Design and development of a planar MEMS microphone array

Damjan Pecioski[1]
Ss. Cyril and Methodius University in Skopje
Faculty of Mechanical Engineering – Skopje
Rugjer Boshkovikj 18, 1000 Skopje, Republic of North Macedonia

Simona Domazetovska Markovska[2]
Ss. Cyril and Methodius University in Skopje
Faculty of Mechanical Engineering – Skopje
Rugjer Boshkovikj 18, 1000 Skopje, Republic of North Macedonia

Viktor Gavriloski[3]
Ss. Cyril and Methodius University in Skopje
Faculty of Mechanical Engineering – Skopje
Rugjer Boshkovikj 18, 1000 Skopje, Republic of North Macedonia

Maja Anachkova[4]
Ss. Cyril and Methodius University in Skopje
Faculty of Mechanical Engineering – Skopje
Rugjer Boshkovikj 18, 1000 Skopje, Republic of North Macedonia

Dejan Shishkovski[5]
Ss. Cyril and Methodius University in Skopje
Faculty of Mechanical Engineering – Skopje
Rugjer Boshkovikj 18, 1000 Skopje, Republic of North Macedonia

Anastasija Ignjatovska Angjusheva[6]
Ss. Cyril and Methodius University in Skopje
Faculty of Mechanical Engineering – Skopje
Rugjer Boshkovikj 18, 1000 Skopje, Republic of North Macedonia

ABSTRACT

In the last few years the ability to visualize an audio source has become a widely used tool in many fields such as speech recognition and enhancement, especially in noise localization in different environments. Acoustic cameras are widely used as a tool for noise source investigation. However, an acoustic image is difficult to achieve in environments with a large amount of noise and reverberation. An effective *approach to overcome this issue is the use of beamforming theory coupled with a microphone array in order to obtain a recording of the desired acoustic signal. The aim of this work is to design and develop a low cost microphone array which can be upgraded with a camera, i.e. acoustic camera, based on digital MEMS microphones and a Raspberry Pi. Different formations of the microphone array are developed, tested and compared. The details on hardware integration as well as the development environment are discussed in this paper.*

1. **INTRODUCTION**

Single microphone systems have their limitations such as they work best when the interference remains stationary and when the signal-to-noise ratio (SNR) stays positive across most of the frequency range. On the other hand, microphone array systems offer a significant advantage as they can handle non-stationary interference or even filter out extremely strong unwanted signals. By using multiple microphones, these systems capture spatial information, such as the location of the sound source, allowing for much better audio clarity. Their primary goal is to extract a high-quality version of the desired signal, even in noisy or challenging environments [1].

A microphone array consists of a set of microphones positioned in a way that the spatial information is well captured. Depending on the nature of the applications the geometry of the microphone array may play an important role in the formulation of processing algorithms. To achieve this, microphone arrays rely on beamforming techniques, which focus on reducing background noise while preserving speech clarity [2] . Essentially, a beamformer acts as a spatial filter, distinguishing between the target signal and interference based on their direction of arrival, even when they share the same frequency range. With the addition of a camera the beamforming can be enhanced to give a special acoustic image of the sound source [3].

Microphone arrays and beamforming techniques are used in a wide range of applications such as speech recognition, machine fault monitoring as well as traffic monitoring. In [4], beamforming techniques are used to modal sound field conditions of a vehicle interior. use beamforming for automotive noise source identification. Within the work of Gerges [5] software as well as a hardware implementation of beamforming has been developed with a sparse planar microphone array for the use in automotive noise source identification. Building of low-cost acoustic cameras is of interest to some researchers as in [6-7]. Most of the design presented in the available literature relies on the use of electret capsule microphones. Where Zimmerman [6] uses an FPGA connected with 32 microphones to achieve sound localization, and Orman [7] creates a handheld acoustic camera for the use in monitoring systems for electric motors. Whereas in the field of voice recognition Yoganathan, [8] employs a dual-microphone setup in order to capture speech and attain a more clear and precise voice capture.

Developing high-performance technical systems is a constant challenge in modern engineering, driving continuous advancements in technology. The introduction of micro-electro-mechanical systems (MEMS) has created new possibilities for the development of sensor units based on MEMS microphones. These systems are not only cost-effective but also enable widespread deployment, making them suitable for applications requiring continuous noise monitoring across various environments. Integrating these microphones into larger sensor arrays and with the application of detection algorithms these arrays can be used to locate acoustic sources as well as estimate their strength [9]. Authors G. Szwoch and J. Kotus uses micro electro-mechanical sensors in their research in sound intensity detection of motor vehicles [10]. In [11] the authors use an acoustic sensor network of a larger number of sensors to monitor vehicle traffic. While the authors Jelena Tomić et all are considering the possibility of using artificial intelligence to improve accuracy and speed of noise data processing[12]. Within the work of

[13] an acoustic camera has been designed using different types of memes microphones and in [14], a Raspberry Pi is used to capture the sound of 24 mems microphones in order to achieve beamforming.

The aim of this paper is the design and development of a planar MEMS microphone array. Multiple variations or the microphone array have been designed using MATLAB phased array toolbox and physically implemented using a raspberry pi. The delay and sum beamforming algorithm has been used in the evaluation of the measurements which allow for accurate noise level detection and noise source localization.

2. THEORETICAL BACKGROUND

Visualization of the origin and intensity of sound waves is done in a similar fashion as thermal cameras visualize origin and intensity of heat sources. A microphone array is focused consecutively on different discrete points in a well-defined area. Superposition of all microphone signals enables the estimation of the sound pressure for each selected spatial point, which translates to a pixel of the intensity-image. A microphone array can be directed to a specific point in space without requiring physical adjustments. Instead, it relies on beamforming, a technique that processes the signals collected by multiple microphones [15]. By applying appropriate time delays to each microphone's input and summing these signals, the array effectively focuses on a desired location while minimizing interference from other directions.

