Microphones' Directivity for the Localization of Sound Sources

Piervincenzo Rizzo[1,*] Mahdi Tajari[2], Antonino Spada[3]

[1] Assistant Professor, Laboratory for NDE and SHM Studies, Department of Civil and Environmental Engineering, University of Pittsburgh, Pittsburgh (PA), USA.

[2] former graduate student

[3] Postdoctoral Scholar, Department of Structural and Geotechnical Engineering, University of Palermo, Italy.

[*] Corresponding author

ABSTRACT

In a recent paper [P. Rizzo, G. Bordoni, A. Marzani, and J. Vipperman, "Localization of Sound Sources by Means of Unidirectional Microphones, Meas. Sci. Tech., 20, 055202 (12pp), 2009] the proof-of-concept of an approach for the localization of acoustic sources was presented. The method relies on the use of unidirectional microphones and amplitude-based signals' features to extract information about the direction of the incoming sound. By intersecting the directions identified by a pair of microphones, the position of the emitting source can be identified.

In this paper we expand the work presented previously by assessing the effectiveness of the approach for the localization of an acoustic source in an indoor setting. As the method relies on the accurate knowledge of the microphones directivity, analytical expression of the acoustic sensors polar pattern were derived by testing them in an anechoic chamber. Then an experiment was conducted in an empty laboratory by using an array of three unidirectional microphones. The ability to locate the position of a commercial speaker placed at different positions in the room is discussed.

The objective of this study is to propose a valid alternative to the common application of spaced arrays and therefore to introduce a new generation of reduced size sound detectors and localizers. The ability of the proposed methodology to locate the position of a commercial speaker placed at different positions in the room was evaluated and compared to the accuracy provided by a conventional time delay estimate algorithm.

Keywords: sound source localization; unidirectional microphones, polar directivity pattern, feature extraction, time-delay estimate.

1. INTRODUCTION

Acoustic sound source localization (SSL) has widespread applications such as teleconference tracking of speakers [1], human-robot interaction [2], and air/ground objects recognition [3]. Two approaches can be pursued for SSL: active and passive. An active system sends out preset signals to the potential target and compares it with the echo signal, similar to how a bat locates its prey using ultrasonic pings. Sonar- and radar-like systems belong to this category. Passive localization systems only receive signals generated by the targets.

Conventional SSL methods use microphones arrays made of physically separated microphones so that the variation in amplitude, delay time, and reverberation at different locations can be exploited. These methods can be loosely divided into two main categories: steerable beamforming and time delay estimation (TDE). In beamforming a full directional scan is undertaken to determine the source direction from the maximum signal power [7, 10]. In TDE the time offset between the received signals for a single sound source is calculated and it represents the difference between the travel times for the signal to each microphone position [4]. For both methods, omnidirectional microphones (OMs) are used. OMs are sound detectors whose output does not depend on the direction of the sound propagation, and the response is generally considered to be a perfect sphere.

In a recent paper Rizzo et al. [5] proposed a passive scheme based on UMs and amplitude-based features. UMs are sound transducers that possess a strong dependence on the incidence angle of the acoustic source. Such directivity is described by the polar directivity pattern (PDP). Although directional microphones provide significantly superior sound quality and can be utilized for adaptive noise cancellation and noise discrimination, only few researchers used directional microphones for SSL [6-8].

In this paper, the work of ref. [5] has been expanded. The directivity of UMs is exploited to estimate the direction of the incoming sound. By intersecting the directions identified by microphone's pairs, the position of the source is determined. The effectiveness of this novel localization approach is tested in an indoor setting. The results are then compared to those obtained by a conventional TDE approach. The objective of the effort is to propose a new generation of sound detectors and localizers that can minimize the spacing between the microphone elements..

2. ANALYTICAL POLAR PATTERN

Three Audio-Technica unidirectional microphones were tested in an anechoic chamber. A portable National Instruments PXI unit running under LabView, a commercial loudspeaker, and a motorized rotational stage were used to perform the experiment.

