



Calibration of acoustic vector sensor based on MEMS microphones for DoA estimation

Józef Kotus*, Grzegorz Szwoch

Gdansk University of Technology, Faculty of Electronics, Telecommunications and Informatics, Multimedia Systems Department, Narutowicza 11/12, 80-233 Gdańsk, Poland

ARTICLE INFO

Keywords:

Acoustic vector sensor
Calibration
Direction of arrival
Sound localization
MEMS

ABSTRACT

A procedure of calibration of a custom 3D acoustic vector sensor (AVS) for the purpose of direction of arrival (DoA) estimation, is presented and validated in the paper. AVS devices working on a p-p principle may be constructed from standard pressure sensors and a signal processing system. However, in order to ensure accurate DoA estimation, each sensor needs to be calibrated. The proposed algorithm divides the calibration process into two stages. First, amplitude calibration is performed in order to compensate for amplitude differences between pairs of microphones situated on each axis. After the pressure and velocity signals are computed from the corrected microphone signals, the second stage is performed in order to correct phase differences between the pressure and velocity signals, which then allows for computing the intensity signals for each axis. In order to validate the calibration method, a reference AVS was constructed from low cost components, namely MEMS microphones and a DSP board. The method was engineered with an assumption that it will be applicable to any AVS working on the same principle. A set of experiments was performed in order to validate the calibration method and to compare the accuracy of the calibrated sensor with a commercial AVS. It was found in the experiments that DoA accuracy of the proposed 3D AVS calibrated with the proposed procedure matches that of a commercial, high cost, factory-calibrated sensor. Therefore, the proposed calibration method fulfills the requirements of accurate DoA estimation, and it is applicable to calibration of custom built, low cost AVS devices, that may be implemented in practical applications for determining the direction of sound sources, such as environmental monitoring, traffic monitoring and public security systems.

1. Introduction

Acoustic vector sensors (AVS) are devices that are capable of measuring acoustic particle velocity as well as pressure, as in a standard microphone. AVS devices working on the ‘p-u’ principle consist of one omnidirectional sensor for measuring pressure, and three directional sensors on the orthogonal axes, measuring acoustic particle velocity [1]. However, AVS can be also practically realized using only acoustic pressure sensors. The majority of available sound intensity measurement systems, including AVSs, are based on the ‘two-microphone’ (p-p) principle, which uses two closely spaced pressure microphones and rely on a finite difference approximation of the sound pressure gradient [2]. The IEC 1043 standard on instruments for the measurement of sound intensity deals exclusively with the p-p measurement principle [3].

Sound probes based on p-u or p-p principles are used for measurements of sound intensity which is a measure of the flow of acoustic energy in a sound field. The sound intensity I is a vector quantity

defined as the time average of the flow of sound energy through a unit area in a direction perpendicular to the area. The intensity in a certain direction is the product of sound pressure (scalar) $p(t)$ and the particle velocity (vector) component in that direction $u(t)$. Typically, intensity metric quantities are used for measuring energy transmission and propagation paths [4,5], as well as for detection of noise source localization [6], determination of acoustic impedance and reflection index of materials [7], although one may find several examples of employing them in audio engineering [8], and in particular in the recording and reconstruction of the acoustic field, e.g. ambisonics [9]. Commercial sound intensity measurement systems were introduced to the market in early 1980s, and the first international standards for sound intensity measurements and related instruments were issued in the mid-1990s [3,7,10,11].

One of the most common applications of AVS devices is direction of arrival (DoA) estimation, i.e. determining the azimuth and elevation angle of the incoming sound, allowing for positioning the sound source

* Corresponding author.

E-mail address: joseph@sound.eti.pg.gda.pl (J. Kotus).

in 3D space [12]. The described sensors are useful in many practical applications, e.g. in sound field measurements, acoustic detection of security threats in surveillance systems [13,14], environmental noise monitoring, etc. However, the cost of commercial AVS devices that provide a satisfactory level of accuracy, is usually high. This is caused not only by the cost of components, but also by the fact that each individual sensor needs to be calibrated at the factory in order to compensate for differences in signals obtained from individual sensors. Therefore, AVS devices are a kind of measurement equipment, not a solution suitable for large scale implementations, e.g. in urban monitoring systems. However, a custom AVS may be constructed from low cost microphones, supported with a signal processing system, employing the principle of p-p sensors. Previously, the authors experimented with a custom sound intensity probe built from seven microphones [15] and experimented with application of a commercial AVS for DoA estimation [6,16]. In this paper, a novel AVS construction, based on miniature digital microphones, is used. In order to obtain meaningful data from this sensor, with errors below some accepted level, a proper calibration method has to be developed.

There are several known techniques that can be employed to calibrate a sound intensity probe based on the pressure gradient principle [17]. The first one requires using specialized microphones for the sound intensity measurements, having similar properties (i.e. sensitivity, frequency response) and which need to be exposed to the same sound pressure by using a coupler [2,18]. In such case, the microphones should have identical frequency responses. A well-known method of sound intensity probe calibration based on the cross-spectral density approach was described by Krishnappa [19]. The phase and gain mismatch errors are corrected by measuring the transfer function between the two microphone systems while exposing them to signals with identical phase and pressure level, over a wide range of frequencies. Calculation of cross-spectral density between two microphones (p-p probe) was then used to determine the sound intensity. Another kind of calibration technique, proposed by the Nagata et al. [20], is called a rotating microphone method. The calibration in this case is performed in an anechoic chamber and it uses a reference intensity probe which was calibrated with a sound-pressure phase-difference calibrator. During the comparative sound intensity measurement, the bias error is determined. Another approach by Zhang [21] aims to compensate for phase differences by means of instantaneous correlation integral. Other techniques are employed for AVS calibration in special applications, e.g. for underwater measurements [22].

The calibration methods presented above focus mainly on correcting the amplitude and phase responses of microphone pairs, by means of processing the acoustic pressure signals acquired from microphone pairs. In this paper, the Authors take a similar approach and propose a two-stage procedure for calibration and correction of a 3D p-p intensity probe [23]. Since the purpose of the AVS described here is DoA estimation, the calibration procedure does not need to ensure that accurate values of sound intensity are measured. It is sufficient to ensure that relationship between sound intensity measured on the orthogonal axes is maintained. Contrary to the calibration methods presented earlier, a reference intensity probe or a calibrator are not needed to perform the calibration with the proposed approach. In order to validate the calibration method, a custom AVS was constructed from MEMS microphones and a DSP board. The method was engineered with an assumption that it will be applicable to any AVS working on the same principle. A set of experiments was performed in order to validate the calibration method and to compare the accuracy of the calibrated sensor with a commercial AVS.

The rest of the paper is organized as follows. Section 2 presents the algorithm used for DoA detection from signals recorded by the microphones. Section 3 shows the proposed calibration method. In Section 4, a reference hardware implementation of the sensor is presented, and some practical issues of software and hardware implementations of the calibration method are discussed. Section 5 presents the results of

experiments related to the calibration of the proposed AVS, the achieved accuracy of DoA detection with the results obtained from a commercial AVS.

2. Acoustic vector sensor

An acoustic vector sensor (AVS) measures velocity of sound particles in a defined direction, as well as pressure [1]. A three-dimensional (3D) AVS described in this paper is able to measure the acoustic particle velocity (referred to as “velocity” further in the text) in three orthogonal directions, as well as pressure in the central point of the sensor. Based on the velocity and pressure measurement results, sound intensity vectors in three orthogonal directions may be obtained. The AVS described here is based on the p-p principle, i.e. on pressure measured at two points, then the velocity vector is computed as a pressure gradient (a difference of pressure values measured at two points), and the pressure scalar at the AVS central point is averaged from these point measurements [3].

In the described AVS, pressure has to be measured at six points located on three orthogonal axes, at identical distances d from the origin. These points are denoted as $x_1, x_2, y_1, y_2, z_1, z_2$, describing their location in the coordinate system, e.g. point y_2 is located at $(0, d, 0)$ and y_1 at $(0, -d, 0)$. Omnidirectional microphones of the same type are used to measure pressure $p_l(t)$ at six locations l . According to the Euler's formula, velocity vectors $\mathbf{u}_i(t)$ alongside axes X, Y, Z may be computed as:

$$\begin{bmatrix} \mathbf{u}_x(t) \\ \mathbf{u}_y(t) \\ \mathbf{u}_z(t) \end{bmatrix} = \begin{bmatrix} a_x & 0 & 0 \\ 0 & a_y & 0 \\ 0 & 0 & a_z \end{bmatrix} \cdot \begin{bmatrix} p_{x2}(t) - p_{x1}(t) \\ p_{y2}(t) - p_{y1}(t) \\ p_{z2}(t) - p_{z1}(t) \end{bmatrix} \quad (1)$$

where a_i are scaling factors. Magnitude of the $\mathbf{u}_i(t)$ vector will be denoted as $u_i(t)$. Pressure $p(t)$ measured at the origin is averaged from two points at the given axis and it has to be equal on all three axes. In practice, pressure is averaged from all six microphones (as in Eq. (2)):

$$p(t) = \frac{p_{x1}(t) + p_{x2}(t) + p_{y1}(t) + p_{y2}(t) + p_{z1}(t) + p_{z2}(t)}{6} \quad (2)$$

Sound intensity I at a given axis may be then computed as [24]:

$$I = \frac{1}{T} \int_T p(t) u(t) dt \quad (3)$$

where T is the integration period.

If a single, omnidirectional sound source is put into the system at polar coordinates (r, φ, θ) , then the angles of the sound received by the AVS may be computed as:

$$\phi = \arctan\left(\frac{I_y}{I_x}\right) \quad (4)$$

$$\theta = \arctan\left(\frac{I_z}{\sqrt{I_x^2 + I_y^2}}\right) \quad (5)$$

In a calibrated AVS, the following conditions have to be fulfilled:

$$\begin{aligned} p_{x1} &= p_{x2} \text{ if } \varphi = \pm\pi/2, \\ p_{y1} &= p_{y2} \text{ if } \varphi = 0 \text{ or } \varphi = \pi, \\ p_{z1} &= p_{z2} \text{ if } \theta = 0. \end{aligned} \quad (6)$$

3. Calibration procedure

In an uncalibrated AVS, Eq. (6) is usually not fulfilled due to imperfections in microphone positioning in the system and differences in parameters between microphones. Therefore, a calibration procedure is required in order to equalize amplitude and phase differences between the microphones. The proposed algorithm divides the calibration

process into two stages. First, amplitude calibration is performed in order to compensate for amplitude differences between pairs of microphones situated on each axis. After the pressure and velocity signals are computed from the corrected microphone signals, the second stage is performed in order to correct phase differences between the pressure and velocity signals, which then allows for computing the intensity signals for each axis. The details of both stages are described in the following Subsections. The calibration is a one-time procedure, although it may be repeated if needed. The computed correction functions are then used for signal processing during the DoA detection.

