Improvement of Range Estimation with Microphone Array

Volodymyr Kudriashov

Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria

Email: Kudriashov.Vladimir@Gmail.com

Abstract: *This paper presents a new approach for the three-dimensional (3-D) localization of sound sources. An acoustic camera uses an angular beamforming to measure the Direction of Arrival (DoA) of an incoming signal, to localize the emission source. The acoustic sensor used in this article is the Brüel & Kjaer acoustic camera transformed to operate in a bistatic mode. The transformation consists in a placing of one of the microphones of the acoustic camera outside of its microphone array. This allows simultaneous estimation of the DoA and the Time Difference of Arrival (TDoA) of the incoming signal(s). Such sensors were not found. The paper proposes emitter localization in range – cross range – elevation coordinates by combining estimates of TDoA and DoA and presents the signal processing method for that purpose. The range resolution of 0.2 m was achieved in an experiment. Experimental results were obtained using different emission sources. A description of resolution cell limitations is presented. The obtained results show acoustic noise source localization without the pre-metering of the range of the imaging plane, i.e., without a need to use the additional range meter which is not a part of the acoustic camera. The latter is important in tasks of non-destructive testing.*

Keywords: *Acoustic noise source localization, acoustic camera, bistatic reception, time difference of arrival, stationary acoustic noise signal.*

1. Introduction

Acoustic cameras are extensively used for acoustic noise source localization [1-3]. An acoustic camera shows the acoustic image overlaid onto the optic image, to assist the user in the identification of noise sources on the acoustic image [1]. The Direction of Arrival (DoA) data of a microphone array of an acoustic camera do not provide the range information, at ranges of the source larger than the array dimensions. The formation of an acoustic image using the DoA estimates requires the range value(s), at a short range span comparable to the array dimensions [2]. The estimation of the range requires the use of additional range meter (a ruler, a laser range finder, etc.) those are not a part of an acoustic camera. The application of the additional meter is connected to additional expenses and time lags. The paper describes an approach for

the formation of the images which allows the range estimation without the additional meter i.e. by the acoustic camera itself.

The acoustic location allows measuring the range of the emission source along its angular coordinates. Spatially diverse receiving units allow measuring DoA of the received signals (triangulation), or Time Difference of Arrival (TDoA) of them. When, however, more than one emission source exists and several of the sources are localized simultaneously, unwanted intersections of a directional lines may generate ghost spatial responses, i.e., the location uncertainty. Some bistatic sensors localize emission sources without generation of the ghost spatial responses [4, 5].

The paper is focused on the formation of acoustic images with the acoustic camera transformed to operate in the bistatic mode [4]. In opposite of known approaches [1, 2], the proposed focalization of the bistatic acoustic camera (onto nodes of the 3-D spatial coordinates) uses both TDoA and DoA information for imaging. The paper presents expressions for the formation of images, dimensions of resolution cell, and experimental results (got with different emission sources).

2. An acoustic camera operation

Numerous acoustic imaging approaches are based on delay and sum beamforming [1]. The beamforming is based on a model of the incoming wavefront (Fig. 1). The wavefront is described by slant ranges from the focusing point to positions of microphones and parameters of propagation media and emission [2]. The beamforming aligns the wavefront, to process incoming signals in phase. In the model, Euclidean distance denotes the slant range as:

$$(1) \quad d_m(\vec{r}) = \sqrt{(x - x_m)^2 + (y - y_m)^2 + (z - z_m)^2},$$

where $\vec{r} = (x, y, z)$ denotes the Cartesian coordinates of the focusing point; $\vec{r}_m = (x_m, y_m, z_m)$ denotes the coordinates of the microphone with index $m = 1, 2, \dots, M$, where M is the amount of microphones in the array.

To form each pixel of the image, the measured pressure signals are delayed and summed [2]:

$$(2) \quad B_i(\vec{r}, f) = \sum_{m=1}^M p_m(f) \exp(-j 2 \pi f d_m(\vec{r})/c),$$

where f is the frequency which is selected for the imaging; $p_m(f)$ is the measured sound pressure; c is the propagation speed of sound; $j = \sqrt{-1}$ is the imaginary unit. The obtained value (2) is to be normalized with respect to the input data.

