t_2 and observe how the sound propagates through the fixed medium to the microphone. Because the sound emitted at t_2 needs less time to reach the microphone than the signal emitted at t_1, the waveform between these events is compressed resulting in an increase of frequency known as the Doppler Effect [6]. Likewise, because the distance decreases, the amplitude increases over time. Conversely, when the microphone moves away from the sound source, the sound emitted at t_2 needs more time to reach the microphone than the signal emitted at t_1, the waveform between these events is expanded resulting in a decrease of frequency.

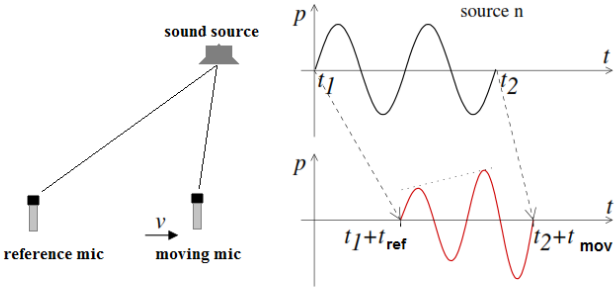

Figure 1. Doppler’s effect when microphone moves closer to the sound source

In the case of moving microphone, two microphones are implemented to record the signals. One microphone in static position acts as reference signal, and another one moves in linear track with a constant speed. Both microphones record the signal in a same time, and have the same dimensions. Both signal then are divided in accordance with the desired number or microphone distance of artificial microphones arrays of beamforming to substitute some channels of linear microphone array. Delay time of each artificial microphone array is obtained from cross-correlation of the signal from moving microphone and the signal from reference microphone at the same time measurement. Then, beamforming is performed by cross-correlation function between signal from moving microphone and reference.

A simple time domain beamforming structure is the delay-and-sum beamformer which implements in Equation (1) and is shown in Figure 2. The structure is repeated *B* times, where *B* is the number of beams desired. The cross-correlation signals are converted to a digital signal at a frequency greater than the Nyquist rate to achieve acceptable beam pattern performance. The signals are then delayed appropriately and the weighting function applied and effectively summed. If same set of weights are to be used for all steering directions, as is often the case, the weights can be applied directly at the sensor outputs, before the delay sum operation.

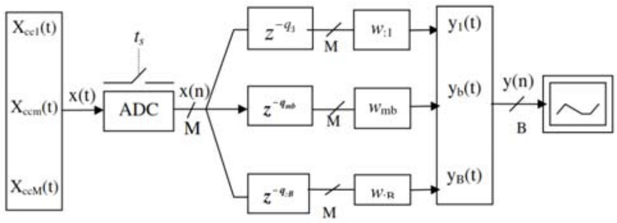

Figure 2. A simple Time domain-Delay sum beamformer using cross-correlation signal

3.0. RESULT AND DISCUSSION

Simulation is performed using configuration shown in Figure 3. The reference microphone is set at the beginning of the track, and the moving microphone moves on the microphone track with a constant speed. First, Microphone track is 2 meters length and sound source is allocated at 1 meter from the center of microphone track in x-axis and 4 meters in y-axis. It means that the sound source position is equal to $\tan^{-1}(4/1) = 75.9^{o}$ from x-axis of microphone track. Noise source is set in coordinate (4 meter,4 meter) from the center of the microphone track.

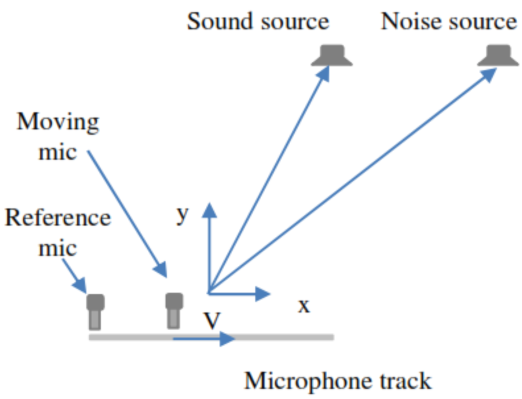

Figure 3. Simulation configuration

First, The simulation is performed with a single sound source and without noise. The beamforming is applied to predict the sound location i.e sound source direction from the center of microphone track with variation of sound frequency and microphone speed. Artificial array is set 1 cm of microphone distance, and frequency signal is varied from 500 to 4000 Hz.

When moving microphone moves at 1 m/s, beamforming is able to detect the position of sound source at 75^{0} at frequency from 500 Hz to 1000 Hz (Figure 4). However aliasing occurs at frequency 2000 Hz or more, where total time delay between artificial microphones along the track is more than a wave length of the sound.

