FOURFOLD MICROPHONES AREA FOR ACOUSTIC SOURCE LOCALIZATION

Olimpiu POP, Lacrimioara GRAMA, Corneliu RUSU
Signal Processing Group, Basis of Electronics Department,
Faculty of Electronics, Telecommunications and Information Technology, Technical University of Cluj-Napoca, Cluj-Napoca, Romania
olimpiupop@yahoo.com, {Lacrimioara.Grama; Corneliu.Rusu}@bel.utcluj.ro

Abstract: In this paper we present a fourfold microphones area for source localization. The area consists of four shotgun microphones arranged at an angle of 90^0 between them. The distance between the microphone's capsules can be changed in a range of 10 ÷ 20 cm. This is needed to study the phase differences between the acoustic signals captured for a given bandwidth. Experimental results based on spectrum and autocorrelation measurements have shown that this area can be used in domestic applications or for wildlife areas intruder detection.

Keywords: *microphone, intruder detection, spectrum, correlation.*

I. INTRODUCTION

Source localization using acoustic emission (voice, moving noise, animals) is an attractive method to determine the position in multiple applications [1]. Indeed, acoustic emission monitoring is a passive localization method, non-invasive, useful in security applications, supervising wildlife, natural disaster prevention.

For instance, in case of wildlife intruder detection systems our experimental results have shown that the normal speech can be detected in open space at a distance of about 20-40 meters, according to the vegetation of the field. However, the loud speech may be heard at few hundred meters. Moreover, the engine of a chainsaw or truck can be detected normally at a distance less than 1 kilometer; usually this distance is around half of kilometer. Finally, the sound of a rifle may be heard from few kilometers (1 to 5 kilometers), depending on the type of the rifle.

Location accuracy involves placing multiple sensors such microphones in area of interest. Wireless acoustic sensors network is an extremely flexible structure and very attractive. This is especially recommended when using physical cable connections is complicated and impractical. However, the big challenge in this case is the level of on board resources (power supply, storage, computing power, etc.).

Applications based on acoustic sensors collect information by capturing the sound signal that is further used by processing algorithms. By adapting the acoustic and electrical sensor's parameters (or of an area of sensors) to certain types of sound signals, the efficiency of the system can increase making possible to integrate the data processing - in situ. By filtering and processing of primary information, the amount of data can be decreased. In this way, it is possible to further transmit data through a collection point for further processing. The transport of an optimized data volume means lower power consumption. This represents an important aspect for developing a smart sensors network, autonomous, sustainable from the power supply point of view.

In this work, we intend to develop an experimental model for laboratory measurements and field tests. With this prototype, we can determine the characteristics of this area of microphones in different configurations and environmental conditions. Few of these characteristics might be directivity, frequency response and signal to noise ratio.

In this paper, we shall analyze the results of the laboratory measurements and of the tests under real conditions. Our secondary goal is to optimize the experimental model: types of microphones, geometry of area, filtering and sampling of signals, optimizing the transfer rates and reducing the volume of transmitted data.

The rest of the paper is organized as follows. In Section II the theoretical background is presented, while the experimental setup is the subject of Section III. In Section IV experimental results are shown. Finally, Section V describes future developments and concludes the paper.

II. THEORETICAL BACKGROUND

A. Why an Area of Microphones Instead of a Single Microphone?

Since the propagation of sound is a three-dimensional (3D) process, its capture in a single point with one microphone does not give us enough data to describe 3D acoustic phenomena such as ambient noise, reverberation, multiple sound sources [2]. On the other hand, the use of an area of microphones allows the estimation of sound signal arrival direction, i.e. its localization.

One microphone captures too much noise and reverberation, which limits the noise reduction algorithms performance. With the help of an area of microphones and by combining signals from all the microphones within the given area, one can get a very good microphone directivity.

Manuscript received August 30, 2017; revised September 22, 2017

A good microphone directivity will capture less noise, thus increasing the noise removal algorithms performance [3].

