Direction of Sound Wave Arrival Determination Using Time-Delay Estimation and Beamforming methods

PETR DOSTÁLEK, VLADIMÍR VAŠEK, JAN DOLINAY
Department of Automation and Control Engineering
Tomas Bata University in Zlín, Faculty of Applied Informatics
Nad Stráněmi 4511, 760 05 Zlín
CZECH REPUBLIC
{dostalek; vasek; dolinay}@fai.utb.cz http://www.fai.utb.cz

Abstract: - This paper deals with direction of sound wave arrival determination using time-delay estimation and beamforming methods with utilization of static microphone array with fixed geometry configuration. Theoretical part describes main principle of both methods and possibility of program implementation. Experimental part proposes hardware design of the sensory systems separately for time-delay estimation and beamforming evaluation methods due to different requirements on microphone array structure. Developed localization system is based on standard personal computer equipped with Advantech PCI-1716 multichannel data acquisition device.

Key-Words: - audio source localization, microphone array, direction of arrival, time-delay estimation, beamforming, signal processing.

1 Introduction

The first device designed by human for audio localization was created in 1880 by Professor Mayer. This instrument for navigation improvement in fog was called by its author Mayer's topophone. On the basis of its construction originate number of similar devices but with questionable practical usage. The biggest interest in audio location systems occurs in the period between World War 1 and World War 2. They were primarily used for detection a localization of the aircraft engine sound. Measured data about aircraft position was directly transferred to air-defense artillery which can aim at target before visual contact. Constructions and dimensions of these systems were very various but the basic concept is based on Mayer's topophone improved with next two horns oriented in vertical plane. Due to state of electronics then minimally two people were required for sound analysis originated from horn system. Since it was impossible to continuously enlarge horn dimensions for better gain achieving, static dishes and walls based on spherical reflection surface was developed. These systems were able to detect aircrafts at longer distances. After radio locator invention in 1934 audio location devices were not further developed in this area because they were completely replaced by RADAR systems with better detection and ranging properties [4].

Nowadays very dynamical development in electronics and computer science enables applying of the sound localization systems in areas where it was impossible due to technical and economical aspects several years ago. These areas include applications in security, teleconferencing, robotic systems and other else where information is coded in audio signal source position.

This paper deals with both design of the input part of the each localization system which is sensory system based on microphone array and software implementation of detection algorithms. First part describes circuit solution of the microphone preamplifiers units, automatic gain control amplifier and output anti-aliasing filter which are necessary for signal conditioning to correct voltage levels before analog-to-digital conversion process in data acquisition card. Second part of the paper describes time-delay estimation and beamforming methods for direction of arrival of the sound wave determination and its software implementation.

2 Overview of the acoustic source localization methods

2.1 Time-delay estimation

This method is very often used in microphone sensory systems due to its simple implementation in hardware and software. It is based on audio samples analysis acquired from each microphone unit in an array. Analysis takes place usually in two basic steps – time-delay determining of each data samples acquired from different microphones following by

computation of the sound source position on the basis of microphone array geometrical organization. On assumption that sound source is in much larger distance than is each sensor spacing d_S, spherical surface of the sound wave can be considered as a flat surface. It considerably simplifies evaluation of the direction of arrival of the sound wave. This assumption and resulting simplification is obvious from Fig. 1.

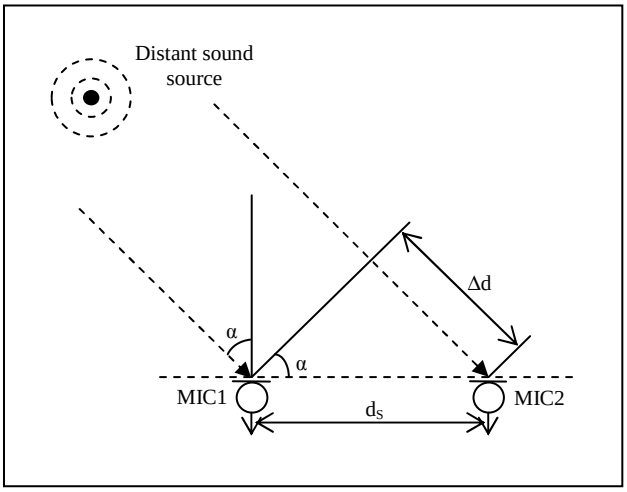

Fig. 1. Sound wave impacting microphone array.

Then sound wave arrival angle α impacting one pair of microphone units can be determined by equation (4).

$$\alpha = \arcsin\left(\frac{\Delta d}{d_S}\right) \tag{1}$$

Difference of the sound wave trajectory Δd can be computed as product of estimated time-delay t and sound wave speed v. Arrival angle can be then computed using following equation:

$$\alpha = \arcsin\left(\frac{t \cdot v}{d_S}\right) \tag{2}$$

For time-delay estimation t can be very advantageously used cross-correlation analysis which is defined by equation (3). Where x_1 is acquired signal from microphone 1, x_2 from microphone 2 and N number of acquired data samples.

$$R_{x1x2}[k] = \sum_{n=0}^{N-1} x_1^*[n]x_2[n+k] \tag{3}$$

In the Fig. 2 and Fig. 3 are for instance depicted waveforms of the audio signal acquired by two microphone units each located at different position against acoustic source. Cross-correlation result with size of 2*N-1 samples is in Fig. 4. Its maximum is achieved at shift of +225 samples corresponding to time delay of 4.69 ms at 48000Hz sampling rate.

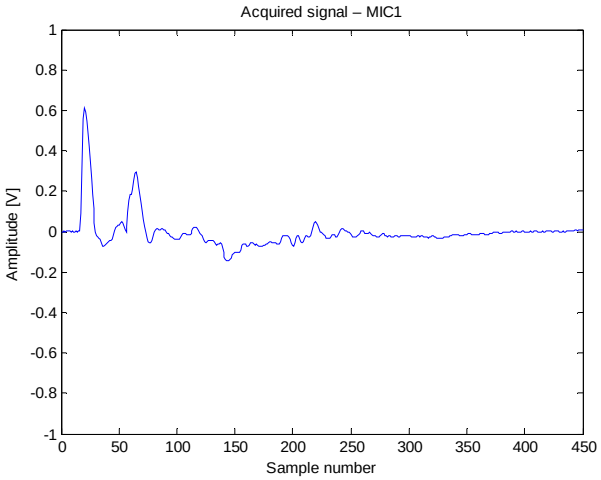

Fig. 2. Audio signal acquired by microphone unit 1.

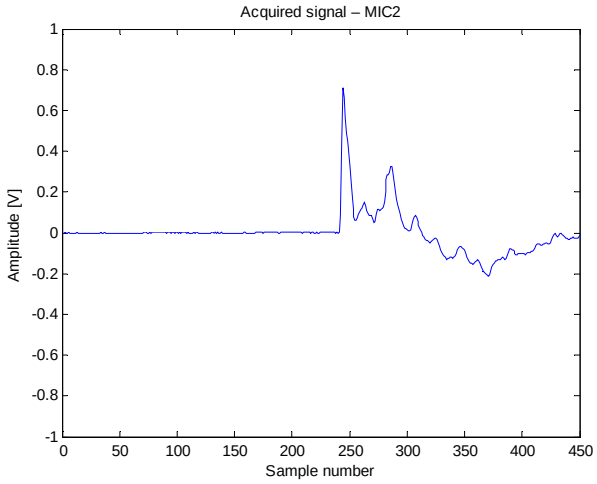

Fig. 3. Audio signal acquired by microphone unit 2.

