Technology of acoustic monitoring, detection and localization of objects in the controlled area

V Y Taranuha[1], Y M Linder[1], D O Volosheniuk[1] and A K Sieriebriakov[1,2]

[1] International Research and Training Center for Information Technologies and Systems NAS and MES of Ukraine, 40 Glushkov ave., Kyiv 03187, Ukraine

[2] E-mail: dep185@irtc.org.ua

Abstract. Given article describes an algorithm for determining the distance to the sound source based on estimating the incoming sound power and frequency variations over distance. In this way, it is possible to locate objects based only on their acoustic characteristics, without any visual information. The audio signal is recorded with the help of a group of four interconnected microphones, located on an immovable platform, which was shown to produce the best measurement results. In practice, the precise distance to the sound source is often unknown, which motivates its estimation with the help of other input parameters, such as the incoming sound intensity and its time of arrival. Using these input parameters allows to calculate distance to the sound source in the case of the impulse acoustic signal. Afterwards, it is processed by applying amplitude and frequency algorithms, which are shown to be robust to noises. Such approach can be applied effectively in the necessity of observing moving or stationary objects, the visibility of which may be restricted.

1. Introduction

Determining the direction to the source of the acoustic signal and estimating the distance to it are representing typical problems that are widely considered in the scientific literature [1-5].

Solution of such tasks according to the mentioned criteria becomes dominant, when alternative methods of object localization are limited or too complex, for example, in the case of restricted visibility.

Acoustic distance determination is used in the field of surveillance and monitoring of many types of objects, and can also be used to build an interface for practical systems of human-machine interaction, for mobile platforms such as surveillance cameras or teleconferencing systems [6], and also can be used for military purposes. Acoustic surveillance systems, in the case of availability of means of direction and distance determination, can be used in the fight against terrorism when the source of a potential threat needs to be identified.

Tasks concerned with the audio signal recording and estimation are usually solved with the help of the group of sensors, or microphones, connected together, in a quantity from 1 to 6 [7-15]. Microphones can be located on an immovable or movable platform. In the latter case, the primary task becomes more complex due to necessity of moving platform noise filtering in addition to the noise of the environment.

The following methods [16-18] for the sound source characteristics determination are distinguished: the application of neural networks, distance estimation based on sound power, by statistical data analysis, sound localization based on cross-correlation functions and others.

This article proposes a method for estimating the distance to the sound source, based on the application of 4 microphones for the impulse signal recording and its further processing using amplitude and frequency algorithms. Each of the algorithms allows to obtain a fairly accurate estimate of the distance to the sound source, even when the received signal is distorted by noise and echoes.

2. Physics of the investigated process

During an explosion, the sound source is a shock wave that turns into a sound wave. There are two phenomena acting as a sound source in the process of shooting: a muzzle wave and a ballistic wave.

A muzzle wave is a powerful sound pulse in the form of a spherical wave. Its center is slightly in front of the muzzle of the weapon. Near the gun (or rifle) the speed of the muzzle wave is slightly higher than the speed of sound, but at a distance of several dozens of meters, this speed does not differ from the speed of sound.

A ballistic wave is a region of increased pressure formed during the flight of a projectile (bullet) at supersonic speed. It has the shape of a cone that separates disturbed and undisturbed air.

Because it is desirable that the method is suitable for determining the direction to different sound sources and because it is not possible to know where the shot was aimed, and therefore estimate at least the azimuth for the axis of the ballistic wave cone, the muzzle wave has been chosen for the study.

It is important to note that the muzzle wave, despite its spherical shape, has an inhomogeneous distribution of energy in different directions at a considerable distance from the sound source. Thus, according to [19], the averaged difference between the volume in the direction of the shot and in the opposite direction reaches 15 dB. In addition, the presence of the muzzle brake and its design significantly affect the distribution of sound energy in different directions. This in turn necessitates the use of several different, and preferably independent, features or criteria in order to determine the distance to the sound source.

There are two features for determining the distance to the sound source:
1) based on the amplitude response,
2) based on the frequency response.

