USE OF MICROPHONE DIRECTIVITY FOR THE LOCALIZATION OF SOUND SOURCES

by

Mahdi Tajari

BS, University of Tehran, 2003

MS, University of Tehran, 2006

Submitted to the Graduate Faculty of

Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This thesis was presented

by

Mahdi Tajari

It was defended on

May 10, 2010

and approved by

Dr. Amir Koubaa, Assistant Professor,
Department of Civil and Environmental Engineering

Dr. Albert To, Assistant Professor,
Department of Civil and Environmental Engineering

Dr. Piervincenzo Rizzo, Assistant Professor,
Department of Civil and Environmental Engineering
Thesis Advisor

USE OF MICROPHONE DIRECTIVITY FOR THE LOCALIZATION OF SOUND SOURCES

Mahdi Tajari, M.S.

University of Pittsburgh, 2010

In a recent paper [1] the proof-of-concept of a novel approach for the localization of sound source was demonstrated. The method relies on the use of unidirectional microphones and amplitude-based signals' features to extract information about the direction of the incoming sound. By intersecting the directions identified by a pair of unidirectional microphones, the position of the emitting source can be identified.

In this study we expand the work presented in that paper by assessing the effectiveness of the approach for the localization of an acoustic source in an indoor setting. As the method relies on the accurate knowledge of the microphones directivity, analytical expression of the acoustic sensors polar pattern were derived by testing them in an anechoic chamber. Then an experiment was conducted in a classroom-type environment by using an array of three unidirectional microphones. The ability to locate the position of a commercial speaker placed at different position is discussed.

It is believed that this method may pave the road toward a new generation of reduced size sound detectors and localizers.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like express my thanks to Dr. Piervincenzo Rizzo for his guidance in this research project. His motivation for me kept me going and he played a key part in getting this document into its final form. I want to thank him for all of his time and input on this project.

I would like to thank my Peers Mr. Mohammad Jamaly, Mr. Giovanni Boemio, , Mr. Sasan Salkhordeh, Mr. Mehdi Bostandoost, Mr. Pooya Naeeni, Mr. Amin Dehghanian, Mr. Ali Tabatabayi, Miss. Bahareh Naghibi, Mr. Vahid Javanbakht, Mr. Xuan Zhu, and Mr. Xiangli Ni for their help and support on this project.

Lastly, I would like to thank my parents and extended family for all of their love, support and encouragement. I could never have gotten here without them and I owe them the world.

1.0 INTRODUCTION

1.1 BACKGROUND

Sound source localization (SSL) based on microphone arrays has widespread applications in many fields, such as teleconference tracking of speakers [2], auditory perception, human-robot interaction [3], patrolling, fault detection in rotating machines [4], and air/ground objects recognition [5-6]. The latter, for instance, is pivotal in military applications to increase battlefield awareness and to locate snipers. In conferencing scenarios, SSL can be used to direct a pan-tilt-zoom camera toward the speaker or the audience such that the viewing experience is more interesting and/or network bandwidth is used more efficiently [7]. Despite years of research, the presence of noisy environment or reverberation maintains SSL challenging [8].

Two approaches can be pursued for SSL: active and passive. An active system sends out preset signals to the potential target and compares it with the echo signal, similar to how a bat locates its prey using ultrasonic pings. Sonar- and radar-like systems belong to this category. Passive localization systems only receive signals generated by the targets. Although they have the disadvantage to detect only activities that generate sounds (which may be intermittent), passive methods have the advantage of not radiating signals that can be used to indicate the presence and identity of the tracking station and they are less subjected to the reception of "false echoes" by virtue of multiple reflections.

Conventional SSL methods use microphones arrays made of physically separated microphones so that the variation in amplitude, delay time and reverberation characteristics at different locations can be exploited [9]. These methods can be loosely divided into two main categories: steerable beamforming and time delay estimation (TDE). In beamforming a full directional scan is undertaken to determine the source direction from the maximum signal power [7, 10-14]. This method is ineffective to determine the position of single events. In TDE the time offset between the received signals for a single sound source is calculated and it represents the difference between the travel time for the signal to each microphone position [15]. The delay is then related to the sound source-to-microphone propagation paths.

For both beamforming and TDE methods, omnidirectional microphones (OMs) are used. OMs are sound detectors whose output does not depend on the direction of the sound propagation, and the response is generally considered to be a perfect sphere; which means that identical signals of equal power are perceived to be of same magnitudes irrespective of the incident angle.

The main challenges for SSL systems are background noise, reverberation (in home environment), echoes (in open fields), intermittency, and source movement. In speech localization for instance accuracy is dependent on several factors like quantity, quality, and positioning of the applied microphones, number, distance, and spectral content of the active sources [8,16,17]. Several microphone array-based sound localization techniques have been proposed with varying degrees of accuracy and computational complexity [7,10-14]. These methods are based on the use of single or multiple OM arrays, linearly [18], triangularly [19] circularly [7,20-21], or squared [14] arranged.

In beamformer based locators, the source is localized by maximizing the output of a steerable beamformer [22-24]. The approach combines delay-and-sum beamforming with statistical analysis to trace the position of the acoustic source. Acoustic localization by means of beamforming is challenging due to its wide-band nature, near- and far-field geometry (relatively near/far distance of the source from the sensor array), and arbitrary array shape [10,15].

In TDE, sometimes indicated as time-difference of arrival (TDOA), the arrival time delay between signals detected by pairs of microphones is estimated to determine the direction of the propagating sound [14,25-26]. In order to measure the time delays, computations based on cross-correlation [17,27-28], eigenvalue decomposition [13], least mean square estimation [29-31], crosspower-spectrum phase [32], or the maximum likelihood [33-34] have been proposed.

Similar to TDE is the Interaural-Time-Delay (ITD) which is simply the time delay of incoming sound between "two ears" [35]. Although ITD is an excellent sound cue in stimulating a lateral perception on the horizontal plane, confusion is often raised when tracking the sound location from ITD alone because each sound source and its mirror image about the interaural axis share the same ITD. The ITD is often used to devise binaural microphone systems for telerobotic applications [36]. Some hybrid methods combine two of the three aforementioned categories [21-22].

Although directional microphones provide significantly superior sound quality [37] and can be used for adaptive noise cancellation [38], only very few researchers have proposed the use of directional microphones for SSL. Rui et al. [21] replaced previously used OMs with directional microphones pointing outwards for audio-visual conferencing. They introduced a hybrid algorithm between steered beam SSL and the one-step TDOA to overcome some of the challenges associated with the use of UMs for SSL: their phase response varies significantly with frequency, direction, and even from microphone to microphone. Gunël et al. [9] used directional microphones and spectro-temporal changes in a signal to find the direction but not the location.

Recently the results of field experiments conducted by NATO in France in the framework of the NATO Task Group SET-093/RTG53/MSE were presented [39-44]. Overall, the presented research shared the common approach of measuring the arrival time of the sound wave. Differences were in the position of the microphones (ground versus aerostats), single or multiple arrays, algorithms for the analysis, and methods to shield microphones from wind noise.

Researchers at the Photonics Center at Boston University developed a biomimetic acoustic detection and localization system that uses acoustic cues including spectral content, ITD, and interaural intensity difference (IID) [45-48]. The biomimetic aspect consists of a mammalian auditory model, which is an electronic implementation of the biological auditory system, with one left audio channel and one separate right audio channel. The channels work as the left and right human ears. The sound signals received by the microphones first pass through a filter bank, which mimics the inner ear filtering functionality done by the cochlea. The characteristic frequency components are extracted and processed through different auditory nerve channels. Both ITD and IID processors, respectively, extract positional information using arrival time and intensity difference cues detected at omnidirectional acoustic sensors. The Shot-Spotter commercial system [49] use acoustic sensors deployed over areas from one square mile up to hundreds of linear miles to locate gunfire and other violent threats. The system uses a network of 8-12 toaster-sized microphones per square mile fastened to rooftops and/or poles. From the available literature it is believed that this system employs OMs and TDOA estimation.

1.2 NOVEL CONTRIBUTION

In a recent paper Rizzo et al. [1] introduced the results of a passive SSL scheme based on the use of unidirectional microphones (UMs) and amplitude-based features. Sound direction was projected by exploiting the unidirectionality of certain features associated with the microphones output level. The incidence value estimated by each microphone was fed into a semi-analytical algorithm aimed at identifying the source location. The study was part of an ongoing effort to pave the road toward a new generation of sound detectors and localizers that can minimize the space between the microphone elements.

In this study, the work initiated and presented in [1] has been expanded. The main novelty of the overall study is the use of UMs and amplitude-based features that are not related to the arrival time of the sound. The proposed scheme is feasible to detect both continuous acoustic sources and single events. The

first achievement of this study is to consider the possibility of finding an analytical function for the properties of unidirectional microphones. Having obtained this function, behavior of applied microphone with respect to the sound source is detected. The method has been applied for localizing the sound sources in an indoor environment. The main contribution of this study is to evaluate the feasibility study of applying this method as an alternative to the common sound source localization methods. According to the obtained results, it may be concluded that this method is considered to be efficient and accuarte.

1.3 THESIS ORGANIZATION

This thesis is organized as follows. Chapter 2 describes the proposed analytical model and the general algorithm. Chapter 3 describes the hardware and software used in this study. Chapter 4 illustrates the results from an experiment conducted inside an anechoic chamber to characterize the microphones used in this study. Chapter 5 presents the results of the SSL algorithm to localize the position of single source indoor events. The conclusions will set some recommendation for future studies which are presented in chapter 6.

2.0 APPROACH

2.1 UNIDIRECTIONAL MICROPHONES

Unidirectional microphones are sound transducers that possess a strong dependence on the incidence angle of the acoustic source. They are used in all those applications or circumstances that require rejection of noise, filtering of unwanted acoustic signals, or elimination of reverberation phenomena. The variation of the detected sound wave amplitude (microphone output level) with the direction of the wave propagation (angle of sound incidence) is described by the polar directivity pattern (PDP). Generally, the PDPs detect a plurality of amplitude values of a propagating wave approaching at different angles. When the microphone's sensitivity (output level) is not function of the incidence angle the microphone is called omnidirectional. Mathematically a plane PDP of an unidirectional microphone may be expressed as [50]:

$$M(\vartheta) = K \times abs\left[\left(a + \cos\vartheta\right)\left(b + \cos\vartheta\right)\right] \qquad (1)$$

where M is the microphone output level, a and b are frequency dependent constants based on the microphone's design and the sound speed, ϑ is the sound's angle of incidence relative to the microphone axis. The constant K depends on the sensitivity of each sensor composing the array microphone, frequency and speed of the incoming sound as well as microphone design parameters. Particularly Eq. (1) describes a microphone of order two, in which the incidence angle ϑ is the argument of the second power of the cosine term. Fig. 1 shows an example of a PDP expressed by Eq. (1) in polar coordinates for a case where a = 0.69 and b = -0.29. As shown by Sessler and West [50], for low-frequency sounds these values provide the highest directivity to second order gradient unidirectional microphones. The PDP has

been normalized with respect to the maximum value expected along the microphone axis ϑ=0° by considering a constant equal to $\hat{K} = K / M(0°)$, instead of K.

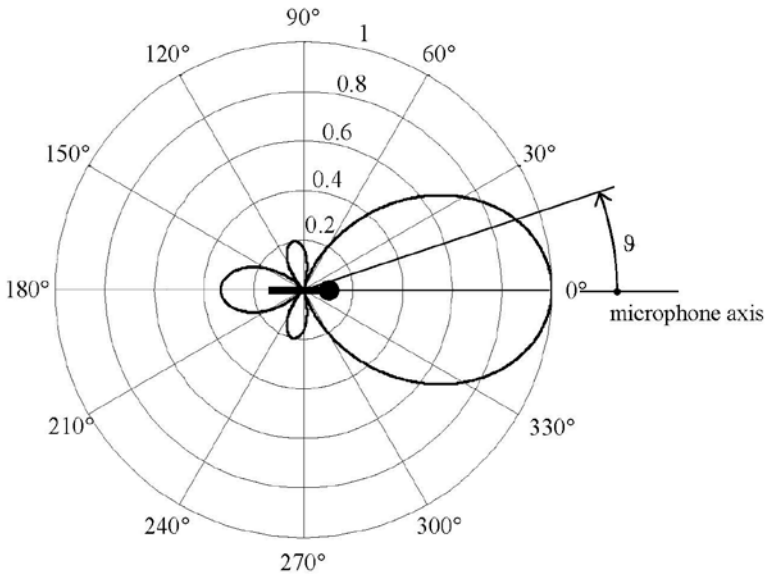


Fig. 1 - Polar directivity pattern of a unidirectional second order gradient microphone.

2.2 SOURCE LOCATION: SEMI-ANALYTICAL FORMULATION

To describe the basic hypothesis, let us consider an array of three unidirectional microphones A, B and C, represent in Cartesian coordinates as (x_i, y_i) (i = A, B, C), Fig. 2(a). The origin O of the x-y coordinate system corresponds to the center of mass of the triangle ABC. The angle δ_i, between the reference y-axis and i-th microphone axis, is introduced to univocally orient each microphone.

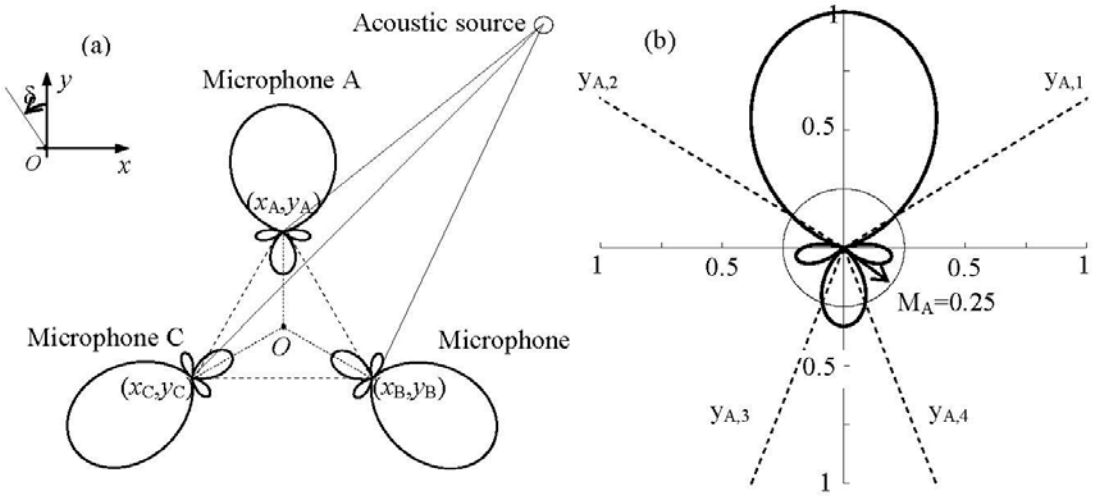


Fig. 2 – (a) Directivity Patterns of the microphones' setting and localization of the source. (b) Angles found and corresponding lines of possible direction.

For instance, in Fig. 2(a) $\delta_A = 0°$, $\delta_B = -120°$ and $\delta_C = 120°$. Let M_A, M_B, and M_C be, respectively, the normalized pulse features recorded by microphones A, B, and C as a result of an emission from the acoustic source. These values were normalized with respect to their maximum values expected at ϑ_i=0. Assuming that each microphone's PDP satisfies Eq. (1), it is obtained:

$$\hat{K} \times abs\left\{[a + \cos(\vartheta_A)][b + \cos(\vartheta_A)]\right\} = M_A \tag{2a}$$

$$\hat{K} \times abs\left\{[a + \cos(\vartheta_B)][b + \cos(\vartheta_B)]\right\} = M_B \tag{2b}$$

$$\hat{K} \times abs\left\{[a + \cos(\vartheta_C)][b + \cos(\vartheta_C)]\right\} = M_C \tag{2c}$$

which can be generalized as:

$$\hat{K} \times abs\left\{[a + \cos(\vartheta_i)][b + \cos(\vartheta_i)]\right\} = M_i \qquad (i = \text{A, B, C}) \tag{3}.$$

In polar coordinates, the value M_i can be geometrically described as a circle of radius M_i.

By solving, numerically, Eqs. (2a), (2b), and (2c) for unknown ϑ_A, ϑ_B, ϑ_C, the intersection between microphone lobes and the relative circle are determined for every microphone. These intersection points identify the possible directions of the incoming sound directions. However, these directions are not univocally determined. In fact, as the *i*-th microphone possesses four lobes, the normalized value M_i may identify up to eight incidence angles, i.e. eight intersections between the PDP and the circle.

For each angle, the algorithm calculates the equation of the semi-straight line having origin in the microphone coordinates (x_i, y_i) and orientation $\vartheta_{i,j} + \delta_i$. In general, for the *i*-th microphone (i = *A, B, C*) the equation will be:

$$y_{i,j} = \tan(\vartheta_{i,j} + \delta_i)\left[x - x_i\right] + y_i \qquad (4)$$

where j = 2, 4 or 8.

Fig. 2(b) presents the case where M_A = 0.25. In this scenario the circle intersects two lobes identifying four potential directions $\vartheta_{A,j}$ (j=1,2,3,4). Replacing these angles in Eq. (4) yields to the lines $y_{A,1}$, …, $y_{A,4}$, shown in Fig. 2 (b). Applying the same approach to microphones B and C, lines $y_{B,j}$ and $y_{C,j}$ are determined. By intersecting all pairs ($y_{k,j}$ and $y_{h,j}$ $k \neq h$, k,h = A, B, C) of direction lines, position of the sound source is univocally identified by the common coordinate resulted from these intersections.

