J. Kotus and G. Szwoch, "Localization of sound sources with dual acoustic vector sensor," 2019 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), Poznan, Poland, 2019, pp. 44-49.
doi: 10.23919/SPA.2019.8936724

Localization of sound sources with dual acoustic vector sensor

Józef Kotus, Grzegorz Szwoch
Gdansk University of Technology, Faculty of Electronics, Telecommunications, and Informatics,
Department of Multimedia Systems
80-233 Gdańsk, Narutowicza 11/12
Gdańsk, Poland
joseph@sound.eti.pg.gda.pl

***Abstract*— The aim of the work is to estimate the position of sound sources. The proposed method uses a setup of two acoustic vector sensors (AVS). The intersection of azimuth rays from each AVS should indicate the position of a source. In practice, the result of position estimation using this method is an area rather than a point. This is a result of inaccuracy of the individual sensors, but more importantly, of the influence of a source size. The proposed method was validated in experiments performed in an anechoic room, using a custom-made setup of two sensors built from digital MEMS microphones, for sound source placed at varying distance and angle from the sensors. The paper discusses the observed variations in the measured angle and distance to the source. The obtained results indicate that the proposed method allows for estimation of the source position with satisfactory accuracy. Errors in the position estimation depend on the source size, the distance to the source and the source angle relative to the sensors. Possible application of the proposed method is estimation of the position of moving sound sources, such as road vehicles.**

Keywords- acoustic vector sensor, sound source localization; sound intensity; acoustics

I. INTRODUCTION

Sound source localization is an important aspect, with practical applications for sound source tracking and automatic event detection based on acoustic signals. Direction of arrival (DoA) of a sound signal is often performed by analysis of phase differences in signals recorded by a system of spatially distributed sensors, called a beamformer [1]. An alternative method is based on measurement of sound intensity in orthogonal directions by means of an acoustic vector sensor (AVS) [2] and determining DoA by examining relations between the measured intensity values. Sound intensity is a vector quantity dependent on sound pressure and the particle velocity [3]. An AVS may either measure the velocity directly (a *p-u* probe) [4] or estimate it from a pressure gradient (a *p-p* probe) [5].

In previous works, the authors researched applications of AVS to DoA estimation in various scenarios, including tracking and classification of sound sources [6], localization of multiple sound sources [7], detection of security threats [8] and estimation of vehicles speed [9]. The authors also proposed a custom AVS designs [10, 11] and a calibration method [12].

A single AVS is only able to measure the azimuth of a sound source. Measurement of distance from the sound source is not possible, hence estimation of the source position also cannot be performed. However, with a setup of two sensors separated by some distance, it is possible (at least in theory) to calculate the source position by finding an intersection of rays, defined by the azimuth measured with both sensors.

It should be noted that contrary to beamformers that measure phase differences between pressure signals [13], the proposed method is based on the intensity signals, and the position estimation is based on the azimuth values measured independently by each AVS. Accuracy of this method is therefore determined by the accuracy of DoA estimation performed by each AVS and by alignment of the sensors with each other. However, the size of a sound source also plays an important role.

This paper presents experiments related to estimation of position of a sound source, performed in an anechoic room with a setup of two custom sensors constructed from digital MEMS microphones. Results obtained for varying distance between the source and the sensors are presented and discussed. The position of sound source was fixed during each measurement. The most important issue presented in this paper is influence of the sound source size on the accuracy of the position estimation.

II. METHODS

A. Acoustic vector sensor

The acoustic vector sensor (AVS) used in the experiments consists of two intensity probes placed on the orthogonal axes of the coordinate system. Each intensity probe is a *p-p* system of two pressure sensors (microphones) placed on the axis at an equal distance from the system origin. Averaged pressure p, measured at the origin point, is equal to the mean of pressure signals p_1, p_2. Acoustic particle velocity $\mathbf{u}$ is a vector oriented along the axis, the magnitude of this vector is computed from the pressure gradient $(p_2 - p_1)$, using a finite difference approximation method [5]. Instantaneous sound intensity is computed as a product of pressure and velocity [3]. The instantaneous intensity is then time-averaged in order to obtain sound intensity I along the measured axis:

$$I(t) = \frac{1}{\rho}\int_0^T p(t)\mathbf{u}(t)\mathrm{d}t = \\ = \frac{1}{2\rho d}\int_0^T \left[\left(p_1(t)+p_2(t)\right)\left(p_1(t)-p_2(t)\right)\right]dt \quad (1)$$

where ρ is the air density, d is distance between microphones, T is integration time. The value of d is typically in the range from 8 to 12 mm and it sets the upper limit of the analyzed frequencies to about 10 kHz. The integration period is usually comparable with the period of the lowest analyzed frequency.