Determining the distance by the amplitude response is based primarily on the effect of spherical sound scattering. The following formula is used for the evaluation:

$$E(r) = \frac{E_0 d_0{}^a}{d_1{}^a},$$
(1)

where d_1 is the distance to the sound source;
a is the degree (parameter to be defined);
E_0 is the initial energy estimate, on the distance d_0.

In addition, it is necessary to take into account longitudinal attenuation of sound, which in general is defined by the following formula [20]:

$$p_t = \mathrm{p_i}\, e^{-0,1151\,\alpha\,d},$$
(2)

where p_t is the sound pressure;
p_i is the initial sound pressure;
d is the length of the sound propagation trajectory, in meters;
α is the coefficient of pure tone sound attenuation due to its absorption by the atmosphere;
Constant $0,1151 = 1/(10 \lg(e^2))$.

The formula, which allows to express the coefficient α depending on the frequency f, is given in the same fashion:

$$\alpha = Kf^2, \tag{3}$$

where K is a coefficient depending on the pressure and temperature of the environment;
f is the frequency of the sound.

To determine the distance based on the frequency response, it is possible to measure the ratio of the amplitudes of the frequencies at the opposite edges of the frequency range, which belongs to the studied fragment of the audio recording containing the shot or explosion. Substituting (3) into (2) and taking into account the ratio between different values, the following dependency for the evaluation is obtained:

$$F(r) = e^{a-rb}, \tag{4}$$

where a is the compensation component (parameter to be determined);
r is the distance to the sound source;
b is the value, which depends on the K coefficient from the formula (3) (it is assumed, that K remains unchanged during the observations), and on the upper and lower bounds of the frequency range, chosen for the experiment and known beforehand.

3. Physics of the investigated process

As it can be seen from formulas (1) - (4), a practically-oriented system should contain a module for the identification of the sound source type, which is motivated by the fact that the key model parameters, including initial sound power, depend on the nature of the sound source and the state of the environment.

The power of the observed sound is a highly variable value depending on: the type of sound source (initial energy), distance, orientation of the sound source, relief, etc.

The distribution of frequency components is a highly variable value depending on such factors as the type of sound source (initial energy), distance, sound source spatial orientation, terrain, etc.

Given that the accumulation of data and training of the identification system for all possible variations of parameters is a non-trivial task, the following approaches are proposed:
- application of the identification system based on neural networks only for determination of the type of sound source,
- assumption, that the impact of the terrain is largely offset by the fact that the microphones will be located on the aircraft,
- the influence of other variables, such as pressure, temperature, and so on are considered the same throughout the observation area.

Under such considerations, the main objective becomes a set of independent tasks aimed at determination of the distance to the sound source for one fixed type of signal.

4. Mathematical model

Let N be the quantity of input signals, which are recorded by four microphones (channels) simultaneously. The task is to estimate the distance to the source of the impulse audio signal with the help of each channel, as well as using the average signal from all channels. The relative error of the i-th observation of the j-th channel will be denoted as r_{ij}. The standard deviation of the relative error on the j-th channel will be denoted as s_j. This is the main index for estimating model's performance, which will be used further.

To express the model of the dependence of the distance towards signal over sound power by amplitude response, the following function was chosen:

$$d(a, v, w, s) = \frac{v}{a^w + z},$$ (5)

where a is the signal power,
d is the distance towards it,
v, w, z are model parameters.

As a model of the dependence of the distance towards signal over sound power by frequency response, serves the following function:

$$d(ff, g, p) = e^{g - p*a},$$ (6)

where a is the frequency response of the signal, calculated by algorithm,
d is the distance towards it,
g, p are model parameters.

5. Signal filtering algorithm

Input data: sound signal X of length divided into M segments with the sampling rate Fs; lower and upper filtering bounds Tr_b and Tr_a respectively.