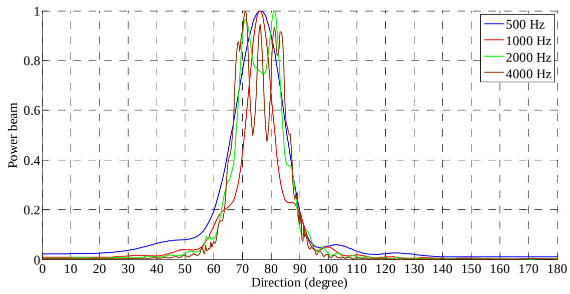

Figure 4. Beam power with microphone speed 1 m/s and 2 meters track length

When the microphone moves at 5 m/s of speed (Figure 5), beamforming is able to detect the sound source at frequency up to 2000 Hz. Aliasing still occurs at frequency 4000 Hz. However, the location prediction is shifted to lower than 75^{o}. It is possibly caused by the Doppler's effect where the frequency and phase is changed when microphone move closer to the sound source in higher speed. Position estimation at higher frequency is distorted to the lower angle of direction.

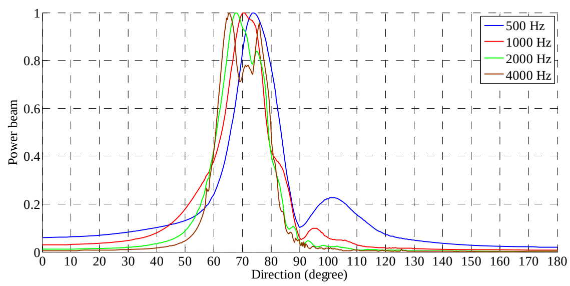

Figure 5. Beam power with microphone speed 5 m/s and 2 meters track length

Furthermore, when the microphone moves at 10 m/s speed (Figure 6), beamforming is able to detect the sound source at frequency up to 2000 Hz. Aliasing still occurs at frequency 4000 Hz. However, the location prediction is shifted to the bigger angle more than 75^{o}. It is same with the previous situation, the doppler's effect causes frequency and phase change when microphone move closer to the sound source at higher speed.

Position estimation at higher frequency is distorted to lower angle of direction

The second condition is the microphone track has 0.5 meters length and sound source is allocated at 1 meter from the center of microphone track in x-axis and 4 meters in y-axis. It means that the sound source position is equal to $\tan^{-1}(4/1) = 75.9^{o}$ from x-axis of microphone track. Artificial array is set 1 cm of microphone distance, and frequency signal is varied from 500 to 4000 Hz.

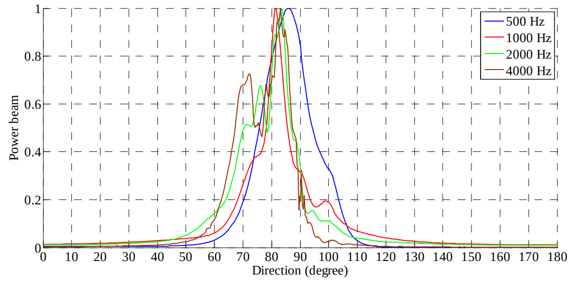

Figure 6: Beam power with microphone speed 10 m/s and 2 meters track length

When the microphone moves at 1 m/s, beamforming is able to detect the position of sound source at 73^{0} - 74^{0} at frequency from 500 Hz to 4000 Hz (Figure 7), without any aliasing. The same condition occurs when the microphone speed is 5 m/s (Figure 8). However, estimation of direction at frequency signal is shifted to higher angle. On the other hand, when microphone speed is 10 m/s (Figure 9), the direction of sound source with higher frequency is able to be estimated well. However, signal at lower frequency is predicted at higher angle.

Above explanation shows that shorter track and higher microphone speed can reduce the possibility of aliasing even. However, higher speed of microphone increases the possibility distortion of angle estimation caused by the Doppler's effect.