The area of microphones can be:

1. Linear area of microphones – the microphones are placed on the same line; the workspace is a semi-plane, i.e. a hemisphere;
2. Circular area of microphones – the microphones are placed concentrically, from 3 to 16; they are omnidirectional;
3. Plane area of microphones – the microphones capture in a hemisphere; may be at least 3, up to 4 or 8;
4. Volumetric area of microphones – is a 3D area which can capture the sound from any direction, in a 3D space; there must be at least 4 microphones; the shape is spherical; a natural shape should be the oval (i.e. the shape of a human head with the microphones placed in the position of ears).

B. Localization Methods of Acoustic Sensors

Sound source localization can be achieved using several methods which are based on the evaluation of the parameters of the signal received from sensors:

- Evaluation of the differentiate amplitude of the received signals;
- Evaluation of the time difference for signals arriving from different sensors with known position (TDOA – Time Difference-Of-Arrivals);
- Determining and comparing the directions of arrival (DOA – Direction-Of-Arrivals);
- Evaluation of acoustic energy.

Depending on environmental conditions and on the conformation of the field which can produce reflections, reverberations, can introduce attenuation. Thus, each of the above mentioned methods can be more or less effective, based on the environment.

C. Microphone Networks

Despite the advantages of areas of microphones compared to systems with a single microphone, the traditional microphones' networks also have their limitations. This because they usually probe the acoustical space mostly in their immediate vicinity, and to a relatively great distance compared to the sound source. Due to limitations of size and power consumption, especially for small portable equipment, the processing power is also limited [4]. To counteract these drawbacks wireless links microphones networks have been proposed.

An acoustic sensor array consists of a set of several sensors placed in the network's nodes. For example, the acoustic sensors may be settled at the corners of the squares or hexagons. The network nodes are interconnected, respectively connected to the central point through wireless technologies [5]. In this way, the limitations related to the covered area is no longer a major problem and also the amount of information collected is considerably increased.

Several considerations must be considered in case of achieving these kind of microphones networks:

- For an ad-hoc distributed network, the microphones from the nodes can be uncalibrated and also their position may be not well established or known. In this case the information related to network geometry cannot be used;
- In some applications, the position of microphones and also their number can vary;
- In most of the cases there are serious limitations in terms of signals bandwidth which are to be transmitted between nodes. Often, the available bandwidth is insufficient, since we are talking about tens or hundreds of microphones whose signal must come from one node to another or to the central point. For this reason, it is very important that the available bandwidth for transmission to be used as efficiently as possible;
- If the case of the distributed networks there is a need for primary processing of data at the nodes level as efficiently as possible without affecting their content. This requirement should be considered due to constraints related to transmission bandwidth, to minimize the amount of information that must be transmitted. In this way one can reduce energy use, often limited, from the central processing node;
- Another issue is the synchronization of transmitted data. Since each sensor from the network's nodes has its own system clock, some differences between sampling rates of signals from microphones will appear. This problem will seriously affect subsequent processing using algorithms based on signals consistency.

These requirements lead to different signal processing approaches and of the algorithms used to implement areas of microphones.

III. EXPERIMENTAL SETUP

The process block diagram is exemplified in Fig. 1. It consists of 4 microphones, an audio acquisition board and a Raspberry Pi development board.

Further, the four audio channels are transferred into an audio signal processing program, called Audacity. It is a free open source digital audio editor and recording software, widely used for applications that require the capture and processing of audio signals.

Via an adapter connected to a C-Media USB port on the Raspberry Pi motherboard, we converted the analog sound signal into a digital signal. This signal was applied to a sound processor that allows us to initiate processing, filtering, sampling, compression and recording, as required.

A. Area of Microphones

To realize the area of microphones we used a total of four ECM-670 shotgun microphones placed on a flat surface of a square, diagonally arranged at an angle of 90^0 between them (Fig. 2).

The distance between the microphones capsules can be changed in a range of $10 \div 20$ cm. This change of the distance between the capsules is needed in order to study the phase differences between the acoustic signals captured for a given bandwidth (maximum detectable frequency without unambiguously phase).

To eliminate confusion during measurements, each microphone is labeled by a cardinal point:

- Mic 1 (North);
- Mic 2 (East);
- Mic 3 (South);
- Mic 4 (Vest).

