Original research paper
UDC 004.382.7:681.32.02
DOI 10.7251/IJEEC2401034R

Realization of adaptive beamformer on open-source hardware

Marija Ratković, Miloš Bjelić

Acoustics Laboratory, Department of Telecommunications, School of Electrical Engineering University of Belgrade, Belgrade, Serbia

E-mail address: rm@etf.rs, bjelic@etf.rs

Abstract—**Realization of an efficient system for extracting useful audio signals in the presence of interferences is an important engineering problem in the field of acoustics. This paper proposes a method for extracting sound signals from a specific direction of incidence. A microphone array consisting of eight microphones was used, coupled with space-time signal processing executed in real time on hardware. The microphone signals are filtered by an adaptive beamformer, whose coefficients are optimized in real time on open-source hardware using the Least Mean Square algorithm. This paper describes the characteristics and limitations of the used microphone array. Testing the system through simulation indicates that extracting narrowband and wideband signals using the presented adaptive beamforming method is theoretically possible. This paper aims to experimentally evaluate the algorithm under real-world conditions, as well as identify the limitations of extracting wideband signals. Adaptation of the filter coefficients and filtering the signals from the microphones was implemented on Bela hardware, specialized for processing audio signals. This experiment highlights the limitations that arise from extracting a wideband signal using this type of hardware in real conditions. By comparing the useful and filtered signals in both time and frequency domains, the quality of the filtering was analyzed. Our experiments suggest that the current hardware specifications are a limiting factor for successful wideband signal filtering in real conditions, as they limit the maximum possible order of the adaptive filter, which proved to be insufficient.**

Keywords- adaptive beamforming, Bela hardware, microphone array, real time processing, signal extraction

I. INTRODUCTION

Microphone arrays are systems consisting of multiple microphones designed to amplify signals from a specific direction and have greater directivity than traditional microphones [1]. Their operating principle is based on the summation of signals from all microphones, which results in the amplification of the signal that reaches the microphone array along its axis. Signals coming from other directions can also be amplified through additional signal processing. One practical application of such systems is in conference audio setups, where it is necessary to increase directivity towards the active speaker and suppress signals from all other directions [2]. Another application is the localization of sound sources using an acoustic camera, which can distinguish dominant sound sources in a visual image [3-4].

The main characteristics of a microphone array are the shape of its directivity diagram and the direction of the axis of maximum radiation, around which the main lobe on the directivity diagram is formed [5]. These properties depend on the spatial distribution of the microphones and the processing of the collected signals [6]. The shape of the directional diagram can be altered by multiplying the signal with complex coefficients, known as spatial filter coefficients. This process is known as spatial filtering or beamforming [5-6]. Algorithms for optimizing spatial filter coefficients originate from the theory of antenna arrays (array processing). Applying these algorithms to acoustic signals is complex because acoustic signals are wideband, while the frequency range of radio signals used in antenna arrays is narrowband.

In this paper, we propose an adaptive beamformer for extracting useful acoustic signals by forming the main lobe of the directivity diagram in the direction of the desired signal. Spatial filter coefficients are optimized for each sample of the signal that reaches the microphone using the Least Mean Square (LMS) algorithm [7-8]. To optimize the coefficients for extracting the desired acoustic signal through beamforming, the algorithm must be improved as demonstrated in [9]. Unlike a narrowband spatial filter, which uses one coefficient per microphone, a wideband filter uses multiple coefficients, each responsible for filtering a specific frequency range.

The aim of this paper is to create a system for extracting a useful signal using an adaptive beamformer in real time. This system was designed by implementing the LMS algorithm on open-source hardware and was tested through experiments in real conditions. A linear microphone array consisting of eight equidistant microphones was used. Two sources were employed: one as the sound source of the useful signal and the other as the source of interference. The most complex part of this study involved designing an experiment in which the coefficients of the wideband signal extraction filter are optimized in real time. The filter coefficients are dynamically

This paper is a revised and expanded version of the paper presented at the XXIII International Symposium INFOTEH-JAHORINA 2024.

adjusted as the signal is being processed, which ensures that the filtering adapts to changes in the signal and the acoustic environment. Such a system could be effortlessly used in situations where there is a desired sound source that needs to be separated from noise originating from other directions. The first part of this paper involves testing the algorithm through simulation with narrowband and wideband musical signals as useful signals. The second part involves an experiment using hardware equipment. Real-time filtering experiment of narrowband signals indicates that successful extraction of this type of signal is possible. When the useful signal is a musical sequence, spatial filtering becomes more complex due to the wide frequency range and dynamic nature of the music. The paper examines the limitations of implementing such a system on open hardware, particularly focusing on the challenges posed by real-time adaptive filtering.

The paper is organized into four sections. Section two presents the theoretical basis of microphone arrays and adaptive beamforming. Section three discusses the results obtained from simulations and experimental implementations using Bela hardware [10]. The conclusions of the research are presented in section four.

II. THEORETICAL BASIS

A. *Microphone arrays*

Microphone arrays consist of multiple microphones designed to capture sound signals from various directions. The directionality of a microphone array implies a spatial sensitivity pattern where sound energy is unequally collected from different directions [1]. These arrays are specialized constructions that enhance directivity, thereby isolating signals from desired directions more effectively than standard microphones. If a signal doesn't align with the system's axis, phase differences occur due to delays between microphone signals, resulting in a weaker resulting signal compared to on-axis signals. Signal summation occurs in the electrical domain, allowing for directional sensitivity adjustment via software processing. This adjustment involves introducing selective delays to individual microphone signals to ensure phase alignment. Microphone arrays come in three main geometries depending on the spatial arrangement of microphones: linear, planar, and spatial arrays.