The beamforming is suitable for imaging in the 3-D Cartesian coordinate system, at the range span comparable to the microphone array dimensions. At the range span, the range of the emission source (1) is required.

For stationary incoming emissions, the improved beamforming was proposed [2]. It enables to improve the side lobe level and other features of the pattern of the microphone array. The improvement is achieved by the nulling of the main diagonal of the cross-spectrum matrix.

Fig. 1. The two-dimensional imaging geometry for M – microphone array

It is assumed that the discrete-time complex output signals are under processing. Their spectra estimates are obtained with the discrete time Fourier transform. The frequency band Δf of several frequency nodes f_i , $i = 1, 2, \dots, N$, with the center frequency f_c is selected for the imaging. Hence, the cross-spectrum matrix is expressed as

$$(3) \quad F'(\Delta f) = \frac{1}{M^2} \begin{bmatrix} \sum_{i=1}^N p_1(f_i) p_1^*(f_i) & \sum_{i=1}^N p_1(f_i) p_2^*(f_i) & \dots & \sum_{i=1}^N p_1(f_i) p_M^*(f_i) \\ \sum_{i=1}^N p_2(f_i) p_1^*(f_i) & \sum_{i=1}^N p_2(f_i) p_2^*(f_i) & \dots & \sum_{i=1}^N p_2(f_i) p_M^*(f_i) \\ \dots & \dots & \ddots & \dots \\ \sum_{i=1}^N p_M(f_i) p_1^*(f_i) & \sum_{i=1}^N p_M(f_i) p_2^*(f_i) & \dots & \sum_{i=1}^N p_M(f_i) p_M^*(f_i) \end{bmatrix}.$$

The main diagonal of the matrix (3) contains the auto spectrum part of the latter matrix. Both upper and lower triangular parts of the matrix, above and below its main diagonal, consist of the cross-spectrum part. Absolute values of elements, situated symmetrically relatively to main diagonal of this matrix, are equal. Thus, each pixel of the acoustic image is formed as:

$$(4) \quad B_2(\vec{r}, f_c) = X(\vec{r}, f_c) F(\Delta f) X^H(\vec{r}, f_c),$$

where $F(\Delta f)$ is the cross-spectrum matrix $F'(\Delta f)$ with nulled main diagonal elements; $X(\vec{r}, f_c)$ is the row of $\exp(-j 2 \pi f_c d_m(\vec{r})/c)$ elements, which compensates the propagation delay for the focusing point $\vec{r}$; the superscript symbol H denotes the conjugate transpose.

The latter approach (4) is extensively used on practical conditions for the formation of acoustic maps and spatial filtration of incoming signals [2]. In the paper, it is used as the reference point because the range is required (1) to form the acoustic image. On practical conditions, the approach (4) requires using the additional meter mentioned above.

3. The operation of the transformed acoustic camera

Microwave bistatic systems are used for estimation of parameters of emission sources, in passive mode [4]. The bistatic reception enables measuring TDoA of the signal, additionally to its DoA (Fig. 2). This principle appears applicable to acoustic noise source localization which includes range measurement that is the purpose of the paper. An existing system which uses measurements of both angle(s) of arrival and time difference(s) of arrival for 3D localization was not found. Hence, the approach is examined as the new one.

3.1. The bistatic imaging approach

The bistatic acoustic system consists of the microphone array and single auxiliary microphone (Fig. 2).

Fig. 2. The two-dimensional imaging geometry for the bistatic acoustic system

The TDoA between their signals is expressed as:

$$(5) \quad \tau_m(\vec{r}) = \frac{d_{\text{aux}}(\vec{r}) - d_m(\vec{r})}{c},$$

where $d_{\text{aux}}(\vec{r})$ is propagation path length (the slant range) of the auxiliary microphone, c is the propagation speed of sound, as in (2). The absolute value of the path difference of arrival (PDoA) $d_{\text{aux}}(\vec{r}) - d_m(\vec{r})$ can't exceed the spacing Δ_m between auxiliary microphone and microphone with index m . Locus of possible sound source positions is one-half of the two-sheeted hyperboloid, for estimated TDoA and known coordinates of two microphones. The TDoA is estimated using cross-correlation of signals from the auxiliary microphone and microphones in the array of the acoustic camera. Dimensions of the locus are defined by the spacing Δ_m , the estimated TDoA and the width of the Δf .