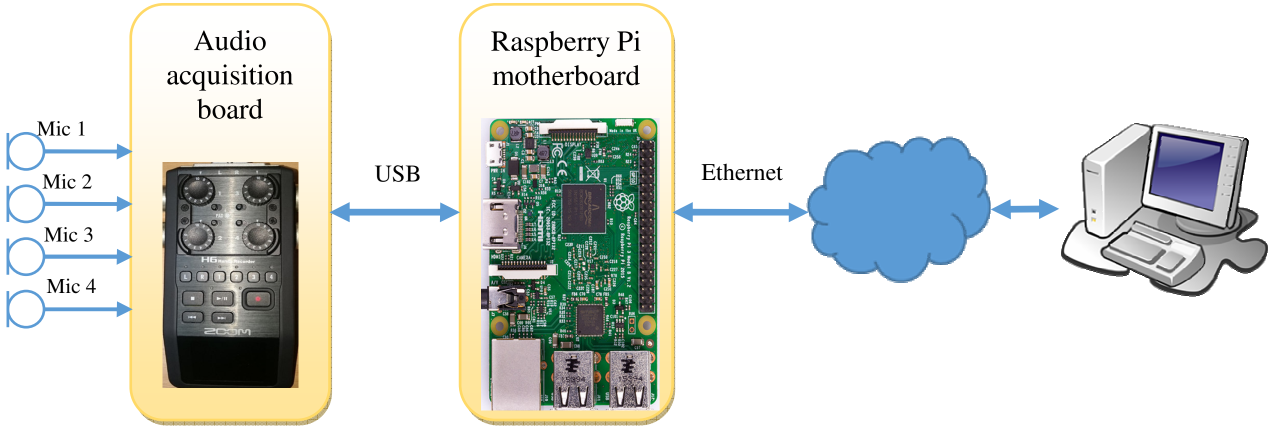

Figure 1. Simplified block diagram.

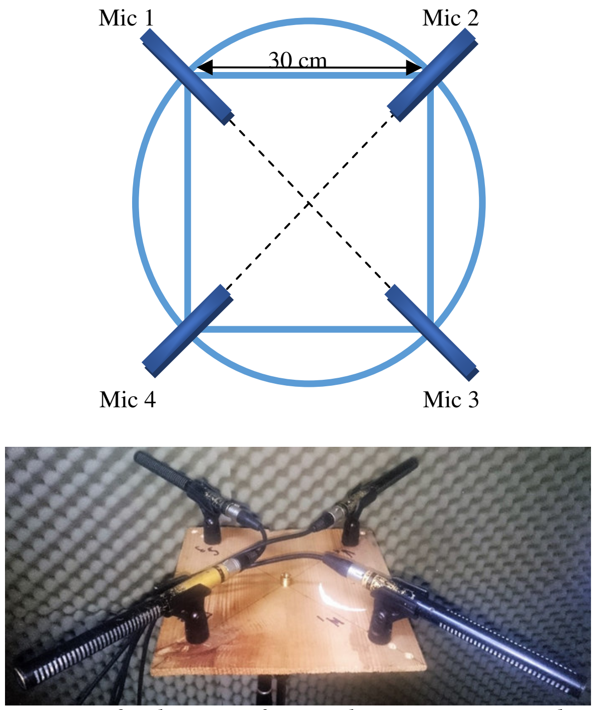

Figure 2. The area of microphones geometry and placement.

ECM-670 shotgun microphone

The ECM-670 is a compact shotgun microphone which has a highly directional response, allowing the microphone to be used for voice pick-up in noisy environments [6]. Here we shall recall several of its specifications, used during the design:

- Capsule type: back electret condenser;
- Frequency response: 70 Hz ÷ 16 kHz;
- Directivity: Uni-directional (super-cardioid);
- External power supply: DC 12÷48 V;
- Current consumption: less than 2.4 mA;
- Output connector: XLR-3-12C type;
- Sensitivity: -44.0 ± 3.0 (dB);
- Dynamic range: equal to or greater than 101 dB;
- Effective output level at 1kHz (0dBm=1mW/1Pa): -42 dBm;
- Signal-to-noise ratio (A-weighted, 1 kHz, 1 Pa): 70 dB or more;
- Inherent noise: 24 dB SPL or less;
- Wind noise: 60 dB SPL or less;
- Max. input sound pressure level: 125 dB SPL;
- Output impedance at 1kHz (balanced): 200 Ω ± 20%.