3.1. Amplitude calibration

It is important to emphasize the assumption that the AVS is constructed using the same type of microphones, produced in MEMS technology. It should guarantee that they have a similar sensitivity and other acoustic parameters. At the beginning of the calibration process, each microphone should be checked using a standard acoustic calibrator, equipped with special attachment which enables proper injection of the acoustic pressure into each microphone. This step is very important and it gives the possibility to detect improper operation of the microphone. For practical reasons, this preliminary step can be limited to checking only one microphone from each pair. This step also enables alignment of the gain for all channels and in consequence, between each direction for the 3D sound probe. The values of the gain factor obtained during the absolute calibration for each pair are stored in the constants: ag_x , ag_y and ag_z (ag means amplitude gain, x , y , z indicate the given microphone pair). The second microphone of the given pair can be checked during the general amplitude calibration step described below. It is also assumed here that signals p_{x1} , p_{x2} are aligned using the gain factor ag_x . In practice, it means that the signals p_{x1} and p_{x2} are multiplied by the factor ag_x . In general, the amplitude alignment process can be described by the formula:

$$p_{ni} = ag_n * p'_{ni} \quad (7)$$

where: n is the axis (x , y or z), i is the index of a microphone in a given pair (1 or 2), p' is the pressure before the amplitude alignment.

The amplitude calibration stage aims to compute correction functions (filters) that modify the microphone signals in such a way that relative amplitude differences are equalized and that Eq. (6) is satisfied. The correction procedure will be described for the X axis, the two remaining axes are calibrated accordingly. If a single point sound source is placed at polar coordinates $(r, \pi/2, 0)$, then the microphone signals $p_{x1}(t)$ and $p_{x2}(t)$ should be equal to each other (Fig. 1). The aim of the amplitude calibration algorithm is to compute correction functions $q_{x1}(t)$ and $q_{x2}(t)$ that fulfill the condition (Eq. (8)):

$$p_{x1}(t) * q_{x1}(t) = p_{x2}(t) * q_{x2}(t) \quad (8)$$

Both the calibration process and further signal analysis in an AVS are performed in the digital domain. The correction functions are computed, thus the following condition (Eq. (9)) is fulfilled:

$$h_{x1}(n) * q_{x1}(n) = h_{x2}(n) * q_{x2}(n) \quad (9)$$

where $h_i(n)$ are impulse responses of the respective microphones, $i = 1, \dots, 6$. The impulse responses were obtained using a swept sine signal and the cross-correlation technique [29]. The whole procedure of the impulse responses calculation is described in details in Section 4.2.

The amplitude correction filter is prepared in the frequency domain. The truncated impulse responses $h_{x1}(n)$, $h_{x2}(n)$ are converted to their spectral forms $H_{x1}(f)$, $H_{x2}(f)$. A difference $E_x(f)$ is computed as (Eq. (10)):

$$E_x(f) = \log_{10}|H_{x2}(f)| - \log_{10}|H_{x1}(f)| \quad (10)$$

In the described AVS operation principle, there is no reference microphone and $E_x(f)$ has to be distributed evenly between the two microphones. The correction functions in the frequency domain may be

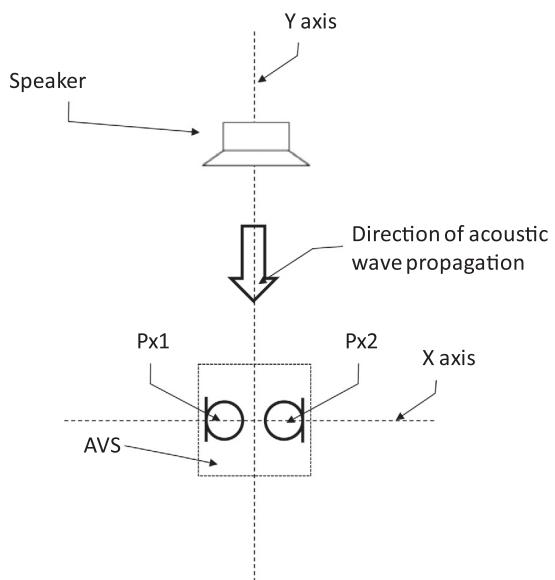


Fig. 1. Setup for the amplitude calibration of the X axis.

computed as (Eqs. (11) and (12)):

$$Q_{x1}(f) = 10^{\left(\frac{E_x(f)}{2}\right)} \quad (11)$$

$$Q_{x2}(f) = 10^{\left(-\frac{E_x(f)}{2}\right)} \quad (12)$$

The results are then transformed to the time domain, resulting in the final correction functions $q_{x1}(n)$ and $q_{x2}(n)$ which will be used in the DoA detection in order to equalize amplitudes of the microphone signals according to Eq. (8).

With the same setup, i.e. sound source at $(r, \pi/2, 0)$, functions $q_{z1}(n)$ and $q_{z2}(n)$ may be calculated. In order to compute $q_{y1}(n)$ and $q_{y2}(n)$, the sound source has to be placed at $(r, 0, 0)$. The calibration procedure remains the same as the one described above.

3.2. Phase calibration

The second calibration stage is performed after the pressure and velocity signals are computed from the amplitude corrected microphone signals (Eq. (8)). Calculation of sound intensity (Eq. (3)) requires that for orientation presented in (Fig. 2), the pressure and velocity signals are phase-aligned, which is usually not a case in an uncalibrated system. The phase correction procedure that aims to equalize phase differences between the velocity and the pressure signals, will be described for the X axis, the remaining axes-related components are calibrated the same way. In this stage, the sound source has to be placed on the calibrated axis, i.e. for calibration of $u_x(t)$, the sound source has

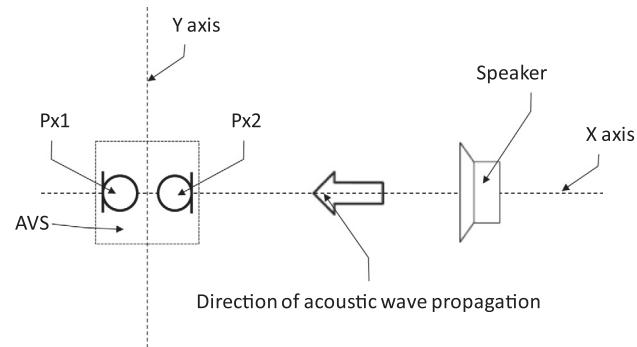


Fig. 2. A setup for phase calibration of the X axis.

to be placed at polar coordinates $(r, 0, 0)$. The aim of the phase calibration stage is to correct the phase of $u(t)$ so that the following condition holds satisfied (Eq. (13)):

$$\frac{p(t)}{u_x(t)} = \frac{p_{x2}(t) + p_{x1}(t)}{2(p_{x2}(t) - p_{x1}(t))} = C \quad (13)$$

where C is a constant (related to acoustic impedance) that is identical for all calibrated axes. The exact value of c is not important for the calibration, $C = 0.5$ was chosen for convenience.

Signals $h_p(t)$ and $h_{ux}(t)$ are computed from the impulse responses $h_{x1}(t)$ and $h_{x2}(t)$, after the amplitude correction:

$$h_p(t) = 0.5(h_{x1}(t) + h_{x2}(t)) \quad (14)$$

$$h_{ux}(t) = h_{x2}(t) - h_{x1}(t) \quad (15)$$

Similarly to the amplitude correction stage (see section 4.2. for details), signals $h_p(t)$ and $h_{ux}(t)$ are aligned so that $h(0) = \max(h(n))$, and they are truncated with a window function of length L , centered at $n = 0$. The same $h_p(n)$ is used for all axes, therefore, the phase correction is performed on $h_{ux}(n)$, so that the following equation is valid (Eq. (16)):

$$\arg(H_{ux}(f) \cdot Q_{ux}(f)) = \arg(H_p(f)) \quad (16)$$

where $H_{ux}(f)$ and $H_p(f)$ are Fourier transforms of $h_{ux}(n)$ and $h_p(n)$, respectively, $Q_{ux}(f)$ is the spectral phase correction function. The following conditions have to be fulfilled:

$$\arg(Q_{ux}(f)) = \arg(H_{px}(f)) - \arg(H_{ux}(f)) \quad (17)$$

$$|Q_{ux}(f)| = 1 \quad (18)$$

where $\arg(H(f))$ is the unwrapped phase of $H(f)$. The phase correction function $q_{ux}(n)$ is then computed by transforming $Q_{ux}(f)$ into the time domain. The same calibration procedure is applied to $u_y(n)$ with sound source at $(r, \pi/2, 0)$, and to $u_z(n)$ with sound source at $(r, 0, \pi/2)$.

After all the calibration functions are computed, the complete signal processing in the AVS is carried out as it was described in Section 2, with two additional steps. Namely, amplitude correction functions are applied to the microphone signals before the velocity and the average pressure computed, and phase correction is applied to the calculated velocity signals before the intensity is computed.

4. Implementation

In this Section, a custom hardware implementation of the AVS, constructed in order to validate the proposed calibration scheme, as well as some practical issues of a software implementation of the calibration algorithm, are presented.

4.1. Hardware realization of a testing system

In order to validate the proposed calibration method, a custom AVS was constructed. The sensor is built from six miniature ($4.7 \times 3.8 \times 1$ mm), digital MEMS microphones, type INMP441 [25], mounted on circuit boards of ca. 10×10 mm size, that together form a sensor cube (Fig. 3). The origin of the coordinate system is situated in the center of the cube. The eZdsp5535 board by Spectrum Digital [26], equipped with Texas Instruments TMS320C5535 digital signal processor (DSP) [27], was employed for the tasks of controlling the microphones, acquisition of sound samples and signal processing [28]. The DSP chipset used in the testing system is a low-power, fixed point processing unit with a sufficient processing power at a disposal for the calibration and DoA detection algorithms, and the development board is equipped with the necessary interfaces (I2S and USB). Any other hardware platform meeting these requirements may be used instead. Communication with the microphones is performed via the Inter-IC Sound (I2S) bus, pairs of MEMS microphones placed on the same axis

are connected to three separate I2S channels. In order to receive synchronous samples from the microphones, I2S channels have to be controlled with a clock signal (the board is an I2S master device). The dimensions of the designed probe limited the frequency range to c.a. 8 kHz, thus the sampling frequency of about 16 kHz was needed. Since the I2S clock is determined by the DSP clock f_{DSP} and divisors that are integer powers of two, $f_{DSP} = 75$ MHz with the divisor equal to $2^{6.64}$ was used, resulting in the I2S sampling rate of 18.31 kHz. With $f_{DSP} = 100$ MHz and divisor $2^{5.64}$, the sampling rate of 48.83 kHz may be achieved, which is close to the standard 48 kHz rate.

Sound samples are acquired from all microphones and they are buffered in the DSP board memory. For the offline calibration with the testing system, the buffered samples (about 800 ms of 6-channel signal sampled with 18.31 kHz rate) were retrieved by an external computer through the USB interface. After the calibration, the system was used for processing the microphone signals in the online mode, as described in Section 2, using the calibration functions stored in the DSP memory, and the computed intensity and angle values were retrieved using the USB interface.