1) Transform signal to the frequency spectrum with the help of fast Fourier transform.
2) Calculate real frequencies, corresponding to X_n values by formula:

$$f_n = \begin{cases} Fs \cdot \frac{n}{M}, & 0 \leq n \leq \left[\frac{M-1}{2}\right] \\ Fs \cdot \frac{M-n}{M}, & \left[\frac{M-1}{2}\right] \leq n \leq M - 1 \end{cases}.$$

3) Calculate Y_n in the manner:

$$Y_n = \begin{cases} X_n, & Tr_b \leq f_n \leq Tr_a \\ 0, & f_n > Tr_a \lor f_n < Tr_b \end{cases}.$$

4) Perform inverse Fourier transform of the spectrum Y_n, and receive filtered signal.

The signal filtering algorithm allows to neglect the low-frequency components of the input audio signal, which in a series of experiments carried out has increased the accuracy of the distance estimation.

6. Signal distance evaluation algorithm

Input data: set of multichannel signals $X_i, i = 1, 2 ..., N$ with sampling rate Fs; set of distances d_i for signals.

1) For each signal X_i from set, and for each channel $j \in \{1,2,3,4\}$ of input signal:
   - Find position of maximum absolute value of the signal pos_{max};
   - Cut-off signal $y_{ij} = x_{ij}[pos_{max} : pos_{max} + 0.6 * Fs]$;
   - Filter signal from each channel y_{ij} with the help of the previous algorithm with the filtering bounds (see the parameters given below).
2) Find respective signal characteristic a_{ij}.
3) Separate all signals into training and test dataset. Training set contains N_{tr} multidimensional signals, and testing set has N_{ts} signals, such that $N_{tr} + N_{ts} = N$.
4) For all channels $j \in \{1,2,3,4\}$ of the input signal find parameters of the curve $d(\cdot)$, which best describes relation of the received values of signals characteristics a_{ij} and distances d_i towards them for all $i = 1, 2 ..., N_{tr}$.

5) For each channel calculate value of standard deviation s_j, and choose channel j^* with lowest value s_j.

6) For each signal form testing set according to formula (1) calculate values $d(\cdot)$, which will be the distance towards signal estimators.

7. Sound data

The algorithm was tested on multi-channel recordings of micro-explosions at distances from 20 to 100 m (4 explosions at distances from 20 to 80 meters inclusive, 3 explosions at a distance of 90 meters and two explosions at a distance of 100 meters). It should be noted that micro-explosions, in contrast to shots, have a much higher variability of sound, because the destruction of the charge shell differs significantly between different explosions, in contrast to shots produced by firearm, in which the same charge of explosive accelerates similar bullet along similar barrel.

Total number of records is 33.

8. Evaluation of the distance towards sound by amplitude response

The following parameters are used for the algorithm:
- Filtering range from 10 kHz till 24 kHz;

- Amplitude response of the signal $a_{ij} = \sqrt{\dfrac{(y_{ij}-\overline{y_{ij}})^2}{0.6*Fs}}$, where $\overline{y_{ij}}$ is the mean value of the signal y_{ij}.

Table 1. Performance indices coefficient values for each channel by amplitude response

| Channel № j	Coefficient v	Coefficient w	Coefficient z	Standard deviation of relative errors s_j	Standard deviation of absolute errors	Mean value of relative errors $\overline{r_j}$
1	254830	1,466	2562,3	0,137	6,767	0,014
2	705320	1,616	7123,6	0,124	6,201	0,005
3	109850	1,274	927,9	0,095	4,482	0,007
4	308720	1,459	3380,5	0,18	10,061	0,03
Signal mean				0,119	6,010	0,014

As can be seen from the previous table, the best results were received for the 3rd channel, with an average deviation of absolute error of 4.5 meters. The average value of relative errors for all channels is close to zero, which means that the given model gives good approximation of the input data.