B. Audio Acquisition Board

The area of microphones is connected to an audio acquisition board (Zoom H6 handy recorder [7]) which have 4 symmetrical XLR inputs with phantom power (48 V) needed to supply the four condenser type microphones. Each channel has an amplifier with adjustable level and frequency filters.

For measurements, we used a constant amplification level and we did not filter the bands, to assess the full spectrum and to determine the signal to noise ratio, in each situation. The sampling frequency for all the recorded signals was 44.1 kHz.

Zoom H6 handy recorder

The H6 offers four main inputs which are combo connectors that can accept either XLR or ¼” balanced or unbalanced phone cables. They can handle both mic- and line-level signals. All connectors use the industry standard Pin 2 hot on XLR connectors. The inputs have a dedicated gain control knob as well as a -20 dB pad, allowing to prevent distortion even when high-level signals are introduced. All capsules utilize higher voltage preamps (5 volts instead of the more commonly used 3 volts) for distortion free recording, even at high volumes. A built-in instrumentation amp allows signals to be transmitted with minimal noise even when long cables are used. If using high-quality condenser (powered) microphones (as in our case - ECM-670 shotgun microphones) a simple menu option allows the unit to provide Phantom Power (either +12, +24, or +48 V) to any or all of the main inputs. The H6 USB port provides a digital output of either a stereo mix or the all individual input signals, depending upon the setting of the “Audio Interface” function in the USB menu.

C. Raspberry Pi 3 Model B Motherboard

Using the USB port of the Zoom H6 handy recorder we connected the Raspberry Pi development board [8]. Raspberry Pi is a SBC (Single-Board Computer) which can be used successfully in many real-time applications, concerning with measurement, monitoring and control. A major advantage of this minicomputer are the integrated auxiliary circuits (inputs outputs, USB ports, Ethernet, etc.). This system has also developed a series of sensors that can be integrated with very good results for environmental monitoring and control. It can be controlled by a TCP/IP protocol.

The audio signal captured from the Zoom H6 audio card sensors is digitized, multiplexed and transferred to a serial port (USB) on the Raspberry Pi board. This microcomputer runs with a Linux-derived operating system called Rasbian. On this microcomputer (SBC) is installed an open-source application Audacity.

The four channels multiplexed by the audio acquisition board are taken over by this application, demultiplexed and displayed on a graphical interface that can be accessed on a computer connected in the same network. Each of the four channels can be processed (filtered, limited, sampled, etc.) using application functions. Also, processed data can be saved in a file that can be accessed or transferred via a mobile data network (Internet protocol) to a destination, a personal computer, for other complex processing.

At the current stage of the work, the processing and operations that will be executed by the application and the microcomputer are not specified. Of course, it will be desirable to streamline the content and reduce the volume of data to be able to carry them using a lower transfer rate, considering also that the system will operate in a hostile environment with independent energy sources.

The Audacity application installed on the microcomputer allows access to the input data from its four microphones in real time. The resources required by this application to be installed on a microcomputer running a Linux operating system are 64 MB RAM and a 300 processor MHz. The Raspberry Microcomputer has 1 GB of RAM and a 1.2 GHz processor far beyond the application requirements.

In the following the project will use Raspberry microcomputer for data processing using Audacity application in location and transfer files, from a central point via a data network

IV. EXPERIMENTAL RESULTS

To model the parameters of the area of microphones we have conducted several measurements in different environments, using various sources. However, in this work we shall present the results obtained into an acoustically designed room (concert hall with a noise level below 30 dB), where we used an amplified sinusoidal signal generator.

We performed a series of measurements using multiple frequencies and we have determined the frequency response characteristic of the area of microphones in these conditions. These experiments were needed for calibrating our prototype for future experiments.

The recordings were imported into Audacity. This software allows simultaneous recording of four or more audio channels. There is also the possibility of processing this information. The measurement results for audio signals are illustrated as follows:

- time domain representation – gray background;
- magnitude of the spectrum – purple;
- enhanced autocorrelation – magenta.

For obtaining the spectrum and the enhanced autocorrelation [9] corresponding to the sound signal captured by each microphone, we have used a Hanning window of length 512.