3.0 HARDWARE AND SOFTWARE

3.1 LAYOUT OF SYSTEM

The experimental set up used in this study will be reviewed in this chapter. The hardware included a PXI unit, a rotational stage, three unidirectional microphones, preamplifiers, an omnidirectional microphone, A/C converter, and a commercial speaker. The layout of the hardware system is schematized in Fig. 3.

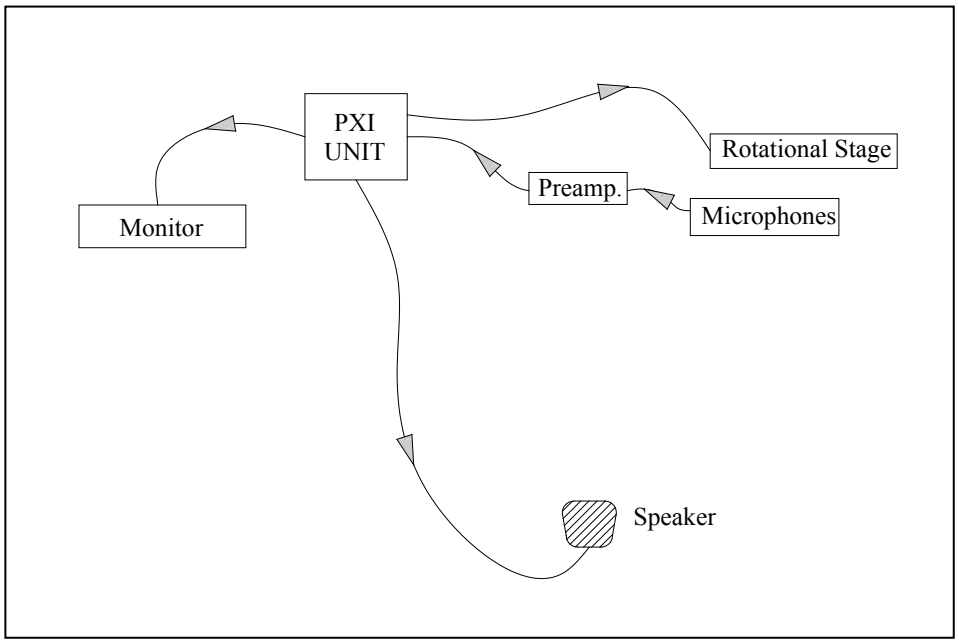


Fig. 3 – Schematic layout of the hardware system used in the experiments.

The PXI is the central processing unit which controls and analyzes the received and sent signals. Other mentioned instruments are connected to the PXI unit either directly or indirectly. PXI unit sends signals to the speaker in which the signals will be converted to a sound pulse. The propagated pulse will

be received by the unidirectional microphones and converted to electrical signals. The signals will be magnified as go through the preamplifier and, eventually, received by the PXI unit. The PXI unit records, analyzes and sends the characteristic of these received signals to an attached monitor, which illustrates the results during the experiment. The commercial Lab View software is applied for programming of the PXI unit. The characteristics of these hardwares along with a brief introduction to the Lab View software are discussed in details in the following sections.

3.2 PXI UNIT

The PXI is a computer which manages data communication among the attached hardwares. Some digitizers are attached to this unit, which are supposed to either send or receive signals, the electrical high frequency currents, upon their respective duties. These digitizers record the magnitude of the instant voltage of these signals. The PXI unit used in this study is presented in Fig. 4.

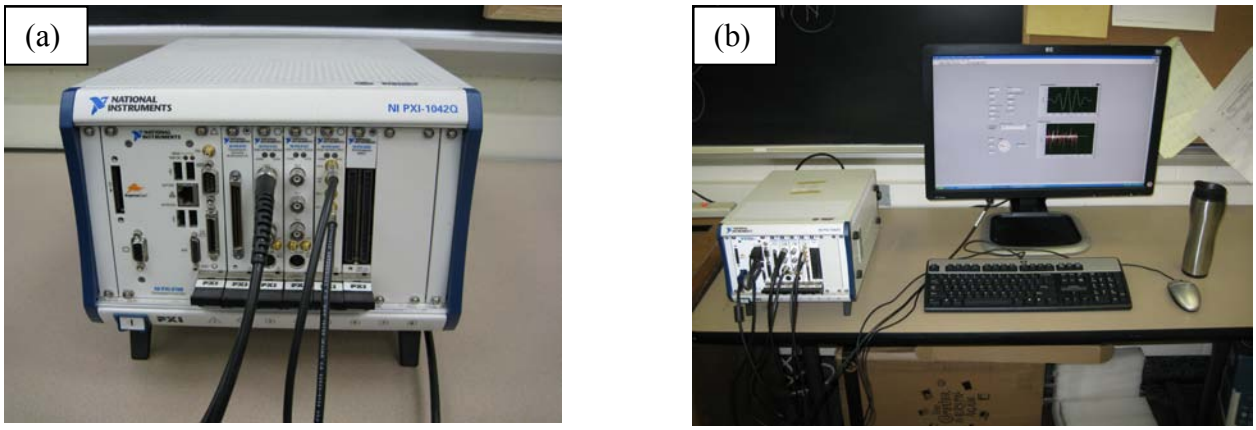


Fig. 4 – PXI unit used in this study.

3.3 LAB-VIEW SOFTWARE

The PXI run under LabView environment. In this research, few programs have been developed to select the type of output waveform, number of cycles, wave frequency, sampling frequency, and other

settings to generate acoustic wave. A typical program consists of *block diagram* and *front panel*. The Block diagram is a command console provided in a graphical mode. Front panel is designed to show the current status of the program during the computational operation. The designer can specify some properties as variable, which can be modified while the program is being used. Fig. 5 shows a schematic of front panel designated for this study.

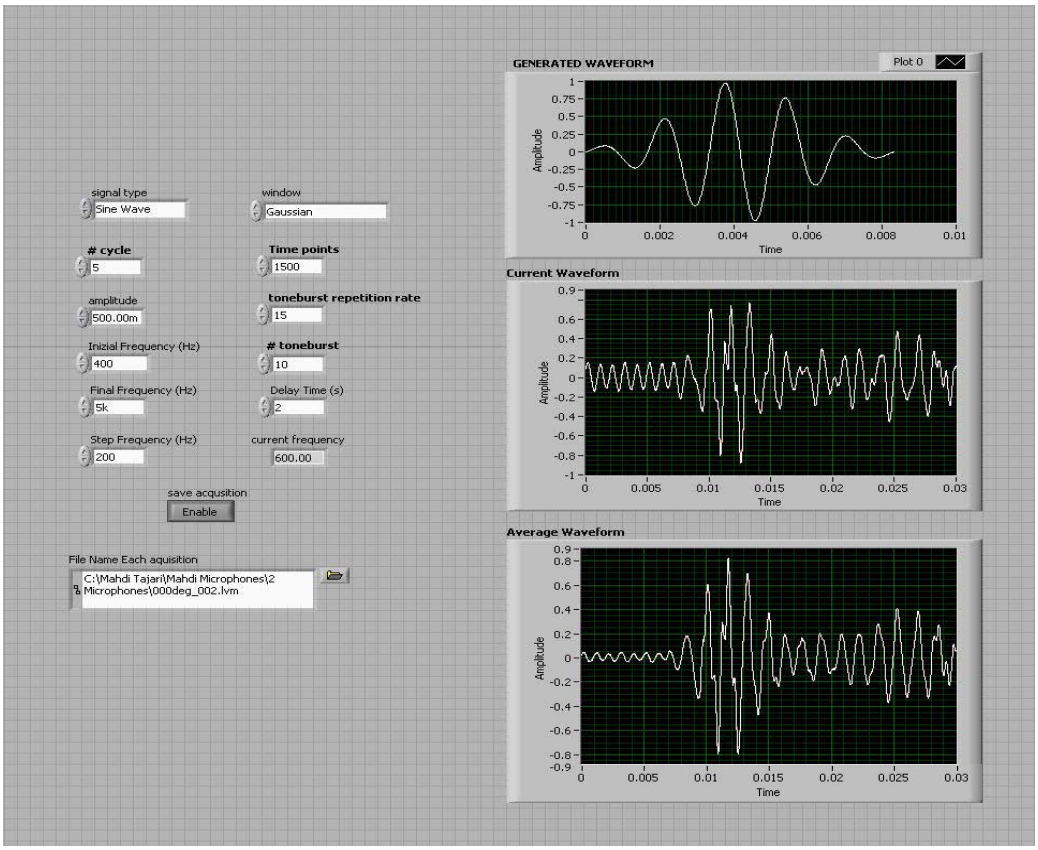


Fig. 5 - Front panel of the program to control PXI for generating and receiving signals for experiment.

As it is shown in this figure, the generated tonebursts are a five cycle Gaussian-affected sine wave tonebursts. This can be achieved by selecting the sine wave in the control box. This toneburst is shown in *generated waveform*, the upper plot in Fig. 5.

As it is depicted in Fig. 5, the amplitude of tonebursts may be specified in *amplitude* control box. The applied frequencies may be selected as a sequent on a pre defined step in the interval between *Initial*

frequency and *Final frequency*. The rate at which tonebursts are generated may be selected in *Toneburst repetition rate* control box. By specifying the properties of the generated five-cycle signal, this toneburst is sent to speaker from output port to generate a similar acoustic pulse.

The sound detected by the microphones is sent to the PXI's digitizer.

The time window of the digitized signal is controlled through the *Time points* control box and can be correlated to the time duration according to the sample frequency. For instance, in Fig. 5, having a sample frequency of 50 kHz and selecting 1500 Time Points, the duration may be calculated by dividing 1500 by 50k which results in 0.03 sec.

The sound data are shown in the *current waveform* plot. These data will are saved in a file (in *.lvm* format) including time in first column and corresponding amplitude in the second one. *Average waveform* plot is an average evaluated using several repeated tonebursts in each frequency, which is specified in *# toneburst* control box.

The block diagram of the user panel in Fig. 5 is shown in Fig. 6. This block diagram includes two loops, represented as two rectangle boxes in Fig. 6, for controlling the frequency range (peripheral box) and toneburst repetition (central box). The *NI-FGEN Express* and *NI-SCOPE Express*, which are shown in light blue boxes, control the digitizers corresponding with sending and receiving signals respectively.

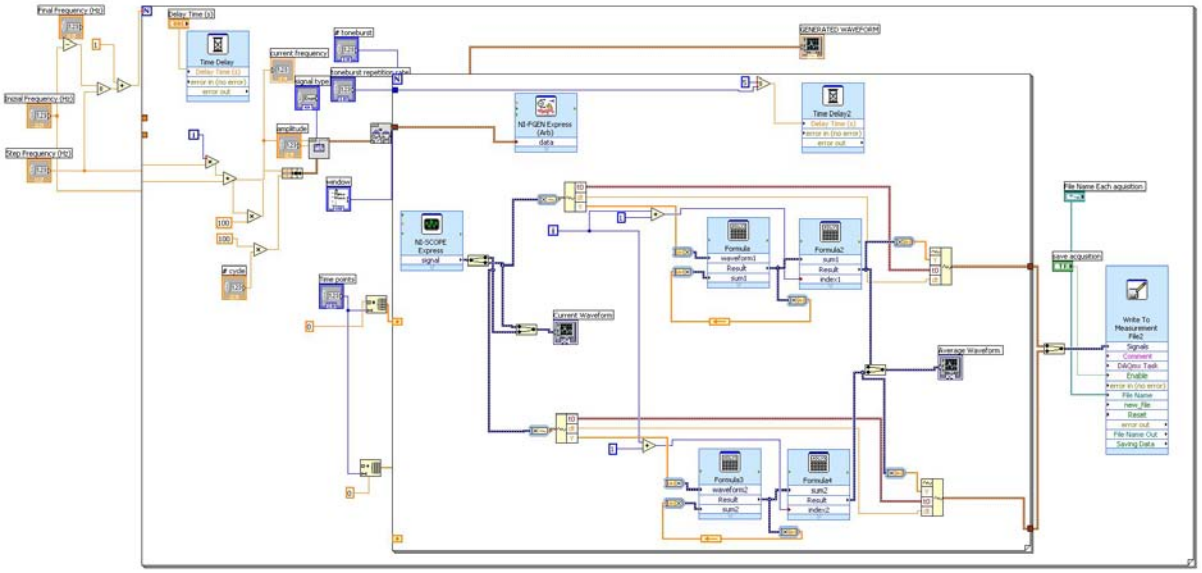


Fig. 6 – Block diagram of the program to control PXI for generating and receiving signals for experiment.

3.4 ROTATIONAL STAGE

The next step in the predefined set of experiments of this study is to evaluate sensitivity of microphones to the incident angle and frequency of sound wave. To collect data over a specified range of angles, the rotational stage shown in Figure 7 was used. This stage is controlled by LabView software.

Fig. 7 – Rotational stage controlled by LabView software.

A program is developed to rotate the rotational stage over 360 degrees as well. The front panel of this program is presented in Fig. 8. In each angle, a five cycle sine wave is generated over a predefined range of frequencies and received waves may be recorded for each angle at each frequency.

As it can be seen in the bottom plot of the front panel in Fig. 8, there are red and white lines. Red line associates with the current waveform and the white one associates with average of several received sound waves. In this program, it may be selected the microphone to be rotated from *Initial Angle* to *Final Angle* on desired *Step Angle*.