4.2. Software implementation of the calibration algorithm

The algorithm for calculating the correction functions was implemented in MATLAB, according to the description in Section 3. The calibration signals were recorded in an anechoic room. In order to obtain impulse responses of all microphones, a linearly swept chirp was used as the excitation signal, preceded with an impulse for the synchronization purposes. The test signal was generated in a digital form, and its samples were played back through a loudspeaker (Genelec 6010A) placed at the required polar coordinates, at the distance of approximately 1 m from the AVS. Six microphones in the AVS recorded the received signals. The duration of the test signal was 700 ms (limited by the available DSP memory), the sampling rate was 18.31 kHz, and the frequency was swept from 1 Hz to 8.5 kHz (sweep rate 12.14 kHz/s). The maximum frequency (about 9 kHz) was limited by the size of the sensor cube. However, the proposed procedure remains valid for longer test signals, slower sweep rates and wider frequency ranges.

Processing of the recorded signals for the calibration purposes was performed in an off-line mode on a personal computer. Computation of the impulse response of each microphone is performed as follows. The signal recorded by each microphone is convolved with the time-reversed excitation signal (a chirp), which is equivalent to computing cross-correlation of the signals [29]. The length of the result, is the sum of the lengths of the input signals. In the next step, samples of the impulse response are located by finding the index of a sample with maximum amplitude ($ind.s.max$). Next, the convolution result is truncated to a window ($ind.s.max - L/2, ind.s.max + L/2$) in order to keep only L most important samples. This truncation is performed with a standard windowing function (e.g. a Hamming Window) in order to reduce distortions. The value of L corresponds to the length of FIR filter used for the correction. Therefore, the choice of L is a compromise between the amount of performed computations and the accuracy of the impulse response. The obtained result is the impulse response used in further calculations.

The signal processing pipeline is presented in Fig. 4. The whole correction procedure is implemented and practically realized in time domain. The procedure is realized in two separate steps. During the first step, the correction of acoustic pressure signals obtained from each microphone is performed. In this step, the amplitude of each signal is corrected. After this step, the main components used for intensity determination are prepared. These signals are: the common acoustic pressure and three components of particle velocity. Next, the phase between the acoustic pressure signal and each particle velocity component are corrected. In practice, only one signal from each pair (either acoustic pressure or particle velocity) needs to be corrected. For better accuracy, a phase correction for each particle velocity signal is

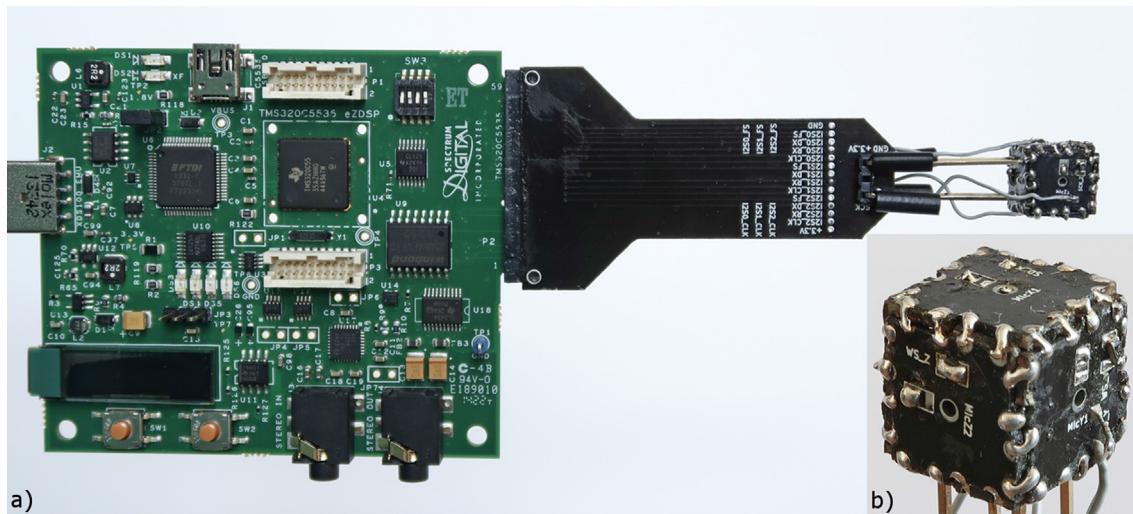


Fig. 3. Picture a) – the custom AVS – a sensor cube attached to the DSP board. Picture b) – close-up of the sensor cube. Microphone ports are visible as holes in the middle of each side of the board surface.

proposed. In practice, FIR filters are used for correction of the particle velocity signals. This process corrects the phase, but it also introduces a delay between not processed acoustic pressure signal and processed particle velocity signals. To avoid synchronization loss between these signals, a delay line is applied for the acoustic pressure signal. After that, synchronized signals of acoustic pressure and phase corrected particle velocity components are obtained.

For each axis, the following operations were performed. Impulse responses of both microphones situated on a given axis, were aligned so that their maximum was at $n = 0$, then they were truncated with a Hamming window of length $L = 128$, centered at $n = 0$. The choice of L was based on the preliminary experiments, described in the next Section. The impulse responses were then transformed with FFT, and spectra of the correction functions were calculated according to Eqs. (11) and (12). After transforming these functions to the time domain with IFFT, discarding the residual imaginary part and circularly shifting the samples by $L/2$, impulse responses of finite impulse response (FIR) filters of order $(L - 1)$, were obtained. These filters will be used for the amplitude correction on the respective microphone signals during the DoA detection. In order to reduce calibration errors, several sets of recordings were performed, and amplitude differences calculated for

the successive recordings for each microphone were averaged in the frequency domain before the correction functions were computed.

For the phase calibration stage, each recorded signal was processed with a filter designed during the amplitude correction stage. The pressure signal was calculated by averaging all six microphone signals, and the velocity signals were computed according to Eq. (1). Signals $h_p(n)$ and $h_u(n)$ were then computed by convolving the pressure and velocity signals with the time-reversed test signal, and they were truncated with a Hamming window of length $L = 128$. The obtained results were then transformed with FFT, and the unwrapped phase difference was calculated. Several recordings were made and the obtained phase differences were averaged. The phase correction functions were then calculated subsequently using Eq. (17), and transformed with IFFT. Thus, the coefficients of phase correction FIR filters of order $L - 1$ were obtained. Since the phase correction filters introduce a delay to the velocity signals, the pressure signal also has to be delayed by $L/2$ samples.

4.3. Software implementation of the DoA detection algorithm

The algorithm for the DoA detection was implemented on the

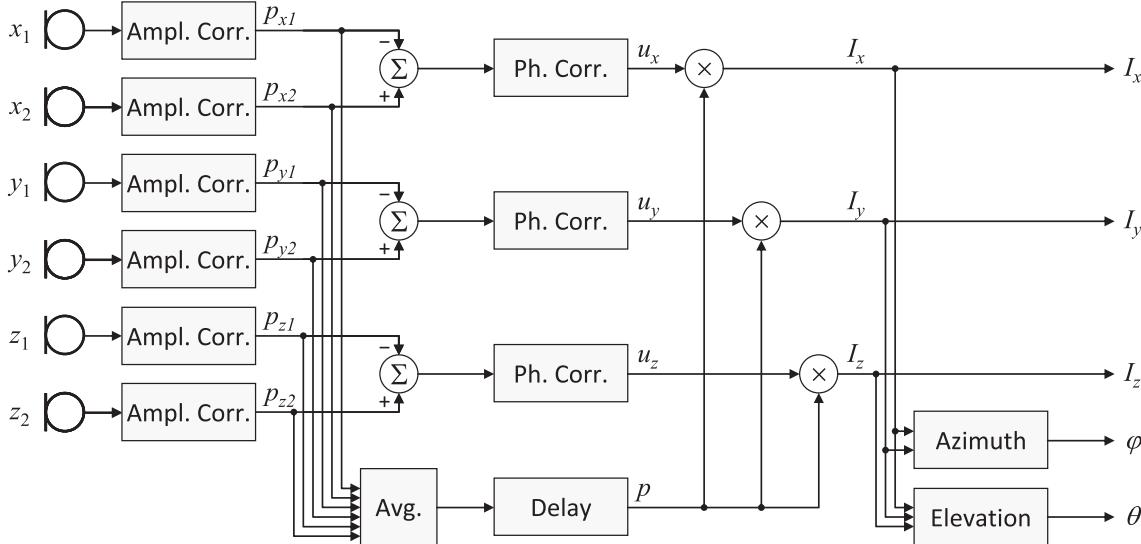


Fig. 4. The processing of microphone signals in the proposed AVS. Amplitude (Ampl. corr.) and phase (Ph. corr.) correction is performed with linear phase FIR filters.

custom DSP platform in a fixed point form. All operations are performed in time domain, with linear phase FIR filters used for amplitude and phase correction. Filter coefficients were quantized to a Q15 format and stored in the DSP memory. The processing algorithm performs amplitude correction, velocity and pressure computation and phase correction. Next, the intensity signals are computed for each axis by averaging the products of pressure and velocity signals. In the described implementation, a sum is computed instead of the average, in order to avoid small values as a results. The averaging period has to be selected according to the expected variability of the sound intensity and to the required temporal resolution of measurement. In the experiments performed by the authors, the averaging period was 1024 samples (56 ms for sampling frequency 18.31 kHz) with 75% overlapping, which allowed to obtain a measurement rate equal to 71.52 Hz (temporal resolution 13.98 ms), with a satisfactory signal to noise ratio. Shorter averaging periods increase the temporal resolution at a cost of increased noise. The azimuth and the elevation of a sound source is then computed from the obtained intensity values (Eqs. (4) and (5)). The output of the algorithm is a vector containing calculated values of intensity and angles (Eq. (19)):

$$\mathbf{y}(n) = [I_x(n), I_y(n), I_z(n), \varphi(n), \theta(n)] \quad (19)$$

where n is the analysed frame index.

5. Results

Signals needed to perform the calibration were recorded in an anechoic room (Fig. 5). The calibrated AVS was mounted on a tripod, 170 cm above ground, and placed on a turntable which allowed for precise control of the AVS orientation relative to the sound source. A loudspeaker (Genelec 6010A [31]) was mounted at the same height, at a distance of 110 cm from the AVS, in a way that the loudspeaker's main axis coincided with the AVS X axis when the turntable was set to zero degree. A digitally generated linear chirp signal (described in Section 4) was played through the loudspeaker and recorded by the sensor into the DSP board memory. The signals were recorded for the X axis with the turntable set to 0 and 180 degrees, and for the Y axis with 90 and 270 degrees. Then the AVS was rotated, so that the Z axis was directed at the loudspeaker, and recordings were made for 0 and 180 degrees. Two recordings were made for each AVS orientation. The recorded signals were retrieved from the DSP board via USB interface and analyzed on a PC with Matlab scripts that implemented the calibration algorithm.

It is important to mention the fact that acoustic energy was not

emitted from single point, but through the speaker, dimensions of which have to be considered for determining the angular accuracy of the setup. In the presented measurement setup, the angular distribution of the acoustic field in relation to the speaker dimensions and distance to the AVS is about 3 deg. It means that the obtained results of the DoA values can differ by about 3 deg. These differences are evoked by the measurement system properties. The angular resolution of the turntable system was 1 deg. The measurements were performed with a 15 deg step. For each angle, the measurement was made twice, in order to check the accuracy of the turntable. The obtained results for both series were consistent.