The normality of relative errors was checked using Anderson-Darling, Jarque-Bera and Liliefors tests. All tests were unable to refute the null hypothesis about the normality of relative errors. That is, the relative errors are likely to have a normal distribution.
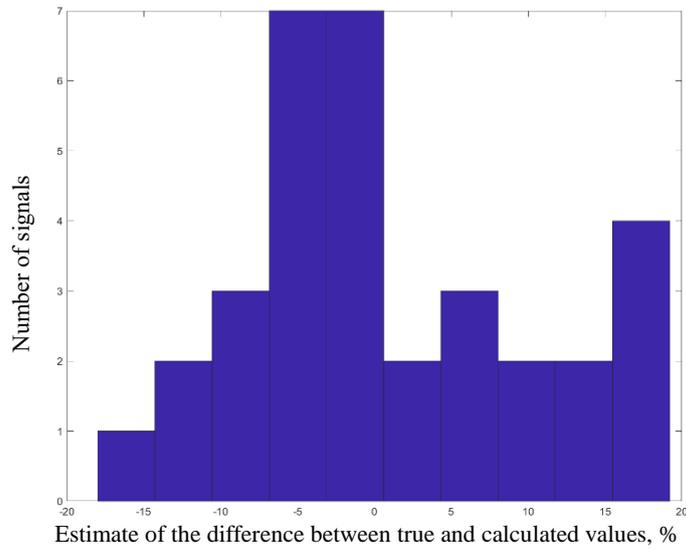
Figure 1. Relative errors distribution of the 3rd channel.

From the previous figure it is seen, that absolute error of the distance estimation doesn't exceed 20%.
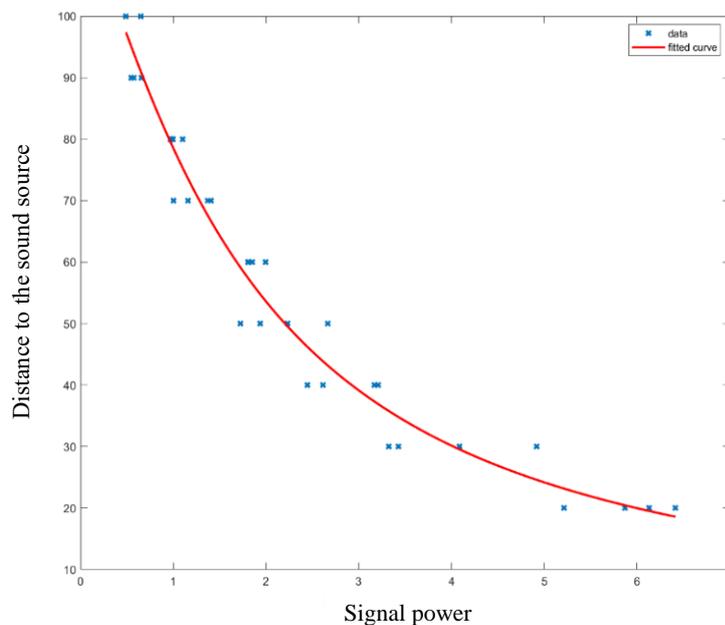
Figure 2. Plot of the comparison quality of experimental data with analytical model by amplitude response.

Distance towards signal estimation is close to the real values of distance.

9. Evaluation of the distance towards sound by frequency response

Input data:
- Filtering range from 12 kHz till 24 kHz;
- Frequency values are smoothed out with the help of the floating window with length equal to 200 units to avoid a sharp collapse at the edges;
- 1500 values are selected from the respective frequency range edges; it should be noted, that quantity of frequencies in this range (9500) doesn't coincide with the range length (12000);

- Frequency response of the signal is calculated as the ratio of sums of frequencies at the edges of the range: $ff_{ij} = \sum_{l=0}^{l=1500} f_l \cdot \left(\sum_{l=8000}^{l=9500} f_l\right)^{-1}$, where f_l is the amplitude with frequency l from the chosen segment.

Table 2. Performance indices coefficient values for each channel by frequency response

| Channel № j	Coefficient g	Coefficient p	Standard deviation of relative errors s_j	Standard deviation of absolute errors	Mean value of relative errors $\overline{r_j}$
1	5,047	0,406	0,5	17,731	0,233
2	4,937	0,351	0,54	16,928	0,139
3	4,949	0,364	0,587	16,383	0,124
4	4,937	0,402	0,276	13,805	0,07
Signal mean			0,399	12,8	0,114

As can be seen from the previous table, the best results were received for the 4th channel, with an average deviation of absolute error of 13.8 meters. The average value of relative errors for all channels is less than 3%, which means that the given model gives good approximation of the input data.