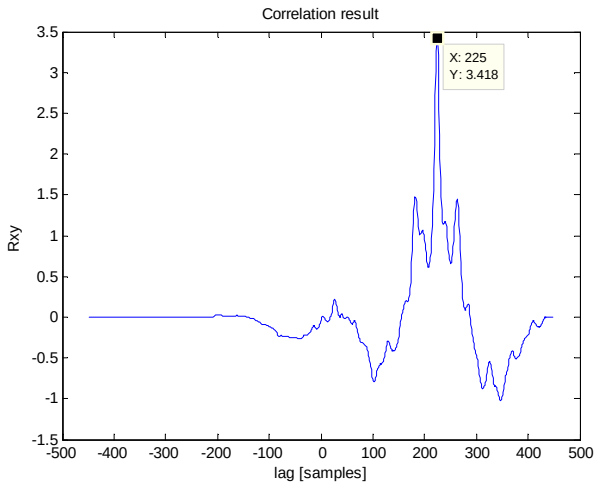

Fig. 4. Cross-correlation result example.

During the computation x_1 is reference signal and x_2 compared signal. Cross-correlation of two same shifted signals resulting in coefficients representing conformity of these two signals. Maximum value of correlation coefficient and corresponding shift of k samples indicates relative shift of these signals. Time-delay t can be computed using formula:

$$t_{12} = T_S \arg\max_k \left(R_{x1x2}[k]\right) \tag{4}$$

where T_S is sampling period. Because of cross-correlation computation (3) for large N takes a lot of computer processing time, it is not possible to use this algorithm directly for systems where is demanded fast response. This problem can be solved by computation in the frequency domain (5) in which time consuming cross-correlation is replaced by multiplication of Fourier transformed signals x_1 and x_2, where * is complex conjugate [5].

$$F\{x_1 * x_2\} = \left(F\{x_1\}\right)^* \cdot F\{x_2\} \tag{5}$$

After inverse discrete Fourier transform of equation (5) we obtain the cross-correlation result (6) in only fraction of computing time needed for computation using (3).

$$R_{x1x2} = F^{-1}\left\{\left(F\{x_1\}\right)^* \cdot F\{x_2\}\right\} \tag{6}$$

In the Table 1 are provided results of our cross-correlation benchmark program which performs correlation computation of the same data in time domain compared to frequency domain. Software was written in MS Visual C with utilization of FFTW library [2] for DFT computations using FFT algorithm. Benchmark was performed on AMD Athlon64 Dual Core processor with clock frequency of 3GHz. The improvement in execution speed is significant.

Table 1. Cross-correlation computation time.

N [samples]	**Rxy computation time [s]**	
	Standard algorithm	FFT based algorithm
12000	1.576	0.006
24000	6.232	0.015
48000	25.08	0.035
96000	100.04	0.074
192000	402.77	0.138

Because it is assumed that detection system will consists of more than two microphone units, it is necessary to evaluate direction of sound wave arrival for each different microphone pair. Number of microphone pairs in the sensory system with n units is 2-combinations (k=2) from a set with n elements, which can be computed by (7).

$$C(n,k) = \frac{n!}{k!(n-k)!} \tag{7}$$

Let suppose that we have linear uniform microphone array with 4 units as depicted in the Fig. 5, where distances between neighboring units are equal. Then direction of arrival we obtain by solving system of 6 equations (8) with one unknown variable α.

$$\begin{aligned}
\sin\alpha \cdot d_{12} &= t_{12} \cdot v \\
\sin\alpha \cdot d_{13} &= t_{13} \cdot v \\
\sin\alpha \cdot d_{14} &= t_{14} \cdot v \\
\sin\alpha \cdot d_{23} &= t_{23} \cdot v \\
\sin\alpha \cdot d_{24} &= t_{24} \cdot v \\
\sin\alpha \cdot d_{34} &= t_{34} \cdot v
\end{aligned} \tag{7}$$

System of equations (7) can be expressed in matrix form (8), where **d** is column vector of microphone pair distances (9) and **t** is column vector of corresponding time delays (10).

$$\sin\alpha \cdot \mathbf{d} = \mathbf{t} \cdot v \tag{8}$$

$$\mathbf{d} = [d_{12} \quad d_{13} \quad d_{14} \quad d_{23} \quad d_{24} \quad d_{34}]^{\mathrm{T}} \tag{9}$$

$$\mathbf{t} = [t_{12} \quad t_{13} \quad t_{14} \quad t_{23} \quad t_{24} \quad t_{34}]^{\mathrm{T}} \tag{10}$$

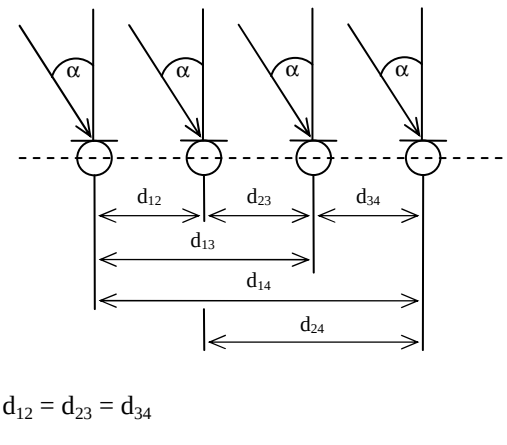

Fig. 5. Linear uniform 4 microphone array.

Because of (8) is overdetermined system, it is useful to solving them approximately by least squares method (11). Resulting direction of sound wave arrival α can be computed using (12).

$$\sin\alpha \cdot \mathbf{d}^T\mathbf{d} = \mathbf{d}^T\mathbf{t} \cdot v \tag{11}$$

$$\alpha = \arcsin\left\{\left(\mathbf{d}^T\mathbf{d}\right)^{-1}\mathbf{d}^T\mathbf{t} \cdot v\right\} \tag{12}$$

2.2 Beamforming

Principle of delay and sum beamformer operation is obvious from Fig. 6. Input signals from microphone array x[k] are delayed by time which depends on sensory system geometrical configuration and demanded angle of directional characteristic main lobe. Because of we can consider that signals from microphone units are same except time-shift, we obtain by setting of appropriate delays H to each audio channel after summing maximum level of useful signal while signal to noise ratio will be improved.

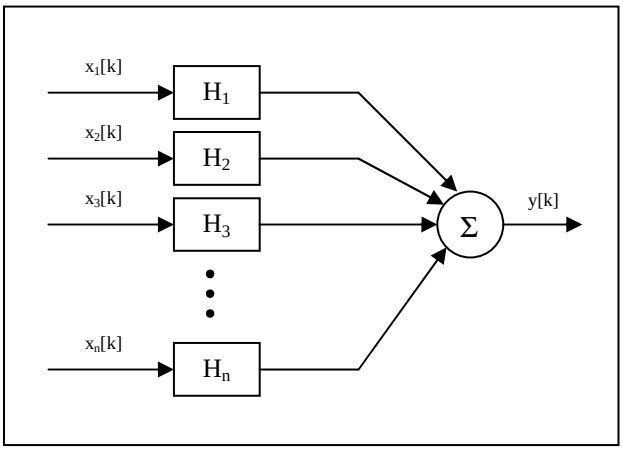

Fig.6. Delay and sum beamformer block diagram.