When using time domain beamforming it is possible to analyze and listen to the data recorded since it is in time and not lost. There are no bandwidth limitations to this type of beamforming which makes it computationally efficient for wide band processing especially for a small beamforming grid. The disadvantage of this beamformer is that the data of all the microphones must be processed for every grid point so it is time consuming for a large dataset. The basic beamforming method for time domain is the delay and sum beamforming shown on Figure 1.

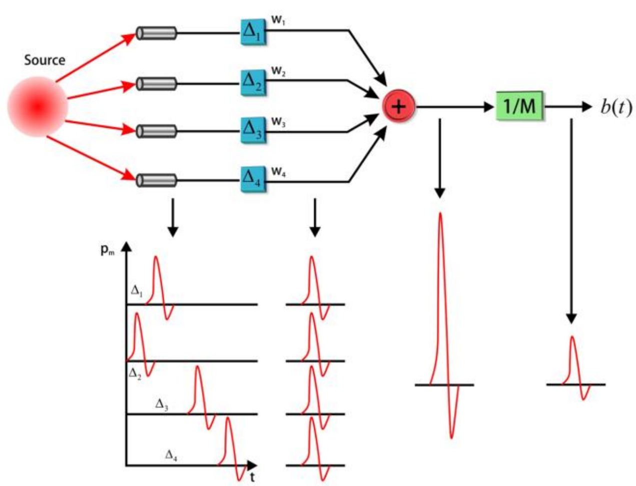


Figure 1. Delay and sum beamforming [5]

This beamforming method works by using the known distance between each microphone and calculating the delay between the time that each microphone picks up the sound. The signal is than summed and normalized to achieve the beamformed result. Considering a sound source at a position $\overrightarrow{x'}$ the pressure field is determined as a function of the distance from the source to an arbitrary microphone and the time f(r,t).

$$f(r,t) = \frac{s(t - \frac{|\vec{x} - \vec{x'}|}{c})}{|\vec{x} - \vec{x'}|} = \frac{s(t - \frac{r}{c})}{r} \tag{1}$$

In the equation the sound source is denoted as s(t), where the expression $s(t - |\vec{x} - \vec{x'}|/c)$ accounts for the time delay due to the finite speed of the sound. The denominator ($r = |\vec{x} - \vec{x'}|$) represents the signal strength decrease dependent on the distance. Considering an array of N microphones with the center as the origin coordinate the classical beam forming delay and sum algorithm is given by:

$$b(t) = \frac{1}{N} \sum_{n=1}^{N} w_n p_n(t - \Delta_n) \tag{2}$$

where $p_n(t) = f(r_n', t)$ is the signal received by the n^{th} microphone, w_n a weight and Δ_n is the time delay and N is the total number of microphones [16].

3. MICROPHONE ARRAY DESIGN

Within this research the MATLAB phased array toolbox has been used in order to simulate different array configuration patterns of omnidirectional microphones and obtain the directivity pattern of each different type of array. MEMS microphones have an omnidirectional response, which means that they respond equally to sounds coming from any direction as can be seen on Figure. 2.

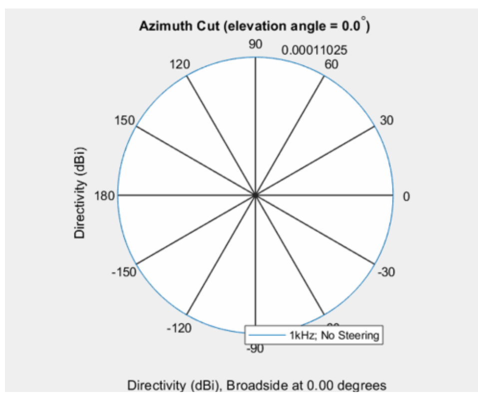
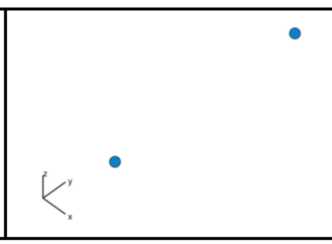
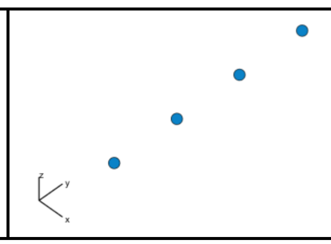
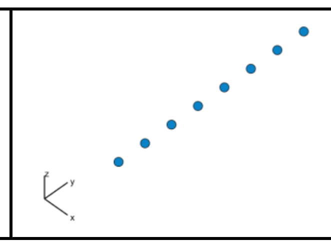


Figure 2. MEMS microphone response

Simulations were performed for different configurations of 4 and 8 microphones as well as one simulation for 2 microphone array which can be seen on Figure 3.