Figure 1. Line microphone array.

A line microphone array used for all experiments, consisting of eight omnidirectional microphones placed at equidistant intervals is shown in Fig. 1 [11]. With line microphone arrays, it is not possible to determine the position of sound source in the vertical plane - the direction can only be determined in the horizontal plane in which the array is located. The directivity diagram of microphone arrays, known as a beampattern [4], depends on the frequency, the spatial arrangement of the microphones, and the coefficients of the spatial filter, w. The directional diagram is defined by the following expression (1).

$$B(w,\phi,\theta)=\sum_{l=0}^{L-1} w(l)\cdot e^{-i\frac{2\pi}{\lambda}a^T p(l)} \tag{1}$$

The vector a represents the unit vector of the direction of arrival of the signal, defined by the azimuth ϕ and elevation θ. The angle θ ranges from 0° to 180°, while ϕ ranges from 0° to 360°. The spatial distribution of the microphones is represented by the variable p. The frequency appears in the expression through the wavelength λ.

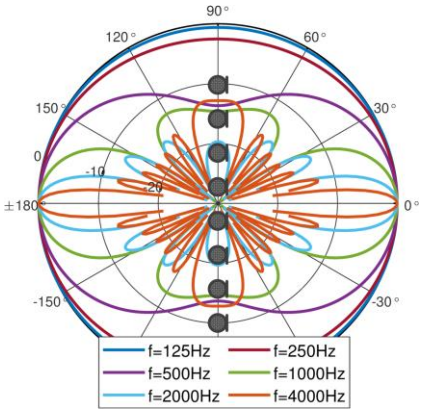


Figure 2. Directivity pattern of the microphone array.

Fig. 2 shows the directivity pattern of the microphone array across different frequency bands, with an elevation of 90°, while the microphone array is oriented in the azimuth as depicted. At low frequencies, such as 125 Hz and 250 Hz, no directivity is present. As the frequency increases, the main lobe narrows, but side lobes appear as well. At higher frequencies, the directivity is greater, resulting in a narrower main lobe; however, the side lobes become more pronounced, as illustrated for the frequency of 4000 Hz.

B. *Adaptive beamforming*

Spatial signal filtering, also known as beamforming, is utilized to extract signals from specific predetermined directions. Processing the signals from the microphone array allows for the extraction of the acoustic signal from the desired direction by effectively directing the main lobe of the directional diagram toward that direction. However, a key limitation of beamforming is its applicability within a restricted frequency range [3]. At lower frequencies, the distance between microphones must be relatively large compare to wavelength to achieve directional sensitivity. Conversely, aliasing issues arise at higher frequencies. The adaptive algorithm employed for directing the microphone array relies on training adaptive filter coefficients based on a predefined desired signal. During each iteration of the optimization algorithm, the filtered signal obtained from the microphone is compared with the desired signal using coefficients from the previous iteration. Subsequently, new coefficients for the adaptive filter are defined based on this comparison. In scenarios where the desired signal coincides with interfering signals, employing an optimized filter to process the

microphone signal redirects the directional diagram of the microphone array towards the direction of the desired signal.

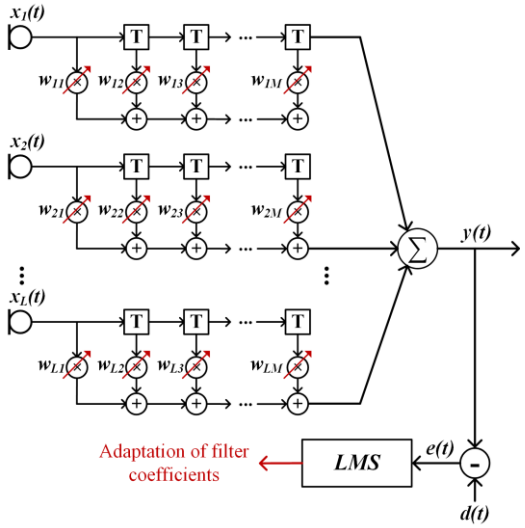


Figure 3. Block diagram of spatial filter coefficients.

In this paper, the LMS algorithm [12] was used as an optimization method. Audio signals on the microphone array are wideband, meaning the signal delay between microphones is greater than the reciprocal of the spectrum width of the observed signal. Beamforming of wideband signals is more complex than that of narrowband signals, where the signal from each microphone is simply multiplied by a single coefficient of the spatial filter. Wideband beamforming involves delaying the signal from each microphone and multiplying it by multiple adaptive filter coefficients. Fig. 3 shows the principal block diagram of the LMS adaptive spatial filter. The number of filter coefficients is $L \times M$, where L is the number of microphones and M is the number of delay lines or coefficients in one branch of the filter. A larger filter order provides more accurate filtering because narrower frequency bands are filtered by individual coefficients. The coefficients of the spatial filter, $w_{i,j}$, are adapted in each iteration of the algorithm. They are adjusted so that after filtering, the resulting signal y closely matches the desired signal d.