A. First Set of Measurements

In the first experiment, no signal was present (almost silence). We can observe the signals (0 amplitude) at all four microphones (Fig. 3).

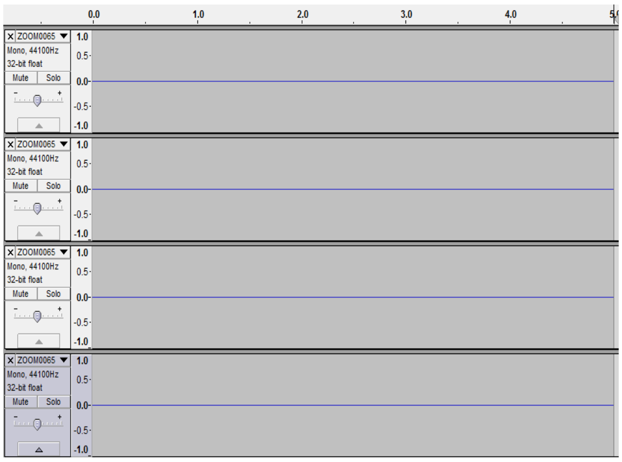

Figure 3. Acoustically designed room – silence (1st row - Mic 1 (North); 2nd row - Mic 2 (East); 3rd row - Mic 3 (South); 4th row - Mic 4 (Vest)).

The spectra corresponding to audio signals present at each microphone are not all illustrated, since they are quite similar. In Fig. 4. We present just the spectrum for Mic 1. The enhanced autocorrelation does not contain significant information to be presented here.

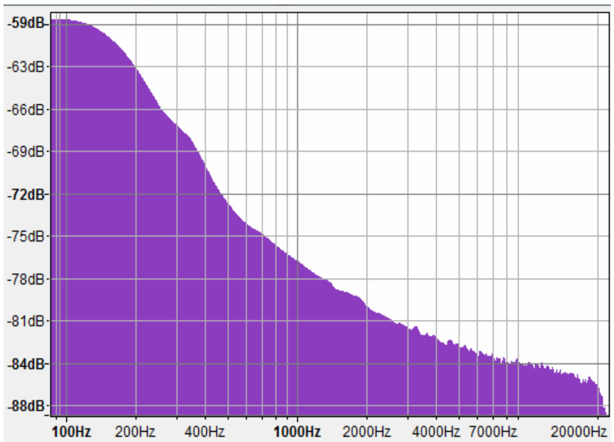

Figure 4. Acoustically designed room – silence – spectra (Mic 1 (North)).

B. The Second Set of Measurements

For the second experiment, we have used a sinusoidal signal, with frequency of 5 kHz, present 3 m far from Mic 1 (North). In Fig. 5 the signals recorded by all four microphones are shown.

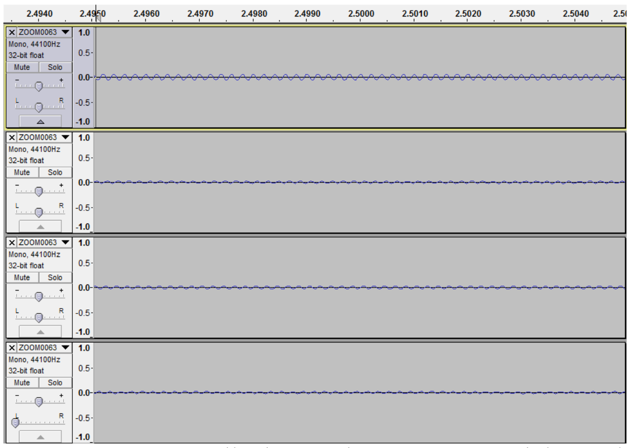

Figure 5. Acoustically designed room – sinusoidal signal with frequency 5 kHz (1st row - Mic 1 (N); 2nd row - Mic 2 (E); 3rd row - Mic 3 (S); 4th row - Mic 4 (V)).