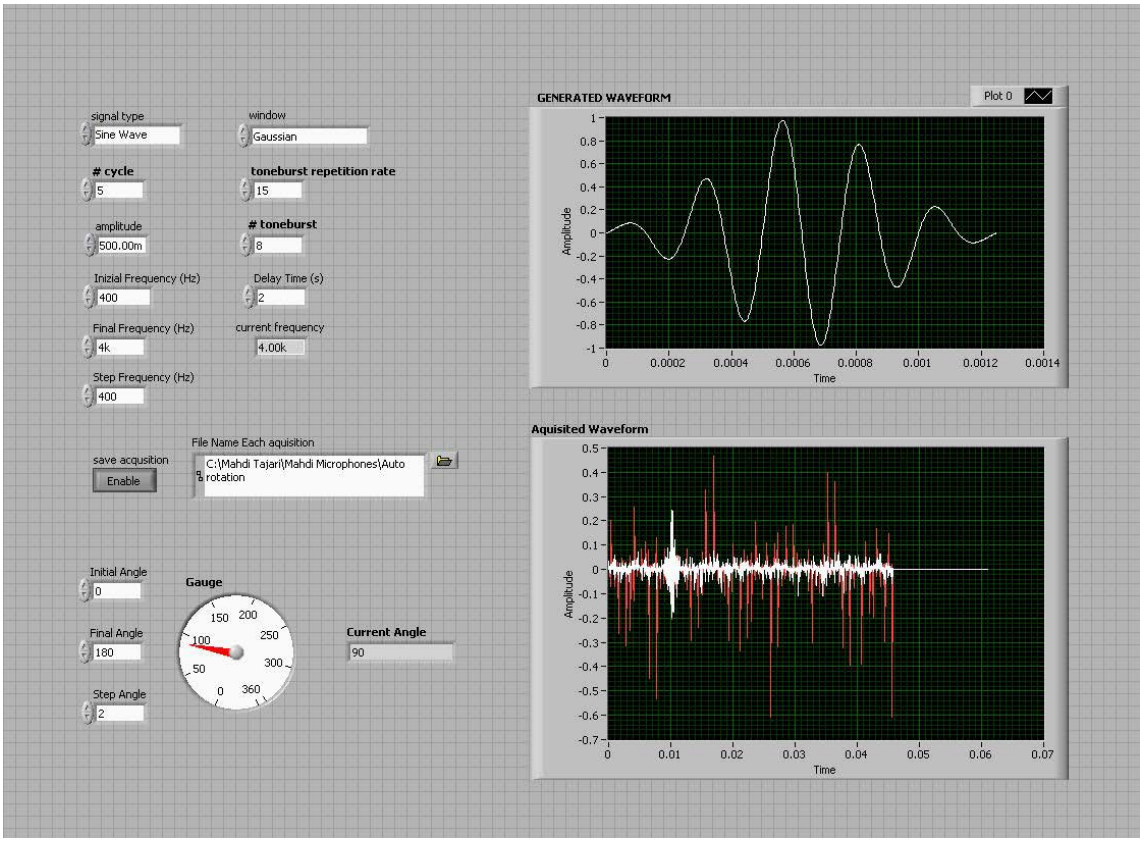
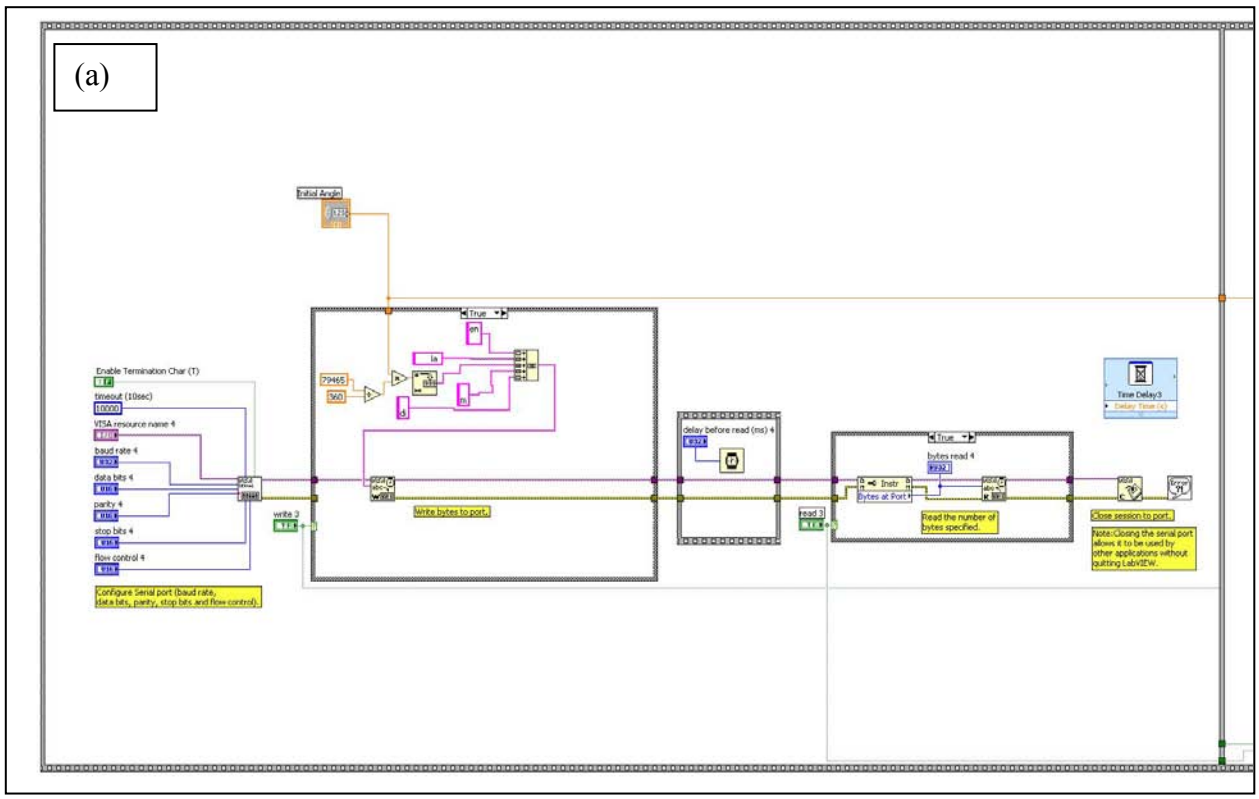
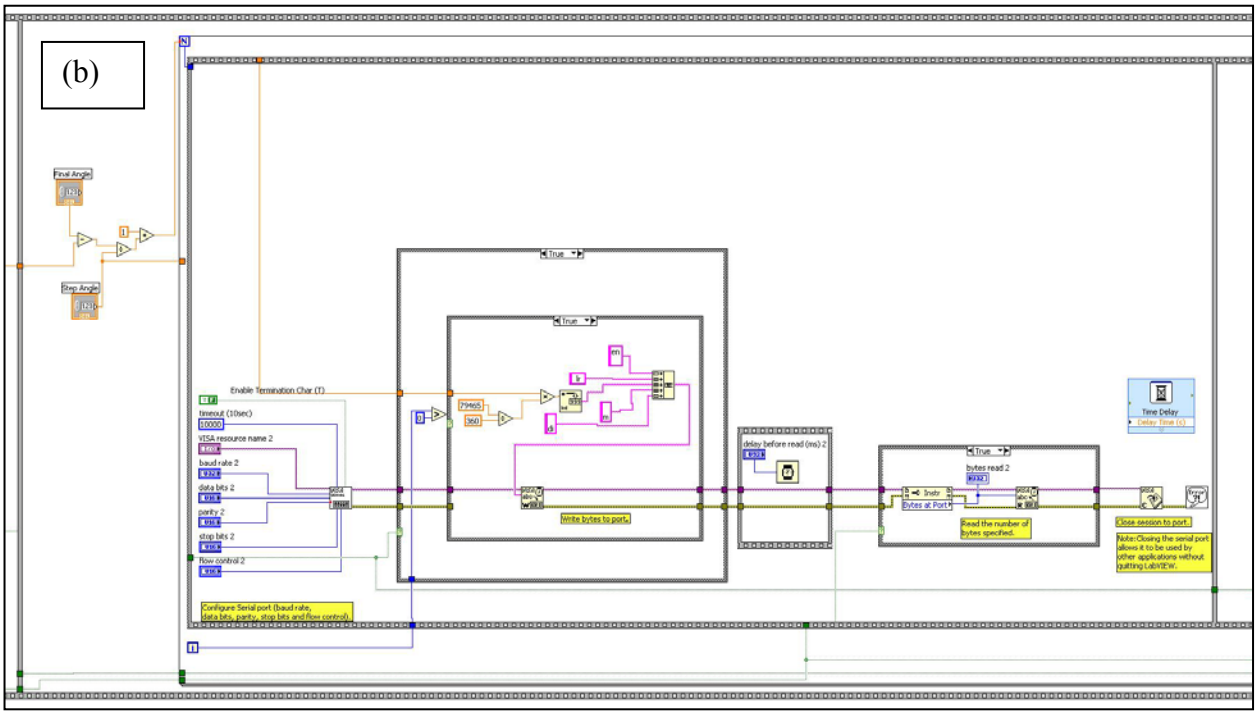


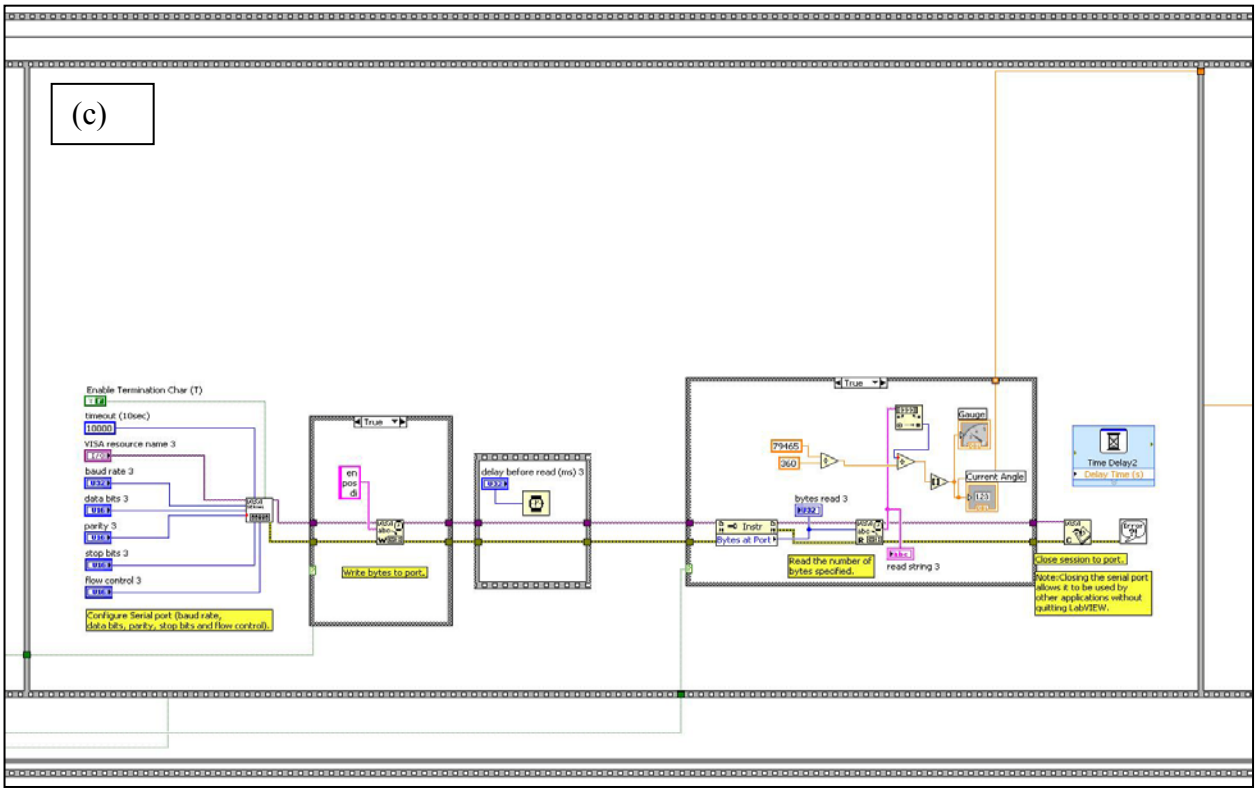
Fig. 8 - Front panel of program designed to rotate microphone and collect data over a range of angles for different frequencies.

The associated Block diagram with this program is shown in Fig. 9. In Fig. 9(a) the designated block diagram for setting the rotational stage at initial angle is illustrated. Fig. 9(b) represents the block diagram

which allows the rotational stage to rotate over the range of angles, step by step. The block diagram for showing the current position of the rotational stage is sketched in Fig. 9(c). Figure 9(d) associated with the block diagram designed to generate and receive signals over the desired range of frequencies.




(b)
Final Angle
Step Angle
Enable Termination Char (T)
timeout (10sec)
10000
VISA resource name 2
baud rate 2
data bits 2
parity 2
stop bits 2
flow control 2
Configure Serial port (baud rate, data bits, parity, stop bits and flow control).
True
79465
360
Write bytes to port.
delay before read (ms) 2
bytes read 2
Instr
Bytes at Port
Read the number of bytes specified.
Time Delay
Close session to port.
Note:Closing the serial port allows it to be used by other applications without quitting LabVIEW.


(c)
Enable Termination Char (T)
timeout (10sec)
10000
VISA resource name 3
baud rate 3
data bits 3
parity 3
stop bits 3
flow control 3
Configure Serial port (baud rate, data bits, parity, stop bits and flow control).
True
Write bytes to port.
delay before read (ms) 3
79465
360
Gauge
Current Angle
bytes read 3
Instr
Bytes at Port
Read the number of bytes specified.
read string 3
Time Delay2
Close session to port.
Note:Closing the serial port allows it to be used by other applications without quitting LabVIEW.

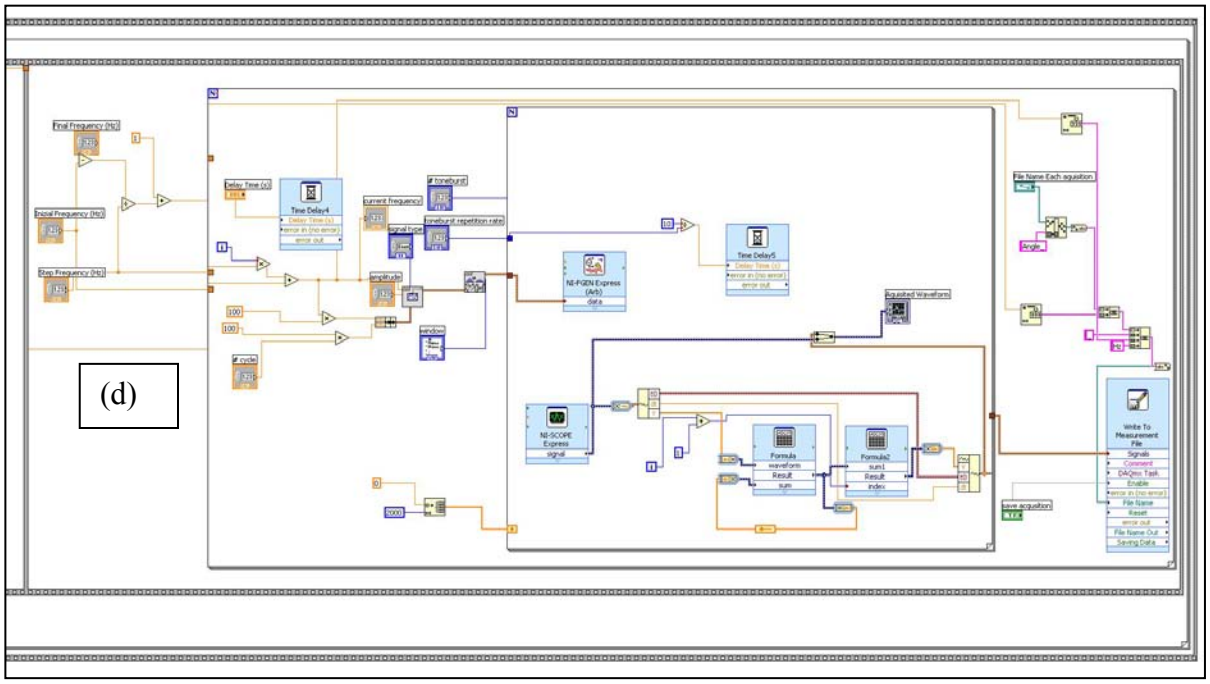


Fig. 9 - Block diagram of program designed to rotate microphone and collect data over the desired range of angles for different frequencies. (a) set the rotational stage at initial angle, (b) allow the rotational stage to rotate over the range of angles step by step , (c) show the current position of the rotational stage (d) generate and receive signals over the desired range of frequencies.

To study the effect of sound pulses with multiple frequencies and amplitudes a program is developed. The key point is to understand the dominant frequencies of generated and received sound. The front panel of this program is shown in Fig. 10.

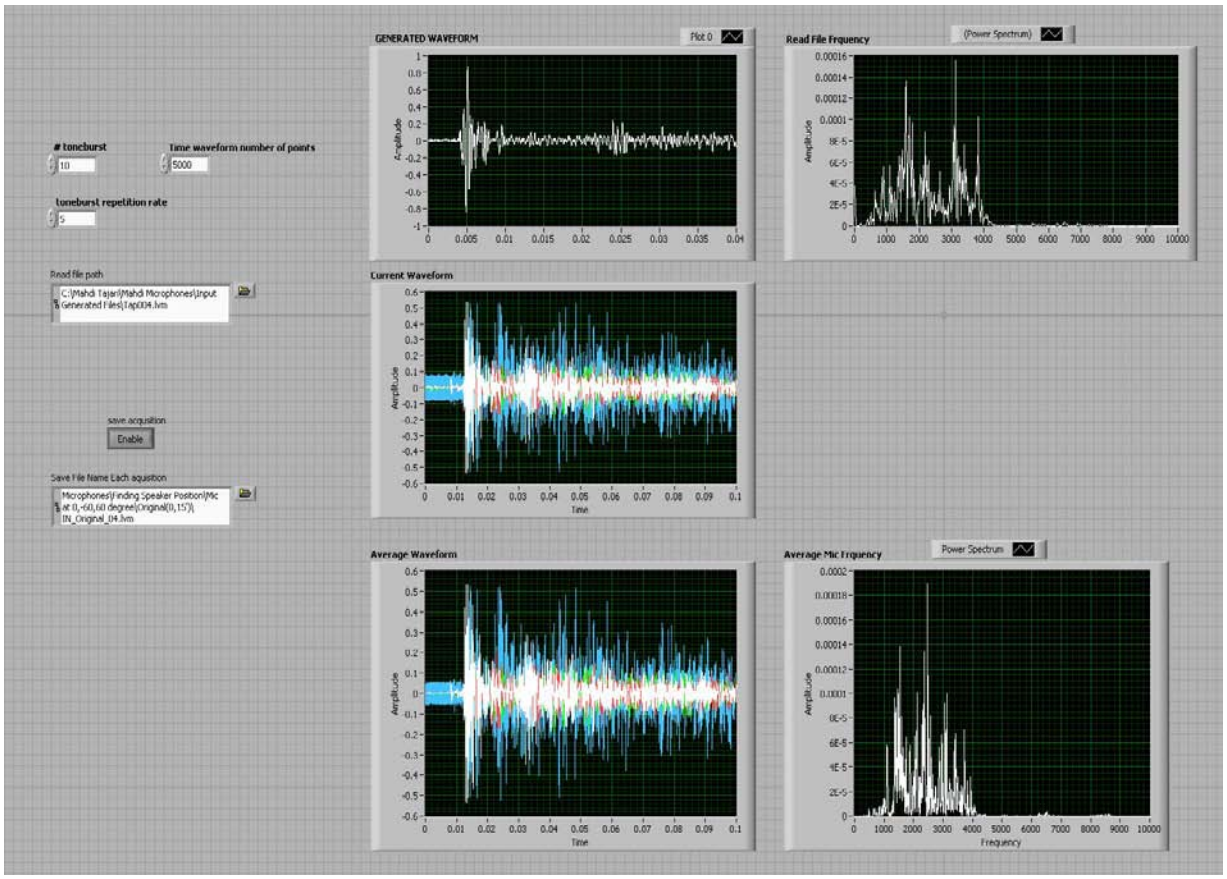


Fig. 10 - Front panel of the program designed to study the effect of multi frequency sound.

This program is designed to read a file associated with a recorded sound and regenerate this signal. The generated signal is sent to speaker to generate a sound pulse which propagates in the air and reaches to the microphone. Received sound from microphone is recorded in PXI. This program is designed to receive data from up to four microphones simultaneously.

Generated waveform in time domain and its corresponding frequency domain is visible in front panel of this program in top left and right plots respectively. The received waveform from each microphone is shown in the middle plot using different color lines (blue, red, green and white) for different

microphones. Average of several tonebursts in time domain and corresponding frequency domain are shown in bottom plots left and right respectively.

3.5 UNIDIRECTIONAL MICROPHONE

It is used three unidirectional microphones for this study as seen in Fig. 11 (a). The model of these microphones is AT8015 produced by Audio-Technica Inc. (Stow, Ohio). The range of frequency in which this microphone can receive data is from 40 Hz through 20 kHz, which covers the required frequency domain of this study. The sensitivity of these three microphones are similar to each other.