It is assumed, that a stationary random noise signal is received. Cross-correlation of output signals of microphones enables estimating their TDoA. A finite

integration time enables estimating of Cross-Correlation Functions (CCF). The amount of the CCF in the passive system may be calculated using the modified expression for combinations without repetitions [6]. The latter corresponds to the amount of elements in one of the triangular parts of (3), mentioned above. The new approach includes the cross-correlation of signals from microphones in the array, at PDoA span up to $\pm \Delta_m$ (5).

The experiment for a microwave radiometric imaging using such bistatic system is described in [4], at the baseline length 3.75 m and the signal coherence length circa 0.65 m, and the PDoA span up to $\pm \Delta_m$ (5). A similar approach is suitable for Earth observation [7]. Respectively to [4], each pixel of the acoustic image is formed in the 3D Cartesian coordinate system as

$$(6) \quad B_3(\vec{r}) = \sum_{m=2}^M \left\{ R_m[\tau_m(\vec{r})] \exp[-j 2\pi f_c \tau_m(\vec{r})] \right\},$$

where $R_m[\tau_m(\vec{r})]$ is the node of CCF of stationary signals of auxiliary microphone and microphone m , defined by mutual delay, introduced to compensate $\tau_m(\vec{r})$. To eliminate measurement error, the image is normalized with respect to the amount of the used CCF.

The applicability of the approach (6) to the measurement of all three spatial coordinates of the source of the incoming acoustic signal is given below.

3.2. The experimental simulation of the resolution cell sizes

The experimental simulation of resolution cells was carried out indoor. The Brüel & Kjaer (B&K) acoustic camera was used as the sensor. A speaker was used as a point-like emitter. A noise signal was fed to the speaker. The speaker was placed at range 1 m, at boresight direction of the acoustic camera. Images of this section 3.2 were formed at the identical frequency range from 6.3-7.4 kHz ($f_c = 6.85$ kHz).

The acoustic camera equipment includes the non-uniform slice wheel microphone array with the diameter circa 0.33 m (B&K type WA-1558-W-021). The array contains 18 microphones (B&K type 4958). Two input modules (B&K type 3053-B-120 and 3050-B-060) fed signals to Pulse LabShop software. Output signals of the equipment were fed to further post-processing. The amplitude calibration was done in advance.

The post-processing was carried out with above approaches (4) and (6). The known distance of the emission source enables to form the acoustic image in the focal plane, according to (4). The response at the center of the image corresponds to the position of the speaker (Fig. 3a). Sound pressure estimate at peak of the response is circa -32.4 dB to 1 Pa, denoted as dB_{Pa} . The threshold is circa -12 dB to the peak value of the image. Thus, pixels with values lower -44.5 dB_{Pa} are not shown. The approach (4) was used to form the image without pre-defining of the range (Fig. 3b). The position of microphones is shown with black circles, at the zero range. A resolution of a sensor may be illustrated with dimensions of its point spread function, at the half power level. The threshold is circa -3 dB to the peak of the normalized image, to illustrate beam width. The approach (4) enables to form the

image at the elevation (0 m) of the center of the microphone array (Fig. 3c). Threshold circa -3 dB to the peak of the image was applied, to illustrate the response dimensions in range – cross range (top view) plane (Fig. 3c). The response shows range span from circa 0.5 m. The estimation of the exact range of the peak of the response is complicated.

Fig. 3. Acoustic images of noise emitter, placed at range 1 m: acoustic image, formed at the pre-defined range (a); normalized acoustic image, formed in the 3-D Cartesian coordinate system (b); normalized acoustic image, formed in range – cross range (top view) plane (c)

Fig. 4. Normalized acoustic images of noise emitter, placed at range 1 m: 17 estimates of CCF were used (a); 153 estimates of CCF were used (b)