Linear array 2, 4 and 8 microphones			

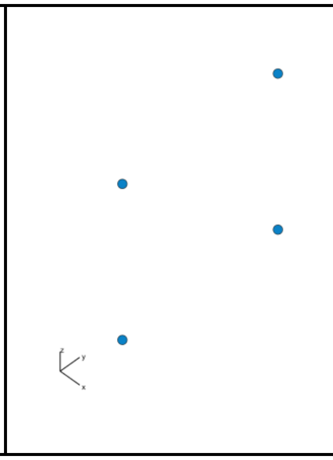
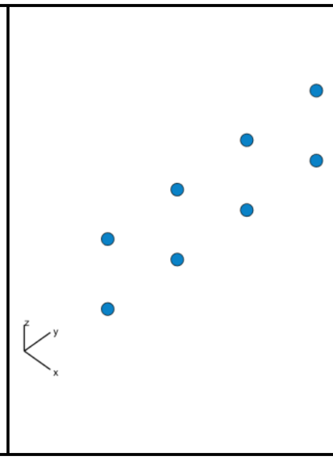
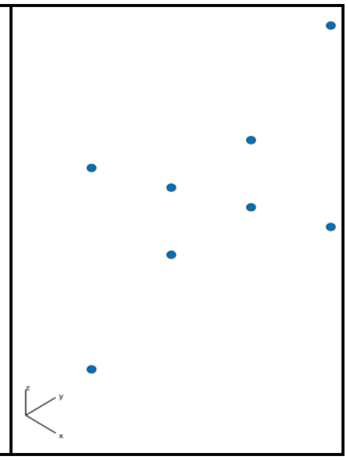
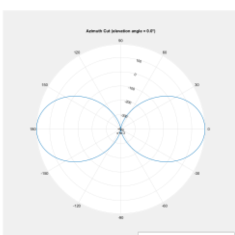
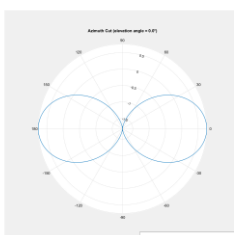
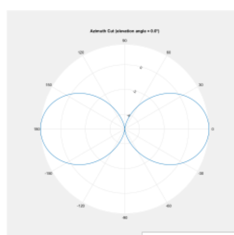
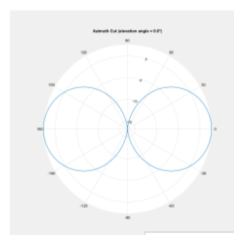
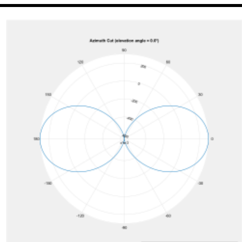
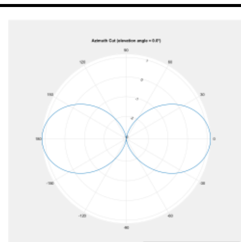
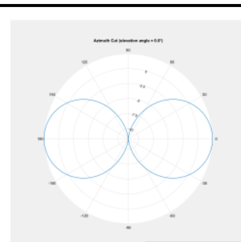
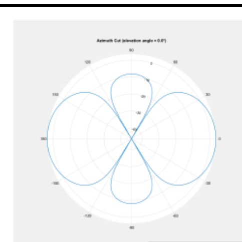
Rectangular array 4 and 8 microphones			

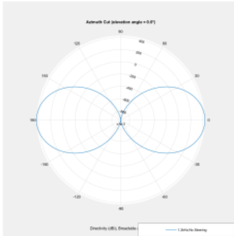
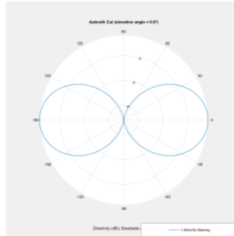
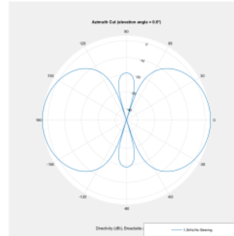
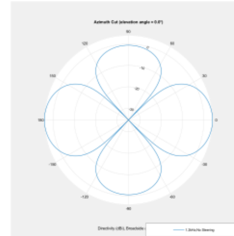
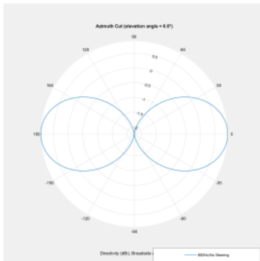
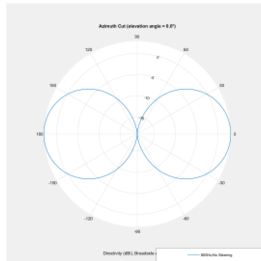