The AVS used in the experiments consists of two intensity probes placed on XY axes. Therefore, sound intensity I_X, I_Y in two parallel directions is measured. The azimuth φ of the incoming sound (DoA) may be computed as:

$$\varphi = \arctan\left(\frac{I_Y}{I_X}\right) = \arctan\left(\frac{\int p_Y u_Y \mathrm{d}t}{\int p_X u_X \mathrm{d}t}\right). \quad (2)$$

It is assumed that the distance d and density ρ are identical for both axes, so all constant factors used for intensity calculation cancel out during the division.

B. Estimation of source position with a dual AVS setup

A single AVS is only able to determine the azimuth of a sound source, which defines a ray connecting the AVS with the source. Provided that the azimuth measured by an AVS is accurate, performing the measurements with two sensors separated by some distance should allow for determination of the source position by finding the intersection of rays from each AVS.

A setup is presented in Fig. 1. Two sensors of the identical construction are placed on the X axis of the common coordinate system, in equal distance s from the system origin. The X axis of each sensor must be aligned with the X axis of the common system. For a sound source at $P = (x, y)$, the azimuth values measured by each sensor are:

$$\varphi_1 = \arctan\left(\frac{y}{x+s}\right) \quad (3)$$

$$\varphi_2 = \arctan\left(\frac{y}{x-s}\right) \quad (4)$$

From (3), (4), the position of the source can be calculated:

$$x = s\frac{\tan\varphi_2 + \tan\varphi_1}{\tan\varphi_2 - \tan\varphi_1} \quad (5)$$

$$y = \tan\varphi_1\left(x+s\right) = 2s\frac{\tan\varphi_1\tan\varphi_2}{\tan\varphi_2 - \tan\varphi_1} \quad (6)$$

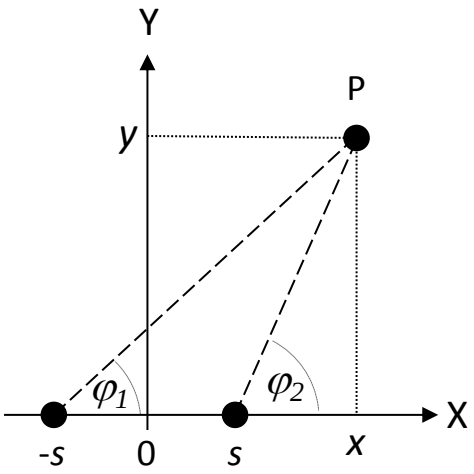

Figure 1. Illustraton of measuring the position P of a sound source with two sensors placed at a distance 2s

The choice of distance between the sensors is made empirically. If the distance is too small, the difference between the measured azimuth values is also small, which leads to numerical errors – the denominator in Eq. (5-6) is small. If the distance is too large, the difference in sound arrival time becomes an important issue, also construction of the setup may be impractical. For the measured distance range up to 6 m, it was assumed that the optimal distance is in range 0.3 to 1.0 m, the distance 0.5 m was selected as a starting point.

C. Errors in position estimation

The method described in the previous Section assumes ideal conditions. In real measurements, imperfection of the sensors and character of sound sources introduce errors to the measured results. The main factors affecting the measurement accuracy are discussed below.

- *Imperfections in sensor construction*. Each AVS must ensure that all microphones are placed in equal distance from the center and that axes determined by the microphones are perpendicular. Due to small size of the AVS, some degree of inaccuracy is inevitable. Small differences in sensor placement may result in significant errors in position estimation. Also, there may be differences between parameters of microphones, even if the same type is used. A calibration procedure is necessary in order to compensate differences in amplitude and phase between the microphones.
- *Misalignment of the sensors*. The coordinate systems of both sensors must be aligned so that their X axes overlap. Otherwise, each AVS introduce its own bias to the measured azimuth values. Setting orientation of two small sensors placed at some distance from each other proved to be a difficult problem which has not been fully resolved yet. Currently, an azimuth correction is computed by comparing the obtained results with known position of the source. However, this method is not sufficiently accurate, and a more sophisticated method needs to be developed in the future research.

