Acoustic Source Localization Based on Time-delay Estimation Method

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: - Paper deals with acoustic source localization using microphone array using time-delay estimation method. It describes main requirements, hardware design of the microphone units with automatic gain control preamplifiers with integrated anti-aliasing filters and finally its connection to standard personal computer equipped with Advantech PCI-1716 multifunction data acquisition card. Next part of the paper describes time-delay estimation method of the direction of arrival of the sound wave and its software implementation for real-time evaluation. Created software application for audio data analysis and direction of arrival computations provides user-friendly graphical user interface which is able to visualize recorded sound waves and frequency spectra in sound analyzer mode and direction of arrival of sound wave in locator mode. All acquired data from microphone units can be saved to the standard user specifiable wav files for further investigation and analysis.

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

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 preamp 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 method of the direction of arrival of the sound wave and its software implementation.

2 Sensory system design

Microphone sensory system was designed with a respect to easy portability, configurability and connectivity with evaluation unit. These requirements best fulfill modular architecture schematic of which is depicted in figure 1. It consists of the following main components: three microphone units with integrated preamplifier, 3-channel automatic gain control (AGC) amplifier with output anti-aliasing filter and evaluation unit – in this case standard personal computer equipped with multifunction Advantech PCI-1716 data acquisition card. Components are connected together with 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.

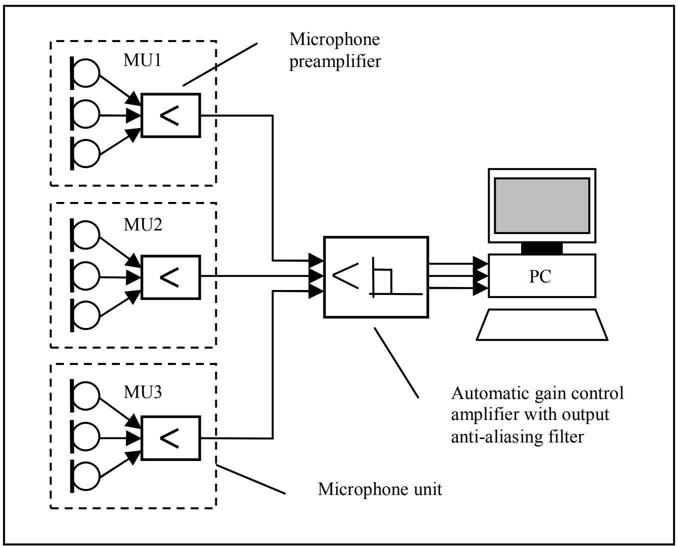

Fig. 1. Block schematics of the sensory system

2.1 Microphone units design

Sound field is measured with three microphone units each equipped with three omnidirectional electret condenser microphone cartridges. They are installed on the small triangular base from fiberglass cooper coated board with the side size of 20mm. This construction improves sensitivity and signal-to-noise ration. Each microphone is connected with summing preamplifier with the gain of 23dB 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 figure 2. 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 electret 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 figure 3.