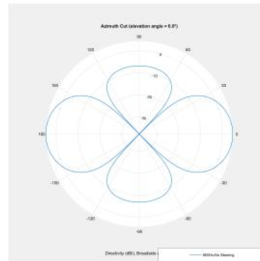
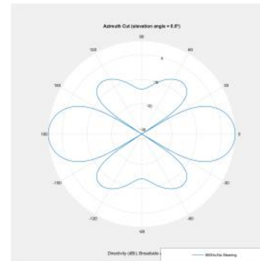
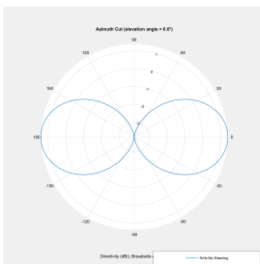
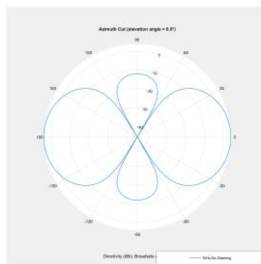
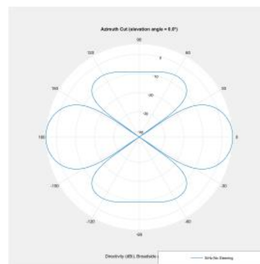
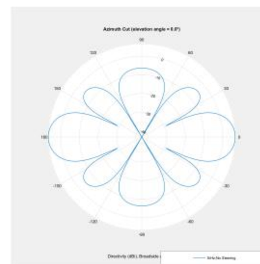
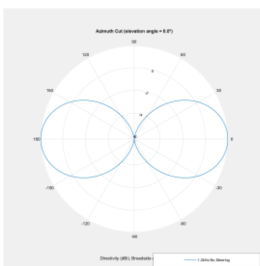
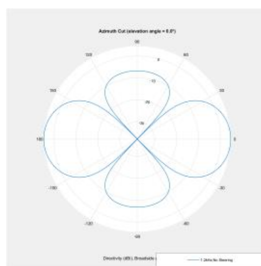
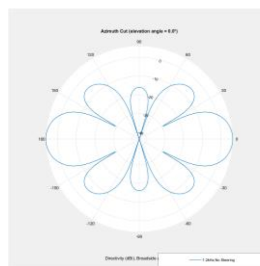
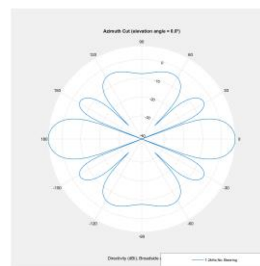
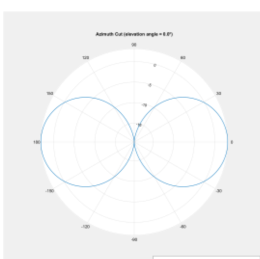
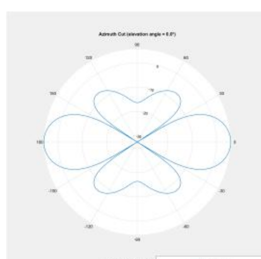
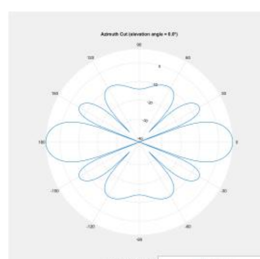
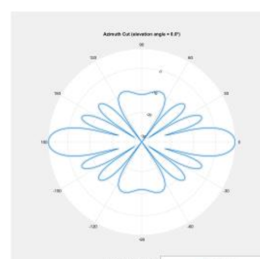
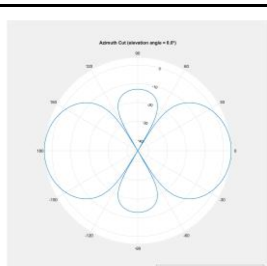
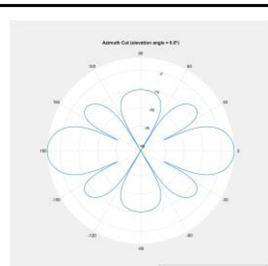
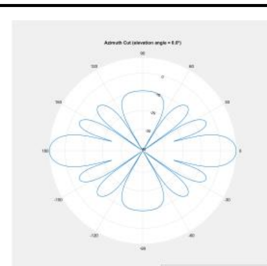
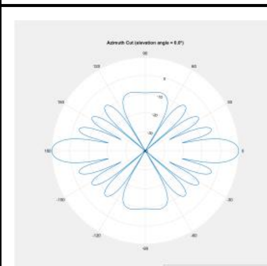
Figure 3. Different configurations for microphone arrays