The beamformer coefficients are adapted in each iteration of the algorithm, and they are adjusted so that the resulting filtered signal closely matches the desired signal. In the block diagram shown in Fig. 3, $d(t)$ represents the sample of the desired signal, which is a sample of the training sequence used to optimize the filter coefficients in each iteration. $y(t)$ is the signal sample obtained by filtering the microphone signal with the adaptive spatial filter, which is calculated as shown in the following expression (2). $x(i, t-j+1)$ is the signal from the i-th microphone delayed by $j-1$ samples. This signal is multiplied by the corresponding filter coefficient $w(i,j)$.

$$y(t) = \sum_{i=1}^{L} \sum_{j=1}^{M} x(i, t-j+1) \cdot w(i, j) \qquad (2)$$

By summing the signals from all microphones, each delayed according to the filter order, $y(t)$ is obtained. The difference between this signal and $d(t)$ represents the error $e(t)$, which is a crucial input parameter for calculating the filter coefficients for the next iteration of the algorithm, $w(t+1)$. The spatial filter coefficients are optimized in each iteration to form a directional diagram towards the desired signal.

$$w(t+1) = w(t) - 2 \cdot \mu \cdot x(t+1) \cdot e(t) \qquad (3)$$

Expression (3) represents how the coefficients of the adaptive filter are calculated for the next iteration. The parameter μ is a constant that defines the convergence speed of the algorithm, with its value ranging between 0 and 1. If the adaptation step μ is too small, many iterations would be needed to reach the optimal filter coefficients. Conversely, if it is too large, the optimal filter coefficients may be skipped, preventing the algorithm from finding the optimal solution. The coefficients are optimized in this manner until the algorithm reaches a predetermined maximum number of iterations. The stopping criterion was empirically chosen at the point where the error, i.e., the difference between the filtered and desired signals, stops improving.

III. METHODOLOGY

This study is divided into two parts. The first part involves evaluating the method for extracting sound from the desired direction through computer simulation. In this phase, the characteristics of the microphone array are examined, the algorithm is tested on different signals, and the optimal results to be pursued during the hardware implementation are defined. The second part involves implementing the algorithm on open hardware and conducting experiments with the hardware equipment.

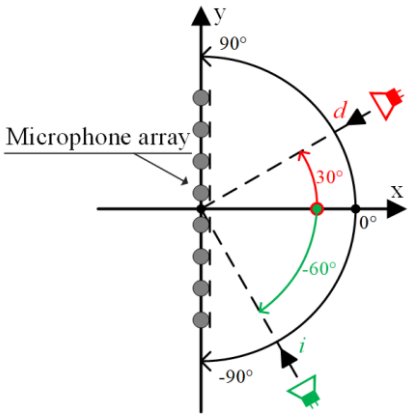


Figure 4. Positions of the sound sources and the microphone array.

A. Simulation

Assessing the system's capabilities is realized by designing a straightforward experiment scenario with two sound sources: one source produces the desired signal that needs to be extracted, and the other produces an interfering signal that should be suppressed as much as possible. The signal sources are located in the horizontal plane of the microphone array, with an elevation of 90°. The azimuth of the sound sources can range from -90° to 90°. The coordinate system that describes the positions of the sound sources and the receiver is shown in Fig. 4. In this figure, an example of the sound source positions is graphically depicted: the source of interference is shown in

green, arriving from an angle of -60°, and the desired signal source is shown in red, positioned at an azimuth of 30°.

The first part of the simulation involves simulating signals from eight microphones. The input parameters for this simulation function are the useful and interfering signals of the same length, as well as the spatial positions of the sources of these signals. The function generates eight signals, each representing the output from one microphone for the given positions of the sound sources. To examine the system's behavior at specific frequencies, a narrowband useful signal was selected in some examples. After analyzing the system at these frequencies, the algorithm was tested with wideband acoustic signals. The next part of the simulation involves selecting the order of the adaptive spatial filter, determining the adaptation step size for the LMS algorithm, and specifying the number of iterations required to extract the useful signal. The hardware implementation requires the adaptive filter order to be as small as possible due to limited execution time. Therefore, it is crucial to define the smallest filter order that still provides optimal results for a given position of the sound sources.

B. Experimental setup

The second part of this work involves the hardware implementation of the algorithm for directing the microphone array. Both narrowband and wideband useful signals were tested. For filtering the signal with an FIR filter and calculating the adaptive filter coefficients in each iteration, it is necessary to store a large amount of information, making execution time a critical factor in hardware implementation [13]. To address this, circular buffers are used to store the N previous samples, which are multiplied during filtering, where N represents the filter order. The experimental setup is shown in Fig. 5. JBL model LSR6325P-1 speakers [14] were used as sources of useful and interfering sounds. Like the simulation, the useful signal extraction system was evaluated with narrowband and wideband useful signals, while the interference was white Gaussian noise. The ratio of useful signal power to interference power was set at -10 dB. The microphone array was mounted on a mechanical holder and powered by a preamplifier, which connected the microphones to real-time data processing hardware. The input signals were read and filtered, and based on this, the filter coefficients were adapted for the next set of input signals.

Figure 5. Experimental setup.

C. Bela processor and hardware implementation

Bela hardware was used during the experiment for real time signal acquisition, spatial filtering, and coefficient adaptation. Real-time signal processing means that there is no pre-defined signal; instead, the signal is read incrementally from the input pins or a file. This type of processing involves block-based signal processing, during which Bela reads the signal in blocks. The block size determines the maximum time within which the samples in one block must be processed. Bela is open-source hardware specializing in audio signals and sensors [10][15]. Its main advantage is low latency, which enables real-time audio signal processing, sound detection, and playback. Bela consists of both hardware and software components. The hardware features eight analog inputs and outputs, as well as 16 digital inputs and outputs, which can be configured depending on the application. However, a limitation of Bela is that it has only two audio inputs and outputs. When it is necessary to detect sound from multiple microphones, as in this paper, additional equipment like an audio expander is required. By using an audio expander and connecting pins for analog and audio inputs, it is possible to read audio signals from eight analog inputs. These eight analog inputs are used to drive the microphone array directional pattern. The audio and digital sampling frequency is 44.1 kHz. The analog sampling frequency depends on the number of analog channels used: it is 88.2 kHz when two channels are used, 44.1 kHz for four channels, and 22.05 kHz for eight analog channels. Due to the use of an audio expander and eight analog inputs, the sampling frequency was chosen to be 22.05 kHz.