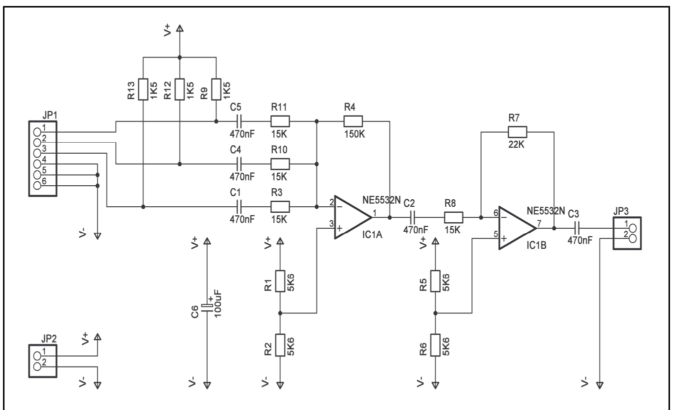

Fig. 2. Microphone preamplifier schematics

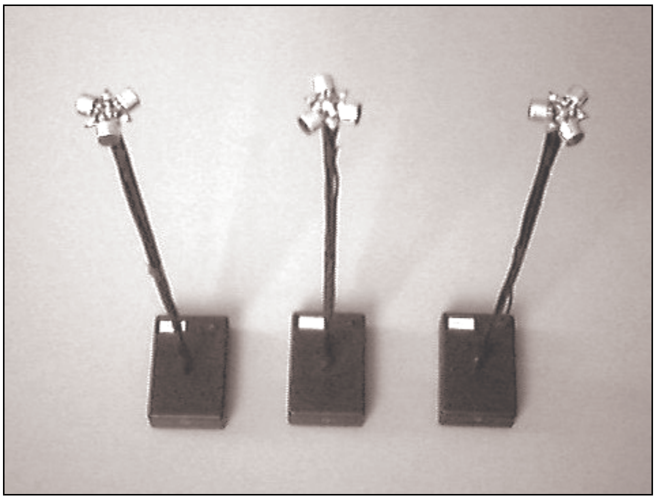

Fig. 3. Completed microphone units

2.2 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 figure 4. 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.

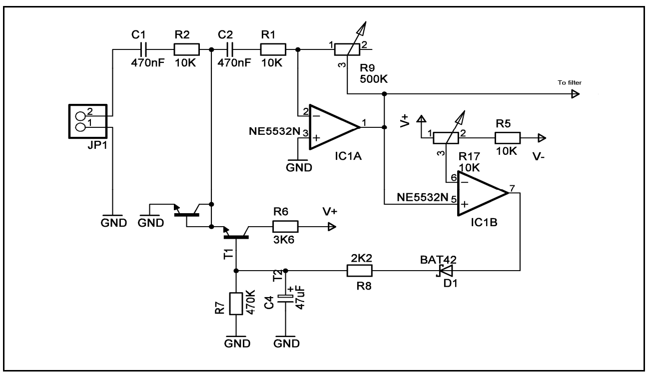

Fig. 4. AGC amplifier – amplifier part

Amplifier stage is followed by 4th order low-pass filter with Sallen–Key topology implemented by operational amplifiers (Fig.5).

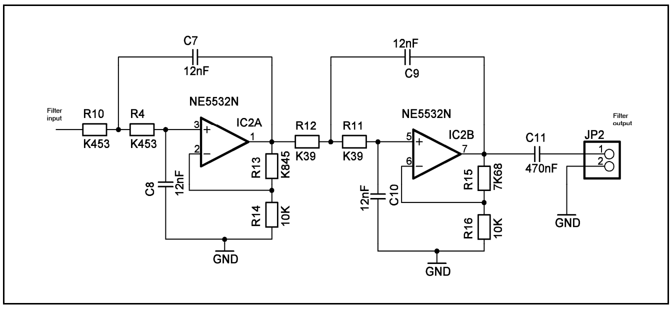

Fig. 5. AGC amplifier – filter part

It is used as anti-aliasing filter which restricts bandwidth of the signal to satisfy sampling theorem. Filter parts was designed using Bessel approximation with cutoff frequency of 20 kHz and gain of 3 dB. 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 (1), (2) and (3) simulated in Matlab 6.5 environment are compared in the figures 6 and 7.

$$G_{Bessel}(s) = \frac{5.258}{s^4 + 4.731s^3 + 10.07s^2 + 11.12s + 5.258} \quad (1)$$

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

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

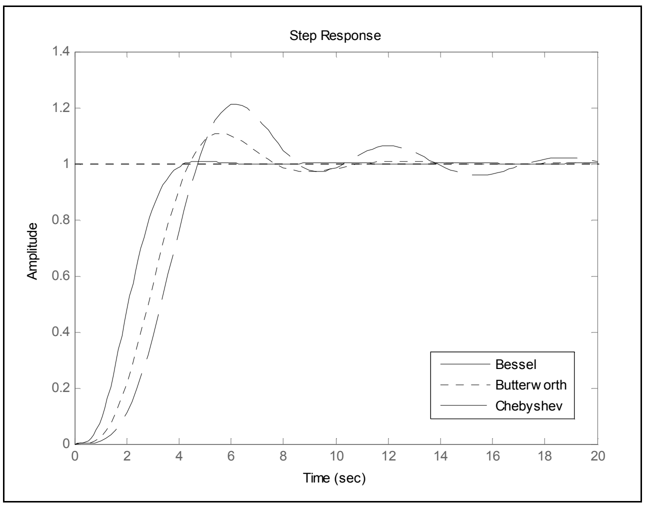

Fig. 6. Filter step response comparison.