The normality of relative errors was checked using Anderson-Darling, Jarque-Bera and Liliefors tests. All tests were unable to refute the null hypothesis about the normality of relative errors. That is, the relative errors are likely to have a normal distribution.

Estimate of the difference between true and calculated values, %
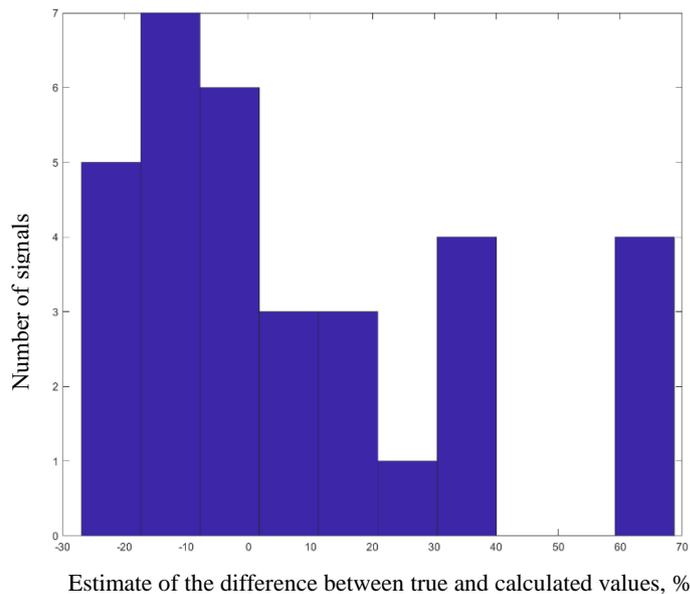
Figure 3. Relative errors distribution of the 4rd channel.

From the previous figure it is seen, that the distance estimation error is mainly located in the 20% range relative to true distance.
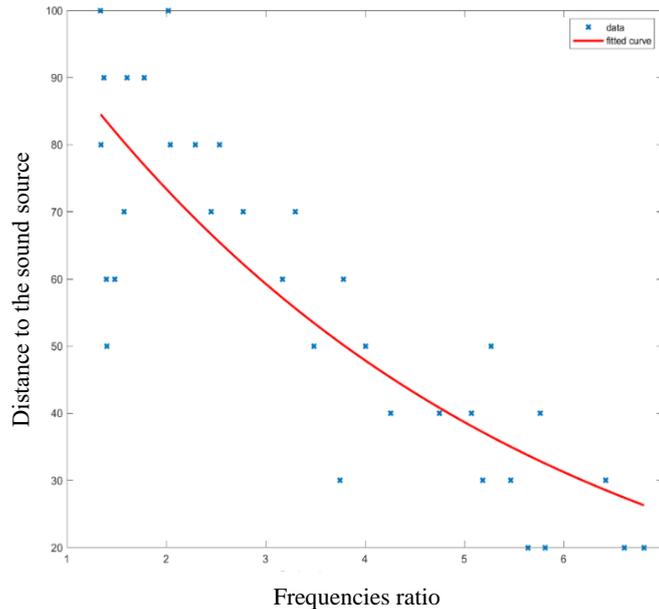
Figure 4. Plot of the comparison quality of experimental data with analytical model by frequency response.

According to the above, distance towards signal estimation is commensurate with the real values of distance.

Conclusions

An algorithm for determining the distance to the sound source based on amplitude and frequency characteristics was developed and implemented. The frequency component of the algorithm was developed because for certain types of sound sources their spatial orientation is critical, which affects the intensity of the observed sound signal, which in turn can mislead the amplitude algorithm and cause significant regular errors (> 30%).

The algorithm involves performing digital low-pass filtering of the recorded signal. Due to this filtering, the approach makes it possible to calculate the distance to the sound source with a relatively small error. The algorithm is implemented in the Matlab software environment.

In general, proposed method is suitable for fairly accurate determination of the distance to the sound sources under the conditions of restricted visibility and noises.