At the beginning of the program, it is necessary to select the order of the adaptive spatial filter. If the filter order is too large, the processing of one block cannot be completed within the allowed time. Therefore, the largest block size, k, is used, which is 4096. This means that the number of samples processed by each of the 8 channels is 2048. The duration of one block is calculated using the following expression (4).

$$\frac{k}{f_s} = \frac{4096}{44.1}\frac{1}{kHz} = 92.88ms \tag{4}$$

The largest block size is used because the execution time of the program is critical due to the complexity of the algorithm for extracting signals from the desired direction. The runtime issue arises with the filtered signal, as shown in Fig. 3, specifically when calculating new coefficients, which requires implementing two loops that increase the program's complexity. To choose the appropriate filter order before implementing the system in real time, the algorithm is tested in a simulation on a hardware platform for selected signals and source positions to determine the smallest possible filter order. However, this analysis of the filter order should not be completely relied upon, because the simulation is conducted under controlled conditions without additional interfering signals.

IV. RESULTS

A. Results of the simulation

The first part of the simulation involves assessing the limitations of the employed microphone array. By analyzing the position and magnitude of the side lobes, we can delineate the constraints of the microphone array. Fig. 6 illustrates the attenuation of the main sidelobe relative to the signal frequency. Notably, the suppression of the side lobes occurs within the frequency range of 650 Hz to 4850 Hz, while no suppression is observed at other frequencies. It's worth noting

that the wavelength for the frequency range of 4850 Hz is approximately 7 cm, corresponding to the distance between two adjacent microphones. In frequency bands where the wavelength is less than 7 cm, significant side lobes emerge at comparable levels to the main lobe. Conversely, at lower frequencies, below 650 Hz, no side lobes are present, indicating a lack of directionality due to the substantially greater wavelength compared to the microphone spacing.

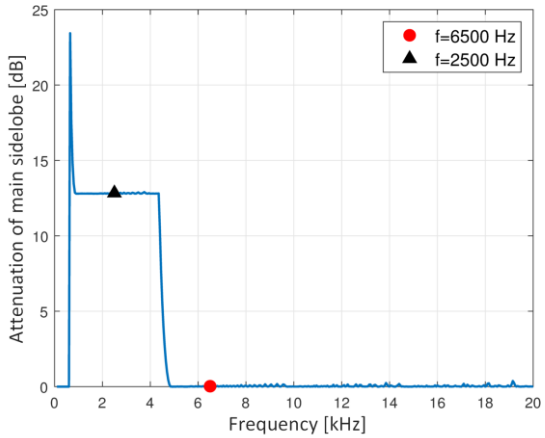


Figure 6. Attenuation of main sidelobe in the directivity pattern.

The first useful signal employed in the simulation is a sinusoidal signal at a frequency of 2500 Hz. This frequency falls within the frequency range where side lobes are effectively suppressed. The signal sources are positioned at an elevation of 90°, with the useful signal located at an azimuth of 20° and the interfering signal at 35°. These positions were chosen to closely test the algorithm's performance under challenging conditions, with the sources in close proximity. The algorithm was evaluated using filter orders of 64 and 128. Similar results were obtained for both filter order values, indicating consistent performance across different levels of complexity.

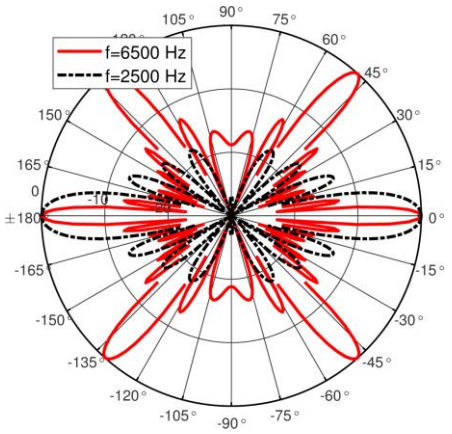


Figure 7. Directivity diagrams comparing frequencies with and without sidelobe suppression .

Fig. 7 displays the directivity diagrams for two frequencies. The black dashed line represents the directivity diagram at 2500 Hz, revealing a suppressed main sidelobe at 20°, attenuated by approximately 13 dB. In contrast, the red line depicts the directivity diagram at 6500 Hz, where the sidelobe at 47° is observed to be at the same level as the main lobe. This comparison illustrates that at higher frequencies, the main lobe becomes narrower, while the side lobes become more prominent. This study evaluated the directionality of the microphone array using an adaptive spatial filter for frequencies within the range of sidelobe suppression, as well as frequencies where the sidelobe intensity matches that of the main lobe.

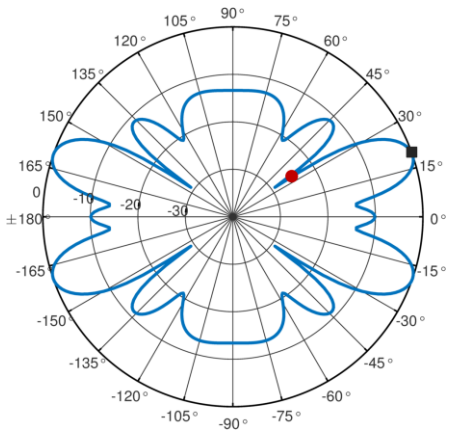


Figure 8. Directivity diagram for frequency 2500 Hz, filter order 64.