An A/D convertor, shown in Fig. 11 (b), provides power which also keeps output signal of this microphone unchanged during experiment. This convertor can support up to 4 microphones simultaneously.

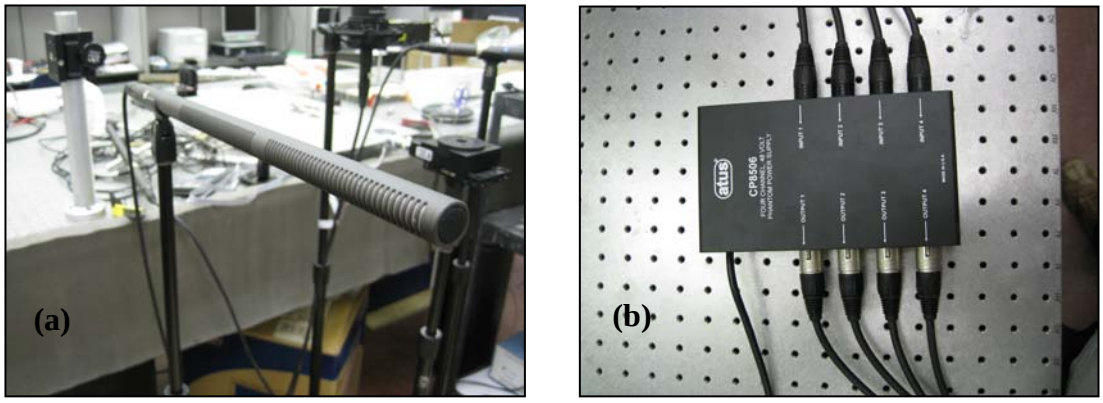


Fig. 11 – (a) unidirectional microphone and (b) converter.

3.6 PREAMPLIFIERS, OMNIDIRECTIONAL MICROPHONES, AND SPEAKERS

The output voltage of the microphones is in the same order of the digitizer error. In fact, the accuracy of digitizer is about 0.65 % of maximum voltage or 1 mV (whatever greater). Therefore, applying a preamplifier to magniify the amplitude voltage is inevitable. Fig. 12 (a) shows a preamplifier used for this

study. Preamplifier has an input and output port which receives and sends signals respectively, in which the output signal is just the amplified of the input one without changing the frequency.

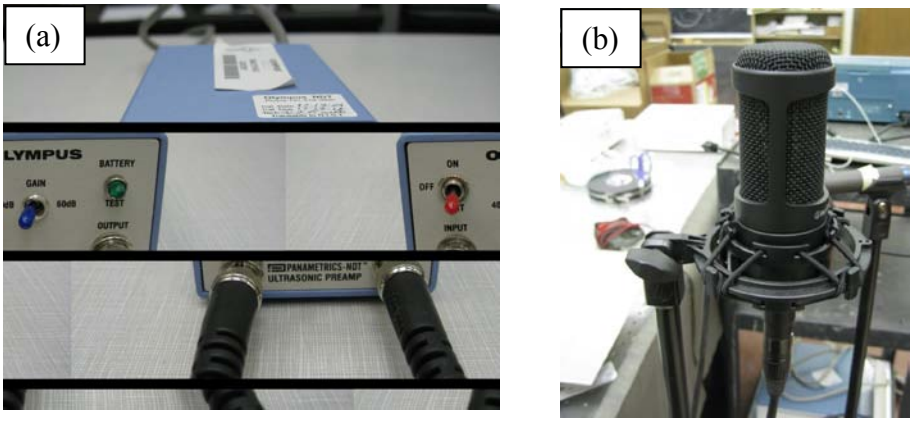


Fig. 12 – (a) Preamplifier and (b) omnidirectional microphone.

In this study, an omnidirectional microphone is implemented which is supposed to have the same sensitivity in all directions**.** Fig. 12 (b) shows a picture of this microphone. The properties of the omnidirectional microphone have been determined in this study. However, analyzing the experimental results related to this type of microphones is beyond the scope of this study.

A typical commercial speaker which covers the required frequency range is used for this study. The output signal of the PXI unit goes into the speaker to generate the corresponding sound pulse. The amplitude and frequency of this sound can be set by changing the signal.

4.0 ANECHOIC CHAMBER TEST

4.1 EXPERIMENTAL SETUP

In order to evaluate the analytical expression of the polar pattern, the microphones used in this study were tested in an anechoic chamber about 4.58 m x 4.28 m in dimension. As discussed in previous chapter, a LabView program was developed to perform signal generation, acquisition, and storage using three unidirectional microphones. This choice provides relatively short duration and small bandwidth. The program was also developed to control the rotational stage.

The experimental arrangement and the relative position of the microphone with respect to the speaker and the floor are shown in Fig. 13.

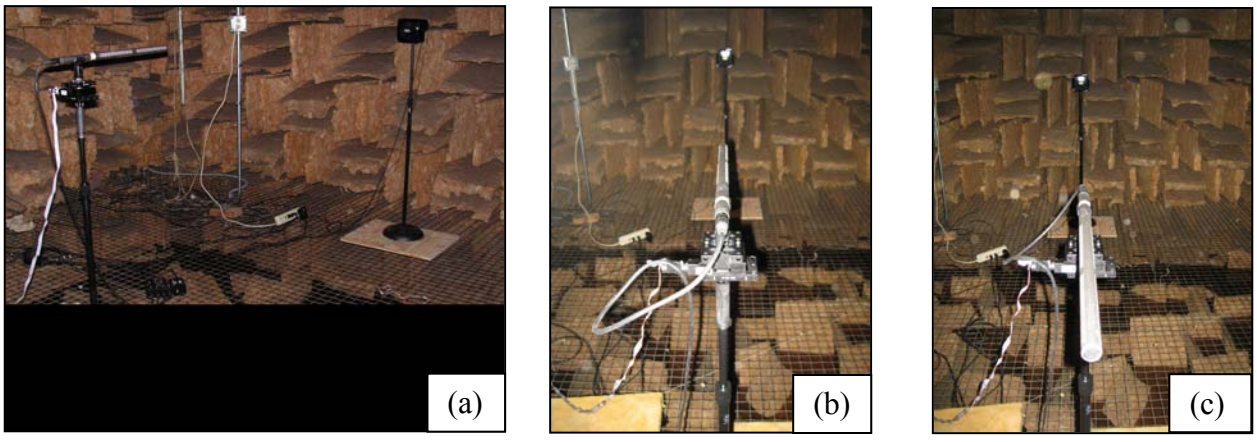


Fig. 13 – Photo of the experimental setup at the anechoic chamber. (a) microphone and speaker position (b) when microphone is in 0 angle to speaker and (c) when microphone is in 180 degree to the speaker.

For the accurate reconstruction of the analytical polar pattern each UM (that is referred to microphones A, B and C) and a omnidirectional microphone was rotated from 0 to 180 degrees at 1 degree step. The rotation was automatically controlled through the data acquisition system. Five-cycle Gaussian-windowed sinusoidal tonebursts from 400 Hz to 4 kHz at 200 Hz step were sent to the speaker. The sound detected by the microphone was sampled at 32.768 kHz.

4.2 EXPERIMENTAL RESULTS

Typical pulse waveforms acquired by microphone A when the excited toneburst was equal to 1200 Hz and the orientation angle was 0º, 45º, 90º, 135 and 180º are presented in Fig. 14. Clearly, as the sensor was rotated away from the direction of maximum sensitivity, the amplitude of the detected pulse diminished significantly.

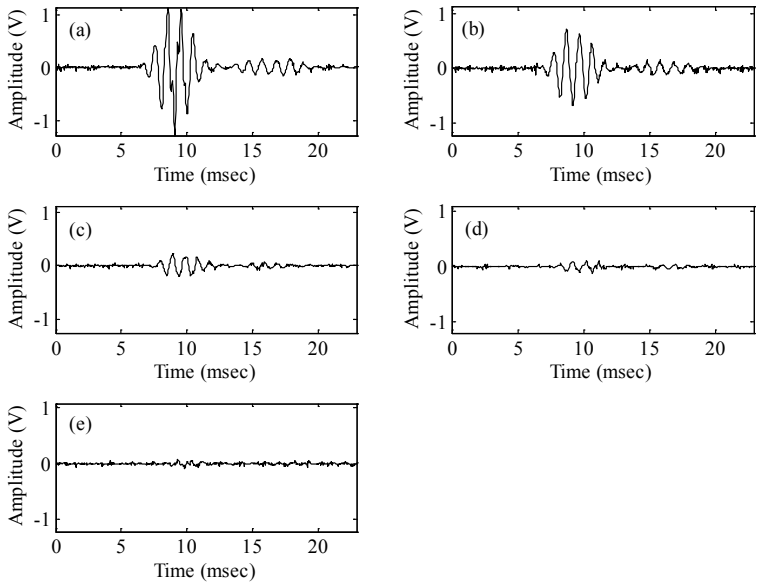


Fig. 14 - Anechoic chamber test. Time waveforms propagating at 1200 Hz and detected at various angles of incidence fro microphone A : (a) 0°, (b) 45°, (c) 90°, (d) 135°, (e) 180°.

The digitized waveforms were post-processed to extract certain features from the time and frequency domain. According to our study, totally nine features are considered in such an analysis including six features from time domain and three features from frequency domain. These features are maximum amplitude (max), peak to peak amplitude (PPK), root mean square (RMS), variance (VAR), peak multiplied to the RMS (KFact), Peak to RMS ratio (CrestFact) which are from time domain and maximum amplitude (max FFT), Area (Area), and root mean square (RM_FFT) from frequency domain.

In the present study two features from the time domain and one feature from the frequency domain have been selected as most promising ones. The approach serves to investigate which signal's features maximize the microphone's unidirectionality. In the time domain the features of peak to peak amplitude (PPK), and K-factor (KFact) were selected. As the frequency spectrum spans from zero to the Nyquist frequency, the interval for the computation of the frequency-domain based features include 1 kHz bandwidth centered at the toneburst frequency.

The selected features were calculated at each acquisition angle. This step resulted in the creation of the PDP relative to each feature. Figs. 16(a), 16(b), and 16(c), respectively, show the experimental PDP associated with the PPK, the KFact, and max amplitude of the Fourier transform relative to the detection of the 1200 Hz pulse for unidirectional microphone A.

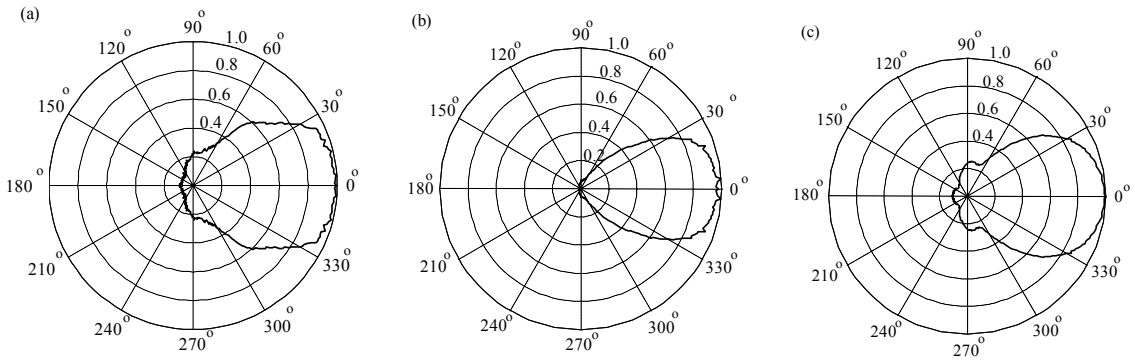


Fig. 15 - Anechoic chamber test. Polar directivity pattern at 1200 Hz for features: (a) peak-to-peak amplitude; (b) K-factor; (c) maximum amplitude of the Fast Fourier Transform for unidirectional microphone A.

Polar pattern at1200 Hz for microphones B, C and Omnidirectional one are presented in Figs. 17 to 19 respectively.

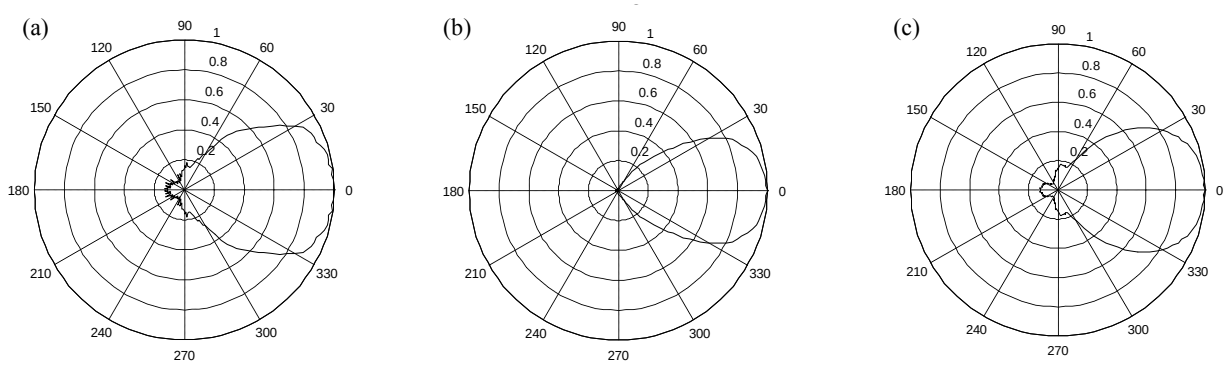


Fig. 16 - Anechoic chamber test. Polar directivity pattern at 1200 Hz for features: (a) peak-to-peak amplitude; (b) K-factor; (c) maximum amplitude of the Fast Fourier Transform for unidirectional microphone B.

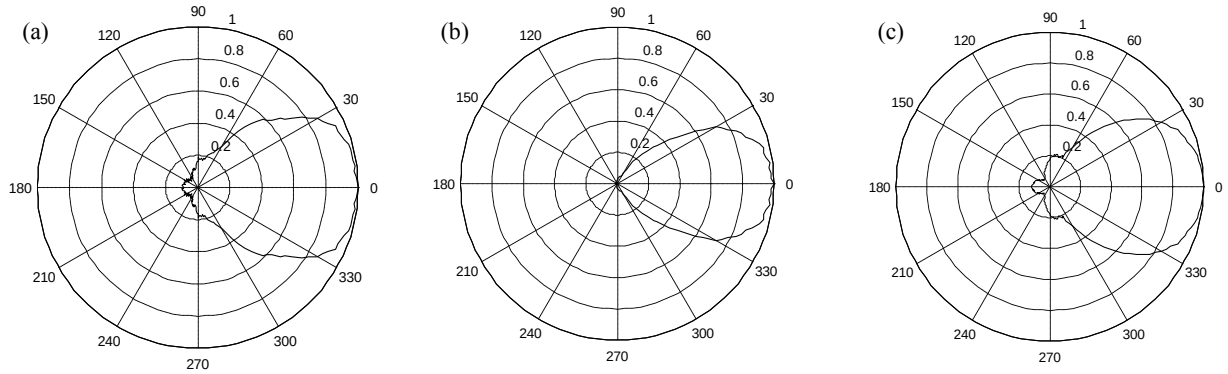


Fig. 17 - Anechoic chamber test. Polar directivity pattern at 1200 Hz for features: (a) peak-to-peak amplitude; (b) K-factor; (c) maximum amplitude of the Fast Fourier Transform for unidirectional microphone C.

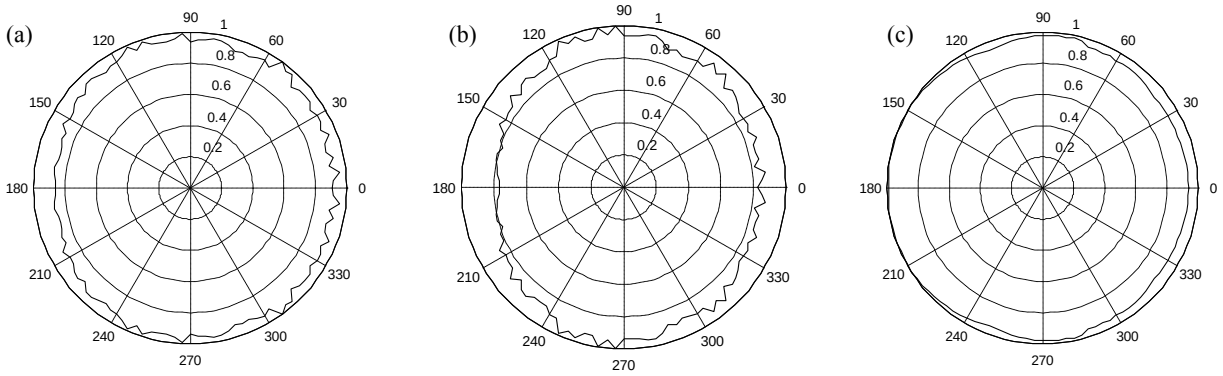


Fig. 18 - Anechoic chamber test. Polar directivity pattern at 1200 Hz for features: (a) peak-to-peak amplitude; (b) K-factor; (c) maximum amplitude of the Fast Fourier Transform for omnidirectional microphone.

Clearly the shape of the polar pattern is feature-dependent. This implies that the directionality of a microphone can be enhanced or mitigated by using certain features instead of others. The strong directionality shown by the KFact is related to its definition. As both the maximum amplitude and the root-mean-square of a signal are dependent upon the sensitivity of a microphone, their product augments the directional characteristics of the sensor.