In the next experiment, the transformation of the microphone array comprises moving one microphone 1 m towards left, from the center of the array (Fig. 2); other experimental setup has not been changed. The transformed array is suitable to form images using approach (6). The image was formed for the zero elevation. The formed image was normalized. Threshold circa -3 dB to the peak of the image was applied, to illustrate the response dimensions (Fig. 4a). Positions of microphones in the array are shown with black circles. The position of the moved microphone (on left) is out of the image bounds. Unwanted responses at the range more 2 m and cross range more 0.2 m are affected by multipath propagation inside the regular office room and other factors. Thus, usage of 17 CCF enabled to localize the emitter onto the formed image. Pairs or acquired signals are same to ones in the second row of the matrix (3), except the element with equal indexes. Index of the row equals to the number of the moved microphone. The amount of the combinations without repetitions is

calculated as [6]: $M(M-1)/2 = 153$, at the $M=18$ microphones. Usage of all the 153 pairs is connected to the processing of signals from nearby microphones. The distances between these microphones are comparable to the coherence length of the processed signal. The distances limit curvatures of locus areas of estimated TDoA, in the experimental setup. Hence, the usage of all the 153 pairs disables achieving range resolution, at the experimental scenario (Fig. 4b).

The latter approach (6) includes estimation of CCF and their addition, after application of corresponding phase adjustment. The shape of a processed spectrum defines parameters of the CCF (width of the main lobe, the level of side lobes, etc.). The equalizing of the spectra may be done, to achieve their rectangular shape. Application of both the latter approach (6) and the equalizing for imaging with the non-transformed microphone array enabled to estimate the exact range of the peak of the response, which corresponds to the range of the speaker (Fig. 5a). The absence of the spectra equalization makes range estimation complicated (Fig. 5b).

Fig. 5. Normalized acoustic images of noise emitter, placed at range 1 m: spectra equalizing (a); no spectra equalizing (b)

Both images were formed for the elevation of the center of the microphone array (0 m). A slight improvement of angular resolution at Fig. 5b comparable to Fig. 3c is connected to experimental conditions and difference in signal processing.

3.3. Limitations of the resolution cell size

The TDoA resolution depends on the width of the main lobe of CCF. The latter depends on the frequency bandwidth, not on the center frequency. At low frequencies (under 3 kHz), TDoA-based images may be more convenient than DoA-based images, in terms of the spatial resolution.

The points of intersection of edges (nulls) of both the array beam and TDoA response are the vertices of the curvilinear quadrangle $ABCD$ (Fig. 2), which limits the resolution of the acoustic system, as in [4]. Projections of the quadrangle on axes of coordinates are calculated as distances between maximum and minimum coordinates of its vertices. The vertices are obtained from cosine theorem, employing estimates of TDoA, azimuth, beam width and others. It enables a brief description of the resolution cell size.

The approach employs following parameters: center frequency (6.85 kHz), speed of sound (343 m/s), azimuth (from 7°-81°) and effective aperture of array at the main lobe position, bandwidth of processed signals (1.1 kHz), bistatic baseline (1 m), and path difference of arrival, expressed as $d_{\text{aux}}(\vec{r}) - d_m(\vec{r})$. One-dimensional microphone array with the aperture of 0.33 m is used. The employed values of the PDoA are -0.1559, 0.1559 and 0.4677 m.

The size of the quadrangle was calculated for the auxiliary microphone at the right of the array. For the first PDoA value (-0.1559 m), the quadrangle is limited by two locus areas (PDoA -0.3118 and 0 m), and by angular resolution of the microphone array. The cross-range size of the quadrangle is from ~0.2 up to ~0.5 m, at the azimuth span defined above. For the second PDoA value, the cross range size is from ~0.3 up to ~9.5 m, for azimuth up to 61°. For the third PDoA value, the size is from ~2 up to ~22 m, for azimuth from 14°-36°. For the first PDoA value, range size of the quadrangle is from ~0.5 up to ~3.5 m, for azimuth from 7°-61°. For the second PDoA value, the range size is from ~0.8 up to ~28.3 m, for the latter values of azimuth. For the third PDoA value, the range size is from ~2.7 up to ~27.5 m, at azimuth from 14°-36° best values of the above dimensions were obtained for azimuth from 20°-30°.

The first value of PDoA corresponds to TDoA response, which is wrapped towards boresight direction of the microphone array (see Fig. 2). Increasing of the PDoA to 0 m leads to the straightening of the TDoA response. Further increasing of the PDoA leads to the wrapping of the response towards opposite direction. The latter enlarges the quadrangle.