Five-cycle Gaussian-windowed sinusoidal tonebursts from 400 Hz to 4 kHz at 200 Hz step were sent to the speaker. The detected signals were sampled at 32.768 kHz and post-processed to extract the peak to peak (ppk) amplitude. This feature was calculated at each acquisition angle, resulting in the creation of the empirical PDP at various frequencies. Figure 1 shows the experimental PDP associated with the ppk amplitude of the detected pulses at four different frequencies. As expected, the shape of the PDP changes with the frequencies and few lobes appear.

In order to apply the semi-analytical approach described in ref. [5], the equation of the PDP is necessary. The values of the normalized ppk at 1200 Hz as a function of the azimuth angle were calculated by using the MATLAB function fit and the Fourier polynomial type. The equation that best fits was found to be the 7th order Fourier equation:

$$M = f(\vartheta,a,b) = a_0 + \sum_{n=1}^{7} a_n \cos(n\vartheta) + \sum_{n=1}^{7} b_n \sin(n\vartheta) \qquad (1)$$
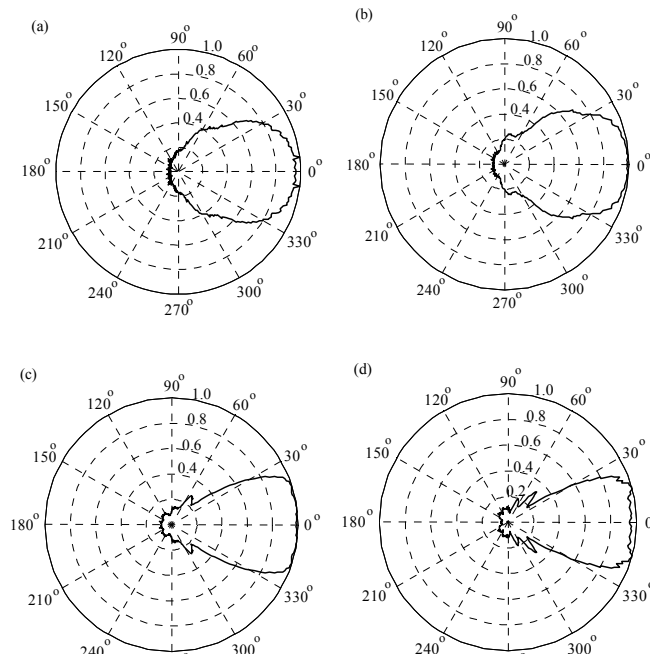

Figure 1 – Anechoic chamber test. Polar directivity pattern associated with the feature of the peak-to-peak amplitude at frequencies a) 800 Hz; b) 1800 Hz; c) 2800 Hz; d) 3800 Hz.

The empirical coefficients a_0, a_n, and b_n are summarized in Table 1.

n	Peak-to-peak amplitude	
	a_n	b_n
0	0.335110476	
1	0.382313665	0.045828754
2	0.178997026	0.043539139
3	0.076988592	0.028795045
4	0.031498699	0.016287085
5	-0.006252076	-0.00424536
6	-0.014339409	-0.01247091
7	-0.003518177	-0.00388666

Table 1 – Empirical coefficients of the Fourier 7 polar directivity pattern for one of the UMs used in this study. Excitation frequency equal to 1200 Hz.

3. SETUP

An array of three UMs was deployed inside an empty laboratory about 7 m long and 6.4 m wide. Photos of the in-house software created for the experiment, of the speaker and the array, and their relative distances are presented in Fig. 2. The photos show the presence of an OM. The role of this microphone is not part of the present research and its presence in the experiments will not be discussed further.

The hardware/software system utilized to conduct this test was identical to the one used for the anechoic chamber tests. A 5-cycles Gaussian-windowed sinusoidal toneburst ranging from 400 Hz to 4 kHz at 200 Hz step was sent to the speaker. The tests were conducted such that the position of the speaker changed.
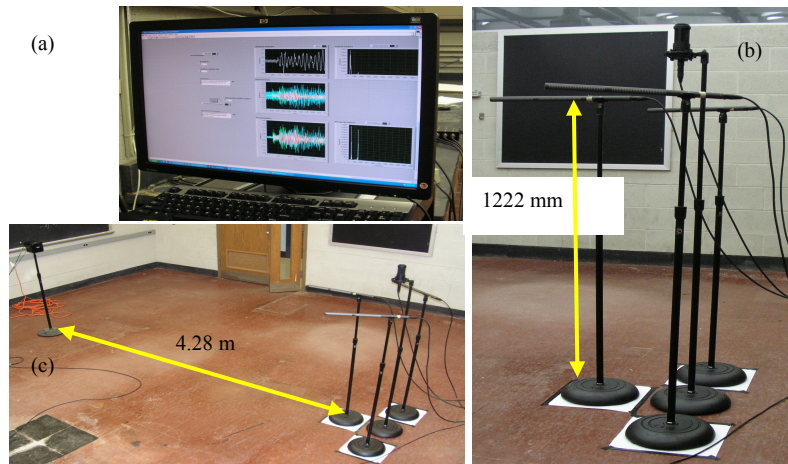