Beamformer output y[k] can be computed by equation (1) where $x_m[k]$ is input signal from microphone array of microphone m, H is delay introduced to signal path by beamformer and M is number of microphone units.

$$y[k] = \sum_{m=1}^{M} x_m[k - H_m] \tag{13}$$

For linear uniform microphone array and on assumption that sound source is in much larger distance than is each sensor spacing d_S, time delay in each microphone unit signal for direction of sound wave arrival α can be computed by equation (14), where k is microphone unit index and v is sound speed in air. Reference unit is microphone with index 1 which has zero time shifts for all source angles.

$$t_k = \frac{\sin\alpha \cdot d_s}{v} \cdot (k-1) \tag{14}$$

Sound source localization using delay and sum beamformer is based on computation of the beamformer output signal level for each directional characteristic angle. RMS value of the n samples length output signal and j azimuth angle is:

$$V_{RMS}[j] = \sqrt{\frac{1}{n}\sum_{k=1}^{n} y[k,j]^2}\ . \tag{15}$$

Maximum RMS value of the beamformer output and corresponding angle indicates sound source azimuth:

$$\alpha = \arg\max_j\left(V_{RMS}[j]\right) \tag{16}$$

3 Sensory system design

This chapter deals with design of the sensory systems for audio source localization using two different approaches: time-delay estimation and beamforming method. Because of there are different demands on microphone array structure and basic electrical properties its description is divided into the two subsections.

3.1 Sensory system for evaluation using time-delay estimation

Microphone sensory system for sound source azimuth evaluation using time-delay estimation was designed with a respect to easy portability, configurability and connectivity with various types of evaluation units (personal computer, notebook, DSP evaluation boards and others). These requirements best fulfill modular architecture schematic of which is depicted in Fig.7.

It consists of the following main components: six microphone units with integrated preamplifier, 3-channel automatic gain control (AGC) amplifier with 4th order low pass output anti-aliasing filter and evaluation unit – in this case of laboratory system standard personal computer equipped with multifunction data acquisition card Advantech PCI-1716. Components are connected together using shielded cables to avoid interference leakage to acquired signal. In case of need system can be upgraded up to 16 audio channels per one data acquisition card.

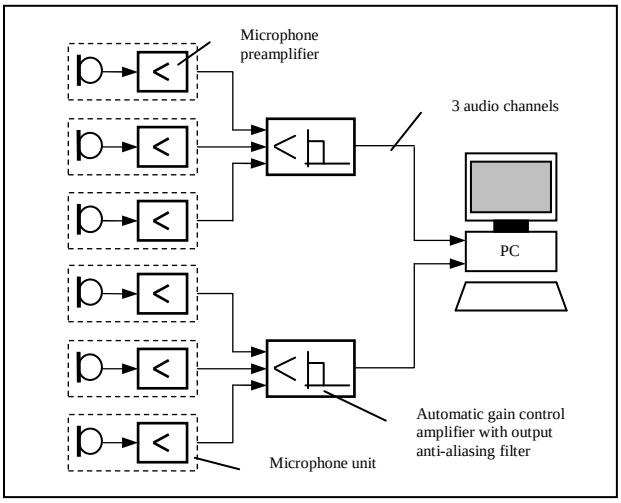

Fig. 7. Block schematic of the sensory system.

3.1.1 Microphone units design

Sound field is measured with six microphone units each equipped with three omnidirectional condenser microphone cartridges. They are installed on the small triangular base from fiberglass cooper coated board with the side size of 20 mm. This construction improves sensitivity and signal-to-noise ratio. Each microphone is connected with summing preamplifier with the gain of 23 dB which is mounted in the base of the microphone unit. Output signal level is sufficient for transmission through shielded cable to distance about 10m. Preamplifier schematic is depicted in the Fig. 8. It is build around low-noise dual operational amplifier NE5532. First stage is 3-channel inverting summing amplifier with gain of 10 followed by inverting amplifier with gain of 1.5. Input part of preamplifier includes simple power supply for condenser microphones MCE 100 which is separated from summing amplifier by coupling capacitors. Due to usage of non-symmetrical voltage source it is necessary to create virtual ground for operational amplifiers using two voltage dividers (parts R_1, R_2 and R_5, R_6). Photograph of the completed three microphone units is in the Fig. 9.

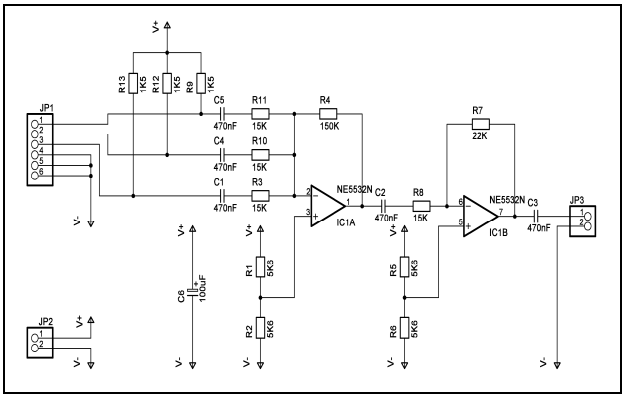

Fig. 8. Microphone preamplifier schematic.

Fig. 9. Completed microphone units.

3.1.2 Automatic gain control amplifier with anti-aliasing filter

This part of the sensory system is composed of two main components: three channel automatic gain control amplifier combined with 4th order anti-aliasing filter. The purpose of this last amplification stage is to adapt signal voltage levels to the appropriate level suitable for analog inputs of the data acquisition card. In the case that input signal has too high voltage level the gain of the amplifier is automatically lowered to prevent overdriving of the card inputs. Function of the amplifier part is obvious from Fig. 10. It uses same dual operational amplifier as microphone preamplifier. First stage is inverting amplifier with user-adjustable gain up to 50. Second stage works as comparator which output is near V+ when output voltage is higher then demanded. In this case diode D_1 is opened and capacitor C_4 is charged through resistor R_8. Rising capacitor voltage opens transistors T_1 and T_2 which drains part of the input signal to ground – gain of the circuit is automatically lowered. On the other hand low input signal level cause comparator output voltage near V-. Then diode D_1 is in reversed polarity and capacitor C_4 is discharged through R_7 to ground – transistors T_1 and T_2 are closing and gain of the circuit is reverting back to its nominal value.

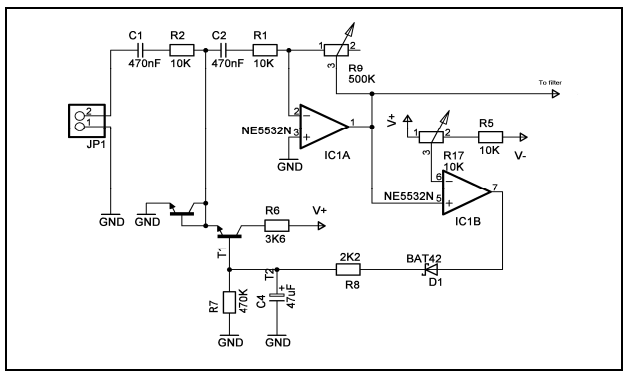

Fig. 10. AGC amplifier – amplifier part.

Amplifier stage is followed by 4th order active low-pass filter with Sallen–Key topology implemented by operational amplifiers (Fig.11) with filter parts designed using procedure published in [3].