5.1. Amplitude calibration

The amplitude correction functions were computed from two sets of recordings, using the procedure described earlier. The length L of the window used to cut the relevant part of the impulse response defines the length of correction filters. For high L , longer part of the impulse response is taken into account, so the correction is more accurate. At the same time, higher filter length L increases the computation time. Therefore, the choice of L must be a compromise between these two criteria. Table 1 presents root mean squared differences (RMSD) between magnitude responses of pairs of microphones on each axis, computed for different window lengths. Ideally, these values should be close to zero. Even a short window ($L = 32$) provided sufficient equalization of magnitude differences. The lowest RMSD was observed for $L = 256$, but $L = 128$ was chosen for the practical implementation as a reasonable compromise between the accuracy and the computational complexity.

Fig. 6 presents magnitude responses of the calculated correction functions (q_{x1} , q_{x2} denote the correction functions for two microphones situated on the X axis). In the examined case, magnitude differences between microphones in each pair were relatively small, below 1 dB for most of the frequency range. Fig. 7 shows the magnitude response of each microphone, calculated for a single set of recordings, before and after the amplitude correction. It is expected that magnitudes of each microphone in a pair situated on the same axis are equal to each other. It can be observed that without correction, there are differences that need to be compensated in order to obtain correct results from the AVS. The correction functions were calculated from averaged differences obtained for two sets of recordings, and applied to the recorded signals, resulting in a set of equalized magnitude responses.

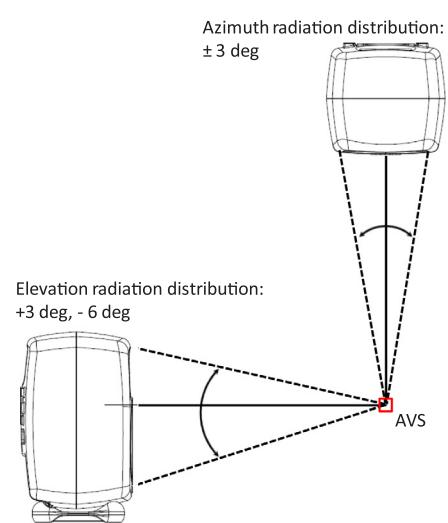
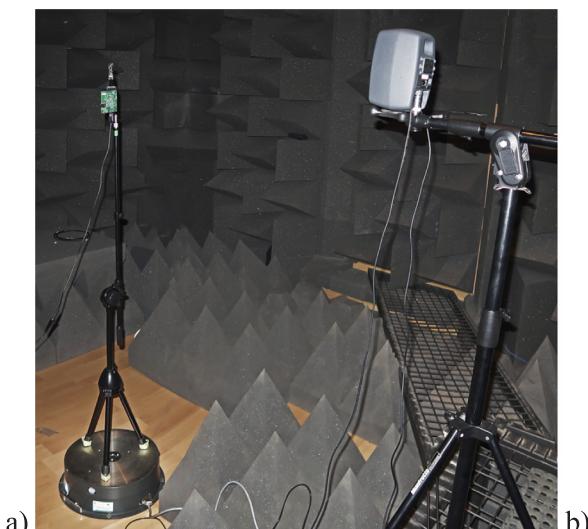
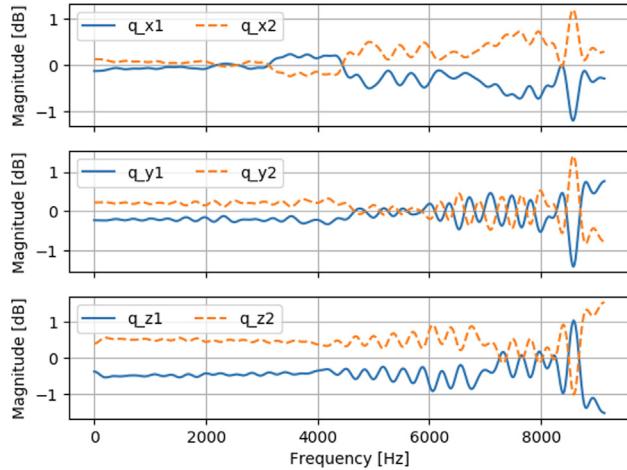


Fig. 5. Calibration setup in the anechoic room (a): the calibrated AVS on the left (standing on a turntable), the test loudspeaker on the right, (b) – the angular radiation distribution for considered directions (azimuth and elevation) in relation to the speaker and AVS distance and dimensions.

Table 1

RMSD values of magnitude responses of microphone pairs (computed for all three axes), without and with amplitude correction, for different lengths L of the window.

| Window | No corr. | $L = 32$ | $L = 64$ | $L = 128$ | $L = 256$ | $L = 512$ | $L = 1024$ |
|--------|----------|----------|----------|-----------|-----------|-----------|------------|
| RMSD | 0.716 | 0.149 | 0.106 | 0.043 | 0.021 | 0.052 | 0.030 |

**Fig. 6.** Magnitude responses of the calculated amplitude correction functions.

It can be observed that in the examined case, the amplitude correction did not need to introduce significant changes to the recorded microphone signals. The main aim of the amplitude correction for the given pair of microphones was to obtain relatively the same amplitude characteristics for each microphone, rather than equalize their responses and reduce ripples. Moreover, even small un conformity in the amplitude responses results in significant errors in particle velocity calculations. It may be concluded that differences in the microphone parameters in the evaluated device were negligible and that placement of microphones in the sensor cube was sufficiently accurate. Nevertheless, there were some differences that needed the correction,

Table 2

RMSD values (in degrees) of phase responses of velocity relative to pressure (computed for all three axes), without and with phase correction, for different lengths L of the window.

| Window | No corr. | $L = 32$ | $L = 64$ | $L = 128$ | $L = 256$ | $L = 512$ | $L = 1024$ |
|--------|----------|----------|----------|-----------|-----------|-----------|------------|
| RMSD | 113.091 | 38.132 | 1.512 | 0.999 | 1.212 | 1.890 | 1.947 |

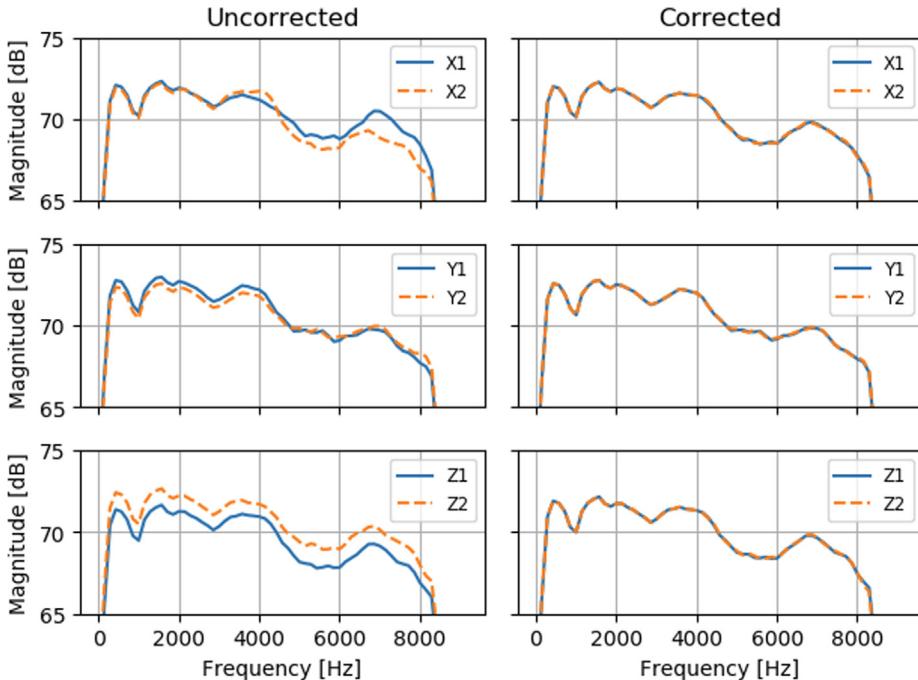
as they might influence the accuracy of the next calibration stage. Furthermore, in other AVS implementations, these magnitude differences may be more significant and the amplitude calibration phase may be much more beneficial in such a case.

5.2. Phase calibration

The phase correction functions were calculated from phase differences between the velocity and the pressure, averaged from two sets of recordings. Similarly to the amplitude calibration stage, the length L of the window had to be chosen. **Table 2** shows RMSD values for phase difference between the velocity and pressure signals. It can be seen that a window of at least $L = 64$ is required to compensate for large phase differences. Increasing L above 128 does not provide an increase in the accuracy. Therefore, $L = 128$ was chosen for the implementation. **Fig. 8** shows group delay plots of the calculated phase correction functions (group delay was plotted instead of phase because it allows for more clear presentation of differences between functions). The largest correction was needed for the vertical (Z) axis.

Fig. 9 shows phase responses of velocity and pressure signals calculated for a single set of recordings, before and after the phase correction (amplitude correction was performed before, in both cases). It is expected that all velocity signals are in phase with the averaged pressure signal. The plots obtained without correction show large differences in phase, while in the phase corrected signals, the plots overlap, except for the extreme sections of the frequency range.

In order to confirm that the phase calibration procedure worked as expected, an additional analysis was performed directly on signals recorded by the AVS microphones (with amplitude correction), but not on the calculated impulse responses as before. **Fig. 10** presents histograms of phase difference between velocity and pressure signals at each axis,

**Fig. 7.** Magnitude response of each microphone in the AVS, without and with amplitude correction. The responses are shown in pairs for each axis: X, Y, Z.

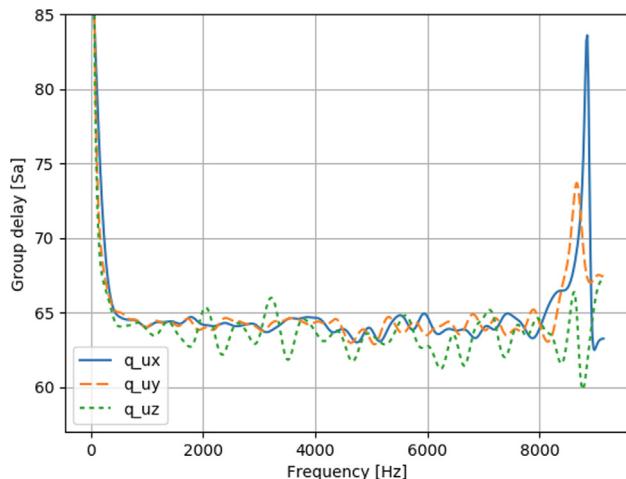


Fig. 8. Group delay of the calculated phase correction functions for each axis.

weighted by magnitude. It can be seen that for signals without phase correction, there is a phase shift between velocity and pressure, and that shift depends on the frequency (especially for the Z axis). The histograms obtained for the same signals after the phase correction show that expected result were obtained – in each case, the phase shift is close to zero for the whole frequency range.