- *Nature of the measured sound source and the measurement space*. Even if the sensors are properly calibrated and aligned, some measurement error is always present in the result. This is caused not only by sensor parameters but also by the character of a sound source (size, variations in the emitted sound energy, etc.), sound reflections in the measured space, and many others.

Influence of the source size (mainly its width) on the distance estimation accuracy is shown in Fig. 2. Assuming ideal sensors and a point sound source, an intersection of DoA rays (shown as dashed lines) indicates the source position *P*. However, practical sound sources have a nonzero size *d*. With the assumption that the whole surface of the source emits sound energy uniformly, and that the DoA sensors are independent, the following cases may be considered. (1) Both sensors indicate the center point of the source, yielding accurate estimations of the DoA (φ_0) and the distance to the source *R*. (2) Sensors indicate different edges of the source (solid red and dotted blue lines in Fig. 2), which results in the intersection point laying beyond (red point) or ahead (blue point) of the source, but the DoA is detected correctly. (3) Both sensors indicate the same edge of the source (intersection of blue and red lines), yielding correct estimations of both the distance and the DoA. Therefore, as shown in Fig. 2, variations in the estimated source position result from the estimation method and they depend on the source size *d*, the distance *R* and the distance between the sensors. In the experiments, it was assumed that the error distribution is Gaussian, with zero mean and standard deviation σ. Therefore, the measured azimuth is in fact not a ray, but a beam that widens as the distance from the sensor increases. Intersection of the beams from both sensors is therefore not a point, but an area in a shape of quadrilateral which is wider in the direction of the source (Fig. 2). The area covered by this figure becomes larger with increase of: (a) standard deviation of the error, (b) distance from the sensor, and (c) distance between the sensors.

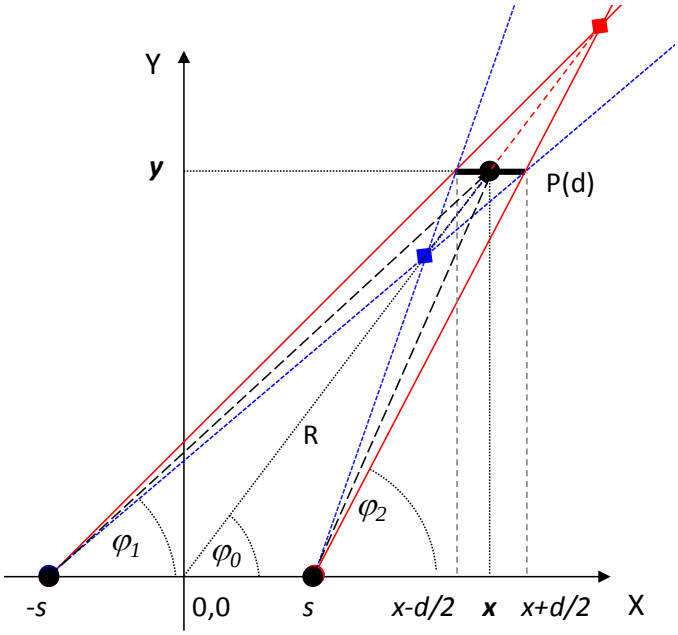

Figure 2. Influence of sound source size on the estimated distance

The most important thing is that the measurement error increases with the distance between the sensor and the source, which affects the practical range of measurements.