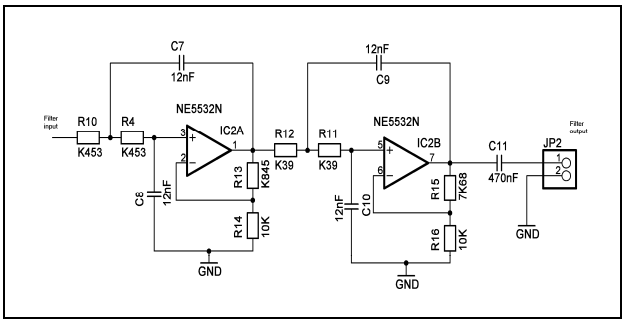

Fig. 11. AGC amplifier – low-pass filter part.

Computation is based on transfer function of Sallen-Key 2nd order low-pass filter (17) where coefficients a_1 and b_1 are equal to (18) and (19).

$$A(s)=\frac{A_0}{1+a_1 s+b_1 s^2} \tag{17}$$

$$a_1=\omega_c\left[C_1\left(R_1+R_2\right)+\left(1-A_0\right)R_1C_2\right] \tag{18}$$

$$b_1=\omega_c{}^2 R_1R_2C_1C_2 \tag{19}$$

Final transfer function of the 2nd order low-pass Sallen-Key filter is:

$$A(s)=\frac{A_0}{1+\omega_c\left[C_1\left(R_1+R_2\right)+\left(1-A_0\right)R_1C_2\right]s+\omega_c{}^2R_1R_2C_1C_2s^2} \tag{20}$$

where ω_c is a cutoff angular frequency, A_0 is gain of the filter in the passband and a_1 and b_1 are filter coefficients determining its properties. After the formulation of R_1 from equation (19) and constituting to (18) we obtain quadratic equation:

$$R_2{}^2C_1{}^2C_2\omega_c{}^2-a_1R_2C_1C_2\omega_c+b_1\left(C_1+C_2-A_0C_2\right)=0 \tag{21}$$

Its solution is equation for computation of R_2 part value (22), R_1 part value can be computed by (23).

$$R_2=\frac{a_1C_1C_2\omega_c+\sqrt{\left(-a_1C_1C_2\omega_c\right)^2-4C_1{}^2C_2\omega_c{}^2b_1\left(C_1+C_2-A_0C_2\right)}}{2C_1{}^2C_2\omega_c{}^2} \tag{22}$$

$$R_1=\frac{b_1}{R_2C_1C_2\omega_c{}^2} \tag{23}$$

In order to obtain non negative value under square root in (22), capacitors values must fulfill (24).

$$C_2\geq C_1\frac{4b_1A_0+a_1{}^2A_0-a_1{}^2}{a_1{}^2A_0} \tag{24}$$

Filter parts was designed using Bessel approximation with cutoff frequency of 20 kHz and gain of 3 dB in the passband. This type of the filter was chosen due to linear curve of the phase characteristic in the wide frequency range and advantageous step response with small overshot. On the other hand its drawback is smaller slope of the stop-band part of the frequency characteristic in comparison with Chebyshev or Butterworth approximations.

Main characteristics of above mentioned 4th order normalized filters with transfer functions (25), (26) and (27) simulated in Matlab 6.5 environment are compared in the Fig.12 and Fig.13.

$$G_{Bessel}(s)=\frac{5.258}{s^4+4.731s^3+10.07s^2+11.12s+5.258} \tag{25}$$

$$G_{Butterworth}(s)=\frac{1}{s^4+2.613s^3+3.414s^2+2.613s+1} \tag{26}$$

$$G_{Chebyshev}(s)=\frac{0.287}{s^4+0.984s^3+1.484s^2+0.775s+0.287} \tag{27}$$

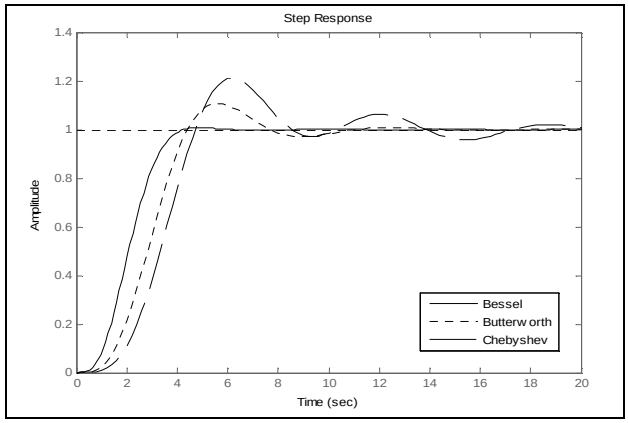

Fig. 12. Filter step response comparison.

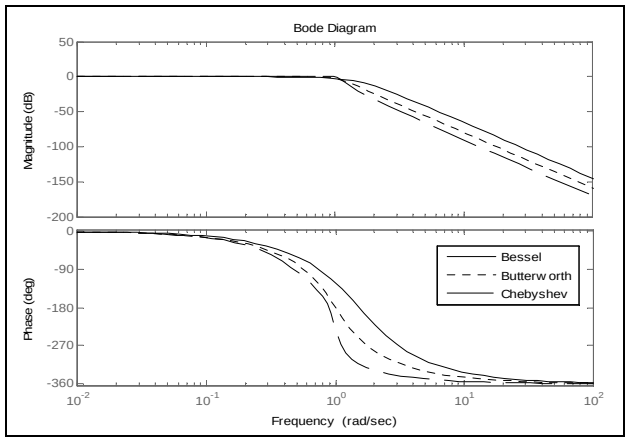

Fig. 13. Filter Bode frequency response comparison.

Filter coefficients for both filter stages are provided in the Table 2.

Table 2. Bessel filter coefficients [3].

Filter order	Stage i	a_i	b_i	Q_i
1	1	1,0000	0,0000	-
2	1	1,3617	0,6180	0,58
3	1 2	0,7560 0,9996	0,0000 0,4772	- 0,69
4	1 2	1,3397 0,7743	0,4889 0,3890	0,52 0,81

Experimentally measured frequency response of the AGC amplifier with low-pass anti-aliasing filter is depicted in Fig. 14.

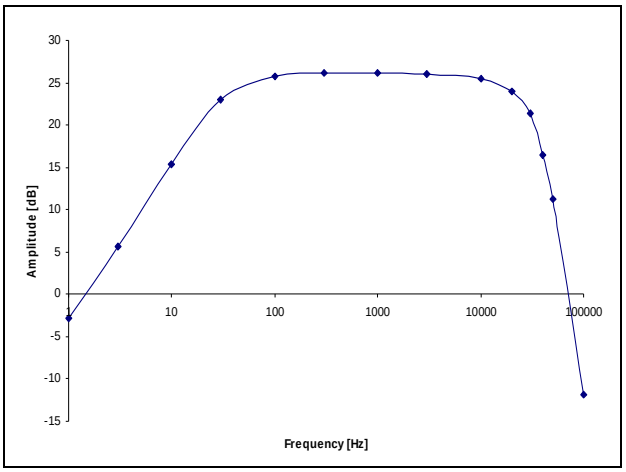

Fig. 14. AGC amplifier frequency response.

3.2 Sensory system for evaluation using beamforming

Sensory system for sound source localization using beamforming was designed on the basis of the same requirements as described in the previous chapter. Due to usage of more acoustic sensors and strict demands on their geometric configuration was chosen different system structure with fixed configuration of microphone units in the array. Another change was applied to the microphone preamplifiers which were integrated with anti-aliasing filters to the one 16-channel centralized amplifier unit. Microphone array is realized on printed circuits board which is connected with amplifier using ribbon cable with 40 wires. This construction minimizes number of cables and connectors causing in practical applications problems with system reliability. On the other hand in case of microphone array configuration change it is needed to develop and make new microphone array. Block schematic of the sensory subsystems and its connection to the evaluation systems is depicted in Fig. 15.