5.3. Detection of DoA

In the next stage of experiments, the accuracy of DoA detection by the reference AVS was evaluated. The purpose of this experiment was not to assess the accuracy of the sensor, but to validate the proposed calibration method by comparing the results with a factory calibrated sensor. Assuming that a single sound source is positioned in the 3D space at a specified azimuth and elevation angle relative to the AVS, the algorithm described in Section 2 should be able to detect angles of the incoming sound precisely. Calculation of angles is performed from intensity values computed for each axis, as described in Section 2. Fig. 11 presents intensity calculated for each axis, with and without each

calibration stage, in a setup in which the sound source was placed on the Y axis at 90 deg azimuth, and at zero elevation. In this case, it is expected that I_y is dominant, while I_x and I_z are both close to zero. It can be clearly seen that this is not the case for signals that did not undergo any correction. The amplitude correction stage sets I_x almost to zero, brings I_z closer to zero, and it does not alter I_y in a significant way. If only phase correction is applied, I_y is corrected as expected, I_x comes closer to zero than in the uncorrected case, but the correction is worse than for the amplitude correction, and I_z deviates from zero even more than for the uncorrected case. Using both calibration stages does not significantly change the results, compared with only the phase correction applied.

In order to perform DoA analysis, a new set of recordings was made. Two excitation signals were used: white noise (uniform random distribution of amplitude and phase, an easier case to analyze) and the same linear chirp signal (continuous-phase signal with increasing frequency) that was used for the calibration (more difficult case, because of variable frequency and because the signal phase has to be taken into account). For the azimuth detection, the turntable was rotated in 15 degree steps, so that 24 angles (-80 to 165 deg) were examined, and the turntable angle settings constituted the ground truth for this experiment. The results were normalized so that the mean azimuth calculated for a sound source at X axis was at zero deg. The results computed for all 24 angles are shown in Table 3. Root mean squared error (RMSE) is an indicator of accuracy of the results (deviation from the ground truth value). It is computed as a square root of a mean of squared differences between the measured and the expected angles, and is expressed in degrees. The STD metric is the standard deviation of results obtained for a given angle, averaged over all angles, and it is an indicator of the repeatability of results.

For white noise, RMSE without correction was about 36 degrees, and applying either of the calibration stages reduced this error significantly. Both stages applied together resulted in a small RMSE decrease compared with the phase calibration only. This is caused by the fact that for white noise, phase difference have less significant influence on the results that tonal signals. For the linear chirp, errors observed when no correction is applied, are very high, about 90 degrees, so an uncalibrated AVS is unusable for DoA detection. Amplitude correction did not provide a satisfactory degree of error reduction, while the phase

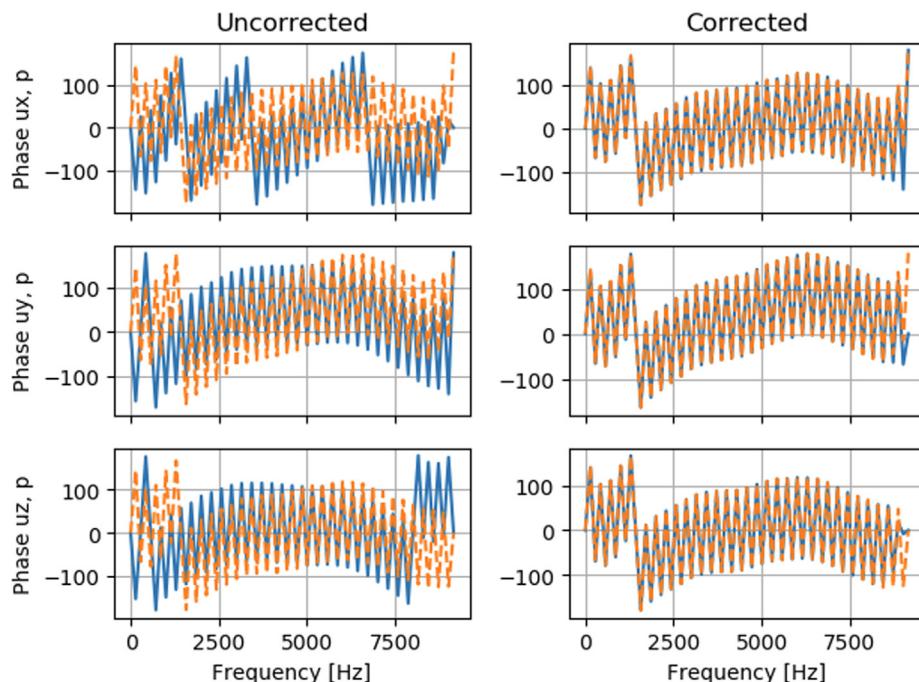


Fig. 9. Phase of velocity (solid line) and pressure (dashed line) for each axis in the AVS, without and with phase correction.

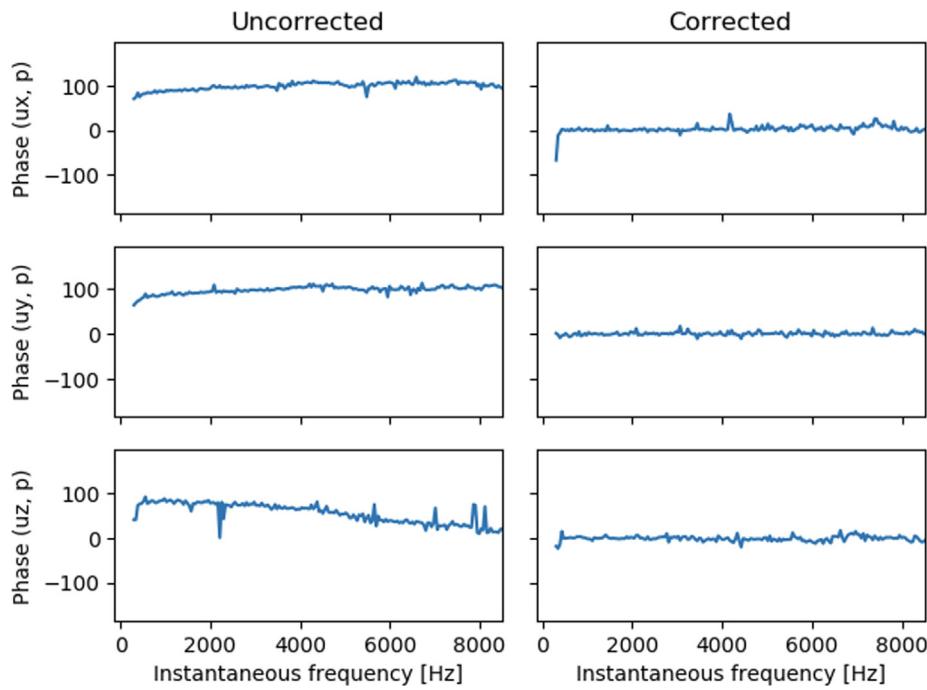


Fig. 10. Mean phase difference between the velocity and pressure signals, as a function of instantaneous frequency of the test chirp signal.

correction reduced RMSE to about 2 degrees. Both calibration stages together resulted in improved accuracy, with RMSE below 2 degrees, slightly larger than for white noise.

Fig. 12 presents the detection results from the uncalibrated and calibrated AVS, obtained for white noise. When no correction is applied, the results are distorted and shifted. Applying both correction stages yields azimuth values that are stable throughout the whole signal, although some small shifts may be observed. Fig. 13 presents the results obtained for a linear chirp, and in this case, it may be observed that the uncalibrated AVS does not provide any meaningful angle

Table 3

RMSE and standard deviation of the azimuth detection (computed from all angles, values in degrees).

| Signal | White noise | | Linear chirp | | |
|----------------------|-------------|-------|--------------|--------|-----|
| | Metric | RMSE | STD | RMSE | STD |
| No correction | 36.793 | 2.788 | 90.275 | 79.523 | |
| Amplitude correction | 2.694 | 0.629 | 36.486 | 49.499 | |
| Phase correction | 1.582 | 0.158 | 2.294 | 2.758 | |
| Both corrections | 1.468 | 0.151 | 1.838 | 1.032 | |

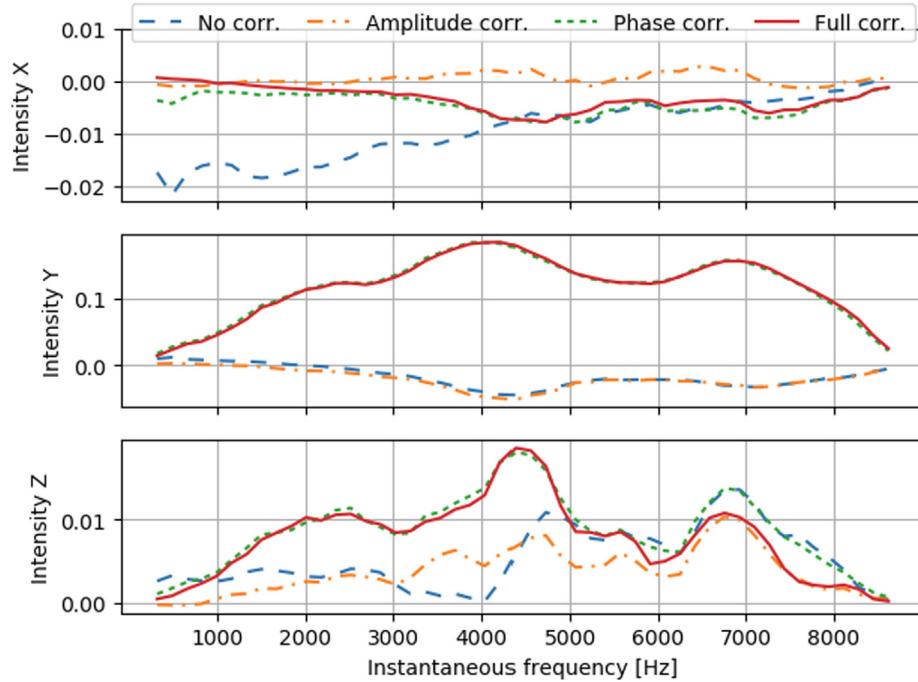


Fig. 11. Sound intensity calculated for each axis, depending on the instantaneous frequency of the test chirp signal, for amplitude and phase correction turned on or off. The sound source was placed on Y axis.