III. Experiments and Results

A. Test setup

The sensor setup consists of two AVS devices designed by the authors. Each AVS was constructed from six omnidirectional digital MEMS microphones, IvenSense INMP441, operating at 48 kHz sampling rate with 24-bit resolution. Each microphone was mounted on a board of ca. 10×10 mm size. The boards were mounted together, forming a cube, with microphone ports facing outwards. Both cubes were connected through an I2S-USB digital interface to an USB port in a computer. Signals from all microphones were recorded on the laptop computer and later they were analyzed on a desktop PC with custom scripts written in Matlab and Python. The functions for correction of amplitude and phase were obtained by measuring the AVS responses in an anechoic chamber [12]. The velocity and the average pressure signals were calculated from the amplitude-corrected microphone signals. The results were then phase-corrected, and the intensity signals were computed. The integration period was 1024 samples (21 ms). Next, the intensity signals were smoothed by applying a Savitzky-Golay filter of length 51, polynomial order 3. Finally, the azimuth values and the estimated position were calculated and analyzed statistically. Comparison of the custom AVS with a reference sensor was presented in [12]. Fig. 3 presents the test setup in an anechoic room – two sensors S1, S2 and the sound source (marked with a circle). A loudspeaker contained in a spherical enclosure having radius 50 mm was used as the test source, emitting pink noise with a constant amplitude. A total of 15 positions of the sound source were used in the recordings, the distance between the sensor setup and the source was measured with a DLE 70 laser rangefinder.

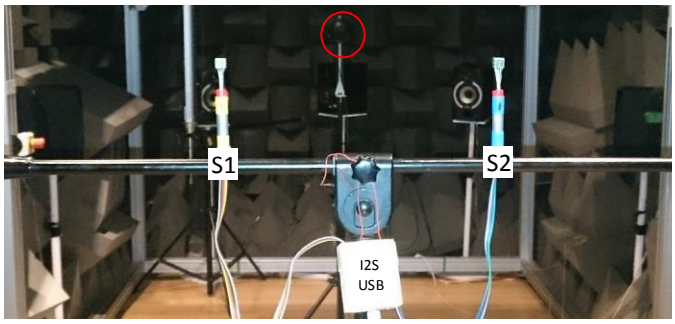

Figure 3. The test setup: acoustic vector sensors S1 and S2, an I2S-USB interface and a test sound source (marked with a circle)

B. Estimation of source position

Reference data were prepared with the assumption that the whole diaphragm of the source emits the acoustic energy. Ranges of angles defining sound emission were calculated from the measured positions of the sensors and the source, as well as the source size. The reference angles are denoted as φ_{g1}, φ_{g1+} for S1, and φ_{g2-}, φ_{g2+} for S2. The reference angles computed for the midpoint of the AVS setup are φ_{g0-}, φ_0, φ_{g0+}.

The angles obtained from the measurements are denoted as φ_1, φ_2. Table I shows the reference values, measured angles and standard deviation for all source positions. From the results it can be observed that the widest range of angles occurs for points closest to the sensors. It can be explained with geometric relations: for small distances between the source and the sensors, relative to the source size, the area of emitted acoustic energy has the largest influence on the results. If the relative distance between the source and the sensor increases, the range of angles decreases. An inverse relationship occurs for the distance estimation. The range of variations in the estimated distance from the source increases with the actual distance. However, the ratio of this range to the distance remains constant. This feature of the presented method may be explained by geometric relations between the sensor setup and the source with a non-zero size (Fig. 2). The range of uncertainty of the measured distance is asymmetrical, from -16.7% to 25.0% in the example presented here. The reference values of the distance between the midpoint of the AVS setup and the source are denoted as R_{0g-}, R_{0g}, R_{0g+}. Table II presents the reference and measured values of the distance R_0 and the angle φ_0. Both sensors operate independently from each other, so they can indicate different points of the sound source. In this case, the accuracy of angle estimation increases, as it can be observed for points 4 and 5 (marked in red in the tables). The value of φ_0 is more accurate than the individual sensor estimates, despite that the sensor S2 exceeded the range of emission angles. On the other hand, error in the distance estimation was the largest for these points. It should be noted that although one of the sensors did not provide an accurate distance estimate, the estimate obtained from both sensors together is within the uncertainty range (case 2 described in Section IIC).