The spectra corresponding to audio signals present at each microphone is also illustrated in Fig. 6. We can observe from the spectra that the frequency of 5 kHz is present in all plots. The nearest microphone Mic 1 (North) recorded -24 dB, while Mic 2 (East) and Mic 4 (Vest) recorded -32 dB, and finally, Mic 3 (South) recorded -29 dB. Besides, we can see that the enhanced correlation has clear periodicity, though the levels of peaks are decreasing. This behavior is well known in signal processing [10].

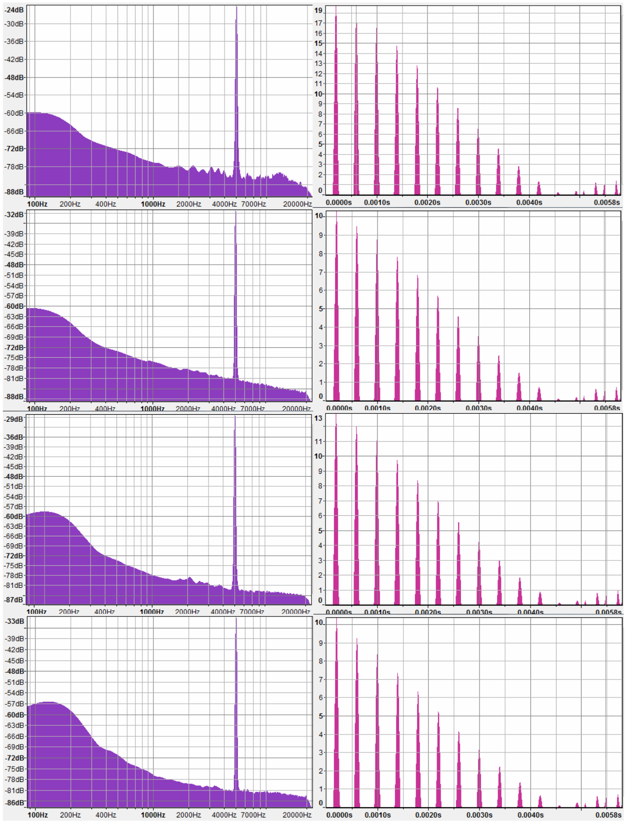

Figure 6. Acoustically designed room – sinusoidal signal with frequency 5 kHz - spectra (1st row - Mic 1 (N); 2nd row - Mic 2 (E); 3rd row - Mic 3 (S); 4th row - Mic 4 (V)).

C. The Third Set of Measurements

For the third experiment, we have used a gunshot signal, present 5 m far from Mic 1 (North). In Fig. 7 the signals recorded by all of the four microphones are shown.

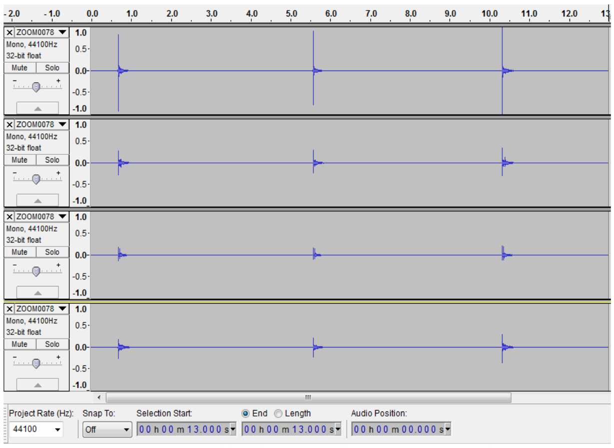

Figure 7. Acoustically designed room – gunshot 5 m (1st row - Mic 1 (N); 2nd row - Mic 2 (E); 3rd row - Mic 3 (S); 4th row - Mic 4 (V)).

The spectra corresponding to audio signals present at each microphone is also illustrated in Fig. 8. We can observe from the spectra that the sound is more prevalent for Mic 1 (-58 dB), thus we can obtain the direction of the sound source. Note the behavior of enhanced correlation, which is specific to an impulse of finite duration.