The configurations for the 4 microphone array were a 2x2 square array and a linear array, while the simulations for the 8 microphone array were a linear 8 microphone array, a rectangular array with 2 rows of 4 microphones each and a custom square array where a square array of 4 microphones were placed in the center and 4 microphone were placed on the corners. These configurations were chosen as possible microphone arrays for a maximum of 8 microphones which is the number supported by the experimental setup.

The simulations were performed for three different frequencies, namely 800Hz, 1000Hz and 1200Hz as well as different spacings between the microphones. The chosen frequencies fall within mid-range frequencies for general acoustic signals in urban environments. The spacings were chosen so that the distance between each adjacent microphone is less than half of the wavelength or more than half of the wavelength (λ =0.429m; 0.343m; 0.286m).

In the following Figure 4 the directional pattern is given for all linear arrays with varying microphone spacing and varying frequencies.

N. mic & frequency	d=0.05m	d=0.1m	d=0.15m	d=0.2m
2 mic 800 Hz				
2 mic 1000 Hz				

2 mic 1200 Hz				
4 mic 800 Hz				
4 mic 1000 Hz				
4 mic 1200 Hz				
8 mic 800 Hz				
8 mic 1000 Hz				

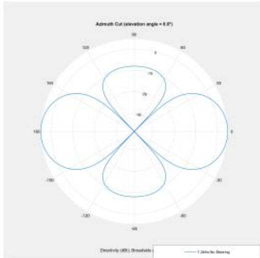
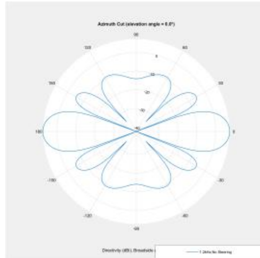
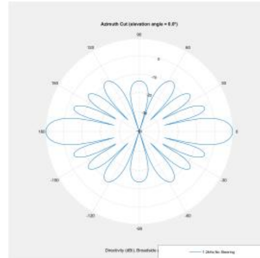
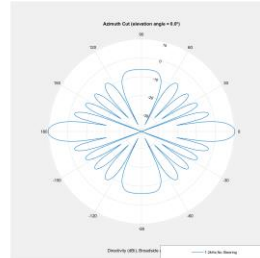
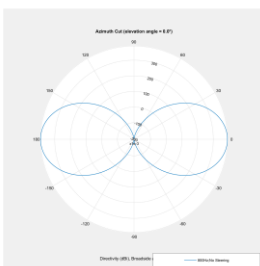
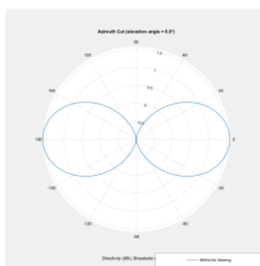
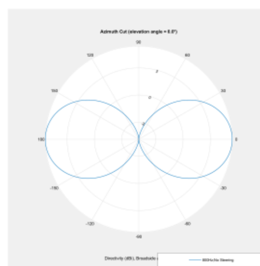
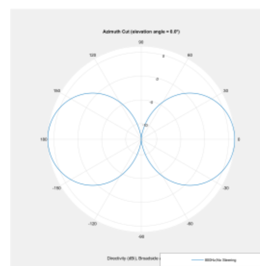
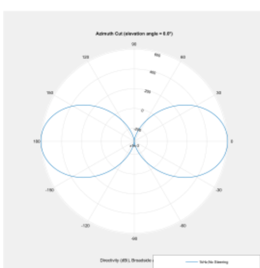
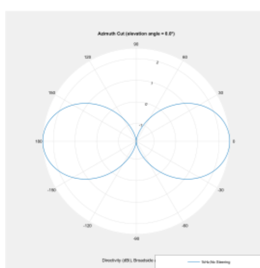
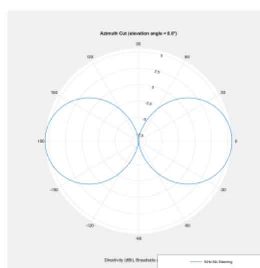
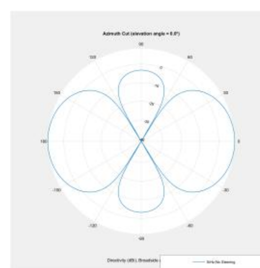
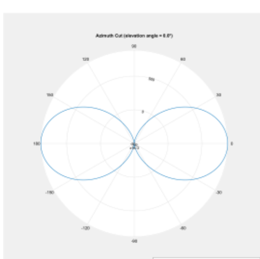
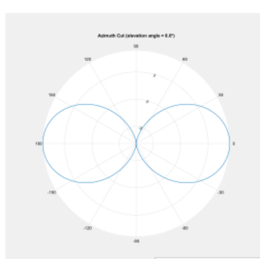
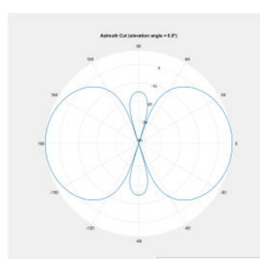
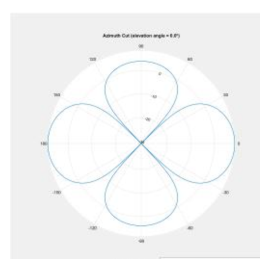
8 mic 1200 Hz				

Figure 4. Directivity pattern of linear array with 2, 4 and 8 microphones for frequency f=[800;1000;1200] Hz and a distance between adjacent microphones of d=[0.05;0.1;0.15;0.2] m

Linear arrays are commonly used for directional beamforming and source localization along one axis. The 2 microphone array is a basic form of beamforming which is often used for noise cancellation and basic direction detection while the 4 microphone array offers better spatial filtering then the 2 mic array and ability to isolate sources in noisy environments. Finally the 8 microphone array has a narrower main lobe which allows for better rejection of unwanted noise from other directions.

In the following Figure 5 the directional pattern is given for the rectangular array with 4 and 8 microphones with varying microphone spacing and varying frequencies.