Fig. 8 displays the directivity diagram for a filter order of 64 at a frequency of 2500 Hz, corresponding to the frequency of the desired signal for extraction. The main lobe on the directional diagram, marked by the black square, aligns with the position of the source emitting the useful signal, situated at an azimuth of 20°. Additionally, the red dot indicates the level of suppression achieved for the interfering signal, located at an azimuth of 35°. The attenuation of the interfering signal is 24 dB.

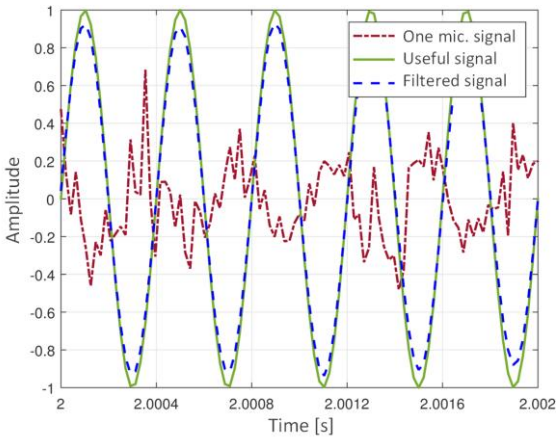


Figure 9. Magnified view of the waveforms of the useful sinusoidal signal, the filtered signal, and the signal from a single microphone.

After presenting the directivity diagrams, which are computed directly from the coefficients of the spatial filter, the next step is to apply these filter coefficients to the signals from all microphones. In Fig. 9, time waveforms are displayed for an illustrative example featuring a useful signal with a frequency of 2500 Hz. The red line represents the signal from one microphone, while signals from the other seven microphones exhibit similar shapes. The green solid line corresponds to the desired useful signal we aim to extract through spatial filtering. The blue dashed line depicts the filtered signal obtained while applying the spatial filter to signals from all microphones. This results in a periodic signal that closely resembles the desired

useful signal. The successful separation of the sinusoidal signal with a frequency of 2500 Hz from interfering signals can be observed from Fig. 9.

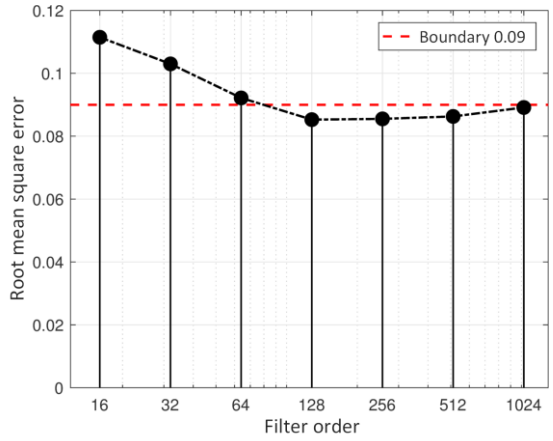


Figure 10. Dependence of the mean square error on the order of the filter.

Audio signals, such as speech and music, typically exhibit wideband characteristics, necessitating thorough testing of the algorithm's performance on such signals. Evaluating the algorithm's efficacy with these useful signals is essential to assess its viability in scenarios where separating the desired signal from interference originating from different directions is critical. In our tests, we examined the algorithm's performance using a musical sequence as the useful signal and white Gaussian noise as interference. For wideband signals, the filter order must be higher than for narrowband signals to achieve a signal that closely resembles the desired output after filtering. While filter order is not a limiting factor in simulation-based testing, with hardware implementation it is necessary to minimize the filter order. We estimated the optimal filter order by calculating the root mean square error (RMSE) between the filtered signal and the desired signal. Testing various source positions revealed that the RMSE reaches a constant value below a certain threshold (0.09 in our case). Fig. 10 illustrates the relationship between filter order and RMSE for source positions at 30° azimuth for the useful signal and -15° for the noise. It's evident that for a filter order of 128, the error remains below the defined limit. Therefore, we determined 128 as the optimal filter order for this scenario.

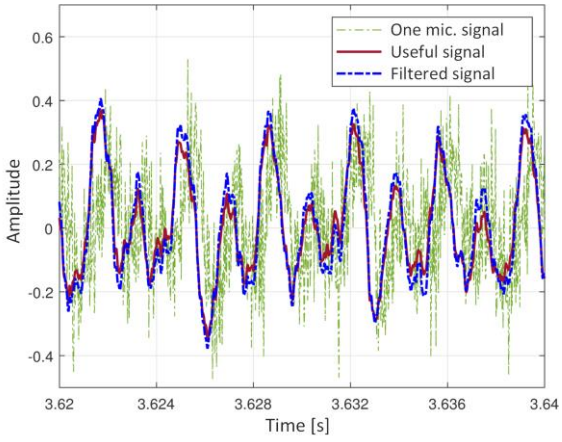
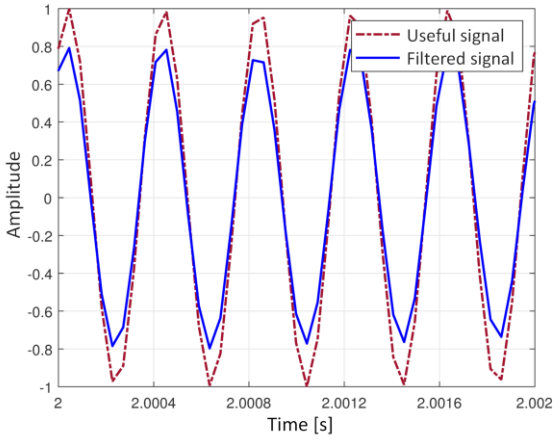
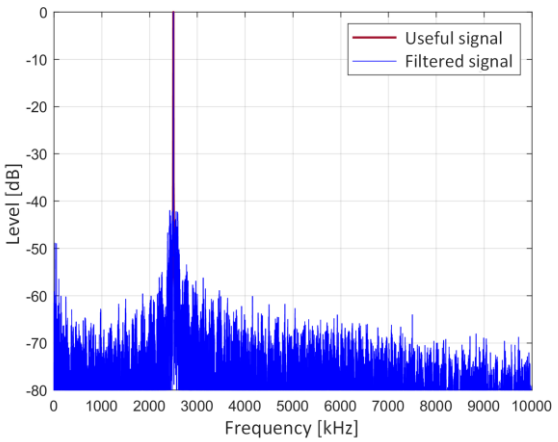