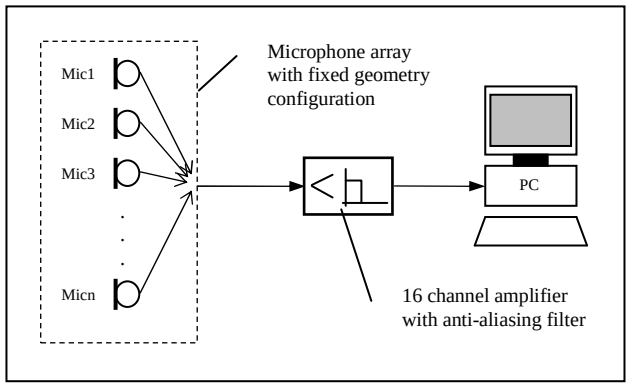

Fig. 15. Evaluation system structure.

3.2.1 Microphone array design

Microphone array consists of 15 omnidirectional condenser microphone units MCE100 soldered on the printed circuit board with dimensions of 200 x 200 mm. On the board is integrated stabilized power supply with output voltage of 5 V which is needed for operation of condenser microphones. Output signal from each microphone is connected to double-row connector where 4 pins are reserved for array supply voltage in the range of 7.5 to 15 V, 20 pins for ground and 16 pins for audio signal output. Because of preamplifiers are not integrated in the microphone array board it is suitable to connect amplifier with shortest possible ribbon cable to avoid interference leakage and useful signal losses. Schematic of the microphone array is in the Fig. 16, photograph of completed array in the Fig. 26.

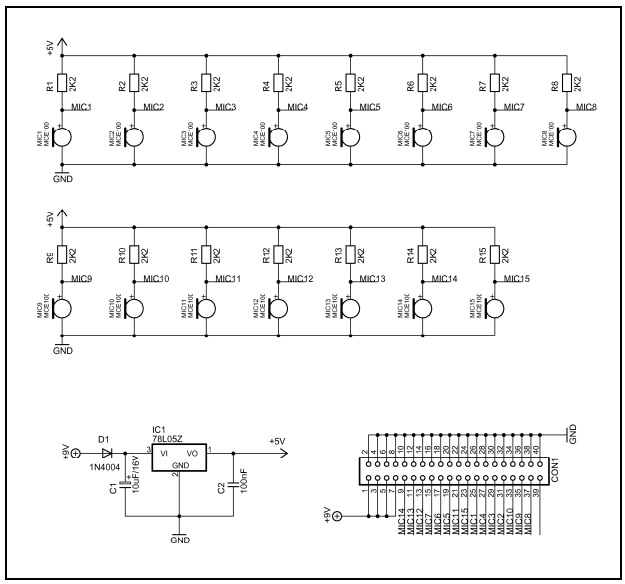

Fig. 16. Microphone array schematic.

For microphone array geometry design was created program equipment working in Matlab 6.5 environment which can compute for given geometry and demanded main lobe angle its directional characteristic. Fig. 17 shows actual array geometry designed for audio source frequencies in the range of 1500 Hz to 3000 Hz, table 3 provides coordinates of each microphone in the array. Operation with wider frequency range is possible but with negative effect on the directional characteristic. Simulated normalized directional characteristic of the electronically steered microphone array for different frequencies and main lobe angles α are shown in the Fig. 18 to 25.

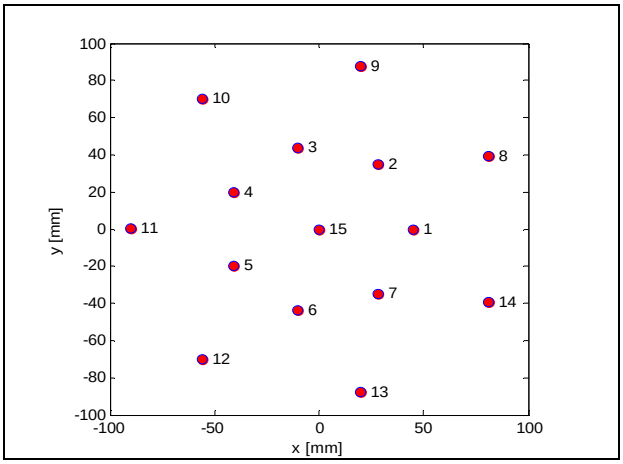

Fig. 17. Microphone array geometry.

Table 3. Microphone coordinates.

Microphone index [-]	x coordinate [mm]	y coordinate [mm]
1	45.00	0.00
2	28.06	35.18
3	-10.01	43.87
4	-40.54	19.52
5	-40.54	-19.52
6	-10.01	-43.87
7	28.06	-35.18
8	81.09	39.05
9	20.03	87.74
10	-56.11	70.36
11	-90.00	0.00
12	-56.11	-70.36
13	20.03	-87.74
14	81.09	-39.05
15	0.00	0.00

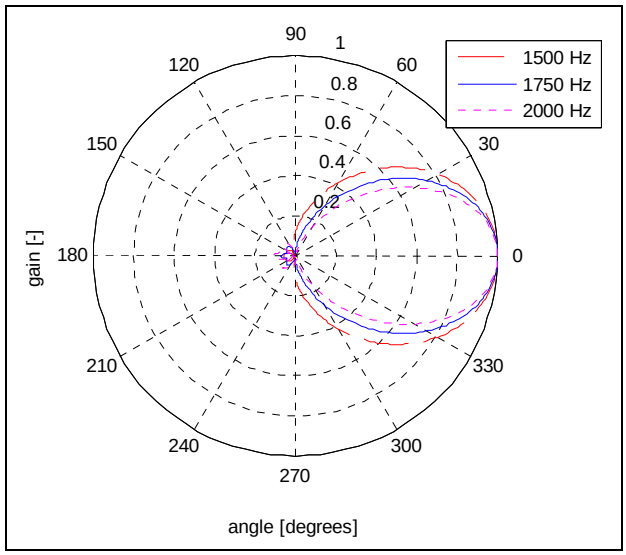

Fig. 18. Directional characteristic, α = 0°.

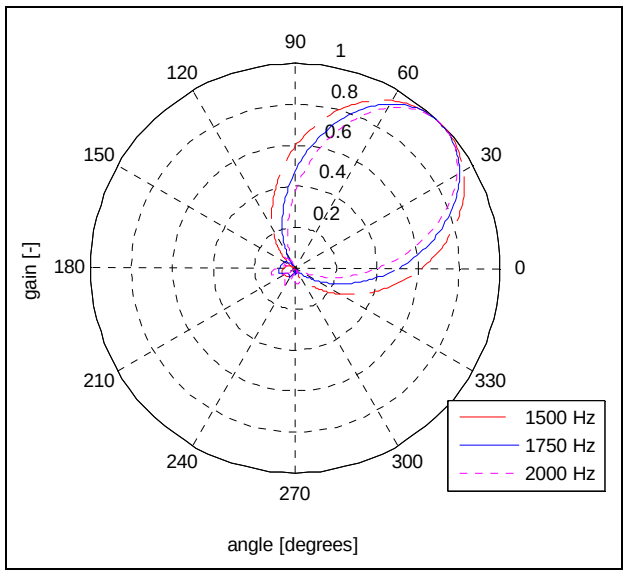

Fig. 19. Directional characteristic, α = 45°.