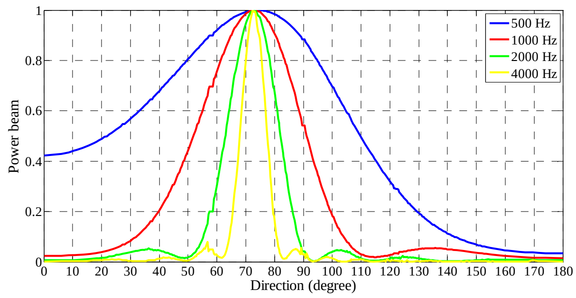

Figure 7. Beam power with microphone speed 1 m/s and 0.5 meters track length

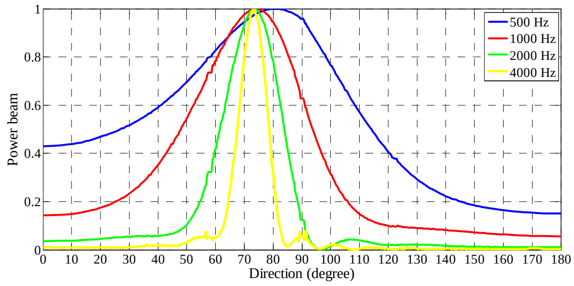

Figure 8: Beam power with microphone speed 5 m/s and 0.5 meters track length

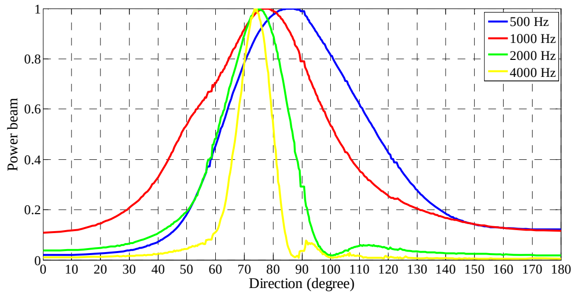

Figure 9: Beam power with microphone speed 10 m/s and 0.5 meters track length

Next, the random noise is applied to the system. Signal to noise ratio (S/N) is varied from 8 to 0.125. Figure 10 shows the example of the signals generated by sound source and noise source. As a result, Figure 11 shows, when sound signal at frequency 1000 Hz and microphone moves on 2 meters track length at 1 m/s and 2 m/s microphone speed, beamforming is able detect the sound location although S/N ratio reach 8 (Figure 11 and Figure 12)

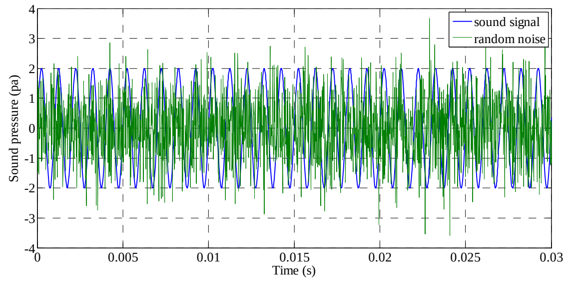

Figure 10: Sound and noise signal with S/N 0.5

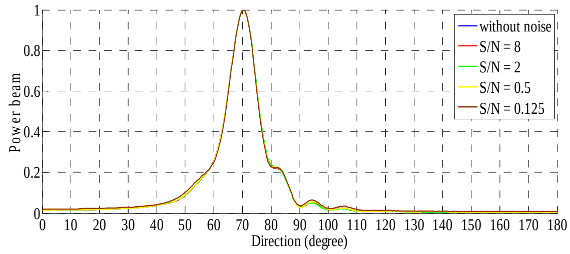

Figure 11: Beamforming of 1000 Hz signal, microphone speed 2 m/s with some levels of noise

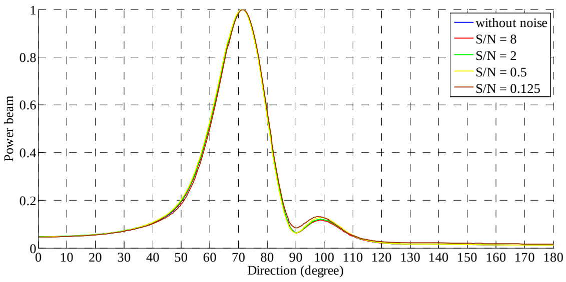

Figure 12: Beamforming of 1000 Hz signal, microphone speed 5 m/s with some levels of noise

The implementation of cross-correlation in beamforming can minimize the effect of random noise to predict the sound source direction.

4.0. CONCLUSSION

This paper has discussed about the implementation of a single moving microphone and another static microphone acts as reference in sound source localization. By this method the sound localization system is modified from single microphone into artificial time delayed linear microphone array. Time delay for each artificial time delayed linear microphone array is obtained from cross-correlation between signal moving microphone and signal from reference static microphone. The Time Domain beamforming method for moving sources of sound is performed by the delay-and-sum algorithm for stationary microphone. It is found that the methods can predict the direction of sound source. The shorter track and higher microphone speed can reduce the possibility of aliasing even at high frequency sound. However, higher speed of microphone increases the possibility distortion of angle estimation by the Doppler's effect. Furthermore, implementation of cross-correlation in beamforming can minimize the effect of random noise to predict the sound source direction.

ACKNOWLEDGEMENTS

This paper is part of the research funded by Hibah Fundamental of BOPTN - Andalas University by contract number 07/UN.16/PL/D-FD/2014. The authors gratefully acknowledge for this financial support.

REFERENCE

1. Lanslots, J.,Deblawe, F.Janssens, K. (2010). Selecting Sound Source Localization Techniques for Industrial Applications, Sound & Vibration, June 2010, pp.6-9
2. Curtis T. and Ward R., "*Digital beamforming for SONAR systems*", IEE Proceedings, vol. 127, August 1980.
3. Barton P., (1980) "*Digital beamforming for RADAR*", IEE Proceedings, vol. 127, August 1980
4. Donnell M. O., "*Applications of VLSI circuits to medical imaging*", Proceedings of the IEEE, vol. 76 September 1988.
5. Zechel, g., zeibig, a. Beitelschmidt, m., Time-Domain Beamforming On Moving Objects With Known Trajectories, 3rd Berlin Beamforming Conference, 24-25 February 2010
6. Ginn, B., Gomes, J.,Hald, J. Recent advances in Rail Vehicle Moving Source Beamforming, Internoise, 15-18 September 2013, Innsbrug-Austria