Figure 2 – (a) Front panel of the software used. (b) Photo of the microphone array. (c) Photo of the room and the array-speaker distance

Figure 3 shows the positions (S1, …, S8) of the speaker with respect to the array location. The arrangement of the UMs is 30º/0º/-30º, where the first angle is relative to microphone C, the second angle to microphone A, and the last angle to microphone B.

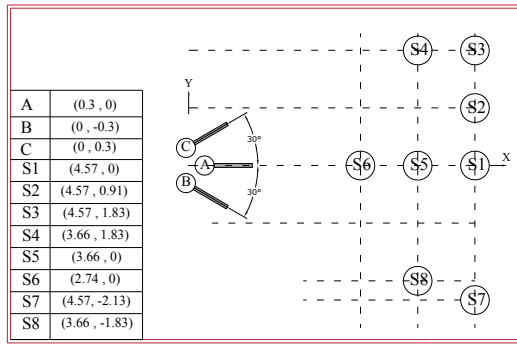|	
A	(0.3 , 0)
B	(0 , -0.3)
C	(0 , 0.3)
S1	(4.57 , 0)
S2	(4.57 , 0.91)
S3	(4.57 , 1.83)
S4	(3.66 , 1.83)
S5	(3.66 , 0)
S6	(2.74 , 0)
S7	(4.57 , -2.13)
S8	(3.66 , -1.83)

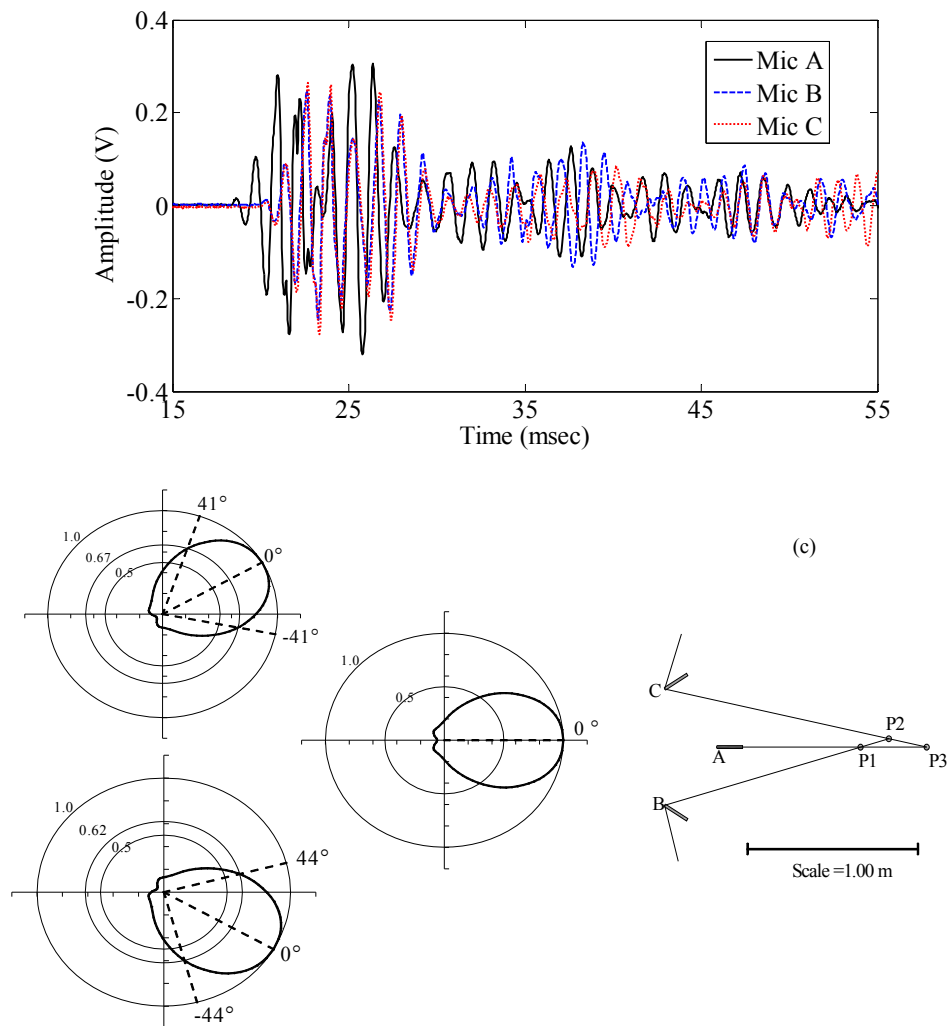Figure 3 - Schematic of microphones' and loudspeaker's position.