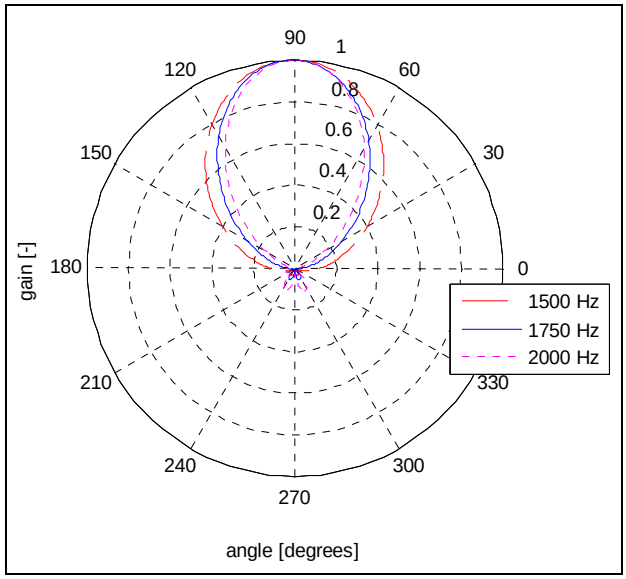

Fig. 20. Directional characteristic, α = 90°.

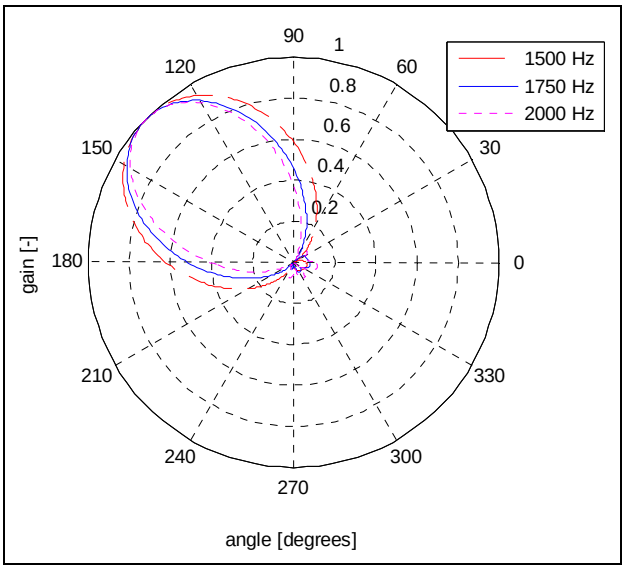

Fig. 21. Directional characteristic, α = 135º.

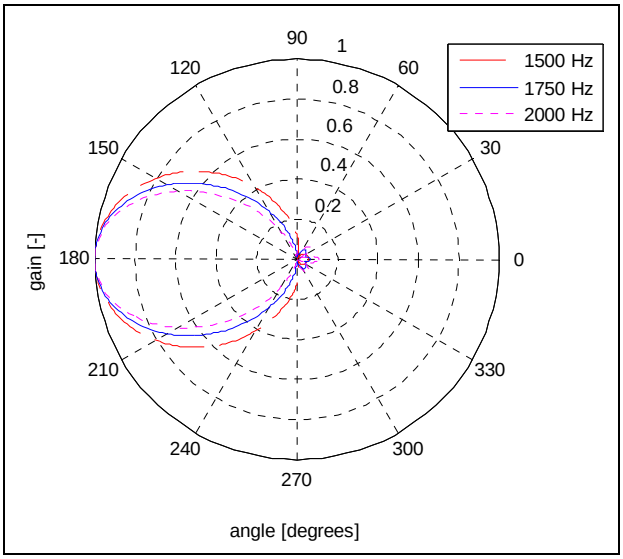

Fig. 22. Directional characteristic, α = 180º.

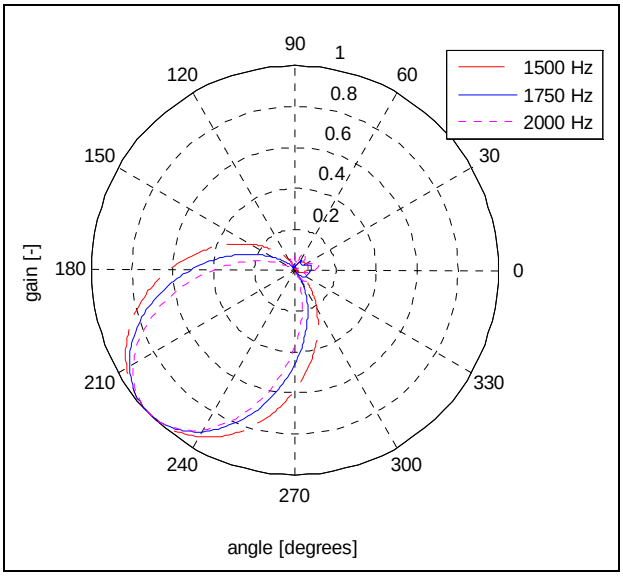

Fig. 23. Directional characteristic, α = 225º.

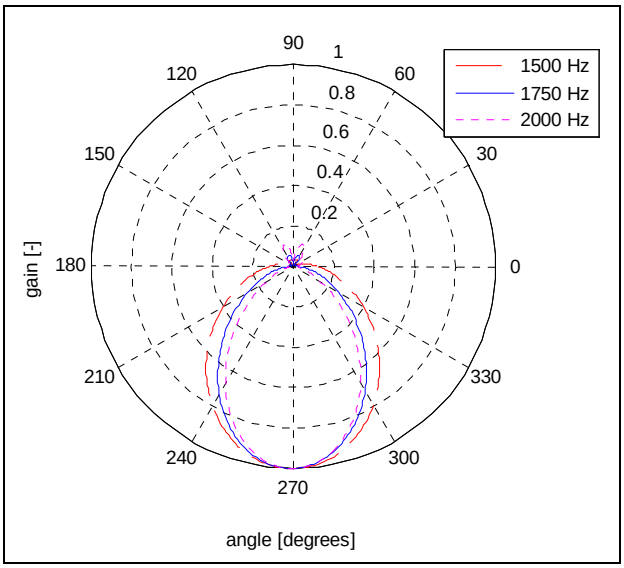

Fig. 24. Directional characteristic, α = 270º.

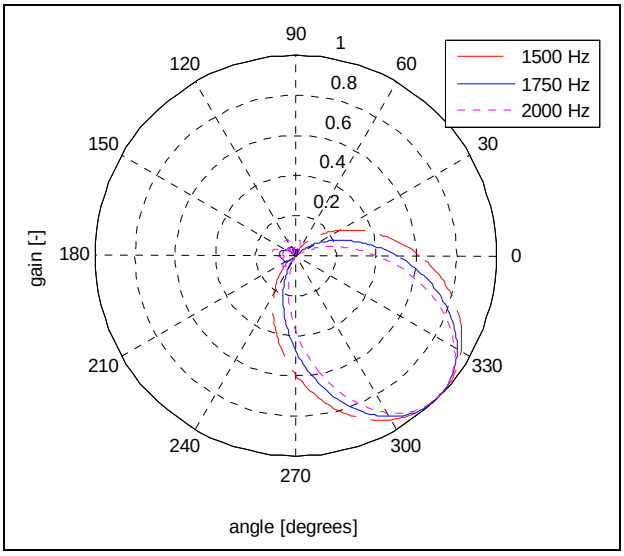

Fig. 25. Directional characteristic, α = 315º.