Points 9 and 14, marked in blue in the tables, represent a case in which the angle estimates from individual sensors were inaccurate (lower than the reference for the point 9 and higher for the point 14), but the distance estimate was correct, despite that the angle estimates were not accurate (case 3 in Section IIC). It may be concluded that if both sensors indicate the same point of the source, then even if the angle estimates are slightly inaccurate, the distance estimate may be correct. Fig. 4 shows the results of position estimation performed in signal blocks of 1024 samples. Small circles denote the reference positions of the sound source. Crosses mark the position of sensors. Colored dots indicate the position estimates computed from signal frames, using the algorithm presented in this paper. The dots obtained for each point form smeared patterns, which is a result of inaccuracies in DoA estimation by the individual sensors. It should be noted that each sensor provides relatively stable DoA estimates, despite the noisy character of the sound. Standard deviation of the angle oscillates around 0.12°, regardless of the source position and distance. Black crosses in Fig. 4 show mean positions computed for each point, computed from averaged angles measured by both sensors. Variations in the estimated position increase with the distance to the source, which is in accordance with the described method, and is caused e.g. by the source size. It should be noted that the patterns are very narrow horizontally, which means that the DoA estimates are correct and stable. The largest variations in the position estimates were observed for points No. 4 and 5 that were situated approximately in front of the sensor setup, 4 and 5 meters away. Section IIC discusses the errors in position estimation in detail. Each of the factors described there may contribute to the accumulated error in estimation of the source position.

TABLE I. RESULTS OF DOA ESTIMATION AND THE REFERENCE VALUES FOR 15 POSITIONS OF THE TEST SOURCE

MP	φ_{g1-}	φ_1	φ_{g1+}	φ_{g2-}	φ_2	φ_{g2+}	stdφ_1	stdφ_2
1	72.1	74.8	77.6	100.6	103.4	106.1	0.08	0.09
2	81.2	82.4	84.0	95.3	96.2	98.1	0.12	0.09
3	84.1	85.3	86.0	93.7	94.1	95.7	0.14	0.11
4	85.7	87.4	87.1	92.9	93.2	94.3	0.15	0.12
5	86.1	87.3	87.2	91.8	92.2	93.0	0.16	0.17
6	40.6	43.4	43.1	56.1	57.8	60.2	0.09	0.09
7	60.6	61.3	62.8	72.4	72.8	74.9	0.09	0.10
8	67.0	67.9	68.5	75.1	76.8	76.8	0.13	0.15
9	73.2	72.6	74.5	79.9	79.7	81.2	0.16	0.09
10	75.5	76.1	76.6	81.0	81.5	82.2	0.11	0.13
11	125.5	129.1	129.2	141.0	142.2	143.2	0.09	0.09
12	111.4	113.4	113.9	122.9	124.7	124.9	0.12	0.11
13	105.0	106.8	106.7	113.5	114.6	115.1	0.11	0.14
14	101.4	103.8	102.7	107.9	109.8	109.2	0.13	0.10
15	100.0	100.4	101.1	105.5	105.7	106.5	0.15	0.16

TABLE II. REFERENCE AND MEASURED VALUES OF DISTANCE TO THE SOUND SOURCE AND THE ANGLE φ_0

MP	R_{0g-}	R_{0g}	R_{0g+}	φ_{g0-}	φ_{g0}	φ_{g0+}	R_0	φ_0
1	0.818	0.981	1.226	86.2	89.1	92.0	0.978	89.0
2	1.684	2.021	2.526	88.2	89.7	91.1	2.067	89.3
3	2.474	2.969	3.711	88.9	89.9	90.8	3.251	89.7
4	3.327	3.992	4.990	89.3	90.0	90.7	4.918	90.3
5	4.163	4.995	6.244	89.0	89.5	90.1	5.783	89.8
6	1.115	1.338	1.672	47.4	49.0	50.6	1.532	49.8
7	1.842	2.211	2.763	66.2	67.4	68.6	2.278	66.8
8	2.757	3.308	4.135	70.9	71.8	72.6	3.055	72.3
9	3.472	4.166	5.208	76.5	77.1	77.8	3.899	76.1
10	4.203	5.043	6.304	78.2	78.8	79.4	5.189	78.8
11	1.147	1.376	1.720	134.2	135.7	137.1	1.524	136.4
12	1.853	2.224	2.780	117.4	118.6	119.7	2.217	119.4
13	2.655	3.187	3.983	109.3	110.2	111.0	3.430	110.8
14	3.542	4.250	5.313	104.7	105.3	106.0	4.565	106.8
15	4.294	5.152	6.440	102.8	103.3	103.8	5.346	103.1