Figure 11. Magnified view of the waveforms of the useful music signal, the filtered signal, and the signal from a single microphone.

Fig. 11 displays a part of the wavelength of various signals, including the filtered signal, the desired music sequence aimed to be extracted, and the signal from one microphone. The blue dashed line represents the filtered signal obtained by processing signals from each microphone with appropriate filters, summing them, and then normalizing. The red solid line depicts the music sequence that is targeted for extraction from the aggregated signal. Upon comparison of these two curves, it becomes evident that the filtered signal closely follows the desired music sequence. This indicates successful extraction of the desired signal from the aggregated input, which is represented by the signal shown in green. The resemblance between the filtered signal and the desired music sequence underscores the effectiveness of the extraction process.



a)



b)

Figure 12. Comparison of the useful and filtered sine waves: (a) time-domain waveform, and (b) frequency spectrum.

B. The results of the experiments

The first signal on which the algorithm was tested was a sine wave with a frequency of 2500 Hz. The chosen azimuth angles are 20° for the useful signal and 35° for the white Gaussian noise. The experiment for these signals was conducted twice: first with the adaptive filter order set to 64, and then with the order set to 128. The suppression of the signal from the direction of the interference was 25 dB and 15 dB, with respect to filter order. Fig. 12a shows a time-domain slice of the useful signal (red dashed line) and the filtered signal (blue solid line) when the filter order is 64. It should be noted that the level of the filtered signal is different from the level of the useful signal, which can be attributed to

the error that occurs during signal normalization to the highest value, as certain peaks occur that are of a higher level than the rest of the signal. Fig. 12b shows the spectrum of the useful signal and the filtered signal for filter order 64. The spectrum of the sine signal exhibits a maximum at 2500 Hz, which is separated from the noise spectrum at levels ranging from −70 to −40 dB.

The real-time adaptation of the beamformer was also evaluated with a sine wave at a higher frequency of 6500 Hz. For this test, the source azimuth angles were set to 25° for the useful signal and 75° for the interfering signal. The power ratio between the useful and interfering signals was −3 dB. Fig. 13 shows the directivity diagram for this example, where the order of the adaptive filter is 128. The diagram indicates that the signal arriving at the microphone array from the direction of the interfering signal source is suppressed by approximately 20 dB. Additionally, the maximum sensitivity is formed at 25°, corresponding to the direction of the useful signal source. This example demonstrates that the algorithm effectively separates a narrowband signal, even when its frequency falls within the range where side lobes are pronounced, as shown in Fig. 6.

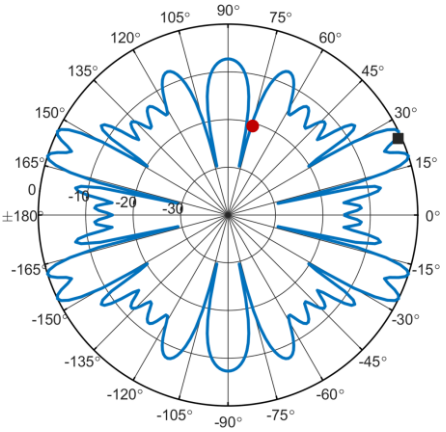


Figure 13. Directivity diagram for frequency 6500 Hz (experiment).

The second part of testing adaptive spatial filtering in real time involves applying the method to wideband signals. As in the simulation, the useful signal is a musical sequence, while the interfering signal is white Gaussian noise. The algorithm was tested for different useful signal-to-interference ratios. The useful signal source is positioned at 30° azimuth, and the noise source is at -15° azimuth. The strength of the useful signal and interference was varied throughout the experiment to determine if increasing the strength of the useful signal resulted in better filtering. Fig. 13 shows the time domain representation of the filtered and useful signals. In this scenario, the filter order is 200, and the power ratio of the useful signal to interference is −4 dB. This filter order represents the maximum achievable when implementing this algorithm in real time.