Figure 4 – Indoor test. Speaker at location S1. Source localization by using the feature of the peak-to-peak amplitude of the 1200 Hz toneburst. (a) Time waveforms detected by the array. (b) Experimental polar directivity patterns given by Fourier 3. (c) Localization of the sound source.

4. RESULTS

The position of the loudspeaker was determined by exploiting the empirical PDP. The results from recording pulses at 1200 Hz are presented herein.

Figure 4a shows the time waveforms captured by the array when the 1200 Hz toneburst was generated. For each detected signal, the main burst is followed by smaller pulses originated by the reflection from the room's boundaries. The ppk amplitudes of the main bursts were equal to 0.625 V, 0.494 V, and 0.542 V for sensors A, B, and C, respectively. The position of microphone C was symmetric with respect to the speaker. This is confirmed by the simultaneous arrival of the pulse. However the ppk amplitudes recorded by the microphones were about 9% different. This might introduce some level of inaccuracy in the determination of the incoming sound direction. To mitigate the effects associated with different sensitivities that may arise from manufacturing differences, a calibration factor computed empirically was applied to the successive analyses.

As a normalized value is necessary to use the experimental polar pattern, we decided to normalize these amplitude values with respect to the amplitude recorded by microphone A (0.625 Volts). The normalized values of the microphone output levels associated with the ppk from detectors A, B, and C were then equal to 1.00, 0.670, and 0.618, respectively. Such values are represented by the circles superimposed in the corresponding PDP of Fig. 4b. The potential incidence angles identified by each microphone were $+ 71°$ and $-11°$ for microphone C, $0°$ for microphone A, and $+ 14°$ and $– 74°$ for microphone B.

The algorithm then traced the lines departing from the microphone coordinate along the direction of these potential incidence angles. The lines, shown in Fig. 4c, intersect into three points P1, P2 and P3 associated with the intersections of microphone pairs A-B, B-C, and A-C, respectively. These points are solutions found by the algorithm, and their planar coordinates are (x = 1.18 m, y = 0.0 m) for microphone pair A-B, (x = 1.37 m, y = 0.05 m) for microphone pair A-C, and (x = 1.62 m, y = 0.0 m) for microphone pair B-C. Ideally, such points should be identical and collapse into a single point coincident with the speaker position. This point shall represent the location
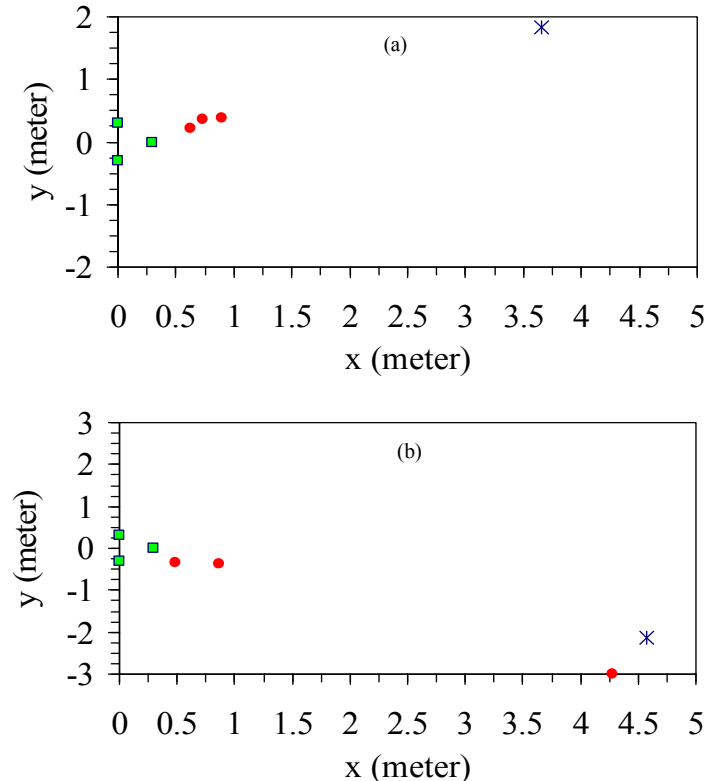