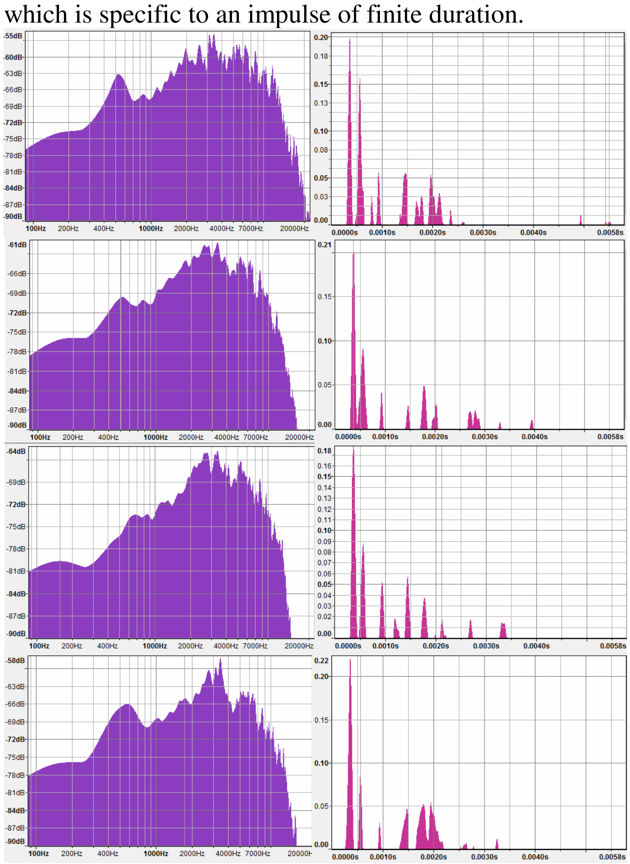

Figure 8. Acoustically designed room – gunshot 5 m – spectra (1st row - Mic 1 (N); 2nd row - Mic 2 (E); 3rd row - Mic 3 (S); 4th row - Mic 4 (V)).

D. The Fourth Set of Measurements

For the fourth experiment, we have used a gunshot signal, present 10 m far from Mic 1 (North). In fig. 9 we can observe the signals recorded by all the four microphones.

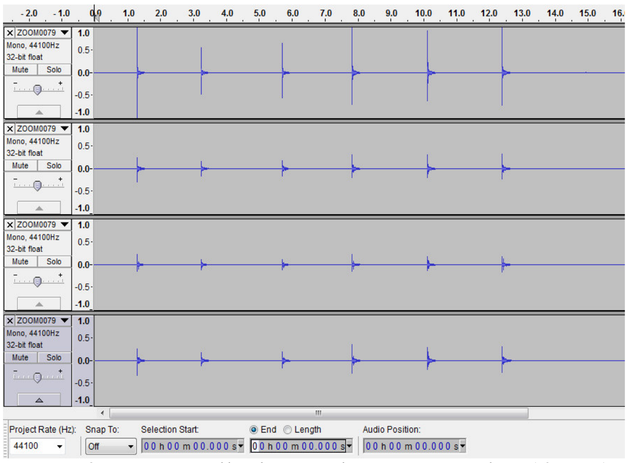

Figure 9. Acoustically designed room – gunshot 10 m (1st row - Mic 1 (N); 2nd row - Mic 2 (E); 3rd row - Mic 3 (S); 4th row - Mic 4 (V)).

The spectra corresponding to audio signals present at each microphone is also illustrated in Fig. 10. We can observe from the spectra that the direction from where the gunshot is coming is Mic 1 (-45 dB).