Similar to Fig. 16, Fig. 19 shows the experimental PDP associated with the PPK amplitude of the detected pulses at four different frequencies for unidirectional microphone A.

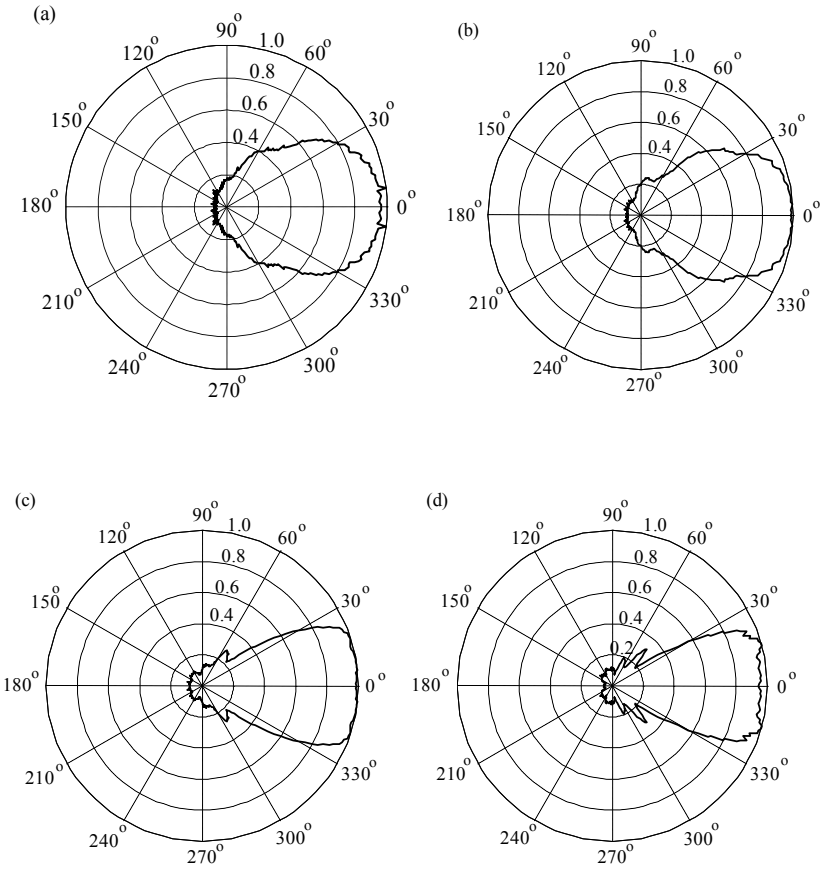


Fig. 19 - Anechoic chamber test. Polar directivity pattern associated with the feature of the peak-to-peak amplitude for microphone A at frequencies (a) 800 Hz; (b) 1800 Hz; (c) 2800 Hz; (d) 3800 Hz.

In Fig. 19(c), the polar pattern shows normalized polar pattern with amplitude equal to one for angles between -20 and 20 de grees. The reason is that for this case, the output signal from preamplifier have maximum voltage more than the maximum acceptable that can be detected by digitizer and in all the angles, the waveform is cut in such way that maximum amplitude is the maximum acceptable voltage from digitizer. In this case, the polar pattern which is normalized with respect to the maximum amplitude is equal to one.

Figures 21 to 23 shows polar pattern for PPK feature for 800 Hz and 4000 Hz for other microphones as comparison.

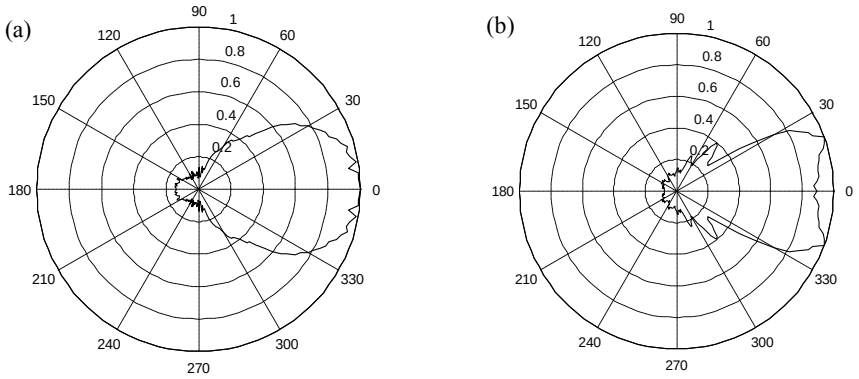


Fig. 20 - Anechoic chamber test. Polar directivity pattern associated with the feature of the peak-to-peak amplitude for microphone B at frequencies (a) 800 Hz; (b) 1800 Hz; (c) 2800 Hz; (d) 3800 Hz.

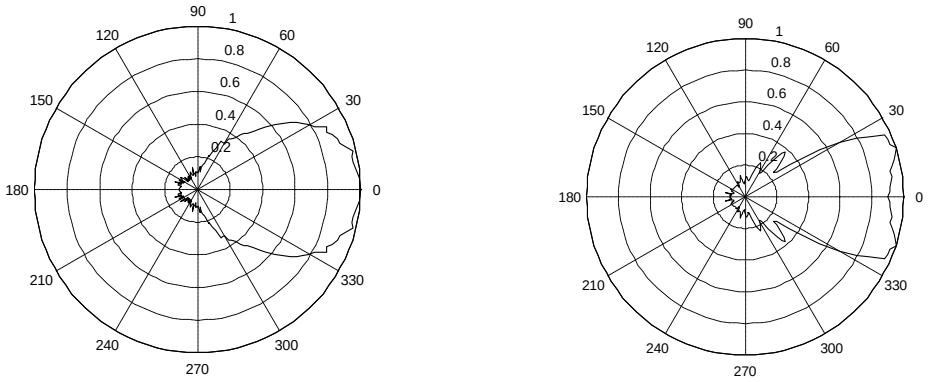


Fig. 21- Anechoic chamber test. Polar directivity pattern associated with the feature of the peak-to-peak amplitude for microphone C at frequencies (a) 800 Hz; (b) 1800 Hz; (c) 2800 Hz; (d) 3800 Hz.

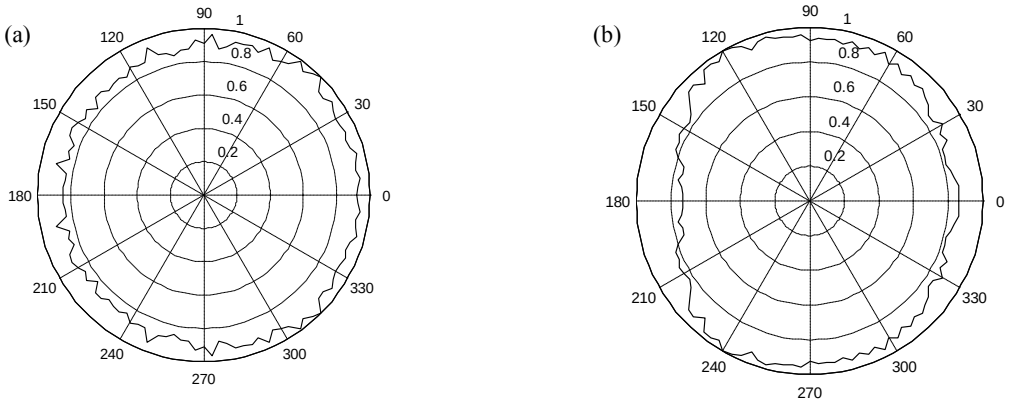


Fig. 22 - Anechoic chamber test. Polar directivity pattern associated with the feature of the peak-to-peak amplitude for omnidirectional microphone at frequencies (a) 800 Hz; (b) 1800 Hz; (c) 2800 Hz; (d) 3800 Hz.

As expected the shape of the PDP changes with the frequency and few main lobes appear as we approach toward higher frequencies. For the localization algorithm, the presence of many lobes may be detrimental as the intersection between the normalized amplitude outputs and the polar pattern may identify several pairs of sound directions. In order to satisfy the symmetry, the experimental data in the range 0º-180º were also used in lieu of the experimental data acquired from 180º-360º. In order to apply the semi-analytical approach described in chapter 2, the equation of the PDP is necessary. The values of

the normalized PPK and KFact at 1200 Hz as a function of the azimuth angle are presented in Figs. 24(a) and 24(b), respectively. By using MATLAB function fit and the Fourier polynomial type, equation that best fits the experimental data was sought and it is superimposed as continuous line for Fourier 7 and dotted line for Fourier 3 in Fig. 23.

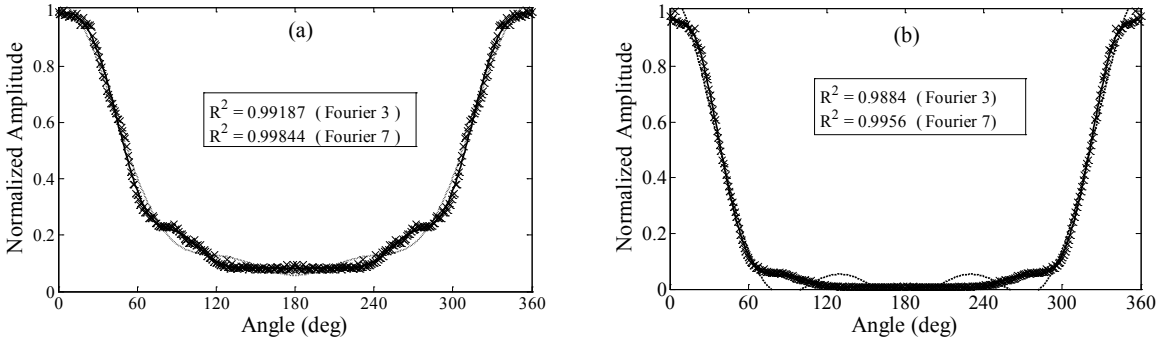


Fig. 23 - Experimental directivity pattern in Cartesian coordinates and curves fit Fourier 3 and Fourier 7 for microhone A: (a) 1200 Hz ppk amplitude; (b) 1200 Hz K-factor.

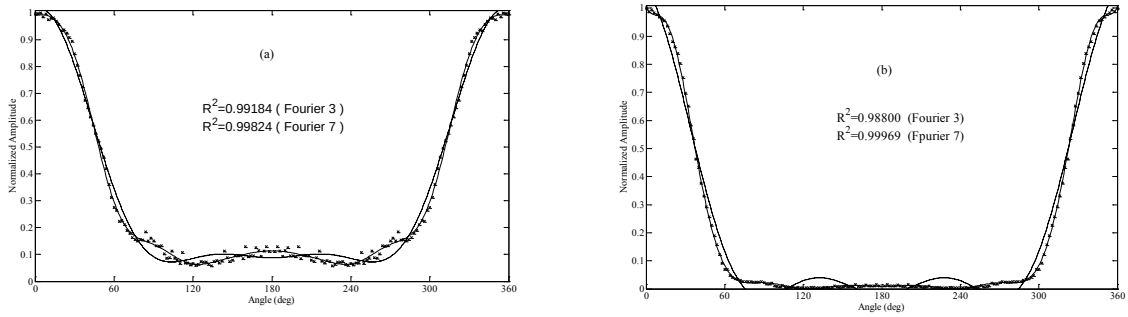


Fig. 24 - Experimental directivity pattern in Cartesian coordinates and curves fit Fourier 3 and Fourier 7 for microphone B: (a) 1200 Hz ppk amplitude; (b) 1200 Hz K-factor.

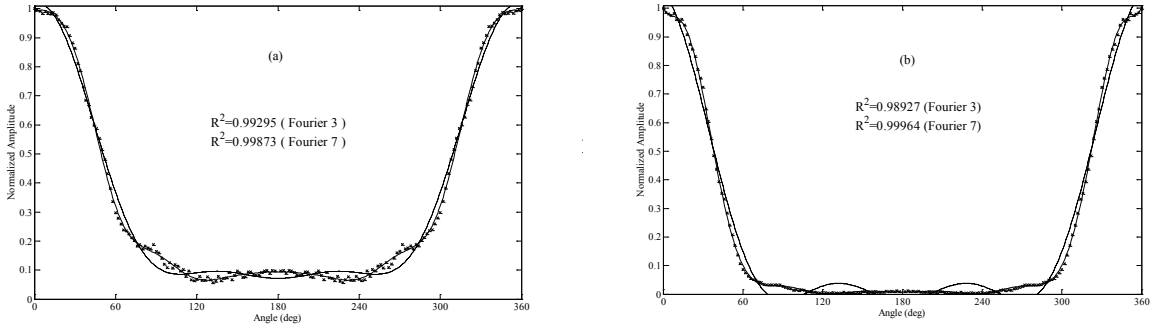


Fig. 25 - Experimental directivity pattern in Cartesian coordinates and curves fit Fourier 3 and Fourier 7 for microphone C: (a) 1200 Hz ppk amplitude; (b) 1200 Hz K-factor.

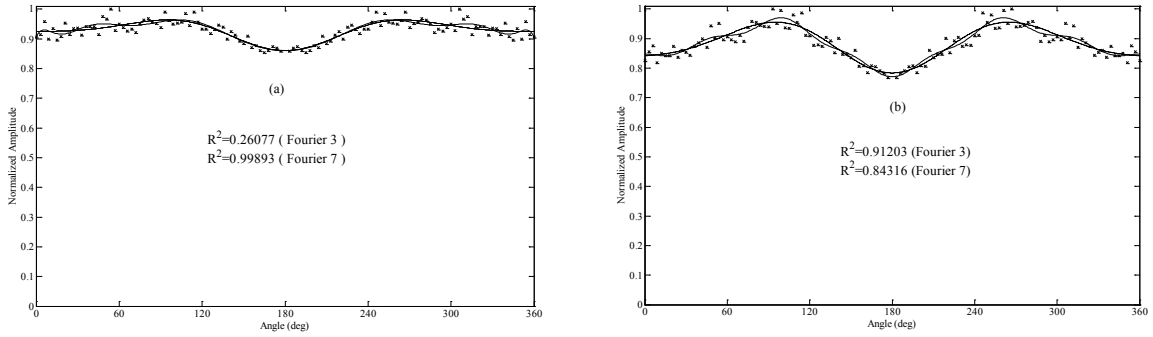


Fig. 26 - Experimental directivity pattern in Cartesian coordinates and curves fit Fourier 3 and Fourier 7 for omnidirectional microphone: (a) 1200 Hz ppk amplitude; (b) 1200 Hz K-factor.

Similar fitting function is evaluated for other microphones as shown in Figs. 24 to 26. By comparing the value of the residual R^2, the Fourier equation of order 3 and 7 were considered in this study. The analytical expression of the k-th order Fourier equation is :

$$M = f(\vartheta, a, b) = a_0 + \sum_{n=1}^{k} a_n \cos(n\vartheta) + \sum_{n=1}^{k} b_n \sin(n\vartheta) \qquad (6)$$

For the lines shown in the above figures, the empirical coefficients a_0, a_n, and b_n are summarized in Table 1. The selection of a higher-order Fourier function yields to more accurate results, i.e. value of R^2 closer to 1, at the expense of larger number of coefficients that needs to be considered.

To complete this analysis, fitting functions using Fourier 3 and 7 are evaluated for the other features. These fitting curves including their corresponding R^2 for each feature are shown in Fig. 27.

As seen in Fig. 27, fitting function is close to unity for all the features except crestfact. This fact illustrates that Fourier series are not a good approximation for creastfact and other equations more likely to equation of circle may be a good approximate for this feature.

Corresponding coefficients for fitting function shown in equation 7 for all other features, is presented in tables 1 and 2 for Fourier 3 and 7 at frequency of 1200 Hz respectively.