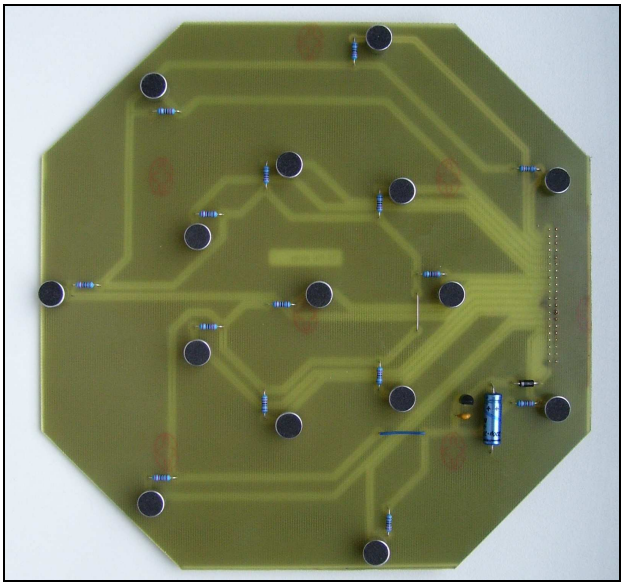
Fig. 26. Completed microphone array board.

3.2.2 16 channel microphone amplifier with anti-aliasing filter

Each microphone cartridge of microphone array is connected with preamplifier units followed by 4-th order active low-pass anti-aliasing filter with Sallen-Key topology based on quad operational amplifiers LM348 (for one audio channel is used half of the operators). Filter parts was designed using Bessel approximation with cut-off frequency of 2000 Hz and gain of 20 dB in the passband. Compared to amplifier used in sensory system for sound source localization using time-delay estimation this one is not equipped with automatic gain control stage. This simplification is possible due to very short distances between each microphone cartridges, so there is no need to adaptively adjust output signal level due to large differences in distance to sound source. Schematic of the amplifier for one audio channel is depicted in the Fig. 27.

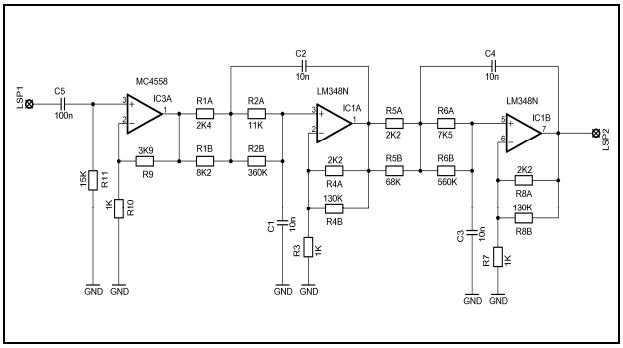

Fig. 27. Microphone amplifier schematic.

3.3 Evaluation system hardware overview

Evaluation system is based on standard personal computer with AMD Athlon64 processor and operating system Windows XP. It is equipped with multifunction data acquisition card Advantech PCI-1716 dedicated for PCI bus interface with full plug and play capability. This card provides sixteen analog inputs in single-ended or eight analog inputs in differential mode with input impedance of 100 MΩ. Each input is through analog multiplexor connected to analog-to-digital converter with 16-bit resolution and maximum sampling rate equal to 250 kHz. Integrated FIFO memory with capacity of 1024 samples enables efficient data transfer from the card to the system memory without excessive CPU utilization. It is also equipped with two analog outputs, sixteen digital inputs and outputs with TTL compatible logic and finally with 16-bit timer with reference frequency of 10 MHz.

Input voltage ranges are fully software programmable in the ranges shown in table 4. Connection with measured object is realized via universal screw terminal module ADAM-3968 suitable for DIN rail mounting. Data acquisition card is connected with ADAM module using 68-pin shielded SCSI-II cable PCL-10168-1.

Table 4. Advantech PCI-1716 input voltage ranges

Mode	Range [V]				
Unipolar	N/A	0 ~ 10	0 ~ 5	0 ~ 2.5	0 ~ 1.25
Bipolar	±10	±5	±2.5	±1.25	±0.625

4 Program implementation

Software application for audio signal analysis and sound source localization was created in Microsoft Visual C++ as Win32 application with utilization of MFC and FFTW library. In the Fig. 28 is depicted internal structure of the developed program application. Input audio signal is acquired with DAQ card and via its device driver and ADSAPI interface transfers digitized raw audio data to the disk writer which in case of need performs input data archiving to the standard WAV files. In the next block data are prepared for frequency analysis by rectangular window function and cross-correlated with utilization of FFTW library. Results of cross-correlation analysis enter direction of arrival computation (DOA) routine and then to the visualization module.

In the case that beamforming localization mode is activated, input signal is 16x upsampled and filtered by bandpass FIR filter. After that follows processing in delay and sum beamformer which weights are set according to array configuration and demanded directional characteristic main lobe angle.

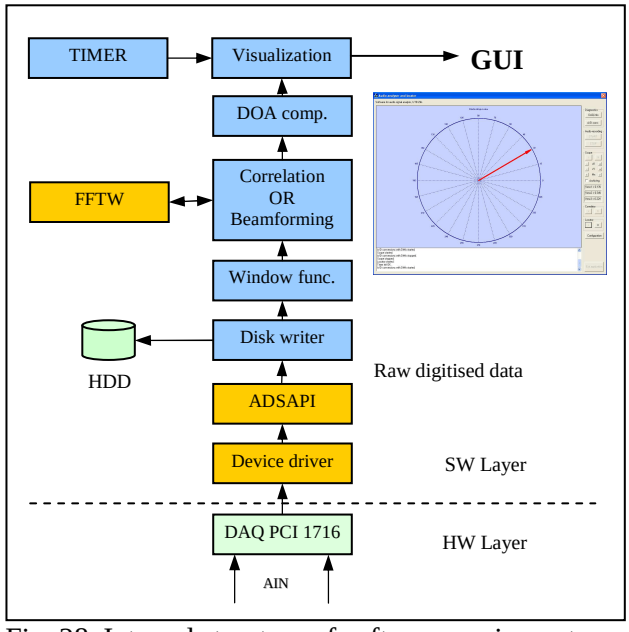

Fig. 28. Internal structure of software equipment.

4.1 Data acquisition layer

Advantech provides software support for variety of programming environments and languages including Visual Basic, Delphi, Visual C/C++, Borland C and C++ Builder. For high speed conversions during which large amount of data are present can be very advantageously used functions utilizing DMA transfers for data acquisition. Due to low CPU utilization during data transfers from buffer to main memory there is enough free computing power for tasks related to FFT and cross correlation computations and data visualization. Acquired audio data are stored in raw format to appropriate buffers corresponding to scanned channels from where they are processed by audio analyzing or localization algorithms. During the processing stage new data are acquired and transferred by driver using bus master DMA transfer from data acquisition board FIFO memory to user buffers. Due to this mode of operation no data are lost even maximum sample rate is chosen.

4.2 Computational layer

Computational core of the program is FFTW library developed by authors Matteo Frigo and Steven G. Johnson which is released under the GNU General Public License. It is used for computations of FFT and IFFT in cross-correlation and FIR filter algorithms in "Audio locator" mode and for frequency spectra analysis of input signals in "Audio analysis" mode.

Library supports both one-dimensional and multi-dimensional transforms with real or complex input data. Due to SSE/SSE2/3dNow! and Altivec support provides very high processing speed. Usage of the FFTW library is very intuitive. It can be split to these main steps:

- Memory allocation for input and output arrays (fftw_malloc).
- Create a plan containing all data needed for DFT computation (fftw_plan_dft_1d for one-dimensional DFT with real input data and complex output data).
- Start DFT computation using created plan (fftw_execute).
- Process computed data by user program, in case of need start another DFT computation repeating previous step.
- Free the plan (fftw_destroy_plan).
- Free allocated memory (fftw_free).