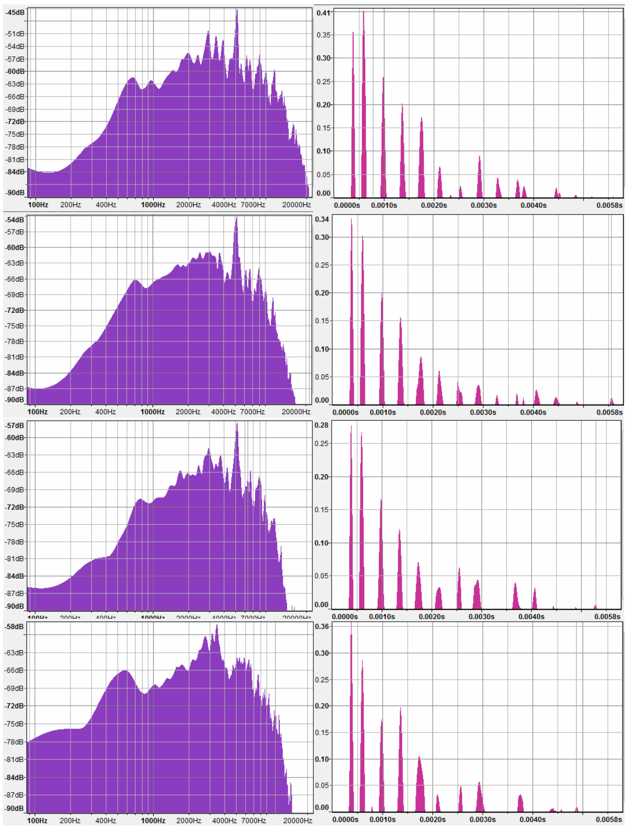

Figure 10. Acoustically designed room – gunshot 10 m – spectra (1st row - Mic 1 (N); 2nd row - Mic 2 (E); 3rd row - Mic 3 (S); 4th row - Mic 4 (V)).

V. CONCLUSIONS

In this paper, we have presented a fourfold microphones area for source localization. The area consists of four shotgun microphones arranged at an angle of 90^0 between them. The distance between the microphone's capsules can be changed to study the phase differences between the acoustic signals captured for a given bandwidth. Experimental results based on spectrum and autocorrelation measurements have shown that this area can be used in domestic applications or for wildlife areas intruder detection.

Another future aim of this work is to develop a device capable to record multiple audio signals (in our case 4 audio signals from 4 microphones of the area) and transmit the file through a network of data transmission (internet, over IP), initiated on request. As a future development, the application can achieve this goal automatically, on the occurrence of a sound source in the surveyed area. Operation of the experimental model (in the current stage of research) requires certain data networks in the geographical surveyed area. This requirement is limiting quite seriously the effectiveness of the application. Another solution for data transmission are related to wireless networks with a central point of communication that have access to a data network.

ACKNOWLEDGMENT

The work of the first and third authors was supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNCS/CCCDI–UEFISCDI, project number PN-III-P2-2.1-PED-2016-1608, 222PED/2017, within PNCDI III. The work of the second author was supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNCS/CCCDI–UEFISCDI, project number PNIII-P2-2.1-BG-2016-0378, 54BG/2016, within PNCDI III.

REFERENCES

[1] A. M. Ali, S. Asgari, T. C. Collier, M. Allen, L. Girod, R. E. Hudson, D. T. Blumstein, "An empirical study of collaborative acoustic source localization," *Journal of Signal Processing Systems*, vol. 57, no. 3, pp. 415-436, 2009.

[2] E. Hänsler, G. Schmidt, G. (2005). Acoustic echo and noise control: a practical approach, vol. 40. John Wiley & Sons, 2005.

[3] I. J. Tashev, "Sound Source Localization and Tracking with Microphone Arrays," *Sound Capture and Processing: Practical Approaches*, pp. 263-305, 2009.

[4] H. B. Wang, C. E. Chen, A. M. Ali, S. Asgari, R. E. Hudson, K. Yao, C. E. Taylor. Acoustic sensor networks for woodpecker localization. Center for Embedded Network Sensing, 2005.

[5] Y. A. Huang, J. Benesty. Audio signal processing for next-generation multimedia communication systems. Springer Science & Business Media, 2007.

[6] Sony ECM-670 shotgun microphone specifications, [Online]. Available: http://www.sony.ro/pro/product/broadcast-products-pro-audio-wired-shotgun/ecm-670/overview/#overview.

[7] Zoom H6 handy recorder features, [Online]. Available: https://www.zoom-na.com/products/field-video-recording/field-recording/h6-handy-recorder/features

[8] Raspberry Pi 3 model B, [Online]. Available: https://www.raspberrypi.org/products/raspberry-pi-3-model-b/

[9] T. Tolonen, M. Karjalainen, "A Computationally Efficient Multipitch Analysis Model," *IEEE Trans. on Speech and Audio Processing*, vol. 8, no. 6, pp. 708-716, Nov. 2000.

[10] J. G. Proakis, D. G. Manolakis. Digital Signal Processing: Principles, Algorithms and Applications. NJ: Pearson, 2014.