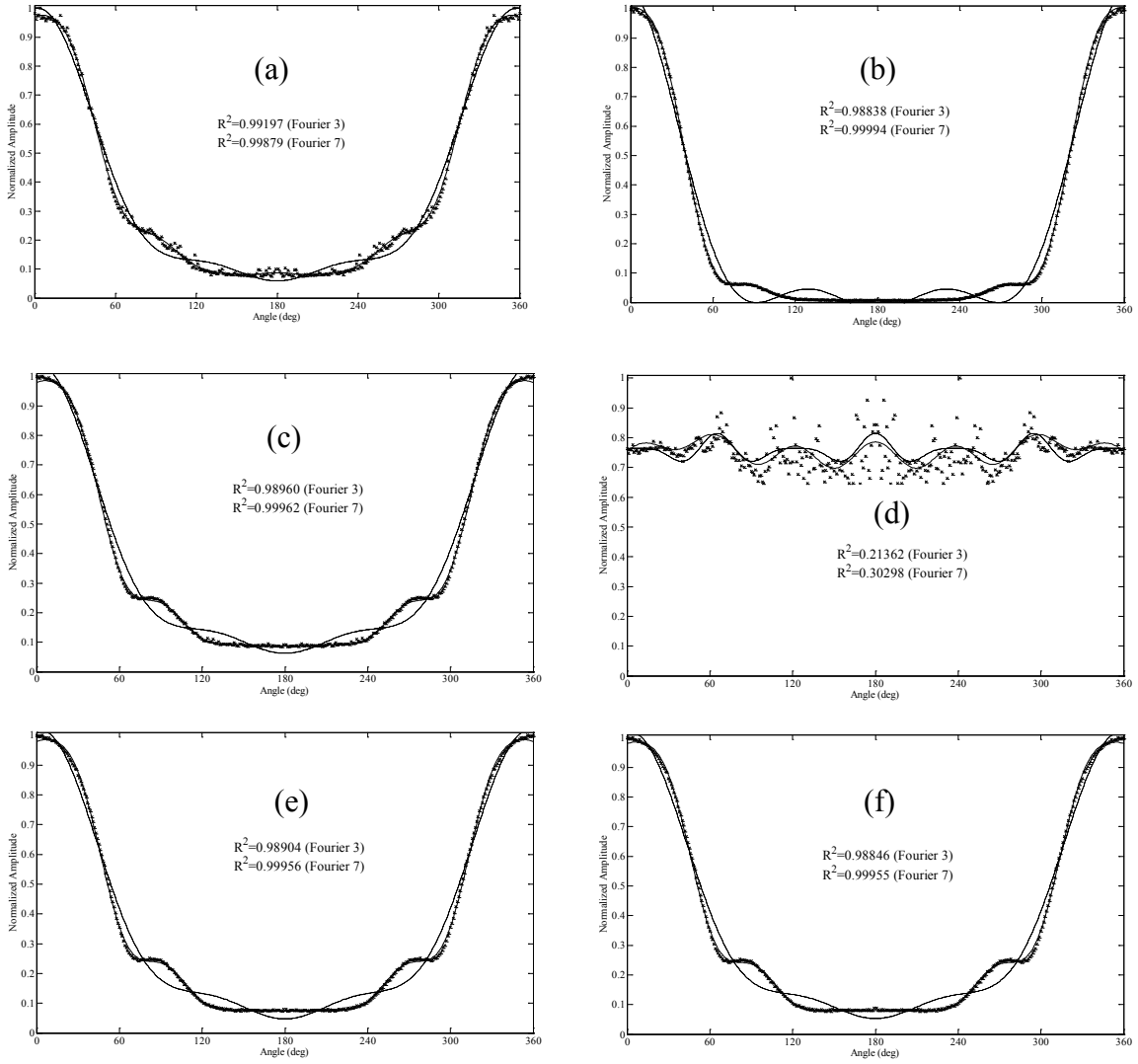


Fig. 27 - Fitting curves for Fourier 3 and 7 at frequency of 1200 Hz for microphone A (a) maximum amplitude of waveform, (b) variance, (c) Root Mean Square, (d) Crestfact, (e) Area of FFT and (f) RM_FFT.

Table 1- Experimental coefficients of the polar directivity patterns reconstructed through the Fourier 3 and Fourier 7 of the unidirectional microphones used in this study at frequencies 1200 Hz and 1600 Hz. for microphone A.

PPK		
n	a_n	b_n
0	0.350959433	
1	0.405327957	0.015254113
2	0.184921949	0.013938447
3	0.072266028	0.008189926

KFACT		
n	a_n	b_n
0	0.256601513	
1	0.417629551	-0.036657573
2	0.25155079	-0.04450278
3	0.111722216	-0.030038096

Maximum		
n	a_n	b_n
0	0.34834378	
1	0.399223361	0.013878591
2	0.181222181	0.012615253
3	0.070964151	0.007424926

Variance		
n	a_n	b_n
0	0.262947623	
1	0.426710689	-0.039012221
2	0.256359527	-0.047270693
3	0.114163595	-0.032028271

Table 1 (Continued)

RMS		
n	a_n	b_n
0	0.359507044	
1	0.405918819	0.010249309
2	0.183430044	0.009268999
3	0.07522899	0.005708224

CrestFact		
n	a_n	b_n
0	0.744927808	
1	-0.010608096	-0.004774996
2	0.021564534	0.02434655
3	-0.003905806	-0.012541224

Max_FFT		
n	a_n	b_n
0	0.364040842	
1	0.403704371	-0.011557852
2	0.186821872	-0.010706006
3	0.069847998	-0.006012283

Area		
n	a_n	b_n
0	0.3535315	
1	0.410765966	0.020669901
2	0.178340622	0.0179939
3	0.073206533	0.011126536

RM_FFT		
n	a_n	b_n
0	0.355612528	
1	0.40988278	0.016222578
2	0.180551835	0.014314393
3	0.072986069	0.008702426

Table 2- Coefficients associated with Fourier 7 for different features at frequency of 1200 Hz for microphone A.

PPK		
n	a_n	b_n
0	0.335110476	
1	0.382313665	0.045828754
2	0.178997026	0.043539139
3	0.076988592	0.028795045
4	0.031498699	0.016287085
5	-0.006252076	-0.004245359
6	-0.014339409	-0.012470905
7	-0.003518177	-0.003886663

KFACT		
n	a_n	b_n
0	10.3754496	
1	7.518045982	-17.57389537
2	-10.51106186	-11.00769982
3	-10.06773496	3.767543376
4	-0.308249232	6.674488332
5	3.091392285	1.49478886
6	0.965973283	-0.840704725
7	-0.078475239	-0.243361738

Maximum		
n	a_n	b_n
0	0.334515795	
1	0.379364911	0.040919755
2	0.176701074	0.038568036
3	0.075692152	0.025280742
4	0.031016975	0.014217291
5	-0.007530207	-0.004485646
6	-0.014564195	-0.010950237
7	-0.006166935	-0.005769783

Table 2 (Continued)

Variance		
n	a_n	b_n
0	0.257410797	
1	0.418065208	-0.031702175
2	0.253440597	-0.038659456
3	0.114565448	-0.026469397
4	0.023125451	-0.007223112
5	-0.026168488	0.010404394
6	-0.031638477	0.015444021
7	-0.013443195	0.007873001

RMS		
n	a_n	b_n
0	0.342097954	
1	0.380990707	0.044165951
2	0.177431086	0.041697372
3	0.080856356	0.029169581
4	0.036337551	0.018077465
5	-0.006450866	-0.004199204
6	-0.020392521	-0.016914951
7	-0.006961937	-0.00728196

CrestFact		
n	a_n	b_n
0	0.742245839	
1	0.013196732	-0.003954714
2	-0.009453715	0.006225102
3	-0.013256502	0.015824426
4	-0.002135014	0.004964214
5	-8.31E-05	0.000719359
6	-0.001580742	-0.00888091
7	0.014443604	0.02862493

Table 2 (Continued)

Max_FFT		
n	a_n	b_n
0	0.341745028	
1	0.373202695	0.031152539
2	0.181891606	0.030579334
3	0.079703663	0.020338215
4	0.041289568	0.014286877
5	-0.010682755	-0.0047246
6	-0.025140485	-0.013723952
7	-0.006160918	-0.004062581

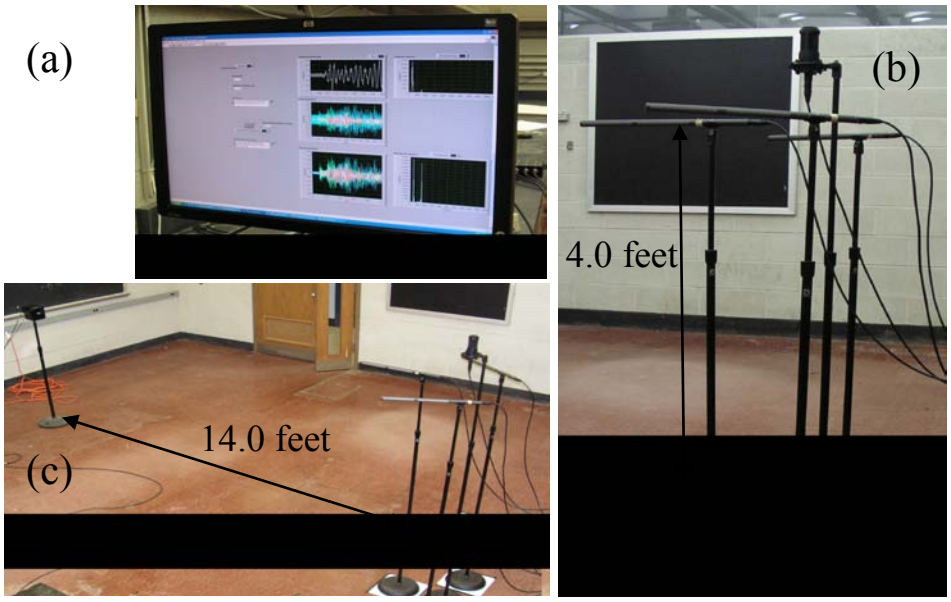
Area		
n	a_n	b_n
0	0.339496313	
1	0.390574402	0.049120465
2	0.173608257	0.044369353
3	0.077579108	0.030566189
4	0.034764073	0.019011175
5	-0.010528623	-0.007604873
6	-0.020940631	-0.019533542
7	-0.005210434	-0.006248668

RM_FFT		
n	a_n	b_n
0	0.339670414	
1	0.386997558	0.048083208
2	0.175157894	0.044208067
3	0.078015936	0.030334918
4	0.036518041	0.01968768
5	-0.009563803	-0.006799808
6	-0.021380836	-0.019588581
7	-0.005760544	-0.006763268

5.0 INDOOR SOUND SOURCE LOCALIZATION

5.1 EXPERIMENTAL SETUP

In order to assess the effectiveness of the proposed SSL method and to investigate the effect of feature selection, an indoor experiment was conducted. An array of three UMs was deployed inside an empty laboratory about 7m long and 6.4 m wide. A speaker was positioned few meters away from the microphones' array. Photos of the speaker and the array and their relative distances are presented in Fig. 28. The photos show the presence of an OM.



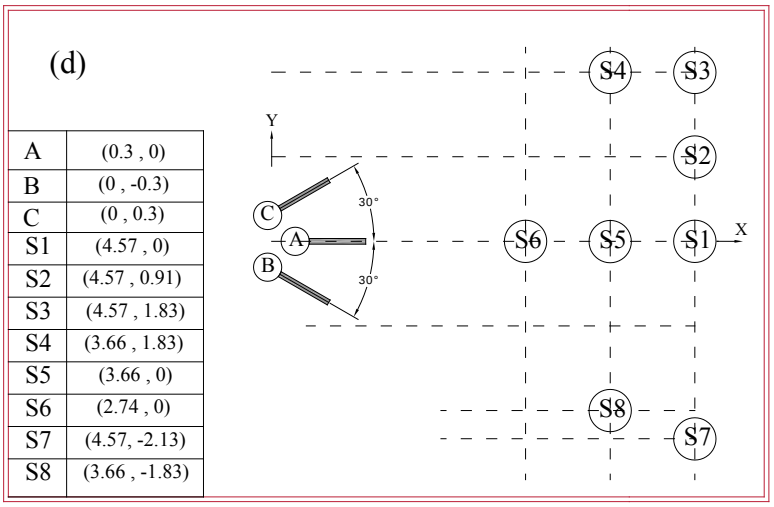


Fig. 28 - Indoor test. (a) Front panel of the software used for the experiment. (b) Photo of the microphone array. (c) Photo of the room and the array-speaker distance (the speaker-array distance is about 18-fold larger than the inter-microphones distance). (d) Schematic of microphone position and orientation and location of the speaker during the experiment.

The hardware/software system utilized to conduct this test was identical to the one used for the anechoic chamber tests. A 5-cycles Gaussian-windowed sinusoidal toneburst ranging from 400 Hz to 4 kHz at 200 Hz step was sent to the speaker.

The tests were conducted such that the position of the speaker changed. Fig. 28(c) shows the positions of the speaker with respect to the array location. Following the angle notation introduced in Fig. 2, the arrangement is 30º /0º/-30º, where the first angle is relative to microphone C, the second angle to microphone A, and the last angle to microphone B. The experimental PDP reconstructed using Fourier 3 was considered first. Moreover, the equation and the coefficients summarized in Table 1 were considered representative of the patterns of microphones B and C as well.

5.2 EXPERIMENTAL RESULTS

In the semi-analytical approach developed to localize the sound source position, Eq. (1) is now replaced by the equation of the fitting curve extracted from the experimental patterns. The formulation

described in section 2.2 was used here to identify the location of the speaker. The approach was tested using the normalized features' values of the PPK and the KFact of the time history, and the maximum amplitude value of the Fourier transform (max FFT). The results from recording pulses at 1200 Hz and 1600 Hz are presented.

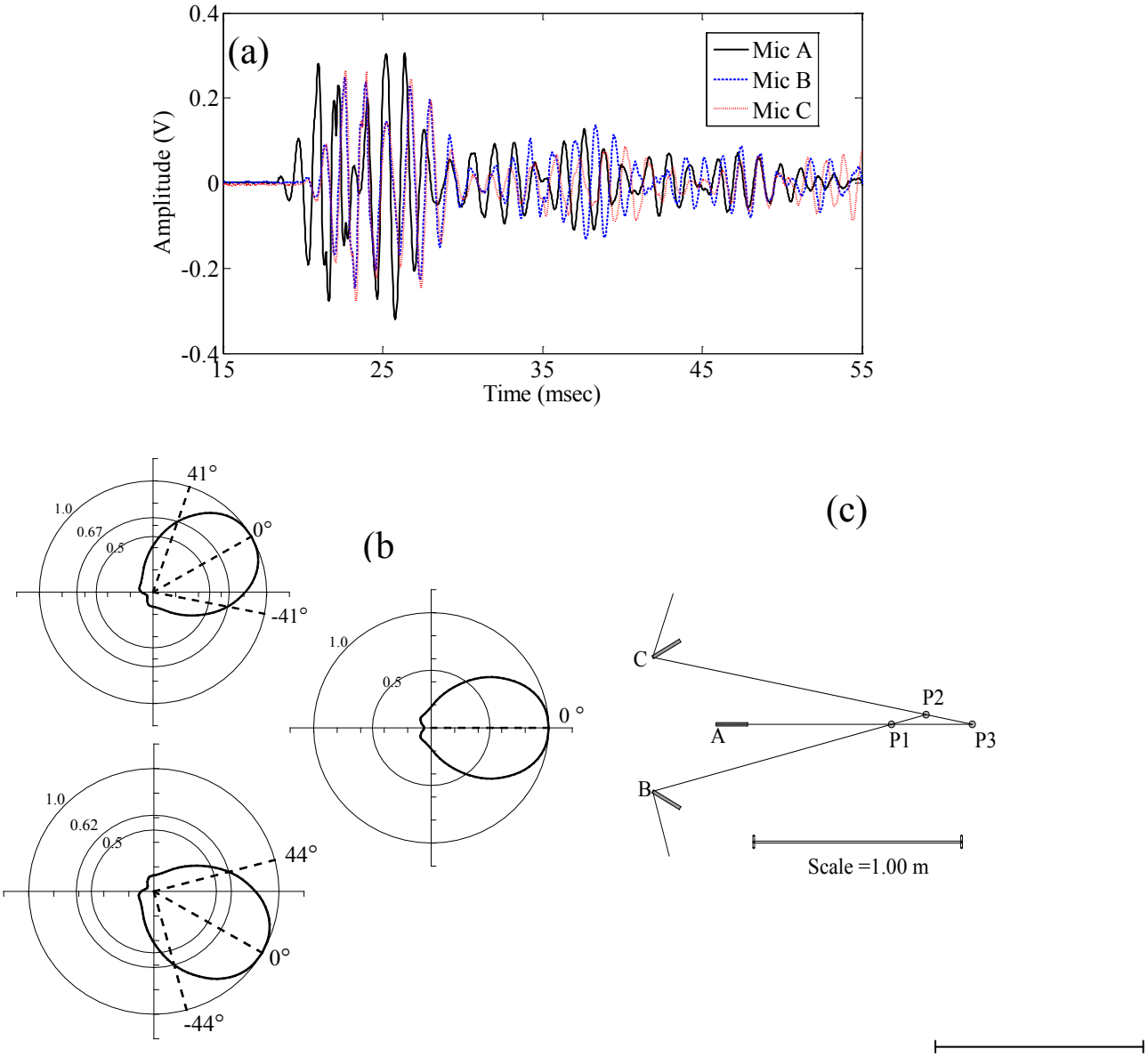


Fig. 29 - Indoor test. Speaker at location S1. Source localization by using the feature of the peak-to-peak amplitude of the 1200 Hz toneburst. (a) Time waveforms detected by the array. (b) Experimental polar directivity patterns given by Fourier 3. (c) Localization of the sound source.

Figure 29(a) shows the time waveforms captured by the array when the 1200 Hz toneburst was generated and speaker was located in position S1. For each detected signal, the main burst is followed by smaller pulses originated by the reflection with the room boundaries. The PPK amplitudes of the main bursts were equal to 0.625 V, 0.494 V, and 0.542 V for sensors A, B, and C, respectively. The position of microphone and sensor C was symmetric with respect to the speaker. This is confirmed by the simultaneous arrival of the pulse. However the PPK amplitudes recorded by microphones were about 9% different. This might introduce some level of inaccuracy in the determination of the incoming sound direction. To mitigate the effects associated with different sensitivities that may arise from manufacturing differences a calibration factor computed empirically was applied to the successive analyses.

Since a normalized value is necessary to use the experimental polar pattern, these amplitudes values were normalized with respect to the amplitude recorded by microphone A (0.625 Volts). Note that a discussion about the normalization process in a general scenario will be provided in the conclusive remarks. The normalized values of the microphone output levels associated with the PPK from detectors A, B, and C were then equal to 1.00, 0.670, and 0.618, respectively. Such values are represented by the circles superimposed in the corresponding PDP of Fig. 29(b). The potential incidence angles ($\vartheta_{i,j} + \delta_i$) identified by each microphone were + 71º and -11° for microphone C, 0º for microphone A, and + 14º and – 74° for microphone B.