N. mic & frequency	d=0.05m	d=0.1m	d=0.15m	d=0.2m
4 mic square 800 Hz				
4 mic square 1000 Hz				
4 mic square 1200 Hz				

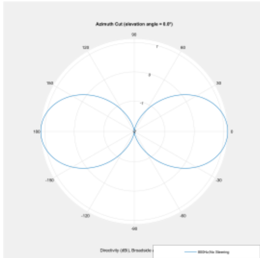
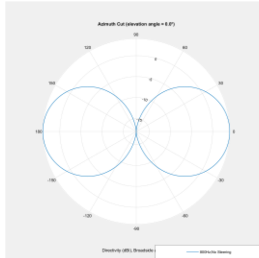
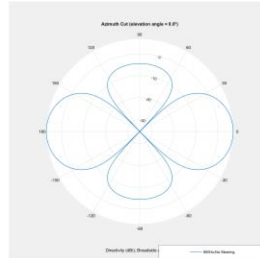
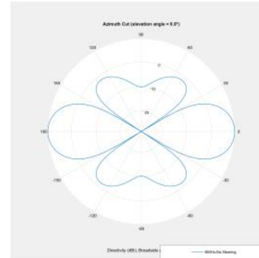
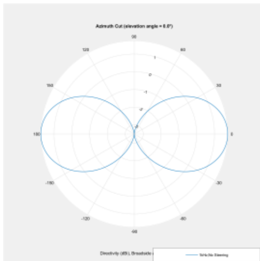
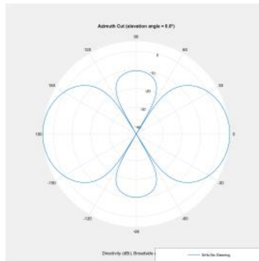
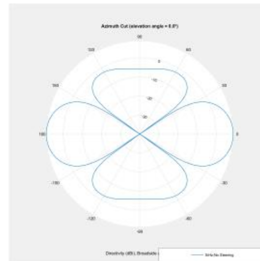
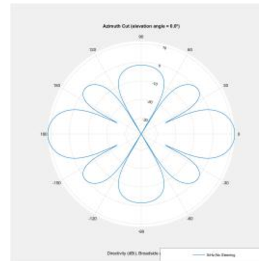
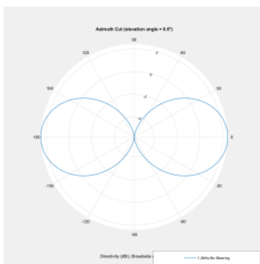
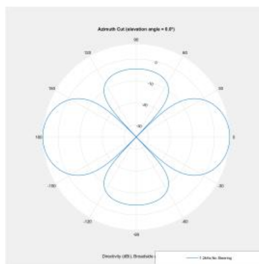
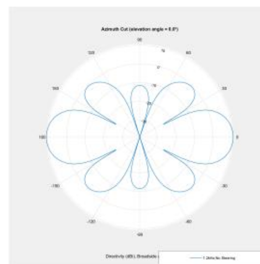
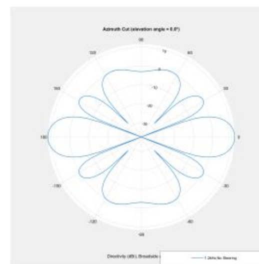
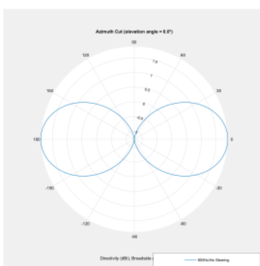
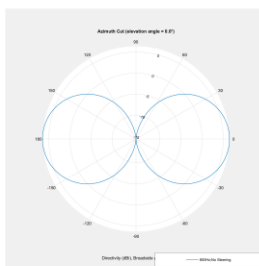
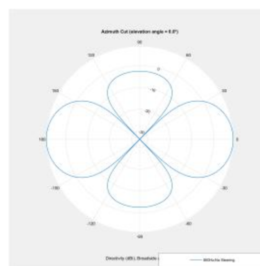
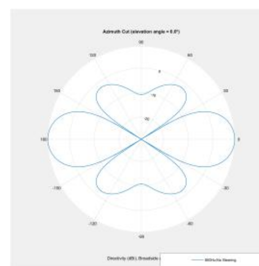
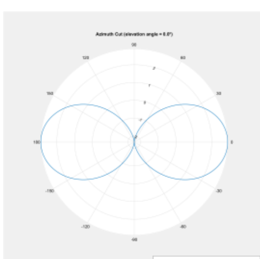
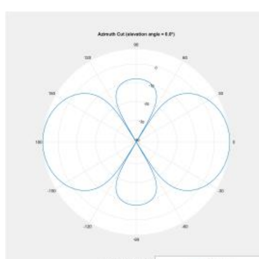
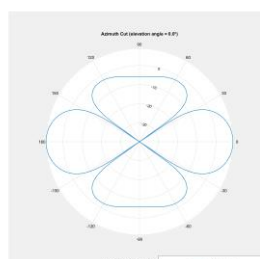
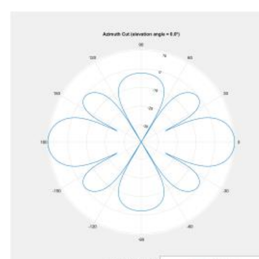
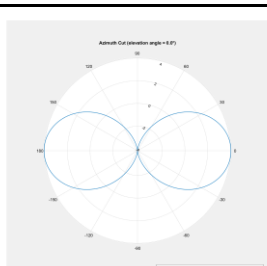
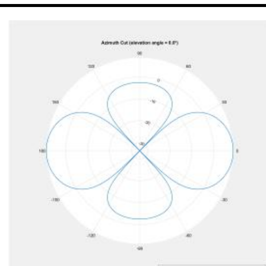
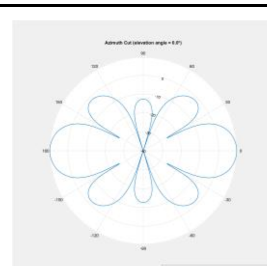
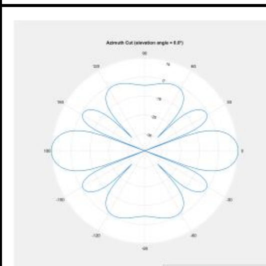
8 mic 2x4 800 Hz				
8 mic 2x4 1000 Hz				
8 mic 2x4 1200 Hz				
8 mic square 800 Hz				
8 mic square 1000 Hz				
8 mic square 1200 Hz				

Figure 5. Directivity pattern of rectangular array 4 and 8 microphones for frequency f=[800;1000;1200] Hz and a distance between adjacent microphones of d=[0.05;0.1;0.15;0.2] m

The uniform rectangular array (URA) with 4 microphone allows for improved sound capture in multiple directions. It is more effective for sound localization than the linear counterpart and suitable for environments where the source may move in a planar direction. The 8 microphone 2x4 is a hybrid between the linear and URA allowing for better rejection of unwanted noise compared to the 4 mic URA. This characteristic can be seen in the 8 microphone custom square configuration which was analyzed.

It can be concluded that from the results gathered from the simulations of linear and square arrays that the increase of the number of microphones produces a larger and narrower main lobe which means that the gain is higher of the signal in the desired direction. Different spacings affect spatial aliasing. Wide spacing works well for lower frequencies whereas narrow spacings allows better high-frequency beamforming but may lead to overlap issues with low frequencies.

4. EXPERIMENTAL SETUP

Multiple experiments were performed mirroring the design parameters from the MATLAB simulations. The experiments were performed using a sound source connected to an amplifier and a signal generator where sine tones were generated at the frequencies 800Hz, 1000Hz and 1200Hz. The speaker has been placed at 3 different locations with respect to the microphone array at a distance of 1.5m. The experimental setup can be seen on Figure 6 where the speaker shown is placed in the center position and the other two locations are 0.5m left and right of this position.