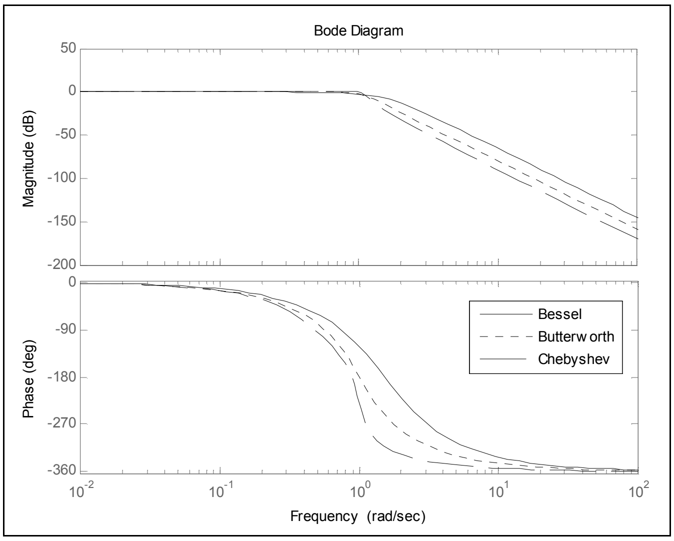

Fig. 7. Filter Bode frequency response comparison.

3 Direction of arrival 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 [3].

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 figure 8.

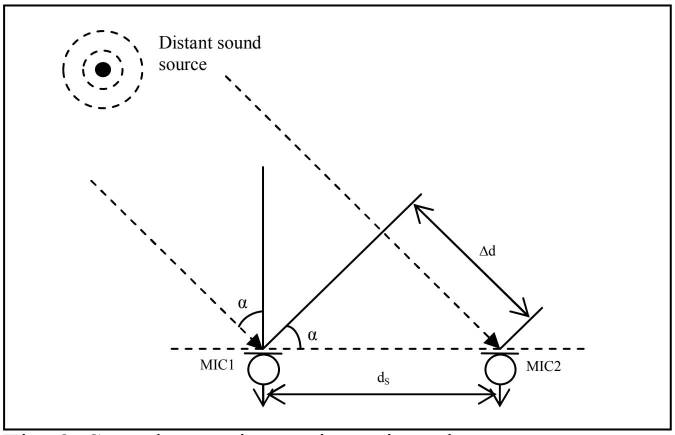

Fig. 8. Sound wave impacting microphone array.

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

$$\alpha = \arcsin\left(\frac{\Delta d}{d_S}\right) \quad (4)$$

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

$$\alpha = \arcsin\left(\frac{\Delta t \cdot c}{d_S}\right) \tag{5}$$

For time-delay estimation Δt can be very advantageously used cross-correlation analysis which is defined by equation (6). Where x_1 is acquired signal from microphone 1, x_2 from microphone 2 and N number of acquired data samples.

$$\hat{R}_{x1x2}[k] = \frac{1}{N}\sum_{n=0}^{N-k-1} x_1[n]x_2[n+k] \tag{6}$$

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 Δt can be computed using formula:

$$\Delta t = \frac{1}{f_{vz}} \arg\max\left(\hat{R}_{x1x2}[k]\right) \tag{7}$$

where f_{vz} is sampling frequency.

4 Experimental verification

Sensory and evaluation system functionality was verified with microphone array consisting of 3 microphone units in triangular configuration with size of d_S = 1m. Microphone array was connected with evaluation unit represented by standard personal computer equipped with data acquisition card. For real-time audio data analysis and direction of arrival computation was created software equipment in MS Visual C++ development environment.

4.1 Evaluation system hardware overview

Evaluation system is based on standard personal computer with processor AMD Athlon64 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 100MΩ. 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 [1].

Input voltage ranges are fully software programmable in the ranges shown in table 1. 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 1. 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

4.2 Program implementation

Software application for audio signal analysis and audio source analysis was created in Microsoft Visual C++ as Win32 application with utilization of MFC library. Program can be divided to the four following logical parts that performs each specialized tasks:

- Data acquisition using Advantech PCI-1716 driver
- Fast Fourier Transform computation using FFTW library
- Correlation and direction of arrival computation
- Data visualization and archiving layer

4.2.1 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 FFTW library. During the processing stage new data are acquired and transferred by driver using bus master DMA transfer. Due to this operation mode no data are lost even maximum sample rate is chosen.