Figure 5 – SSL when the speaker was located at position (a) S4 and (b) S7. The green squares identify the position of the microphones, the asterisk represents the true position of the speaker, and the red circles show the position of the acoustic source found by the algorithm.

of the sound source. The locations P1, P2 and P3 underestimate the true position of the speaker (x = 4.57 m, y = 0.0 m).

The same approach described above was used to localize the speaker placed in different positions in the laboratory according to the scheme illustrated in Fig. 3. Figure 5a compares the results of the ppk feature applied to the 1200 Hz pulse recorded by the microphone array when the speaker was positioned at S4. The red circles represent the solutions found, the asterisk the position of the speaker, and the green square indicates the position of the acoustic sensors.

Similarly, Figure 5b compares the results of the localization process when the speaker was located at position S7 and the feature of ppk applied to the 1200 Hz propagating sound was used. Although positions S3 and S7 were symmetric with respect to the array, the localization of position S7 was more accurate than S3.

The results from these two cases are summarized in Table 2, which lists the results associated with all eight speaker positions.

Speaker position	True coordinates	Solution Mic A-B	Solution Mic B-C	Solution Mic A-C
S1	(4.57 , 0)	(1.14 , 0)	(1.32 , 0.04)	(1.54 , 0)
S2	(4.57 , 0.91)	No intersection	(1.18 , 0.25)	(0.85 , 0.27)
S3	(4.57 , 1.83)	No intersection	(0.70 , 0.24)	(0.62 , 0.25)
S4	(3.66 , 1.83)	No intersection	(0.67 , 0.35)	(0.93 , 0.40)
S5	(3.66 , 0)	(3.60 , 0)	(8.30 , 0.40)	
S6	(2.74 , 0)	No Intersection		
S7	(4.57 , -2.13)	(0.70 , -0.34)	(0.88 , -0.36)	(0.35 , 0.04)
S8	(3.66 , -1.83)	(6.09 , -3.81)	(1.64 , -1.25)	(0.36 , -0.04)

Table 2 – Summary of the localization process using the peak-to-peak amplitude of the sound propagating at 1200 Hz using the Fourier 7 fitting function.

4. COMPARISON WITH CONVENTIONAL APPROACH

For comparative purposes, the localization of the speaker was also attained by comparing the differences among the time of arrival of the waves at the microphones. The authors have used this approach in the nondestructive evaluation method called acoustic emission [9]. The time of arrival ts,i of the wave at each microphone (s) can be written as:

$$t_{s,i} = \frac{|\mathbf{r}_{s,i} - \mathbf{r}|}{v_p} + T = \frac{\sqrt{(x_{s,i} - x)^2 + (y_{s,i} - y)^2}}{v_p} + T \qquad (2)$$

where r = $(x_{s,i}, y_{s,i})$ is the position vectors of microphones (i = 1,2,3), r = (x, y) is the position vector of the source, and v_p is the velocity of sound assumed to be equal to 343 m/s. T identifies the instant at which the sound was emitted.