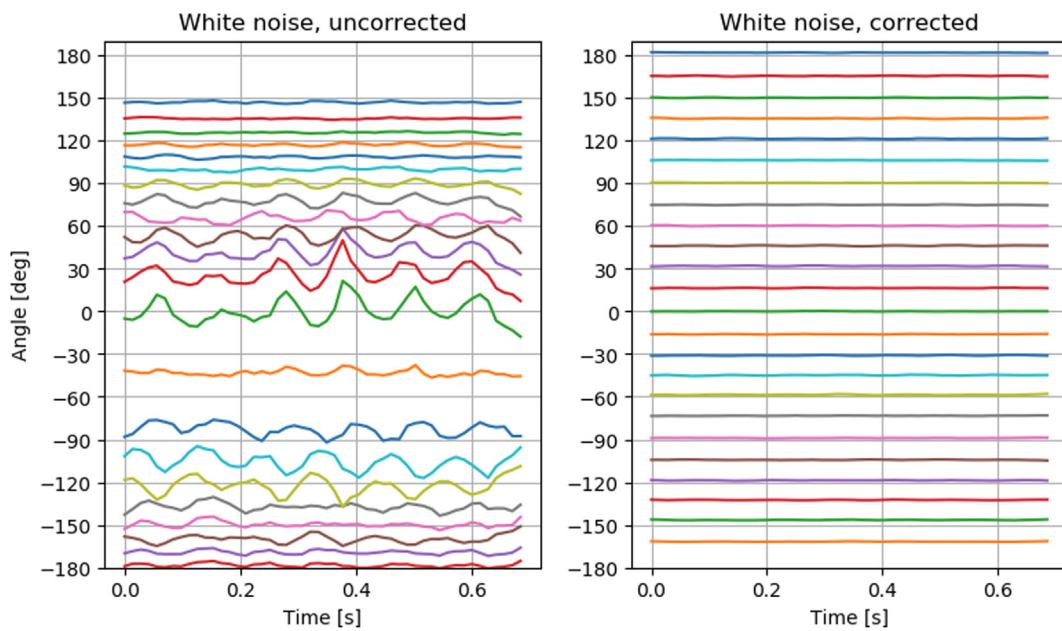


Fig. 12. Results of the azimuth detection with white noise as the excitation signal, without and with amplitude and phase correction.

detection results. The complete calibration procedure allowed for error reduction, although larger variations in calculated values for different signal frames (that correspond to different frequencies) are observed.

A similar procedure was applied to the elevation detection. The AVS was mounted in a position rotated by 90 degrees relative to the orientation used in the previous experiment, and the turntable was rotated within the $(-90, 90)$ deg range, in 10 deg steps. The metrics were calculated as before and they are shown in [Table 4](#). In this case, RMSE values obtained with the calibrated AVS are significantly higher than for azimuth detection. However, the calibration procedure still reduces the detection errors. The elevation detection for the chirp signal was more problematic than for white noise, phase correction is essential in this case. [Figs. 14 and 15](#) show the elevation detection results for the AVS without and with correction. The accuracy obtained without any correction is very low, especially for the chirp. If the correction is applied, the accuracy for white noise are improved, but there is a much

Table 4

RMSE and standard deviation of the elevation detection (computed from all angles, values in degrees).

| Signal | White noise | | Linear chirp | | |
|----------------------|-------------|--------|--------------|--------|--------|
| | Metric | RMSE | STD | RMSE | STD |
| No correction | | 18.820 | 1.553 | 39.849 | 23.134 |
| Amplitude correction | | 6.386 | 1.030 | 34.088 | 26.583 |
| Phase correction | | 4.936 | 0.510 | 7.773 | 5.230 |
| Both corrections | | 4.977 | 0.505 | 7.886 | 5.529 |

larger variation in results than for the azimuth detection, and the results for elevation angles larger than zero are shifted. The results obtained for the chirp signal are less accurate than for white noise, large errors are noticeable in the right part of the plot, which corresponds to higher

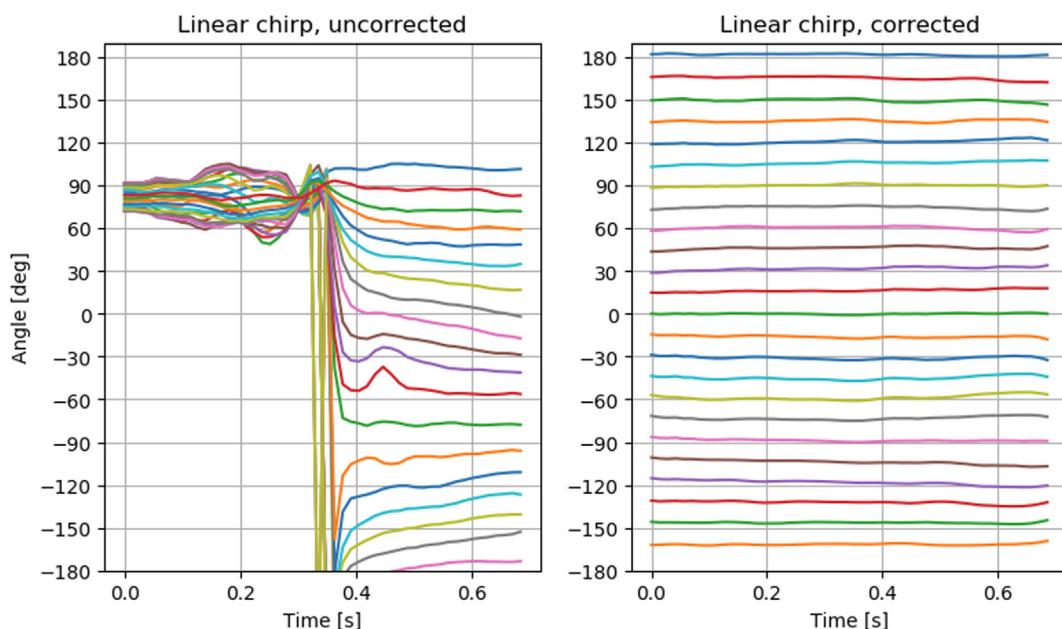


Fig. 13. Results of the azimuth detection with linear chirp as the excitation signal, without and with amplitude and phase correction.

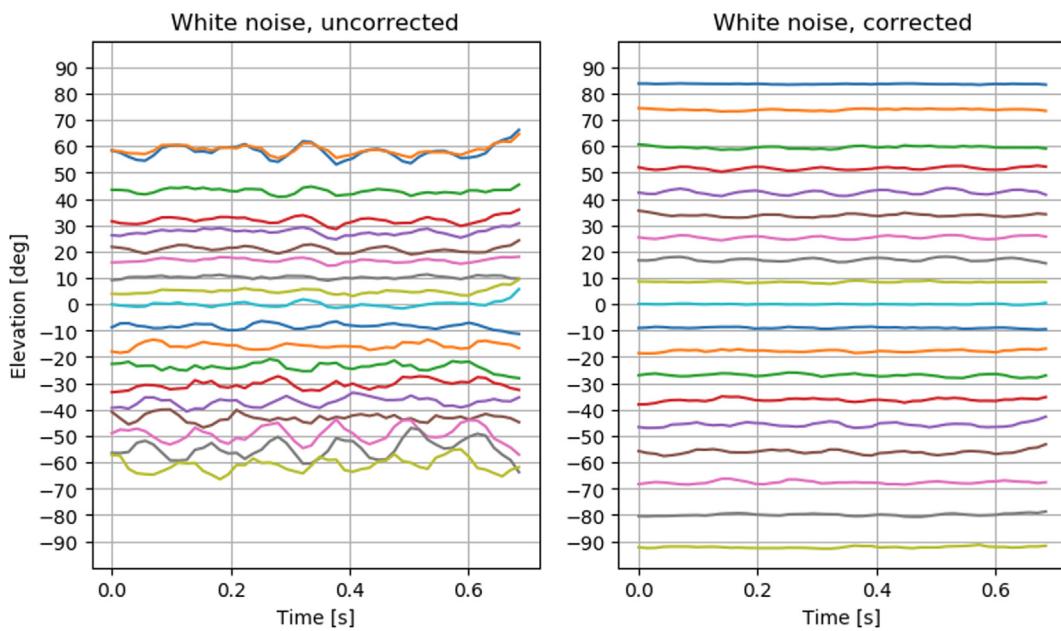


Fig. 14. Results of the elevation detection with white noise as the excitation signal, without and with amplitude and phase correction.

frequencies in the signal. The observed errors may be caused by several factors. Most importantly, the lower microphone on the Z axis was directed at the mounting board (Fig. 3) that may distort the arriving sound waves. The processor board was set crosswise to the sound radiated through the speaker. Such kind of spatial configuration of the sensor and processor board may disturb the flow of acoustic energy especially for tonal components. The directivity of the loudspeaker for higher frequencies, as well as the measurement procedure (i.e. inaccuracies in positioning the Z axis of the AVS relative to the loudspeaker when the AVS was mounted in a horizontal position and rotated on a turntable) may also attribute to the observed errors.

For better understanding of the obtained DoA results, two types of uncertainty should be considered. Uncertainty type A is related to the accuracy of the measured values and it can be achieved using statistical analysis. Type B is related to the method, equipment and conditions of

the measurements [32]. As it was mentioned at the beginning of Section 5, the DoA angular distribution evoked by the physical dimensions of the applied loudspeaker with relation to distance to the sensor, is about 3 degrees. This factor can be associated with the uncertainty type B. The second factor that is used to characterize the proposed calibration technique and the sensor is the standard deviation of DoA values. This indicator can be related to uncertainty type A. The RMSE of the DoA values is ca. 3 degrees, which means that the accuracy of the presented calibration and correction method is comparable to the accuracy of the applied measurement setup. Low values of the standard deviation indicate that the accuracy of the sensor calibration and correction method is high.

Resolution of the measurement results should also be considered. In described experiments, resolution of the measurement can be defined as the smallest change of the measured input value that results in the

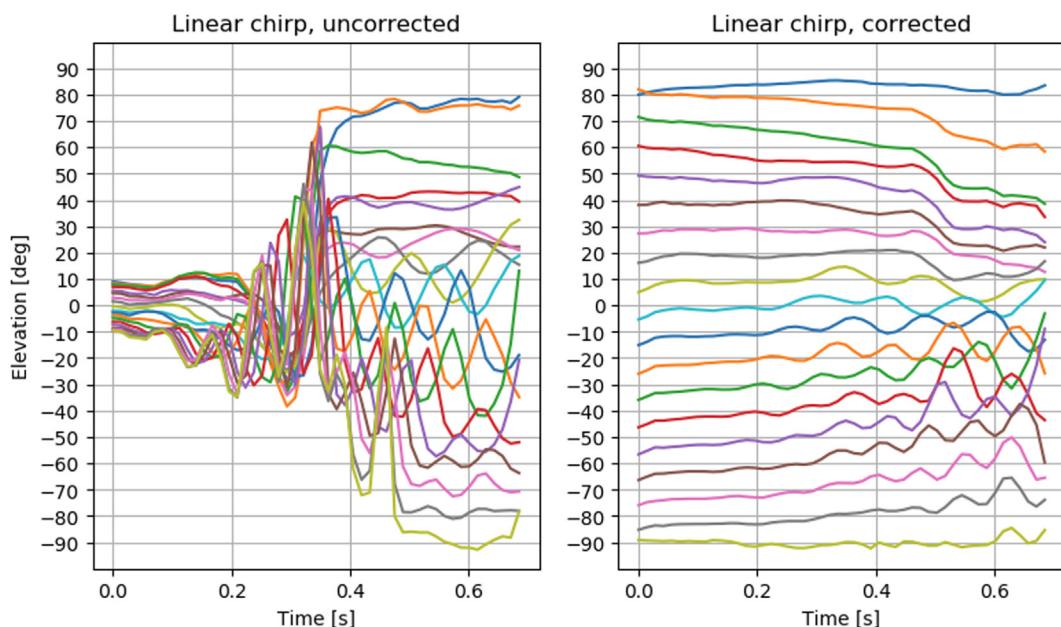


Fig. 15. Results of the elevation detection with linear chirp as the excitation signal, without and with amplitude and phase correction.

measureable change of the value at the output. Typically, this parameter is related to the resolution of the measurement scale. In the considered case, resolution can be defined as the smallest change of the sound source position relative to the sensor that results in a measurable change of the DoA value. Based on the obtained results (especially standard deviation values), the resolution can be estimated to be about 0.1 degree. It is important to consider the resolution in relation to the measurement setup. In the presented case, a turntable was used which enabled positioning of the sensor with 1 degree step. It means that the resolution of the measurements setup is lower than the estimated resolution of the sensor.