4.2.2 Computational layer

In sound analyzer mode program use for computing the discrete Fourier transforms FFTW library developed by authors Matteo Frigo and Steven G. Johnson which is released under the GNU General Public License. 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 – 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 computation using created plan (fftw_execute)
- Process computed data
- 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].

In audio locator mode program computes using formula (6) cross correlation between all microphone pairs in the array. Then using equation (7) are computed time-delays which are inputs for direction of arrival computation algorithm.

4.2.3 Graphical user interface

Main window of the developed application "Audio analyzer" is depicted in the figure 9 (signal analysis mode activated) and figure 10 (locator mode activated). Main dialog window integrates all necessary program components. In the main part are seven systems of coordinates for data analysis visualization originating from three independent 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.

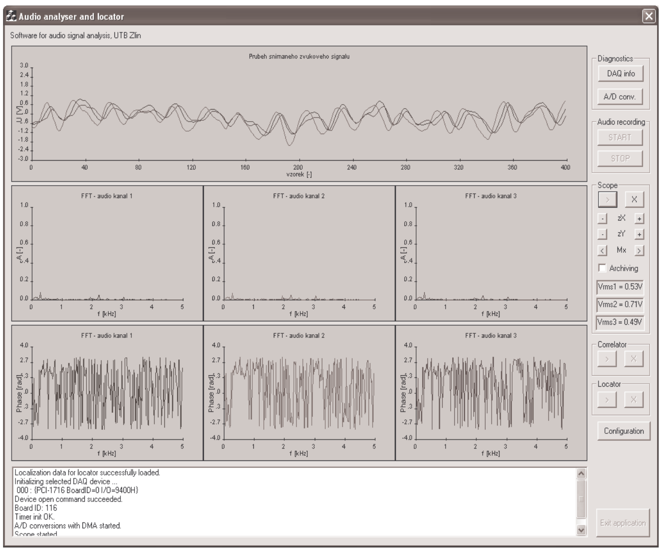

Fig. 9. Audio analyzer SW – signal analysis mode

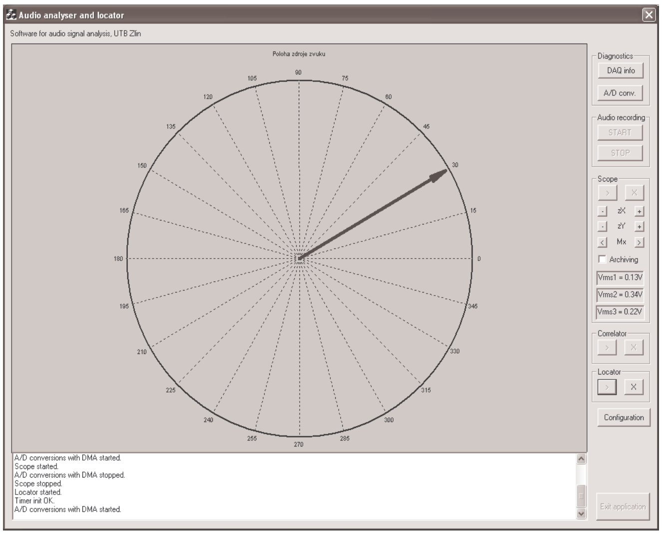

Fig. 10. Audio analyzer SW – locator mode

5 Conclusion

Paper deals with acoustic source localization using microphone array using time-delay estimation method. It describes main requirements, hardware design of the microphone units, preamplifier 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 program computes time delays between each microphone pairs and resulting direction of arrival presents in polar graph.

6 Acknowledgment

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

References:

[1] Advantech, *PCI-1716 250kS/s, 16-bit High-Resolution Multifunction Card*, 2001. Available from WWW: http://www.advantech.com

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

[3] Rajmic, P., *Prostá metoda časových posunů a modifikovaná metoda časových posunů pro detekci směru přicházejícího zvuku*, 2002. Available from WWW: http://www.elektrorevue.cz

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