In Eq. (2) the location of the microphones is known and the arrival time $t_{s,i}$ is measured on the base of a fixed threshold for waveforms' amplitude (0.015 Volts). A non-linear least squares approach was used to solve the unknowns. The method uses an iteration algorithm where the estimation of the solution is computed by minimizing the errors of the unknown parameters. The results are summarized in Table 3. The results show an overall fairly good agreement with the true ones. However in some cases differences on the order of 30 cm and in the S8 case on the order of about 80 cm can be recognized. These values are too high if we compare them and the relative short distance of the speaker with respect to the microphones. A possible explanation can be associable with the choice of a fixed value of the threshold that provides not exact arrival times. Summarizing, this method gives only in some cases good results, but they depends on the microphones' orientation, sensitivity and type, and probably on the threshold chosen. Therefore it is not much reliable for practical applications.

Speaker position	True coordinates	Solution found	
S1	(4.57 , 0)	(4.58	-0.04)
S2	(4.57 , 0.91)	(4.57	1.20)
S3	(4.57 , 1.83)	(4.66	1.62)
S4	(3.66 , 1.83)	(3.68	1.81)
S5	(3.66 , 0)	(3.65	0)
S6	(2.74 , 0)	(2.73	-0.35)
S7	(4.57, -2.13)	(4.59	-2.14)
S8	(3.66, -1.83)	(3.13	-2.60)

Table 3 – Localization of the speaker location by computing the time delay estimate and the Levenberg-Marquardt method.

5. CONCLUSIONS

This paper presents the results of a sound source localization scheme based on the use of UMs and amplitude-based sound features to determine the planar position of an acoustic pulse source. In the method proposed here the PDP is exploited to associate the detector output to a certain sound direction. By intersecting the directions provided by microphones' pairs, the position of the emitter is identified.

In this study, the directivity of the microphones was analytically determined in an anechoic chamber. The PDP associated with one microphone was considered representative of all sensors. Then the performance of the algorithm to localize the position of a speaker in a closed environment was tested and compared to the results found by means of a conventional time-delay estimate method. The outcomes of this study make it obvious that the accuracy of the localization algorithm is closely related to the precision of the experimental polar pattern. As the method relies upon amplitude-based features, the sensitivity of the hardware system to capture the same amplitude pressure irrespective of the microphone and the data acquisition channel used is critical. Future studies must take into account these aspects by, for instance, introducing calibration factors empirically determined.

ACKNOWLEDGEMENTS

REFERENCES

[1] Y. Huang, J. Benesty, and G.W. Elko, "Passive Acoustic Source Localization for Video Camera Steering, Proc. of the IEEE ICASSP-00 2 II909-II912.

[2] H-D. Kim, J-S. Choi, M. Kim, and C-H. Lee, "Reliable Detection of Sound's Direction for Human Robot Interaction, Proc. of the IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS-04) 3, pp. 2411-2416, 2004.

[3] L. Liu, and D.G. Albert, "Locating a Sound Source in an Urban Environment," 147th ASA Meeting, New York, NY, 2006.

[4] E. Martinson, and A. Schultz, "Discovery of sound sources by an autonomous mobile robot", Auton Robot, 27: 221–237, 2009.

[5] P. Rizzo, G. Bordoni, A. Marzani, and J. Vipperman, "Localization of Sound Sources by Means of Unidirectional Microphones, Meas. Sci. Tech., 20, 055202 (12pp), 2009.

[6] B. Fazenda, F. Gu, A. Ball, L. Guan, "Noise source Localisaton in a Car Environment", 8th Intl. Conf. on Damage Assessment of Structures, 3-5 August 2009, Beijing, China.

[7] B. Fazenda, "Localisation of Sound Sources Using Coincident Microphone Techniques Proc. of the Institute of Acoustics, 29(7), 106-115, 2008.

[8] T. Ohta T, H. Yano, S. Yokoyama, H. Tachibana, "Sound Source Localization by 3-D Sound Intensity Measurement using a 6-channel microphone system Part 2: Application in room acoustics Internoise", 37th Intl. Congr. Exp. Noise Control Engr., 26-29 October 2008, Shanghai, China.

[9] P. Rizzo, A. Spada, S. Degala, and G. Giambanco, "Acoustic Emission Monitoring of Chemically Bonded Anchors", J. of Nondestr. Eval., vol. 29, pp.49-61, 2010.