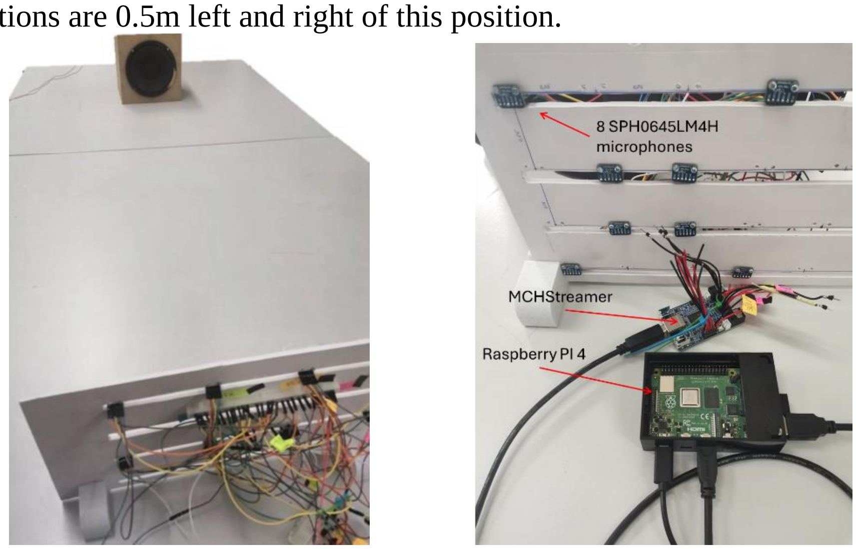


Figure 6. Experimental setup

The microphones used for the recording are I2S SPH0645LM4H MEMS microphones. Inter-IC Sound (I2S) is a standard communication protocol for digital microphones. I2S devices can be simply interfaced with any embedded system that provides I2S interface such as Raspberry Pi 4. Unfortunately, most available embedded systems have maximum capability to connect up to 2 channels of I2S device and that is not adequate for microphone array application. For this reason another device has been used together with the raspberry , MCHStreamer Lite USB interface which has 4 dedicated I2S channels where two microphones can be connected per channel in stereo configuration allowing for a total of 8 channel audio recording. The recorded audio is sent to the raspberry via the USB connection. The microphones for the testing were mounted on a board and their configuration was adapted to match the array configurations from the MATLAB simulations.

5. RESULTS AND DISCUSSION

The performance of the microphone array is measured in terms of accuracy in estimation of the degree of arrival, namely the azimuth and elevation angle as well as the noise intensity that is measured and attenuated using the beamforming method. In the following Figure 7 and Figure 8 two variations of the tests performed are given.

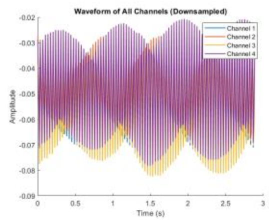
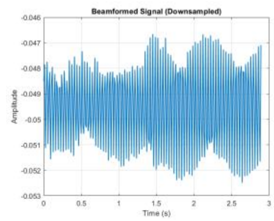
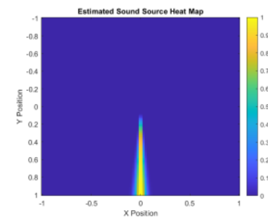
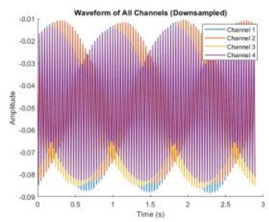
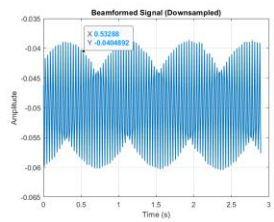
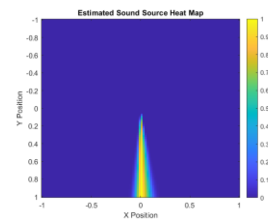
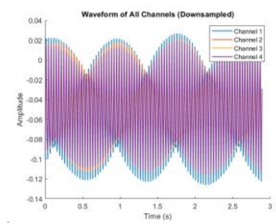
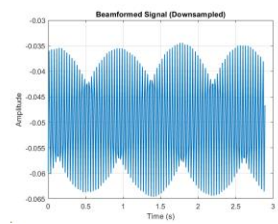
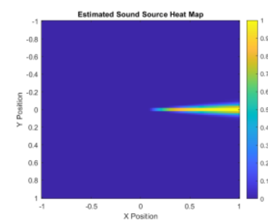
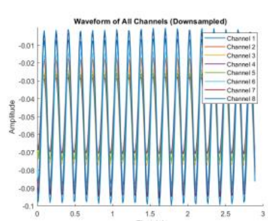
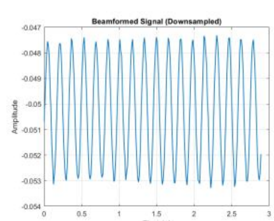
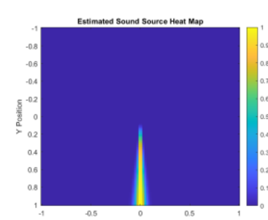
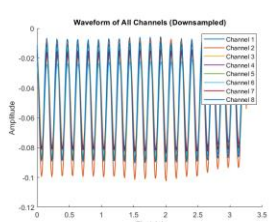
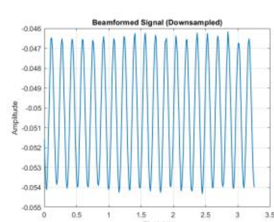
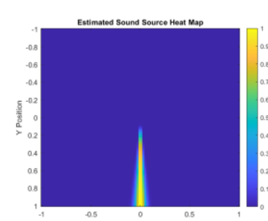
Speaker position	Waveform of recorded signal	Waveform after beamforming	Sound localization
Left			
Center			
Right			

Figure 7. Delay and sum beamforming results of 4 microphone linear array

Speaker position	Waveform of recorded signal	Waveform after beamforming	Sound localization
Left			
Center			

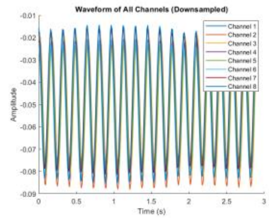
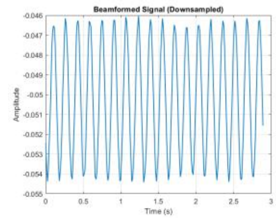
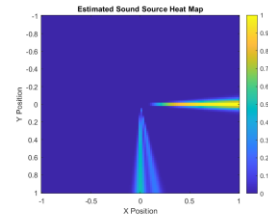
Right			