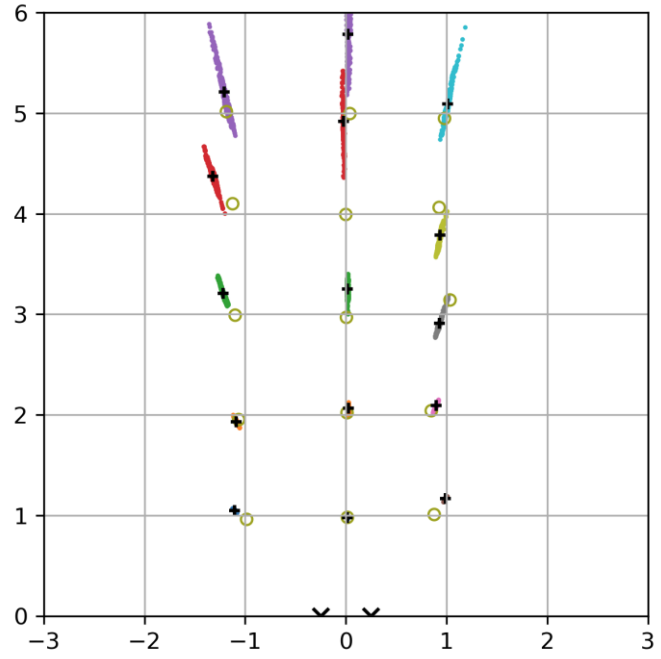

Figure 4. Spatial distributions of sound source position estimates for each source location (colored dots), mean position estimates (+), reference positions of the source (o) and sensor positions (×). Units are distances in meters

C. *Results in relation to the measurement uncertainty*

In this Section, differences between the reference positions of the source and the obtained estimates are analyzed. Fig. 5 shows a distribution of differences in the estimated angle φ_0, Fig. 6 – in the estimated distance R_0. Dashed lines in these figures indicate the uncertainty ranges due to the source size. The measured estimates, shown as dots, are ordered by the distance between the source and the midpoint of the sensor setup. The blue dots indicate points No. 9 and 14, for which the angle estimates are outside the uncertainty range (Fig. 5), but the range estimates lie within the range (Fig. 6).

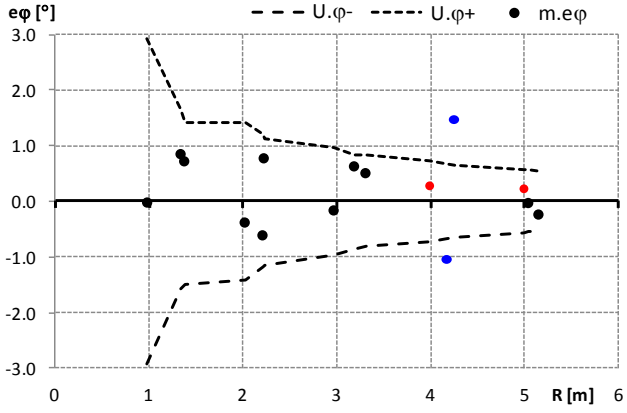

Figure 5. Differences between the estimated and the reference values of angle φ_0. The dashed lines indicate the uncertainty range due to the source size

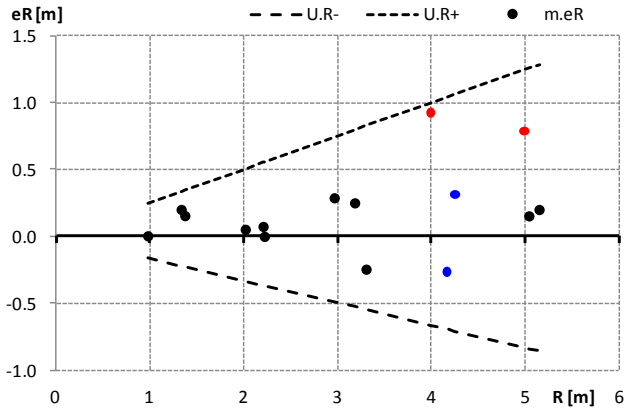

Figure 6. Differences between the estimated and the reference values of distance R_0. The dashed lines indicate the uncertainty range due to the source size

The red dots indicate points No. 4 and 5, for which the angle estimate is within the range, but the error in distance estimation is the largest. For most of the evaluated source positions, despite the wide range of uncertainty, difference between the estimated and the reference distances does not exceed 0.3 m.