The algorithm then traced the lines departing from the microphone coordinate along the direction of these potential incidence angles. The lines, shown in Fig. 29(c), intersect into three points P1, P2 and P3 associated with the intersections of microphone pairs A-B, B-C, and A-C, respectively. These points are solutions found by the algorithm and their planar coordinates are (x = 1.18 m, y = 0.0 m) for microphone pair A-B, (x = 1.37 m, y = 0.05 m) for microphone pair A-C, and (x = 1.62 m, y = 0.0 m) for microphone pair B-C. Ideally, such points should be identical and collapse into a single point coincident with the speaker position. This point shall represent the location of the sound source. The locations P1, P2 and P3 underestimate the true position of the speaker (x = 4.57 m, y = 0.0 m).

Waveforms asso ciated with other pos itions o f spe aker ar e show n in Fig. 30. A ll w aveforms ar e obtained simultaneously from different microphones.

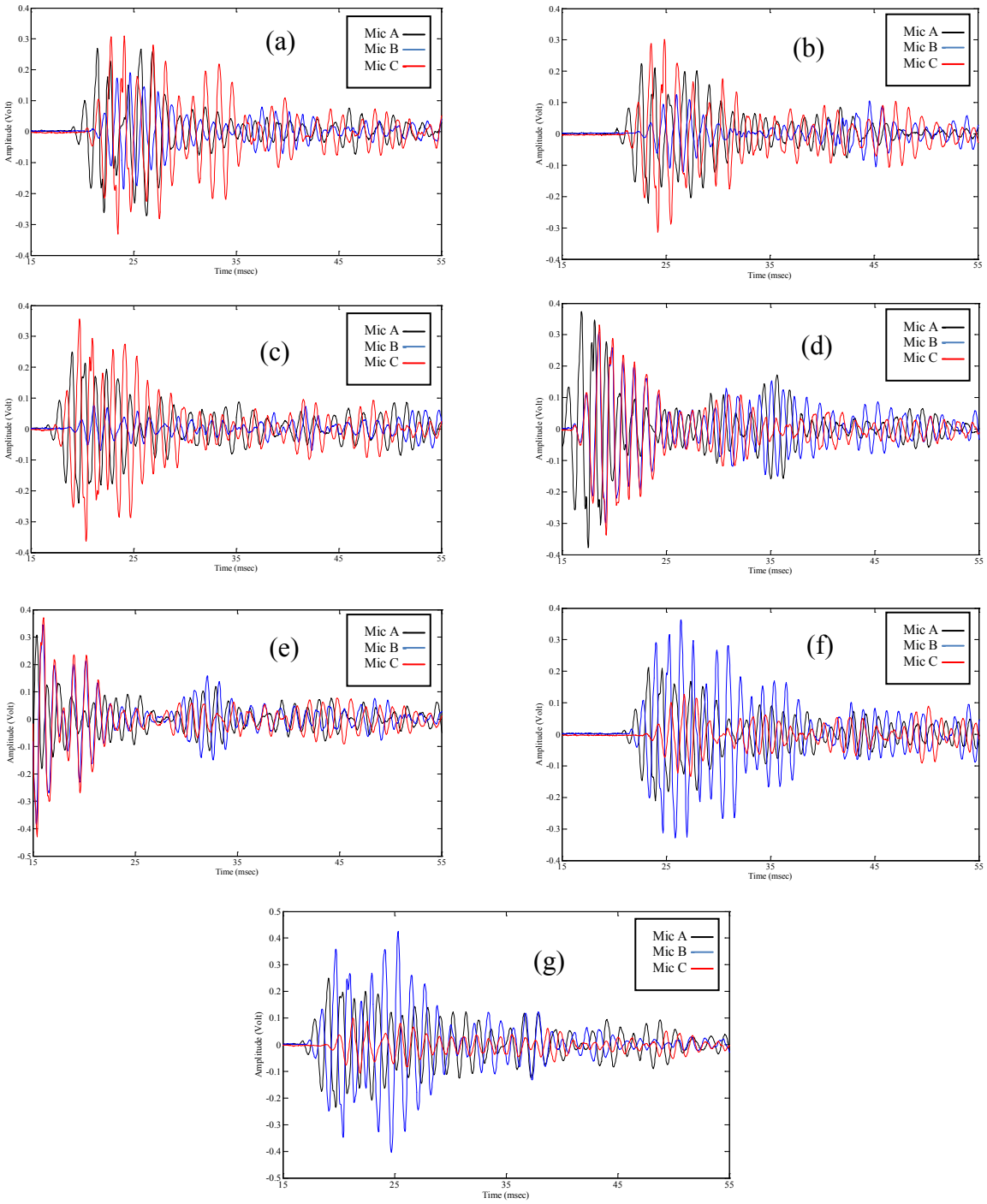


Fig. 30 - Waveform of all microphones at 1200 Hz for Indoor experiment positions (a) S2, (b) S3, (c) S4, (d) S5, (e) S6, (f) S7 and (g) S8.

The localization of the 1200 Hz source was pursued by using the KFact as well. The results are presented in Fig. 31. As the feature's values associated with the microphones A, B, and C were, respectively, equal to 1.0, 0.55, and 0.60; the possible incoming sound directions are those shown in Fig. 31(a). By intersecting the semi-straight lines the planar coordinates of points P1, P2, and P3 were obtained, namely x = 2.60 m y = 0.0 m for microphone pair A-B, x = 3.28 m, y = 0.08 m for microphone pair B-C, and x = 4.47 m y = 0.0 m for microphone pair A-C. The latter location is in excellent agreement with the locations true position of the speaker (x = 0.00 m, y = 4.58 m).

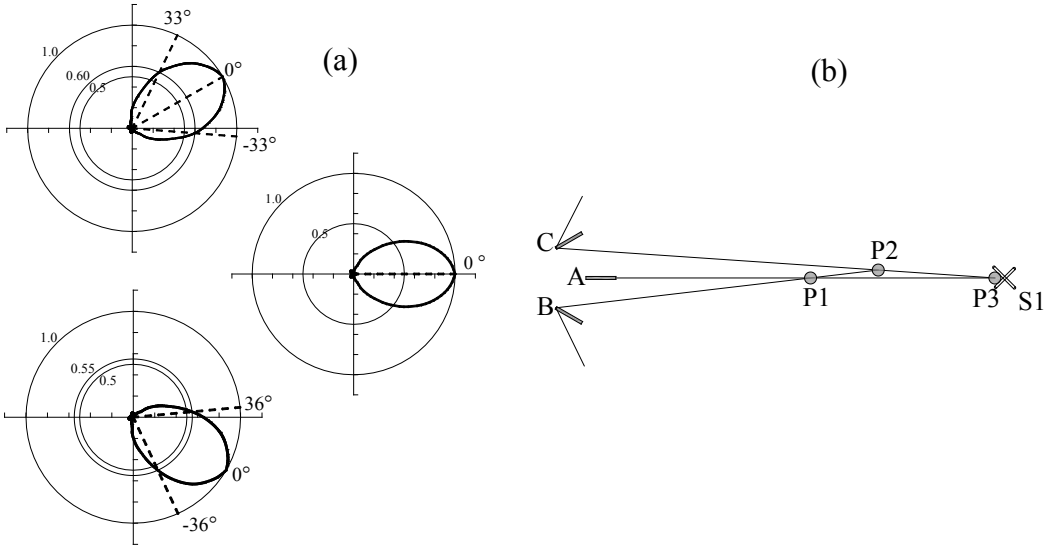


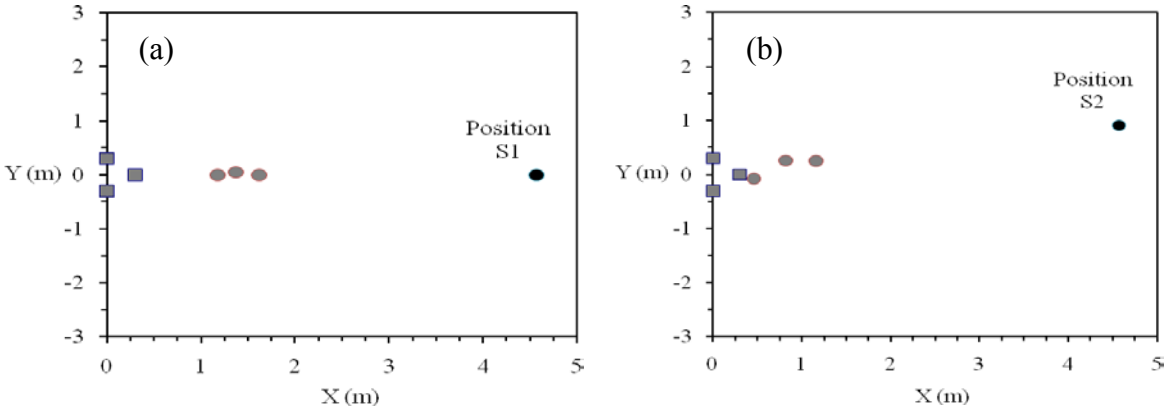
Fig. 31 - Indoor test. Speaker at location S1. Source localization by using the feature of the K-factor of the 1200 Hz toneburst. (a) Experimental polar directivity patterns given by Fourier 3. (b) Localization of the sound source.

The outcomes illustrated in Fig. 29 and 32 are summarized in Table 3, which also includes the results of the algorithm when a 1600 Hz pulse was detected and the feature of Max_FFT was applied. Overall, it can be seen that the position of the speaker was underestimated, i.e. the speaker was located closer than it really was. Microphones A and C identified the position of the source with high order of accuracy when the KFact and the frequency of 1200 Hz was employed.

Table 3 – Localization performance of some frequency and feature when the acoustic source was located at (x=4.57 m, y = 0.0).

(4.57 ,0)				
Feature	Frequency	Solution Mic A-B	Solution Mic B-C	Solution Mic A-C
PPK	1200 Hz	(1.18 , 0.0)	(1.37 , 0.05)	(1.62 , 0.0)
KFACT		(2.60 , 0.0)	(3.28 , 0.08)	(4.47 , 0.0)
Max_FFT		(1.10 , 0.0)	(0.65 , -0.12)	(0.46 , 0.0)
PPK	1600 Hz	(2.76, 0.0)	(2.06, 0.08)	(1.65 , 0.0)
KFACT		(2.89 , 0.0)	(2.26, -0.06)	(1.86, 0.0)
Max_FFT		No Intersection		

The same approach as described above was utilized to localize the speaker placed in different position in the laboratory according to the scheme illustrated in Fig. 28 (d). Figures 32(a) to (h) compares the results of the PPK feature applied to the 1200 Hz pulse recorded by the microphone array when the speaker was positioned at S1 to S8 respectively. The *gray circles* represent the solutions found, the *black circle* the position of the speaker, and the *gray square* the position of the acoustic sensors.



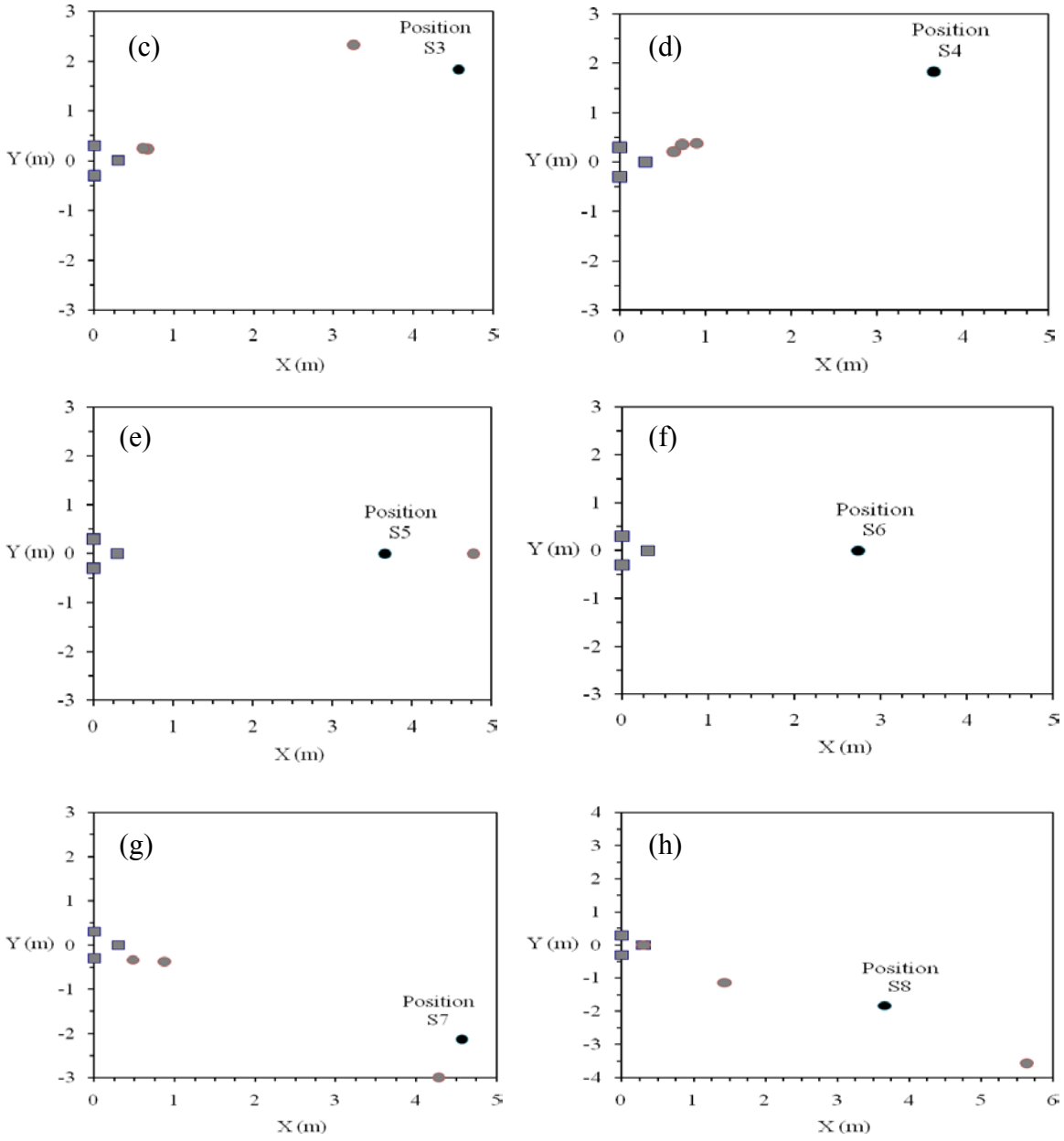


Fig. 32 - Location of the speaker obtained by using the feature of the peak-to-peak amplitude at 1200 Hz when the speaker was located at position (a) S1, (b) S2, (c) S3, (d) S4, (e) S5, (f) S6, (g) S7 and (h) S8 . The gray squares identify the microphones, the gray circles identify the calculated position, the black circle represents the true position of the speaker.

In Fig. 32 the results shows that in some cases the solution is far from the exact position of speaker (32(a), (b), (d) and (h)) while in some cases the solution is close to the real position (32(c), (e) and (g)). I

32(f), there is no intersection resulted in no solution found. In addition, although positions S3 and S7 were symmetric with respect to the array, the localization of position S7 is more accurate than S3.

The results from all these cases are summarized in Table 4, which lists the results associated with all eight speaker positions. It can be readily found that the vertical position (Y-axis) was found in good agreement with the true position. However, the estimate of the x-distance was in many cases unacceptable. A small variation in the selected angle results in large displacements in estimation of speaker position. For completeness the results obtained by using the Fourier 7 are provided in Table 5. It can be seen that the results are overall similar, which therefore imply that increasing the computational costs of the algorithm doesn't alter the performance of the localization algorithm significantly.

In this table, the position of speaker and corresponding coordinate and intersection of semi-straight lines for microphones A, B and C are presented. When the semi-straight lines are not converging, there is no intersection as shown in this table.

Table 4 – Localization of different position of speaker by interpolation of Fourier3 fitting curve for PPK at frequency of 1200 Hz.

PPK - Fourier3 - 1200Hz				
POSITION	Coordinate	Solution Mic A-B	Solution Mic B-C	Solution Mic A-C
S1	(4.57 , 0)	(1.18 , 0.0)	(1.37 , 0.05)	(1.62 , 0.0)
S2	(4.57 , 0.91)	(0.46 , -0.08)	(1.16 , 0.25)	(0.82 , 0.26)
S3	(4.57 , 1.83)	(3.25 , 2.33)	(0.67 , 0.24)	(0.61 , 0.25)
S4	(3.66 , 1.83)	(0.63 , 0.21)	(0.73 , 0.36)	(0.90 , 0.38)
S5	(3.66 , 0)	(4.77 , 0)	No Intersection	
S6	(2.74 , 0)	No Intersection		
S7	(4.57, -2.13)	(0.82 , -0.37)	(0.87, -0.37)	(1.31, -0.71)
S8	(3.66, -1.83)	(5.64 , -3.56)	(1.43 , -1.13)	(0.3 , 0.0)

Table 5 – Localization of different position of speaker by interpolation of Fourier7 fitting curve for PPK at frequency of 1200 Hz.