Figure 8. Delay and sum beamforming results of 8 microphone linear array

The tables show the original recorded waveform and the waveform after implementing the beamforming algorithm as well as the estimated location of the sound. Figure 7 shows the results for testing with a 4 microphone linear array while Figure 8 shows the results achieved while testing with an 8 microphone linear array. The waveforms are given with a down sampled visualization to have a better view of the results. It can be noted that in both figures a reduction can be seen in the amplitude between the original signal and the beamformed one. This reduction is due to the localization and focus of the directionality of the beam which allows for attenuation of outside reverberations on the sound. The algorithm for source localization shows errors as can be seen when the speaker is placed on the left the algorithm give the source from the central position. It can be seen that the amplitude of the signal is higher when the speaker is in the center position compared to the left and right positions, this is due to the directivity pattern of the microphone array. This is the case in both 4 mic and 8 mic linear arrays.

6. CONCLUSION

Within this work the design of multiple variations of microphone arrays were conducted using MATLAB’s sensor array analyzer and the results were presented. Varying the number of microphones, the spacing between them as well as their configurations, it can be concluded that the higher number of microphones leads to greater attenuation of sound coming from the sides of the array. Additionally it can be concluded that the main parameters for a successful beamforming array is the number of microphones and the distance between them. Increasing the number of microphones provides a narrower beam meaning better focus on the desired sound source which leads to improved accuracy, enhanced noise suppression and better estimation of the sound source. Having wider spacing between the microphones allows for capturing lower frequencies while narrower spacing has better high frequency performance.

Working from the designs and simulations a proof of concept for a low budget MEMS microphone array was produced and tested for the use in sound level detection and sound source localization. Multiple configurations of the array were made and the sensor was tested on multiple frequencies and on multiple locations for the sound source. The array shows possible implementation for sound recording and sound source localization. The positions for mounting the microphones were made by hand and the connections for the pins were made with jumper cables which means that some error in the distance between the microphones is to be expected as well as some noise in the data transfer. For future research a specially made PCB will be produced ensuring a low noise connection and precise mounting for the position of the microphones. Another addition will be the use of a camera in order to achieve an acoustic image which will provide information on the conditions of the environment.

REFERENCES

1. Benesty, J., Chen, J. and Huang, Y., 2008. *Microphone array signal processing* (Vol. 1). Springer Science & Business Media.
2. Dougherty, R.P., 2008, February. What is beamforming. In *Berlin Beamforming Conference (BeBeC)* (Vol. 63, pp. 2852-2860).
3. Merino-Martínez, R., Sijtsma, P., Snellen, M., Ahlefeldt, T., Antoni, J., Bahr, C.J., Blacodon, D., Ernst, D., Finez, A., Funke, S. and Geyer, T.F., 2019. A review of acoustic imaging methods using phased microphone arrays: Part of the "Aircraft Noise Generation and Assessment" Special Issue. *CEAS Aeronautical Journal*, *10*, pp.197-230.
4. Nau, C. and Ag, D., 2016. Beamforming within the modal sound field of a vehicle interior. In *6th Berlin Beamforming Conference*.
5. Gerges, S.N., Fonseca, W.D. and Dougherty, R.P., 2009, July. State of the art beamforming software and hardware for applications. In *Proceedings of the 16th International Congress on Sound and Vibration*
6. B. Zimmermann and C. Studer, "FPGA-based real-time acoustic camera prototype," in Proceedings of 2010 IEEE International Symposium on Circuits and Systems, 2010, pp. 1419–1419
7. M. Orman, P. Rzeszucinski, and C. T. Pinto, "Low cost, hand held acoustic camera," in 2014 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), 2014, pp. 398–402.
8. Yoganathan V., 2005. *Real-time implementation of a dual microphone beamformer: a thesis presented in partial fulfilment of the requirements for the degree of Master of Engineering in Computer Systems at Massey University, Albany, New Zealand* (Doctoral dissertation, Massey University
9. Nelson PA. Source identification and location. Advanced Applications in Acoustics, Noise and Vibration, 114-167, September 2018
10. Szwoch G, Kotus J. Acoustic detector of road vehicles based on sound intensity. Sensors, 21(23), 7781, November 2021
11. Barbagli, B., Manes, G., Facchini, R. and Manes, A. Acoustic sensor network for vehicle traffic monitoring. Proceedings of the 1st international conference on advances in vehicular systems, technologies and applications, 24-29, June 2012
12. Tomić, J., Bogojević, N. and Pavlović, D.Š., Application of Soft Computing Techniques in Prediction of Road Traffic Noise Levels. 26th Telecommunications Forum (TELFOR), 1-7, November 2018
13. Grubesa, S., Stamac, J. and Suhanek, M., 2019. Acoustic camera design with different types of MEMS microphone arrays. *Am. J. Environ. Sci. Eng*, *3*, pp.88-93.
14. Tontiwattnakul, K., Hongweing, J., Trakulsatjawat, P. and Noimai, P., 2019, July. Design and build of a planar acoustic camera using digital microphones. In *2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST)* (pp. 1-4). IEEE.
15. Brandstein, M. and Ward, D. eds., 2001. *Microphone arrays: signal processing techniques and applications*. Springer Science & Business Media.
16. Pecioski, D., DOMAZETOVSKA MARKOVSKA, S., Gavriloski, V., Anachkova, M. and IGNJATOVSKA ANGJUSHEVA, A., 2024, October. Development of low-cost MEMS microphone array for sound localization in urban environments. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings* (Vol. 270, No. 7, pp. 4641-4651). Institute of Noise Control Engineering.