FFTW use its own data type fftw_complex which is defined as array of two double type elements. Element with index 0 is real part and with index 1 imaginary part of complex number [2].

4.3 Graphical user interface

Main window appearance of the developed application is depicted in the Fig. 29 (audio analysis mode activated) and Fig. 30 (audio locator mode activated). Main dialog window integrates all necessary program components. In "Audio analysis" mode main part of the window contains seven systems of coordinates for data analysis visualization (wave shapes and corresponding frequency spectra) originating from user selectable three analog channels. In the bottom part of the window is list-box which informs about current program status and its settings. In the right part are placed buttons for program control and configuration. In "Audio locator" mode is activated DOA computational algorithm and visualization module is switched to display polar graph with red arrow pointing to the computed direction of audio source location. Microphone array configuration is automatically loaded from external file which can be selected in program configuration window.

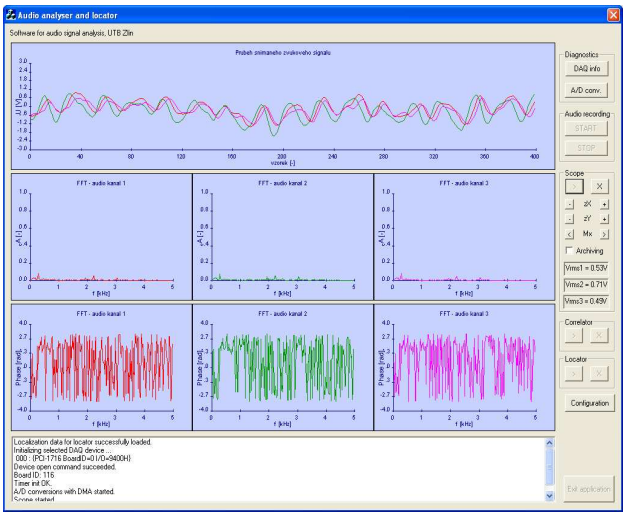

Fig. 29. Audio analyzer SW – signal analysis mode.

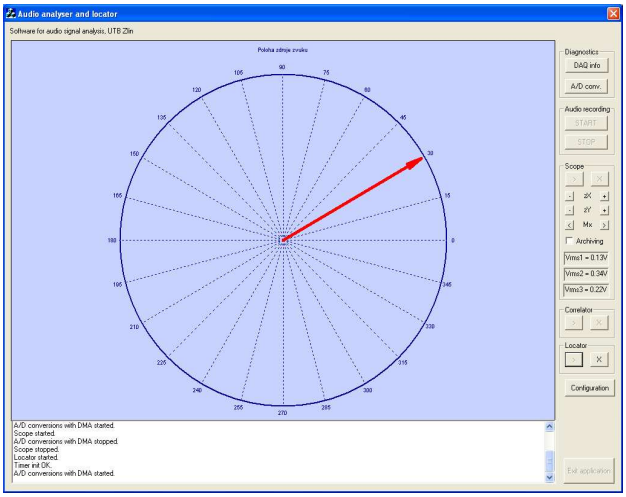

Fig. 30. Audio analyzer SW – locator mode.

5 Experimental verification

Evaluation system functionality was verified in laboratory environment by connecting of first two audio channels directly to the outputs of precision digitally controlled arbitrary waveform generator ETC M631. For our testing is very important phase shift introduction possibility between two output channels. Signal generator thus substitutes one microphone pair of the localization system with distance of 5m. Phase shift modifications in the output signal then simulate changes in the acoustic source position against microphone units.

Table 5. Direction of arrival computation results.

Sound source angle [deg]	**Direction of arrival [deg]**		
	f_s=22kHz	f_s=48kHz	f_s=96kHz
-89	-86.74	-87.36	-86
-75	-75.34	-75.51	-75.35
-60	-60.1	-60.17	-60.22
-45	-44.97	-45.11	-45.12
-30	-30.11	-30.14	-30.08
-15	-15.11	-15.05	-15.05
0	0	-0.02	-0.02
15	14.93	14.97	15.01
30	29.88	30.01	30.03
45	44.72	45	45.03
60	59.74	60	60.08
75	74.58	75.18	75.18
89	83.68	85.85	85.45

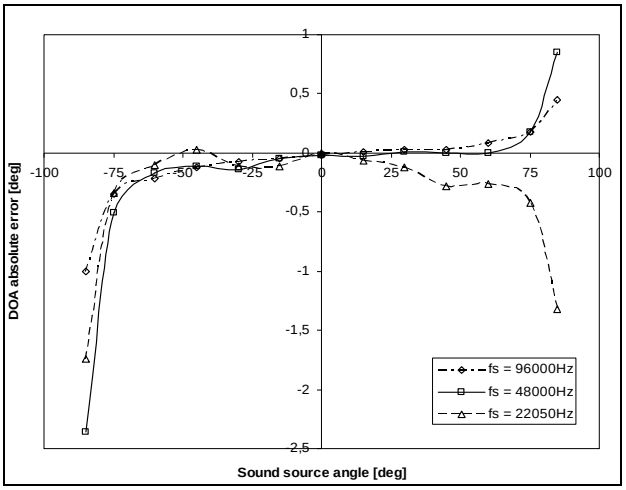

Fig. 31. DOA evaluation absolute error.

Results of the evaluation system accuracy verification for three different sampling rates are provided in table 5. First column contains exact sound source positions generated by signal generator. Next three columns represent evaluation system response acquired using different DAQ sampling rates. Measured dependency of direction of sound wave arrival evaluation absolute error on sound source angle and evaluation system sampling rate is provided in Fig. 31.

4 Conclusion

Paper deals with acoustic source localization using microphone array using time-delay estimation method and beamforming technique. It describes main requirements, hardware design of the microphone units, preamplifiers with anti-aliasing filters and finally its connection to standard personal computer equipped with Advantech PCI-1716 multifunction data acquisition card. Designed sensory system has modular architecture which enables easy portability, configurability and connectivity with evaluation unit. Developed software application for audio signal acquisition and analysis can work in two modes of operation. In analyzer mode it performs frequency analysis of all active audio channels and visualizes results in graphical form in the main window. In locator mode can compute direction of sound wave arrival by two different methods – time-delay estimation and beamforming technique using delay and sum beamformer.

6 Acknowledgment

The work was performed with financial support of research project MSM7088352102. This support is very gratefully acknowledged.

References:

[1] Czerniawski, J., Czyzewski, A., Krolikowski, R., Neural Computation of Direction-Of-Arrival of Sound, *Proceedings of the 3rd WSEAS International Conference on Neural Networks and Applications*, 2002, pp. 4891-4896.

[2] Frigo, M., Johnson, S. G., *FFTW3*, 2006. Available from WWW: http://www.fftw.org/ #documentation

[3] National Semiconductor, *Op Amps for Every One - Design Reference*, 2002. Available from: http://www.ti.com.

[4] Self. D., *Acoustic Location and Sound Mirrors*, 2004. Available from WWW: http:// www.dself.dsl.pipex.com/museum/comms/ear/ ear.htm#steer

[5] Smith, S. W., *The Scientist and Engineer's Guide to Digital Signal Processing, second edition*. California Technical Publishing, 1999.