The range size of the quadrangle varies on paths difference of arrival, the microphone array aperture, baseline, and others. For the system with a single microphone array, the range size decreases with the shortening the range and with the widening of the bandwidth, and for some azimuth values. The enlarging of both baseline and microphone array aperture improve the size too. Defining limitations on the size may be suitable for in order to estimate some limitations for the field of view of the particular measurement task. Nevertheless, accuracies of estimation of azimuth and time difference of arrival depend on signal-to-noise ratio and other factors. Thus, the position and dimensions of the resolutions cell can't be equal to the described ones.

4. Experimental setup and results

The localization of sound emissions of rolling ball bearing(s) is the current application niche of the Acoustic Camera. The bistatic imaging was used for the localization. The obtained range estimates may be suitable for the reception of the sound emission from particular spatial coordinates of the bearing. That is suitable for a testing of the bearing.

The microphone array was placed in front of the equipment for vibration diagnosis of rolling ball bearings (Fig. 6). Manufacturer of the equipment is NSK Ltd. (NSK type: MS-730A). SKF bearing type 6205 was used. Parameters of the

bearing consist of the amount of rolling elements is 9, the rolling element diameter is 7.938 mm, the pitch diameter of the bearing is 39.04 mm, and the contact angle is 0° . The seat for the inner ring of the bearing is the spindle, which rotates with the shaft frequency circa 30.4 Hz. The pressing screw loads the outer ring of a bearing installed on the seat. The screw contains a vibration sensor. The signal of the sensor is used for diagnosis of the bearing, in the equipment. The acoustic emission of the installed bearing is under study.

Fig. 6. Picture of the experimental setup

The microphone array was transformed to the bistatic configuration (Fig. 7). The experimental setup includes acoustic noise source – the speaker. Three experiments were carried out. At the first experiment, the NSK equipment was turned on. The main sound emission source is its motor. At the second experiment, the interfering noise source was added. At the third experiment, the bearing was mounted to the NSK equipment. Approximate positions of these acoustic emission sources and microphones are given at Fig. 7.

Fig. 7. Scheme of the experimental setup (top view)

Spectra of the output signal of microphone channel 1 were estimated, at the three experiments. The spectra were averaged over 64 signals realizations. The duration of the realization is circa 1 ms. Resolution between frequency nodes of the spectra estimates is 256 Hz (Fig. 8).

Fig. 8. Spectra estimates at the output of microphone channel 1

Range-cross range images were formed corresponding to (6). The pixel size is 1.25 cm. The amount of processed CCF is 17. The spectra estimates were equalized to the rectangular shape, at the selected frequency bands. For first two experiments, the frequency band was from 6.8-8.6 kHz. The normalized image, formed for the first experiment is given in Fig. 9, after application of -3 dB threshold. The position of the response corresponds to the position of the motor behind the plexiglass panel (Fig. 7).

At the mentioned above frequency range, sound pressure estimate was circa -43.8 dB_{Pa} and -30.1 dB_{Pa}, in the first and second experiments, respectively. The 13.7 dB difference between these values is several dB bigger than the side lobe level of the microphone array. The normalized image formed for the second experiment is given (Fig. 10). The coincidence in the position of the peak value of the image and the position of the noise emitter (the speaker at Fig. 7) is satisfying.

Fig. 9. Normalized acoustic image for the first experiment

Fig. 10. Normalized acoustic image for the second experiment

For the third experiment, the frequency band from 5.8-7.9 kHz is used. To decrease the level of side lobes of both CCF estimates and pattern of the microphone array, bell-shaped weighting functions were applied. Thus, the frequency band is from 6.3-7.4 kHz (as in section 3.2). At the band, the sound pressure estimates were circa $-43.3 \text{ dB}_{\text{Pa}}$, $-29.5 \text{ dB}_{\text{Pa}}$ and $-22.6 \text{ dB}_{\text{Pa}}$ in experiments 1-3, respectively (Fig. 8). Corresponding normalized image with the threshold is given in Fig. 11.