On the other hand, the angular resolution can be also considered as the ability to distinguish between two closely spaced sound sources. In order to avoid confusion about meaning of the term “angular resolution”, it is important to emphasize that we used a single sound source in our experiments. For that reason, we can only consider the resolution in relation to the measuring scale.

In the next section, a comparative analysis of the presented calibration and correction method with a commercial sound probe equipped with an original correction unit, are described. It is important to emphasize that both sensors were examined using the same measurement setup and in the same condition.

5.4. Comparison with a commercial AVS

The results of DoA detection obtained from the calibrated AVS were compared with a professional device, in order to assess the accuracy of the proposed calibration method and the DoA detection algorithm. An Ultimate Sound Probe Regular (USP UR) by Microflown was used for the comparative analysis. The USP is a professional grade, factory calibrated measurement device, constructed from three MEMS sensors for the acoustic particle velocity measurement and a sound pressure microphone [30]. The device is operating in the 20 Hz – 10 kHz range. The same experiments as before were repeated for the USP sensor, in the same testing conditions. The velocity and pressure signals were recorded with a B&K Pulse analyzer with 65 kHz sampling rate, the sound intensity components were calculated for the same time length of the frame as for the custom AVS (56 ms). Computation of DoA was performed using the same algorithm as before. The results (RMSE and STD) are shown in Tables 5 and 6. For both test signals, the observed RMSE is smaller for the proposed AVS than for the commercial USP sensor. Both devices exhibit higher errors for the elevation detection than for the azimuth.

Figs. 16 and 17 present comparison of DoA detection for selected angles, for both examined devices. For white noise, the results from USP have small variations, but there is a noticeable shift from the expected values, especially for the elevation. The results from the proposed AVS have higher variations for some angles in the elevation detection, but overall, deviations from the expected values are smaller. The situation changes when a linear chirp is used. The USP results exhibit much larger variations for the azimuth detection than the proposed AVS, especially for higher frequencies. For the elevation, both devices distort the results for higher frequencies, but errors are also higher for the USP.

The observed variations in sound intensity, affecting the DoA values of azimuth and elevation, depend on the test signal character. In general, lower RMSE values were observed for the noise than for the chirp.

Table 5

Comparison of RMSE and standard deviation of azimuth detection (computed from all angles) for the custom AVS and the commercial USP AVS.

| Signal | White noise | | Linear chirp | | |
|------------|-------------|-------|--------------|-------|-----|
| | Metric | RMSE | STD | RMSE | STD |
| Custom AVS | 1.468 | 0.151 | 1.838 | 1.032 | |
| USP AVS | 3.521 | 0.167 | 4.219 | 2.858 | |

Table 6

Comparison of RMSE and standard deviation of elevation detection (computed from all angles) for the custom AVS and the commercial USP AVS.

| Signal | White noise | | Linear chirp | | |
|------------|-------------|-------|--------------|-------|-----|
| | Metric | RMSE | STD | RMSE | STD |
| Custom AVS | 4.977 | 0.505 | 7.886 | 5.529 | |
| USP AVS | 9.992 | 0.244 | 10.152 | 5.074 | |

signal. For the noise, variations of the angle result from frame length used for intensity calculation. Shorter frames result in larger variation in the measured intensity. The choice of frame length is a compromise between the accuracy and the computation time. In order to compare the custom sensor with the commercial AVS, frames of equal length (in milliseconds) were used. For the tonal (chirp) signal with varying frequency, variations in DoA angles result from propagation of sound waves between the source and the sensor. During the experiments, a two-way speaker was used, therefore, the actual location of the source was different for low and high frequencies. Taking into account the distance between speakers, and between each speaker and the sensor, differences of up to 3 degrees may be observed with increasing frequency of the test signal. Such an effect was observed during the experiments. Also, interferences of sound waves with the sensor mounting board may be the cause of variations of the results.

Comparison of RMSE calculated per angle for each device is shown in Figs. 18 and 19. It can be observed that the measured detection errors, visualized as an area defined by each plot, are significantly lower for the custom AVS than for the commercial USP. For the USP, lowest RMSE was observed for 0 and 180 deg azimuth, and for 0 deg elevation, while for other angles, large lobes may be observed, with maxima at about 60 and 120 deg for azimuth, and ± 60 deg for elevation. For the custom AVS, the observed RMSE values are lower, and the shape of the azimuth plot is more or less symmetrical, with the highest error for 150 deg. For the elevation, maxima of RMSE are at about 70 deg and -50 deg. The observed RMSE plot shapes may be caused by the sensor construction (interferences between microphones, and with the enclosure) and also some differences between microphones may not be fully compensated during the calibration procedure.

6. Conclusions

A procedure of calibration of a custom 3D AVS was presented and validated. A reference sensor was constructed from low cost elements. Six digital, low cost MEMS microphones were used for the actual sensor, and an inexpensive DSP board was employed for signal analysis. Any processing board with at least six channel I2S interface and sufficient processing power may be used instead. If the sensor is not calibrated in any way, the obtained results of DoA detection are meaningless. The proposed calibration procedure involves only recording the microphone signals in two sensor orientations while a linear chirp is emitted from a test source, whereas performing an analysis of these signals with simple digital signal processing operations, such as Fourier transform and FIR filters. The DoA results obtained with the calibrated AVS have satisfactory accuracy. Additionally, it was found in the experiments that performance of the proposed 3D AVS with the custom calibration procedure matches that of a commercial, high cost, factory-calibrated sensor. In fact, the DoA errors observed for the evaluated AVS were lower than in the sensor used for comparison. However, it should be noted that the experiments were performed solely with the aim of validating the calibration procedure. In order to perform a more elaborate assessment of the custom, calibrated AVS, a much larger set of measurements (e.g. at various distances from the sound source, and with more test signals) is still needed, providing a topic for future research. At this point, it may be concluded that the calibration algorithm presented here works as it was expected, hence it may potentially be

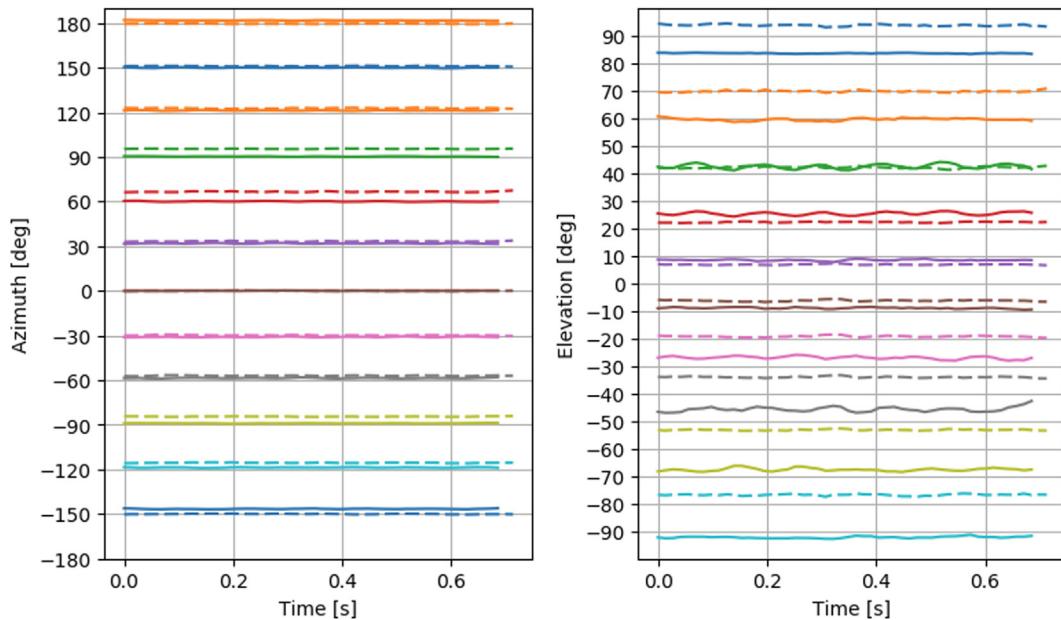


Fig. 16. Comparison of azimuth and elevation detection (selected angles) for the custom AVS (solid lines) and the commercial USP AVS (dashed lines), with white noise as the test signal.

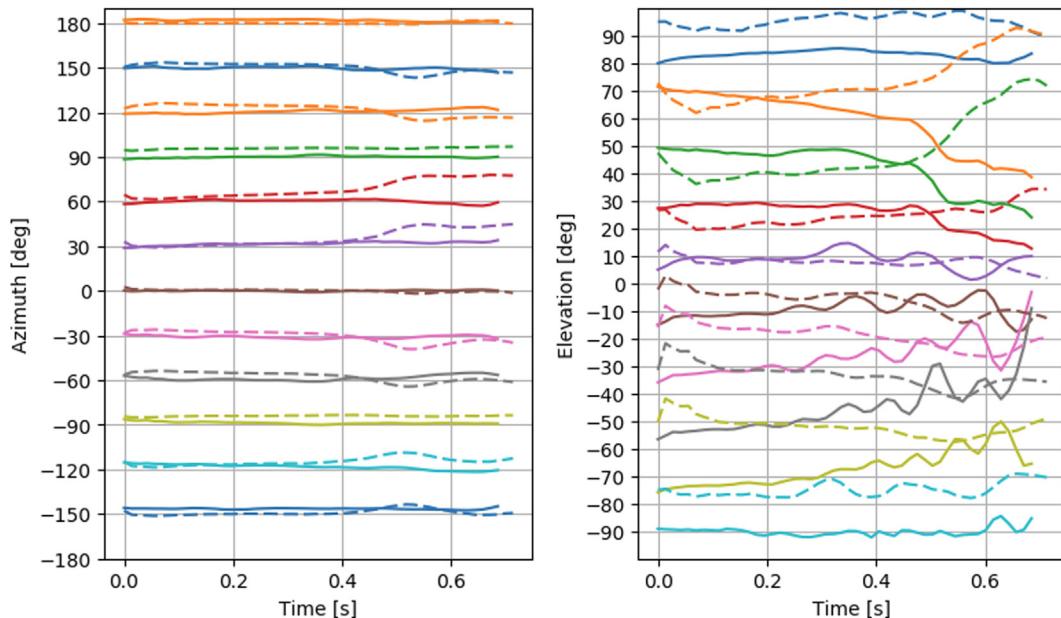


Fig. 17. Comparison of azimuth and elevation detection (selected angles) for the custom AVS (solid lines) and the commercial USP AVS (dashed lines), with a linear chirp as the test signal.

used for low cost AVS devices in practical applications.

It should also be noted that the proposed calibration procedure is applicable to other AVS configurations. For example, the authors used the same algorithm for calibration of a 2D sensor constructed from four measurement microphones, needed for a specific experiment, and it worked as expected. Moreover, the example presented here used a frequency band limited to about 9 kHz, but the procedure works the same way for wider frequency ranges. Also, if the frequency range needs to be limited, bandpass filters may be easily incorporated into the amplitude and phase correction functions.