References

[1] Sree D C, Soura D and Zhi D 2009 Distance estimation from received signal strength under log-normal shadowing: bias and variance *IEEE Signal Proc. Letters* **16** pp 216-218

[2] Pourmohammad A and Ahadi S M 2012 Real time high accuracy 3-D PHAT-based sound source localization using a simple 4-microphone arrangement *IEEE Systems Journal* **6** pp 455-468

[3] Asano F, Asoh H and Nakadai K 2013 Sound source localization using joint bayesian estimation with a hierarchical noise model *IEEE Trans. on Audio, Speech, and Lang. Proc.* **21**

[4] Kim D H and Chung Y 2010 *Accurate Position Detection of Sound Source by LabView* (Sydney:Proceedings of 20th International Congress on Acoustics)

[5] Wind J, de Bree H-E and Xu B 2012 *3D Sound Source Localization and Sound Mapping using a PU Sensor Array* (Stockholm:16th AIAA/CEAS Aeroacoustics Conference)

[6] Al-Sheikh B, Elshebli A, Assaf A, Almashaqbeh S, Alrawabdeh W, Al-Tahat A, Abu baker

A B, Batayneh J and Jabra F 2013 *Sound Source Direction Estimation in Horizontal Plane Using Microphone Array* (Amman: IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies)

[7]   Saxena A and Ngueng A Y 2009 *Learning Sound Location from a Single Microphone* (New York:IEEE International Conference on Robotics and Automation)

[8]   Fan J, Luo Q, Ma D 2010 *Estimation of Sound Source by Microphones Array* vol. 7 (New York:Symposium on Security Detection and Information Processing) pp 312-317

[9]   Bhanu P R, Lavanya H, Lavanya G, Althuru P and Girish G K 2017 Direction and distance finding using microphone array of a sound source *Int. Journal of Cur. Eng. and Tech.* **7**

[10]   Jiang W, Cai Z, Luo M and Yu Z L 2011 *A Simple Microphone Array for Source Direction and Distance Estimation* (Kristiansand:IEEE Conference on Industrial Electronics and Applications)

[11]   Valin J-M, Michaud F, Rouat J and Letourneau D 2003 *Robust Sound Source Localization Using a Microphone Array on a Mobile Robot* (Las Vegas:IEEE/RSJ International Conference on Intelligent Robots and Systems) pp 1228-1233

[12]   Pavlidi D, Griffin A, Puigt M and Mouchtaris A 2013 Real-time multiple sound source localization and counting using a circular microphone array *IEEE Trans. on Audio, Speech, and Lang. Proc.* **21** pp 2193-2206

[13]   Pourmohammad A and Ahadi S M 2013 N-dimensional N-microphone sound source localization *J Audio Speech Music Proc.* **27**

[14]   Zhong X., Sun L and Yost W 2016 Active binaural localization of multiple sound sources *Robotics and Auto. Sys.* **85** pp 83-92

[15]   Cobos M, Antonacci F, Alexandridis A, Mouchtaris A and Lee B 2017 A survey of sound source localization methods in wireless acoustic sensor networks *Wireless Com. and Mobile Computing*

[16]   Yiwere M and Rhee E J 2017 Distance estimation and localization of sound sources in reverberant conditions using deep neural networks *Int. Journal of Applied Eng. Research* **12** pp 12384-12389

[17]   Georganti E, May T, Par S, Härmä A and Mourjopoulos J 2009 *Single Channel Sound Source Distance Estimation Based on Statistical and Source Specific Features* vol. 1 (Munich:126th Audio Engineering Society Convention) p 1

[18]   Wan X and Wu Z 2013 Sound source localization based on discrimination of cross-correlation functions *Applied Acoustics* **74** pp 28-37

[19]   Pater L and Shea J 1981 *Techniques for Reducing Gun Blast Noise Levels: An Experimental Study* vol. 61 (Dahlgren:Naval Surface Weapons Center)

[20]   ISO D.I.N. 9613–1 1993 *Acoustics. Attenuation of Sound During Propagation Outdoors. Part 1: Calculation of the Absorption of Sound by the Atmosphere* (Geneva:International Organization for Standardization) p 50