Fig. 11. Normalized acoustic image for the third experiment

The images (Figs 9-11) show a slight difference between positions of the peak value of the image and position of the most powerful noise source. It is described by both accuracies of estimation of coordinates and interaction of spatial responses.

The equalization of the shape of spectra enabled significant suppression (over 20 dB) of the level of side lobes of estimated CCF of acoustic signals. Such values are larger than both side lobes and grating lobes levels of many modern transducer arrays. The sample of CCF estimated in the latter experiment is given in Fig. 12. Corresponding spacing Δ_{18} is circa 0.83 m. The level at PDoA 0.55 m is circa 10.3 dB. The level is larger than side lobe level for the applied Gaussian weighting function. The level corresponds to the multipath propagation and the presence of several signal sources, in the field of view. Such levels outside the main peak were estimated, in the latter experiment. The delays between main peak and corresponding spacing Δ_m were examined, for all CCF. Average value of the highest level is circa 9.5 dB to the average peak value. The latter is better than such values for spatial

responses of some of the commercially-available microphone arrays. The approach may be promising to decrease the unwanted penetration outside the main lobe.

Fig. 12. The sample of the CCF estimate for the third experiment

The above experiments show the opportunity to measure the range of the point-like emitter of the stationary noise signal, additionally to its angle(s) of arrival, without the use of the additional meter(s). It is achieved by the usage of the transformed acoustic camera.

Future work may be focused on a more detailed comparison of the proposed approach to different localization approaches, and on refinement of both resolution cell size and a field of view.

5. Conclusions

The new approach is proposed for measurement of all three spatial coordinates (range, cross range, and elevation) of acoustic noise emitters. The approach enables measuring both range and angular coordinates of emitters, in range span comparable to the spacing between the auxiliary microphone and the microphone array. Thus, an acoustic image formation may be started simultaneously, i.e., without the time lags required to measure the range value and to input it into acoustic camera software. The approach alleviates the need to use an additional range meter(s) to measure the range to the imaging plane.

The approach is based on bistatic reception. The approach provides estimation of the time difference of arrival of the acoustic signals, additionally to estimation of their angles of arrival. The experimental comparison of the approach to the extensively used beamforming method was carried out.

Limitations on achievable spatial resolution are given using geometrical considerations. In the experiment, the range resolution was up to 0.2 m, for the frequency range from 6.8-8.6 kHz. The achievable resolution depends on parameters of both the imager and the emitter.

Implementation of the approach is promising for acoustic noise source localization.

Acknowledgments: The research work reported in the paper was partly supported by the Project AComIn “Advanced Computing for Innovation”, grant 316087, funded by the FP7 Capacity Programme (Research Potential of Convergence Regions) and by FP-7 Project SCOUT (Grant 607019).

References

1. Michel, U. History of Acoustic Beamforming. – In: 1st. Berlin Beamforming Conference, 21-22 November 2006, Berlin, p. 17.
2. Christensen, J. J., J. Hald. Beamforming. – Brüel & Kjaer Technical Review, Vol. **1**, 2004, pp. 1-48.
3. Chyrka, I. D. A New Approach for Acoustic Field Interpolation. – Cybernetics and Information Technologies, Vol. **16**, 2016, No 1, pp. 135-145.
4. Lukin, K. A., V. V. Kudriashov, P. L. Vyplavin, V. P. Palamarchuk, S. K. Lukin. Coherent Radiometric Imaging Using Antennas with Beam Synthesizing. – International Journal of Microwave and Wireless Technologies, Vol. **7**, 2015, Special Issue 3-4, pp. 453-458.
5. Kudriashov, V. V., A. Yu. Garbar, S. K. Lukin, V. P. Palamarchuk, K. A. Lukin. Mapping of Acoustic Noise and Microwave Radiation. – Cybernetics and Information Technologies, Vol. **16**, 2016, No 1, pp. 126-134.
6. Thompson, A. R., J. M. Moran, G. W. Swenson Jr. Interferometry and Synthesis in Radio Astronomy. Springer, 2017.
7. Fernandez, D. M., A. Camps. Measurements and Simulations of a Doppler Radiometer in an Anechoic Chamber. – In: Proc. of Geoscience and Remote Sensing Symposium, 2004, pp. 3295-3298.