To conclude, the proposed method of calibrating a 3D AVS, together with a sensor constructed from low cost elements, provide a solution that may be implemented in practical systems for environmental monitoring (noise level measurements), surveillance systems (acoustic event detection, including sound source positioning), traffic monitoring (vehicle counting, etc.), teleconference systems (directing cameras at a speaker) and many more. Such systems require a large number of AVS devices, which may be realized by constructing custom AVS as described here and applying the proposed calibration method.

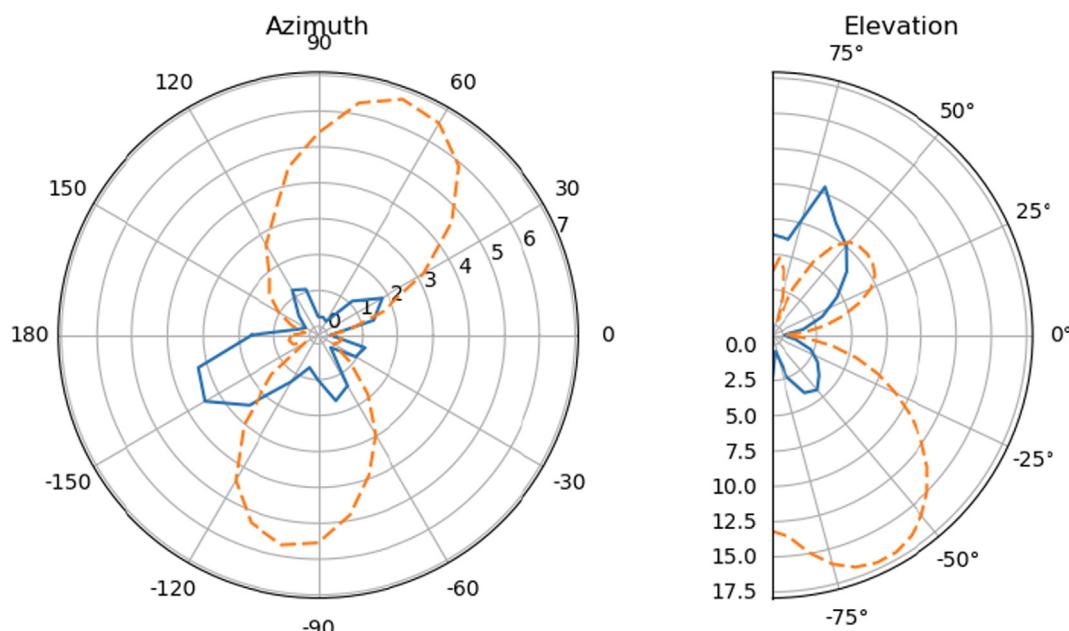


Fig. 18. RMSE values depending on the angle, for the custom AVS (solid lines) and the commercial USP AVS (dashed lines), with white noise as the test signal.

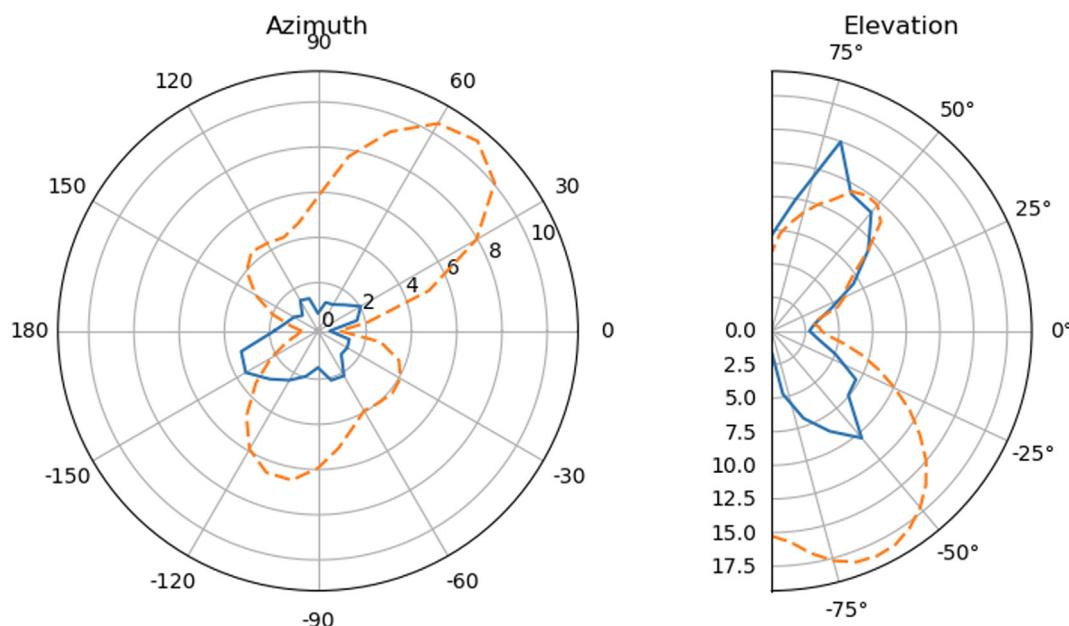


Fig. 19. RMSE values depending on the angle, for the custom AVS (solid lines) and the commercial USP AVS (dashed lines), with a linear chirp as the test signal.

Acknowledgements

Project co-financed 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] Cao J, Liu J, Wang J, Lai X. Acoustic vector sensor: reviews and future perspectives. *IET Signal Proc* 2017;11:1–9.
- [2] Fahy FJ. Sound intensity. E & F.N. Spon; 1995.
- [3] International Electrotechnical Commission: 1043 Electroacoustics – Instruments for the Measurement of Sound Intensity – Measurements with Pairs of Pressure Sensing Microphones, 1993.
- [4] Weyna S. Identification of reflection and scattering effects in real acoustic flow field. *Arch Acoust* 2003;28(3):191–203.
- [5] Weyna S. An acoustics intensity based investigation of the energy flow over the barriers. *Acta Phys Pol A, Acoust Biomed Eng* 2010;118:172–8.
- [6] Kotus J. Multiple sound sources localization in free field using acoustic vector sensor. *Multimedia Tools Appl* 2015;74(12):4235–51. <https://doi.org/10.1007/s11042-013-1549-y>.
- [7] De Bree H-E. The microflown: an acoustic particle velocity sensor. *Acoust Aust* 2003;31(3):91–4.
- [8] Woszczyk W, Iwaki M, Sugimoto T, Ono K, de Bree H-E. Anechoic measurements of particle-velocity probes compared to pressure gradient and pressure microphones. 122nd Audio Eng. Soc. Convention: May 2007, Paper Number: 7107. 2003.
- [9] Cengarle G, Mateos T, Bonsi D. A second-order ambisonics device using velocity transducers. *J Audio Eng Soc* 2011;59(9):656–68.
- [10] American National Standards Institute: S1.9-1996, Instruments for the Measurement of Sound Intensity, 1996.
- [11] Jacobsen F. Handbook of signal processing in acoustics. Intensity techniques. New York: Springer; 2008. p. 1109–27.
- [12] Tervo S. Direction estimation based on sound intensity vectors. Proceedings of the 2009 European signal processing conference (EUSIPCO 2009). 2009. p. 700–4.
- [13] Kotus J, Łopatka K, Czyżewski A. Detection and localization of selected acoustic

- events in acoustic field for smart surveillance applications. *Multimedia Tools Appl* 2014;68:5–21.
- [14] Kotus J, Łopatka K, Czyżewski A, Bogdanis G. Processing of acoustical data in a multimodal bank operating room surveillance system. *Multimedia Tools Appl* 2016;75:10787–805.
- [15] Kotus J, Czyżewski A, Kostek B. 3D acoustic field intensity probe design and measurements. *Arch Acoust* 2016;41(4):701–11.
- [16] Szczodrak M, Kurowski A, Kotus J, Czyżewski A, Kostek B. A system for acoustic field measurement employing Cartesian robot. *Metrol Meas Syst* 2016;23(3):333–43.
- [17] Novo IJ, Seoane MA. The influence of calibration on the estimation of the DOA with acoustic vector sensors. 46th International Congress and Exposition on Noise Control Engineering Inter-Noise. 2017. p. 5919–28.
- [18] Jacobsen F, Jaud V. A note on the calibration of pressure-velocity sound intensity probes. *J Acoust Soc Am* 2006;120(2).
- [19] Krishnappa G. Cross-spectral method of measuring acoustic intensity by correcting phase and gain mismatch errors by microphone calibration. *J Acoust Soc Am* 1981;69:307.
- [20] Nagata S, Furuhata K, Wada T, Asano DK, Yanagisawa T. A three-dimensional sound intensity measurement system for sound source identification and sound power determination by In models. *J Acoust Soc Am* 2005;118(6).
- [21] Zhang Y, Gu H, Li G. A novel calibration method for acoustic vector sensor. OCEANS 2016 MTS IEEE Monterey; 2016. <https://doi.org/10.1109/OCEANS.2016.7761045>.
- [22] Donskoy D. Very low frequency acoustic vector sensor calibration. *J Acoust Soc Am* 2014;135:2396. <https://doi.org/10.1121/1.4877934>.
- [23] Kotus J, Szwoch G, Czyżewski A. Sound intensity probe with correction system and calibration system and the method of correction and calibration of this probe. Patent pending, PL. P. 422287, 20.07.2017 (in Polish), 2017.
- [24] Jacobsen F. Sound intensity and its measurements and applications. *Curr Top Acoust Res* 2003;3:87–91.
- [25] IvenSense INMP441: <https://www.invensense.com/products/digital/inmp441/> (retrieved: 2018-06-01).
- [26] Texas Instruments: C5535/C5545 eZdsp USB Stick Development Kit, <http://www.ti.com/tool/tmdx5535ezdsp> (retrieved: 2018-06-01).
- [27] Texas Instruments: TMS320C5535 Fixed-Point Digital Signal Processor. <http://www.ti.com/product/TMS320C5535/description> (retrieved: 2018-06-01).
- [28] Szwoch G, Kotus J. Detection of the incoming sound direction employing MEMS microphones and the DSP. In: Dziech A, Czyżewski A, editors. Multimedia communications, services and security. Communications in computer and information sciences, vol. 785. Springer; 2017. p. 186–98.
- [29] Swen M. Measuring transfer functions and impulse responses. In: Havelock D, Kuwano S, Vorländer M, editors. Handbook of signal processing in acoustics. Springer Verlag; 2008. p. 65–85 Chapter 5.
- [30] Microflown: Ultimate Sound Probe Regular. <http://www.microflown.com/products/standard-probes/usp-regular.html> (retrieved: 2018-06-01).
- [31] Genelec: 6010a Studio Monitor. <https://www.genelec.com/support-technology/previous-models/6010a-studio-monitor> (retrieved: 2018-06-01).
- [32] Ratcliffe C, Ratcliffe B. Doubt-free uncertainty in measurement. Switzerland: Springer International Publishing; 2015. https://doi.org/10.1007/978-3-319-12063-8_2.