IV. CONCLUSIONS

A method of acoustic localization of sound sources using a dual AVS setup was presented in the paper. The main features of the proposed method were presented and discussed, with a focus on the influence of the sound source size on errors in estimation of the source position. It was shown that, apart from sound localization errors introduced by the acoustic vector sensors, the size of the sound source is an important factor that affect the measured estimates of the source position.

Based on the performed experiments and the analysis of the obtained results, it may be concluded that the proposed setup of two sensors provides an increased accuracy of sound source localization, expressed by the angle φ_0. This is caused by the fact that both sensors are independent from each other and errors in DoA estimation from the individual sensors are compensated in many cases. Moreover, using a setup of two sensors allows for estimation of distance between the sensors and the source. It was shown that the error in distance estimation is the largest when each sensor indicates a different point in the source. It was also demonstrated that the uncertainty ranges resulting from the source size are asymmetric. The experiments show that errors in DoA estimation by the individual sensors may not cause large errors in the distance estimation, provided that both sensors indicate the same point in the source.

The distance estimates R_0 obtained in the experiments were all within the uncertainty range which results from the source size and the estimation method. In most of the evaluated cases, the estimation error was in the range ±0.3 m, regardless of the distance to the source, and the error was within the uncertainty range. It is also worth noting that the standard deviation of DoA estimation from the individual sensors was relatively low, about 0.12° on average.

The presented results are promising. In the next research stage, the proposed method will be verified in the outdoor conditions. Experiments related to estimation of position of moving sound sources, such as road vehicles, and comparison of the obtained estimates with the data obtained using reference sensors, are planned.

ACKNOWLEDGMENT

Project co-financed by the by the Polish National Centre for Research and Development (NCBR) from the European Regional Development Fund under the Operational Programme Innovative Economy No. POIR.04.01.04-00-0089/16 entitled: INZNAK – "Intelligent road signs".

REFERENCES

[1] S. Gade, J. Hald, J. Gomes, G. Dirks, and B. Ginn, "Recent advances in moving-source beamforming," Sound & Vibration 49, pp. 8–14, 2015.

[2] J. Cao, J. Liu, J. Wang, and X. Lai, "Acoustic vector sensor: reviews and future perspectives" IET Signal Processing 11, pp. 1–9, 2017

[3] F. Jacobsen, "Sound Intensity and its Measurement and Applications," Acoustic Technology, Department of Electrical Engineering Technical University of Denmark, 2011.

[4] H. E. de Bree, "The Microflown: an acoustic particle velocity sensor," Acoust. Aust. 31, pp. 91–94, 2003.

[5] F. Jacobsen, H-E de Bree, "A comparison of two different sound intensity measurement principles," J. Acoust. Soc. Am. 118 (3), pp. 1510–1517, 2015.

[6] J. Kotus, "Application of passive acoustic radar to automatic localization tracking and classification of sound sources", Information Technologies 18, pp. 111-116, 2010.

[7] J. Kotus, "Multiple sound sources localization in free field using acoustic vector sensor," Multimedia Tools and Applications 74 (12), pp. 4235-4251, 2015.

[8] K. Lopatka, J. Kotus, and A. Czyżewski, "Detection classification and localization of acoustic events in the presence of background noise for acoustic surveillance of hazardous situations," Multimedia Tools and Applications 75, pp. 1-33, 2015.

[9] J. Kotus, "Determination of the vehicles speed using acoustic vector sensor," Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), Poznan, 2018, pp. 64-69, 2018.

[10] J. Kotus, A. Czyżewski, and B. Kostek, "3D acoustic field intensity probe design and measurements," Archives of Acoustics 41, pp. 701–711, 2016

[11] G. Szwoch, and J. Kotus, "Detection of the incoming sound direction employing MEMS microphones and the DSP," In: Dziech A., Czyżewski A. (eds) Multimedia Communications, Services and Security. MCSS 2017. Communications in Computer and Information Science, vol 785, pp. 186–198, Springer 2017.

[12] J. Kotus, and G. Szwoch, "Calibration of acoustic vector sensor based on MEMS microphones for DOA estimation," Applied Acoustics 141, pp. 307–321, 2018.

[13] J. Yin, C. Xiong, and W. Wang, "Acoustic localization for a moving source based on cross array azimuth," Appl. Sci. 2018, 8, 1281, 2018.