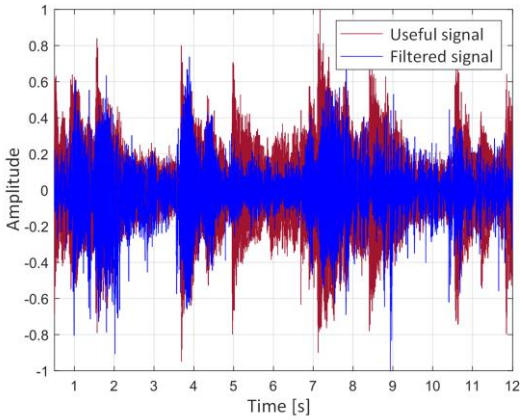
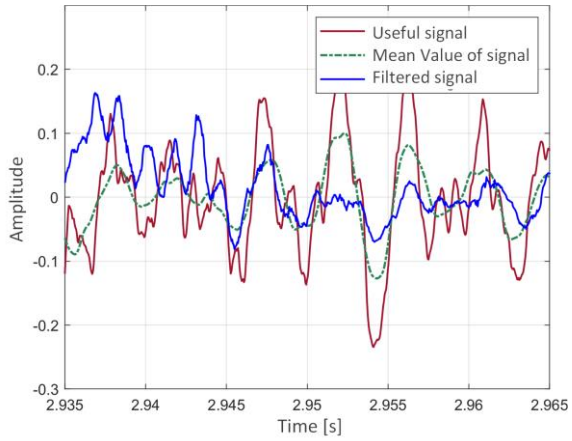
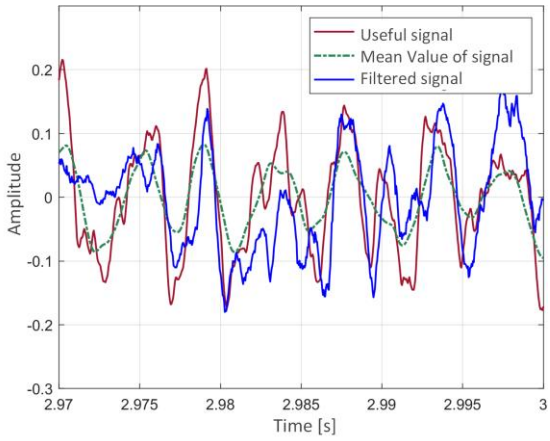


Figure 14. Comparison of the time-domain waveform useful and filtered music signal.

The filtered signal closely follows the amplitude changes expressed in the useful signal. Across different power ratios of the useful and interfering signals, similar results were obtained. In some parts, the filtered signal closely matches the useful signal, while in other parts, discrepancies were more pronounced. To enhance the filtering performance, it would be necessary to significantly increase the filter order. However, due to the limitations of the Bela hardware, increasing the filter order beyond a certain point is not feasible.



a)



b)

Figure 15. Magnified waveform of useful music signal and filtered signal a) bad filtering, b) good filtering.

Fig. 15 shows an enlarged view of sections of the waveforms. Fig. 15a) displays a section where poor filtering is observed, meaning the filtered signal does not follow the shape of the useful signal. In contrast, Fig. 15b) shows a section where the filtered signal closely follows the shape of the useful signal. The green dashed line represents the mean value of the useful signal, calculated using a moving average filter with a window size of 50 samples. By comparing the green line (mean value of the useful signal) and the blue line (filtered signal) in Fig. 15b), it is evident that the filtered signal tracks the temporal changes in the useful signal. Although Fig. 15.a) highlights a section with worse filtering, it can still be observed that the filtered signal approximates the mean value to a certain extent.

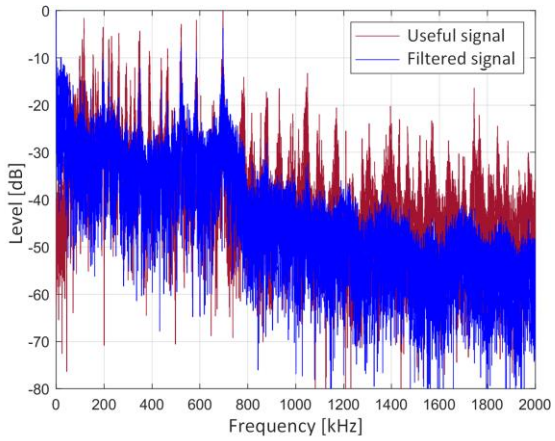


Figure 16. Comparison of frequency spectrum of the useful and filtered music signal.

Fig. 16 illustrates the frequency spectrum of the musical sequence used as the original (useful) signal alongside the spectrum of the filtered signal. This musical sequence exhibits distinct components at lower frequencies, enabling the extraction of the music sequence to a reasonable extent through filtering in the desired direction. Noticeably, the spectrum of the filtered signal up to 800 Hz closely mirrors the shape of the useful signal's spectrum. This alignment in the frequency domain results in a corresponding similarity in the time domain, where the filtered signal follows the shape of the useful signal. However, at higher frequencies, the level of the filtered signal is consistently 10 to 20 dB lower than that of the useful signal. This disparity leads to a less precise match between the two signals in the time domain.

Through this real-world testing example, it was demonstrated that the implementation of the spatial adaptive filter on Bela hardware is not sufficiently effective. The constraints imposed by the maximum filter order are a limiting factor. The number of coefficients directly impacts the width of the frequency band filtered by each coefficient, which proved to be insufficient. In simulations, the wideband signal was successfully isolated using the employed filter orders. However, in real conditions, ambient noise, along with reflections of both the useful signal and interference, complicates the scenario. These additional components, originating from all directions, necessitate larger spatial adaptive filter orders.

To address the shortcomings observed when extracting wideband signals in real conditions, several strategies can be explored in future research. The initial step involves testing the method in controlled conditions, such as an anechoic chamber, to assess the feasibility of extracting wideband signals by analyzing the results obtained. Further research should involve testing the method with various types of audio signals and analyzing their spectrum to understand better the algorithm's performance across different scenarios. Additionally, employing different hardware that supports larger filter orders could potentially enhance the system's performance. Utilizing a larger microphone array or an array with a different geometry could also lead to improved interference suppression. Developing and testing different microphone array configurations represents a promising direction for future research to overcome the current limitations and enhance the effectiveness of adaptive spatial filtering in real-time applications.