PPK - Fourier7 - 1200Hz				
POSITION	Coordinate	Solution Mic A-B	Solution Mic B-C	Solution Mic A-C
S1	(4.57 , 0)	(1.14 , 0)	(1.32 , 0.04)	(1.54 , 0)
S2	(4.57 , 0.91)	No intersection	(1.18 , 0.25)	(0.85 , 0.27)
S3	(4.57 , 1.83)	No intersection	(0.70 , 0.24)	(0.62 , 0.25)
S4	(3.66 , 1.83)	No intersection	(0.67 , 0.35)	(0.93 , 0.40)
S5	(3.66 , 0)	(3.60 , 0)	(8.30 , 0.40)	No intersection
S6	(2.74 , 0)	No Intersection		
S7	(4.57, -2.13)	(0.70 , -0.34)	(0.88 , -0.36)	(0.35 , 0.04)
S8	(3.66, -1.83)	(6.09 , -3.81)	(1.64 , -1.25)	(0.36 , -0.04)

Comparing tables 4 and 5, it can be found that using Fourier7 in some cases does not change the result significantly, as presented for position S1 and S8. On the other hand, in some other cases like S3, there is a large changes obtained by using Fourier 7.

Table 6 illustrates the results of using features Kfact and Max_FFT in addition to PPK for frequencies of 1200 Hz and 1600 Hz when speaker is located at position S1. As shown in this table for each feature at each frequency, results are far from each other.

Table 6 – Results of expected position of speaker associated with PPK, KFACT and Max_FFT at frequencies of 1200 Hz and 1600 Hz when speaker is at position S1.

(4.57 ,0)				
Feature	Frequency	Solution Mic A-B	Solution Mic B-C	Solution Mic A-C
PPK	1200 Hz	(1.18 , 0.0)	(1.37 , 0.05)	(1.62 , 0.0)
KFACT		(2.60 , 0.0)	(3.28 , 0.08)	(4.47 , 0.0)
Max_FFT		(1.10 , 0.0)	(0.65 , -0.12)	(0.46 , 0.0)
PPK	1600 Hz	(2.76, 0.0)	(2.06, 0.08)	(1.65 , 0.0)
KFACT		(2.89 , 0.0)	(2.26, -0.06)	(1.86, 0.0)
Max_FFT		No Intersection		

6.0 CONCLUSIONS

6.1 COMMENTS

In this research, the results of a sound source localization scheme based on the use of unidirectional microphones (UM) and amplitude-based sound features to determine the planar position of an acoustic pulse source. UMs are sensors whose output level strongly depends on the direction of the sound propagation. The profile of the microphones' directivity is typically expressed by the polar directivity pattern (PDP), which correlates the microphone output level with the direction of the wave propagation. In the proposed method, PDP is exploited to associate the detector output to a certain sound direction. By intersecting the directions provided by microphones' pairs, the position of the emitter is identified. The potential novelty of the research presented in this paper with respect to existing technologies is multifold: 1) use of UMs rather than OMs; 2) use of signal processing associated with statistical energy-based features from the sound's time, frequency, or joint time-frequency (wavelet) domain instead of the signal's arrival time; 3) use of empirical equations instead of beamforming algorithms; 4) ability to capture and locate single events.

In this study the PDP of UMs were obtained by testing the sensors in an anechoic chamber. The PDP associated with one sensor was considered applicable for all other microphones. After that, the performance of the algorithm to localize the position of a speaker in a closed environment was tested. The results of some pure tonebursts and few amplitude-based features have been presented. In general, the results show that the performance of the algorithm degrades when the accuracy about the relative sensitivity among the microphone is determined fairly. Moreover, the outcomes of this study make

obvious that the accuracy of the localization algorithm is related closely to the accuracy of the experimental polar pattern. The accuracy of the localization might have suffered from different sensitivity of the microphones in terms of directivity profile, i.e. different microphone may have similar but not identical PDP. In addition, as the method relies upon amplitude-based features, the sensitivity of the hardware system to capture the same amplitude pressure irrespective of the microphone and the data acquisition channel used is critical. Future studies must take into account such factors by, for instance, introducing calibration factors empirically determined.

The study, that has demonstrated the feasibility of the proposed method as an efficient, inexpensive, and compact alternative to the current technologies based on omnidirectional microphones and the estimation of the sound arrival time, is part of an ongoing effort to pave the road toward a new generation of sound detectors and localizers that can minimize the spacing between the microphone elements. The scheme shows computational advantages over the existing techniques and it may offer a suitable tool for further developments and field deployment.

Two key elements of the proposed approach are the normalization of the microphones' outputs and the absence of the time variable. In the study presented here, the analytical polar pattern for the three microphones is considered identical. The normalization is accomplished by knowing the output that would be detected by any unidirectional microphone in the array if oriented along the direction of maximum sensitivity ($\vartheta = 0^{\circ}$). By estimating the relative output values among the array's microphones, the approach becomes immune from the amount of acoustic energy emitted by the source and from the attenuation along the emitter-detectors paths.

6.2 RECOMMENDATION FOR FUTURE WORK

Future solution for the signal normalization may consider an omnidirectional microphone located in the center of mass of the array, for instance the triangle ABC of Fig. 2(a) may be used. The omnidirectional microphone's output $M_{OM} = g(\lambda, d)$ is only dependent on the source-detector distance d

and wavelength λ. Given $M_{UM} = f(0, a, b, d)$ the output of the unidirectional microphones when oriented toward the source, a sensitivity factor $K = M_{UM} / M_{OM}$ can be calculated. Such factor can be easily determined in simple laboratory testing prior to assemble the sensing array. Once the array is assembled and deployed a new measurement will record $M'_{OM} = g(\lambda, d')$, $M'_{UM} = f(\vartheta, a, b, d')$, and $K' = M'_{UM} / M'_{OM}$. Such ratio will be equal to K if the source is oriented at ϑ=0 or lesser than K. Therefore the ratio K'/K becomes the normalized value comprised between 0 and 1 that can be exploited for the sound source localization.

For comparing the results, it is suggested to perform an experiment in a usual room; similar to the experiment for anechoic chamber described in chapter 4. In this regard, the effect of echo and reverberation of sound with respect to the walls, floors and ceiling may be detected. It is seemed that, most probably, the echo and reverberation of sound affects the results drastically. The applied sound source localization method may be enhanced by recognizing this contribution and filtering it.

It is also recommended to perform a test to record data for different positions of speaker in anechoic chamber, similar to the experiment described in chapter 5. By analyzing these data, with the absence of reverberation and echo, there is a possibility to localize the sound source more accurately.

BIBLIOGRAPHY

1. Rizzo, P., Bordoni, G., Marzani, A., and Vipperman, J. (2009). "Localization of Sound Sources by Means of Unidirectional Microphones," Measurement Science and Technology, 20, 055202 (12pp).

2. Huang Y., Benesty J., and Elko, G.W. (2000). "Passive acoustic source localization for video camera steering," Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-00) 2 II909-II912, doi: 10.1109/ICASSP.2000.859108.

3. Kim, H-D., Choi J-S, Kim, M., and Lee, C-H, (2004). "Reliable detection of sound's direction for human robot interaction," Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-04) 3 2411- 2416, doi: 10.1109/IROS.2004.1389769.

4. Choi, Y.-C., and Kim Y.-H. (2007). "Near field impulsive source localization in a noisy environment," J. Sound Vib., 303, 209-220.

5. Wilson, D.K, Sadler, B.M. and Pham, T. (2002). "Simulation of detection and beamforming with acoustical ground," sensors Proceedings of SPIE, Seismic and Acoustic Unattended Ground Sensors I 4743 50-61, doi: 10.1117/12.443525.

6. Liu, L., and Albert, D.G. (2006). "Locating a sound source in an urban environment," 147th ASA Meeting, New York, NY (http://www.crrel.usace.army.mil/sid/ASA/SoundSource.htm)

7. Merimaa, J. (2002). "Application of a 3-D microphone array," Proc. of the 112th Convention of the Audio Engineering Society, 1-10, Paper 5501.

8. Benesty, J., Chen, J., Huang, Y.: Microphone Array Signal Processing. Springer, Berlin (2008)

9. Gunel, B., Hacihabiboglu, H., and Kondoz, A.M. (2007). "Wavelet packet based analysis of sound fields in rooms using coincident microphone arrays," Applied Acoustics, 68, 778-796.

10.

11. Brandstein, M.S., and Silverman, H.F. (1997). "A practical methodology for speech source localization with microphone arrays," Computer Speech & Language 11 91-126.

12. Bourennane, S., and Bendjama, A. (2002). "Locating wide band acoustic sources using higher order statistics," Applied Acousics 63 235-251, doi: 10.1016/S0003-682X(01)00039-1.

13. Reid, G.L., and Milios, E. (2003). "Active stereo sound localization," The Journal of the Acoustical Society of America 113, 185-193, doi: 10.1121/1.1518469.

14. Benesty, J. (2000). "Adaptive eigenvalue decomposition algorithm for passive acoustic source localization," The Journal of the Acoustical Society of America 107 384-391, doi: 10.1121/1.428310.

15. Liu, H. and Milios, E. (2005). "Acoustic positioning using multiple microphone arrays," The Journal of the Acoustical Society of America 117 2772-2782, doi: 10.1121/1.1895005.

16. Martinson, E., and Schultz, A. (2009). "Discovery of sound sources by an autonomous mobile robot," Auton Robot, 27: 221–237, DOI 10.1007/s10514-009-9123-1.

17. Bian, X., Abowd, G. D., and Rehg, J. M. (2005). "Using sound source localization in a home environment, Pervasive Computing," 3468, 19–36, doi: 10.1007/b136550.

18. DiBiase, J.H. Silvermann, H.F., and Brandstein, M.S. (2001). "Robust Localization in Reverberant Rooms," in Microphone Arrays: signal processing techniques and applications. M. Brandstein and D. Ward (Eds.), Chapter 8, 157-180, Springer.

19. Huang, Y., Benesty, J., Elko, G.W. (1999). "Adaptive eigenvalue decomposition algorithm for real time acoustic source localization system," Acoustics, Speech, and Signal Processing, 1999. ICASSP '99. Proc. 1999 IEEE International Conference on, 15-19 Mar 1999, 2, 937-940. doi: 10.1109/ICASSP.1999.759826.

20. Nakadai, K., Matsuura, D., Okuno, H.G., and Kitano, H. (2003) "Applying Scattering Theory to Robot Audition System: Robust Sound Source Localization and Extraction," Proc. of the 2003 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems (IROS-03), 2, 1147-1152.

21. Birchfield, S., and Gillmor, D. (2001). "Acoustic source direction by hemisphere sampling," Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-01), 5, 3053-3056, doi: 10.1109/ICASSP.2001.940302

22. Rui, Y., Florencio, D., Lam, W. and Su, J. (2005) "Sound source localization for circular arrays of directional microphones," Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-05), 3, 93-96, doi:10.1109/ICASSP.2005.1415654.

23. Ikeda, A., Mizoguchi, H., Sasaki, Y., Enomoto, T., and Kagami, S. (2007). "2D Sound source localization in azimuth & elevation from microphone array by using a directional pattern of element," Proc. of IEEE Sensors 2007, 1213-1216, doi: 10.1109/ICSENS.2007.4388627.

24. Elko, G.W., Chou T.C., Lustberg R.J. and Goodwin, M.M. (1994). "A constant-directivity beamforming microphone array (A)," The Journal of the Acoustical Society of America, 96, 3244, doi: 10.1121/1.411096.

25. Choi, Y.C., and Kim, Y.H. (2007). "Near field impulsive source localization in a noisy environment," Journal of Sound and Vibration, 303, 209–220.

26. Kobayashi, T., Kameda, Y., and Ohta, Y. (2008). "Sound source localization with non-calibrated microphones," Advances in Multimedia Modeling, 4903, 134-143, doi: 10.1007/978-3-540-77409-9.

27. Kwon, B., Park, Y. and Park, Y.-S. (2007). "Sound source localization using the compensation method in robot platform," Proc. of the IEEE International Conference on Control, Automation and Systems (ICASSP-07), 1911-1914, doi: 10.1109/ICCAS.2007.4406659.

28. Atmoko, H., Tan, D.C., Tian, G.Y. and Fazenda, B. (2008). "Accurate sound source localization in a reverberant environment using multiple acoustic sensors," Meas. Sci. Technol. 19 024003 (10pp), doi: 10.1088/0957-0233/19/2/024003.

29. Brutti, A., Omologo, M., and Svaizer, P. (2008). "Comparison between different sound source localization techniques based on a real data collection," Proc. of the Hands-Free Speech Communication and Microphone Arrays (HSCMA-08) 69-72, doi: 10.1109/HSCMA.2008.4538690.

30. Smith, J.O., and Abel, J. S. (1987). "Closed-form least-squares source location estimation from range-difference measurements," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-35, 1661–1669.

31. Brandstein, M.S., Adcock, J.E., and Silverman, H.F. (1997). "A closed-form location estimator for use with room environment microphone arrays," IEEE Trans. Speech Audio Processing, 5, 45–50.

32. Yao, K., Hudson, R.E., Reed, C.W., Chen, D., and Lorenzelli, F. (1998) "Blind beamforming on a randomly distributed sensor array system," IEEE J. Select. Areas Commun., 16, 1555–1567.

33. Omologo, M., and Svaizer, P. (1996). "Acoustic Source Location In Noisy And Reverberant Environment Using CSP Analysts," Proc. ICASSP, 921-924.

34. Brandstein, M.S., Adcock, J.E., and Silverman, H.F. (1995). "A Closed-Form Method For Finding Source Locations From Microphone-Array Time-Delay Estimates," IEEE International Conference On Acoustics Speech And Signal Processing, 3019-3022.

35. Chen, J.C., Hudson, R.E., and Yao, K. (2002). "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," IEEE Trans. Signal Processing, 50, 1843–1854.

36. Nix, J., Hohmann, V. (2006). "Sound source localization in real sound fields based on empirical statistics of interaural parameters," J. Acoust. Soc. Am., 119(1), 463-478.

37. Pu, C.J., Harris, J.G., and Principe, J.C. (1997). "A neuromorphic microphone for sound localization," 1997 IEEE International Conference on Systems, Man, and Cybernetics, 2, 1469-1474.

38. Tashev, I. and Malvar, H., (2005). "A new beamformer design algorithm for microphone arrays", Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP'05), 3, iii/101-iii/104, Digital Object Identifier: 10.1109/ICASSP.2005.1415656.

39. Elko, G.W. (2001). "Spatial coherence functions for differential microphones in isotropic noise fields," in Microphone Arrays: signal processing techniques and applications. M. Brandstein and D. Ward (Eds.), Ch. 4, 61-85, Springer.

40. C. Marty, F. Bruel, D. Prieur, P. Naz, and L. S. Miller (2009). "NATO SET-093 joint field experiment at Bourges, France", Proc. SPIE Vol. 7333, 73330K.

41. Reiff, C.G.(2009). "Acoustic source localization and cueing from an aerostat during the NATO SET-093 field experiment", Proc. SPIE Vol. 7333, 73330M.

42. Xiao, D., Wagstaff, R.A., Anderson, J.D., and Gilbert, K.E. (2009). "Sound source localization using distributed elevated acoustic sensors", Proc. SPIE Vol. 7333, 73330O.

43. P. Naz, Ch. Marty, S. Hengy, and L. S. Miller (2009). "Acoustic detection and localization of weapons fire by unattended ground sensors and aerostat-borne sensors", Proc. SPIE Vol. 7333, 73330N.

44. W.E. Prather, R.C. Clark, J. Strickland, W.G. Frazier, and J. Singleton (2009). "Detection of impulsive sources from an aerostat-based acoustic array data collection system", Proc. SPIE Vol. 7333, 73330Q.

45. Ash, J.N., and Moses, R.L. (2009). "Time-delay estimation in time-warping environments", Proc. SPIE Vol. 7333, 73330T.

46. Deligeorges, S., Xue, S., et al. (2009). "Biomimetic Smart Sensors for Autonomous Robotic Behavior I: Acoustic Processing," Bio-Inspired/Biomimetic Sensor Technologies and Applications, edited by Nicholas F. Fell Jr., Venkataraman S. Swaminathan, Proc. of SPIE Vol. 7321, 732107 · © 2009 SPIE · CCC code: 0277-786X/09/$18 · doi: 10.1117/12.820935.

47. Xue, S., Deligeorges, S., et al. (2009). "Biomimetic Smart Sensors for Autonomous Robotic Behavior II: Vestibular Processing" Bio-Inspired/Biomimetic Sensor Technologies and Applications, edited by Nicholas F. Fell Jr., Venkataraman S. Swaminathan, Proc. of SPIE Vol. 7321, 732108 · © 2009 SPIE · CCC code: 0277-786X/09/$18 · doi: 10.1117/12.820936.

48. Hubbard, A., Cohen, H., et al. (2008). "Biologically Inspired Circuitry that Mimics Mammalian Hearing," Bio-Inspired/Biomimetic Sensor Technologies and Applications, edited by Nicholas F. Fell Jr., Venkataraman S. Swaminathan, Proc. of SPIE Vol. 7321, 732109 · © 2009 SPIE · CCC code: 0277-786X/09/$18 · doi: 10.1117/12.821282.

49. Pu, Y., Kelsall, S., et al. (2009). "Back-End Algorithms that Enhance the Functionality of a Biomimetic Acoustic Gunfire Direction Finding System," Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland

Defense VIII, edited by Edward M. Carapezza, Proc. of SPIE Vol. 7305, 730513 · © 2009 SPIE · CCC code: 0277-786X/09/$18 · doi: 10.1117/12.818811.

50. http://www.shotspotter.com/index.html (Last date accessed December 27, 2009).

51. Sessler G M and West J E 1975 Second Order Gradient Uni-Directional Microphones Utilizing an Electret Transducer The Journal of the A coustical S ociety of A merica 5 8 273 -278, doi : 10.1121/1.380657.