V. CONCLUSION

This paper demonstrates the implementation of a real-time system for extracting signals from a desired direction. The system utilizes a linear microphone array consisting of eight microphones and open-source Bela hardware. The extraction of the useful signal is achieved by filtering with a spatial adaptive filter. The implemented adaptive algorithm is the LMS algorithm, which adapts the filter coefficients in each iteration. The system was tested through both computer simulation and practical experimentation, including the implementation of the algorithm on hardware.

The simulation demonstrated that adaptive beamforming could theoretically suppress interference and extract a useful signal. In this scenario, the order of the adaptive filter is not a limiting factor because the coefficients are not adapted in real time. Different useful signals were used to evaluate the signal extraction system, with white Gaussian noise as the interference in all examples. It was concluded that narrowband signal extraction is possible regardless of the selected filter order. Additionally, the simulation showed that the extraction of wideband signals is successful.

The main part of this work involves the implementation of a system for extracting a desired signal on open hardware. Bela was chosen due to its affordability, digital signal processor and sampling frequency adequate for working with audio signals. The useful signals tested include narrowband signals and musical sequences. Testing the algorithm with narrowband signals in the experiment was successful. The experiment also examined the extraction of a wideband useful signal, more specifically a musical sequence. In this case, the useful signal was partially separated, with the filtered signal following the shape of the useful signal. The frequency range up to 800 Hz for both the filtered and useful signals coincided, allowing for partial separation through filtering. However, in real conditions with ambient noise and reflections, the adaptation of the coefficients was insufficiently effective. The maximum filter order that can be implemented on Bela is 200, which was inadequate for this example. Increasing the filter order beyond the hardware limit could potentially yield better results. Additionally, reducing ambient noise and reflections would improve the filtering results. It has been demonstrated that extracting a useful wideband signal with this algorithm is possible, but improvements are needed for the method to function effectively in real-world systems.

ACKNOWLEDGMENT

This work was financially supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia under contract numbers 451-03-47/2023-01/200103.

REFERENCES

[1] M. Mijić, „Audio sistemi“, Akademska misao, Belgrade, 2011.

[2] Y. Tamai, S. Kagami, H. Mizoguchi, K. Sakaya, K. Nagashima, T. Takano, „Circular Microphone Array for Meeting System, Sensors“, vol. 2, pp. 1100 - 1105, 2003.

[3] M. Eric, „Some Research Challenges of Acoustic Camera“, 19th Telecommunications Forum TELFOR, 22-24 Nov. 2011, Belgrade, Serbia, pp. 1036 – 1039.

[4] Henry E. Heffner and Rickye S. Heffner, „The Evolution of Mammalian Sound Localization“, Department of Psychology, University of Toledo, 2016.

[5] M. Erić, N. Vukmirović, „Uvod u obradu signala sa antenskih nizova“, Akademska misao, Belgrade, 2019.

[6] H. L. Van Trees, „Optimum Array Processing", John Wiley and Sons, New York, USA, 2002.

[7] S. Haykin, „Adaptive Filter Theory“, Prentice Hall, New Jersey, USA, 2002.

[8] B. Widrow, J.M. McCool, M. Larimore, C.R. Johnson, „Stationary and nonstationary learning characteristics of the LMS adaptive filter“, Proceedings of the IEEE, Vol. 64, No. 8, 1976, pp. 1151 – 1162.

[9] M. Bjelić, M. Stanojević, „Comparison of LMS Adaptive Beamforming Techniques in Microphone Arrays“, Serbian Journal of Electrical Engineering, Vol. 12, No. 1, pp. 1-16, Feb. 2015.

[10] McPherson and V. Zappi, “An environment for submillisecondlatency audio and sensor processing on BeagleBone Black,” Proc. Audio Engineering Society Convention 138, pp. 1-7, 2015.

[11] M. Bjelić, M. Stanojević „Određivanje pravca nailaska pomoću mikrofonskog niza na osnovu vremenskih kašnjenja“, 58. ETRAN konferencija, 2-5. jun 2014., Vrnjačka banja, Serbia.

[12] W. Liu, „Wideband Beamforming“, John Wiley and Sons, New York, USA, 2002.

[13] Mitra, Sanjit K. „Digital Signal Processing: A Computer Based Approach“, New York: McGraw-Hill, 1998.

[14] Datasheet and tehnical documentation, available at: https://jblpro.com/en/products/lsr6325p-1, (30.08.2023.)

[15] Bela Knowledge Base, Available at: https://learn.bela.io/ (04.09.2023.)

Marija G. Ratković was born in Belgrade, Serbia, in 1999. She received aBachelor degree in electrical engineering from the University of Belgrade, School of Electrical Engineering, Belgrade, Serbia, in 2022, and an M.Sc. degree in electrical engineering from the University of Belgrade, Belgrade, Serbia, in 2023. In 2024, she joined the Faculty of Electrical Engineering at the University of Belgrade. Currently, she is a Research Assistant at the University of Belgrade, School of Electrical Engineering. Her research interests include digital signal processing, space-time signal processing, microphone and speaker arrays, electroacoustics and room acoustics.

Miloš R. Bjelić was born in Priboj, Serbia, in 1989. He received the Bachelor degree in electrical engineering from the University of Belgrade, School of Electrical Engineering, Belgrade, Serbia, in 2012. The M.Sc. and Ph.D. degrees in electrical engineering from the University of Belgrade, Belgrade, in 2013 and 2018, respectively. In 2013, he joined the Faculty of Electrical Engineering at the University of Belgrade. Currently, he is an Assistant Professor at the University of Belgrade, School of Electrical Engineering. His research interests are in audio systems, electroacoustics, room acoustics, digital signal processing, space-time signal processing, microphone array, sound insulation, physical and technical measurement. Serbia