Measurement Automation Monitoring, Dec. 2016, no. 12, vol. 62, ISSN 2450-2855

Mirosław ŁAZORYSZCZAK, Cezary WERNIK
WEST POMERANIAN UNIVERSITY OF TECHNOLOGY SZCZECIN
FACULTY OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
ul. Żołnierska 52, 71-210 Szczecin

An implementation of binaural sound source localization in programmable devices

Abstract

In this paper an example of hardware implementation of binaural sound source localization is presented. Using only two microphones, which correspond approximately to binaural hearing, limits the possibility of exact sound source localization. In contrast to human auditory system (HAS), only the angle of arrival determination is possible in implemented system. Moreover, the angle of arrival (AoA) could be determined here in a limited range of values located on a half-plane. First, the base formulas used by implemented algorithm are shown. Next, selected hardware platforms and peripheral modules are described. The VHDL tools for synthesis and implementation are used. Finally, resources consumed by hardware CPLD/FPGA implementation and selected test results are presented.

Keywords: CPLD, FPGA, programmable devices, binaural localization, time delay estimation.

1. Introduction

Sound source localization is an important task in many applications. It covers areas such as autonomous robots, hearing aids, widely understood multimedia, etc. There are a lot of different approaches solving sound localization problems. Some of them are biology inspired, while others are technologically innovative [1]. Humans can recognize and process several sound parameters by the auditory system. One of the abilities of a human auditory system (HAS) is directional hearing. Such capability is typical in an animal world. However, the magnitude of sensitivity may vary within a wide range of values [5]. In addition, mainly due to the pinna, human has the ability to locate sounds in the vertical plane [12]. Such an effect is not achievable by using only a set of two “ordinary” microphones. Sound localization, as a typical computational task, can be implemented in software or hardware. The software solutions like Sound Localizer [2] or LibLaura [9] are available either as a ready-made module or an open library. There are hardware implementations designed for specific platforms [10] and/or operating in special environment [11]. A special group of solutions includes an embedded system support for selected platform as a software extension, offered by microcontroller manufacturers [13].

The main motivation for the current work is a need of hardware module implementation that meet various criteria like power consumption and computational efficiency, on one hand, and potential scalability on the other hand. Scalability is understood here as a possibility of system expansion with additional microphones and their algorithmic integration, allowing parallel processing of signals from selected microphone pairs or from the set of all microphones (to be addressed in future work).

2. Sound source localization by two microphones

The basic assumption of a system presented in this work is the use of two microphones only. However, the architecture of the hardware system can be expanded in further. Sound localization by two sensors means that only the angle of a sound wave front can be determined. The CPLD and FPGA development board are considered as a platform for hardware implementation.

The calculation of the angle is based on the time delay of an arrival (TDOA), which corresponds to the interaural time difference (ITD) of the HAS. We assume that the distance between a sound source and the center of section connecting both microphones is much greater than the distance between microphones, Fig. 1. In contrast to HAS, the limitation here is the inability to distinguish by omnidirectional microphone set the direction of arrival between sources located on front or back half-planes.

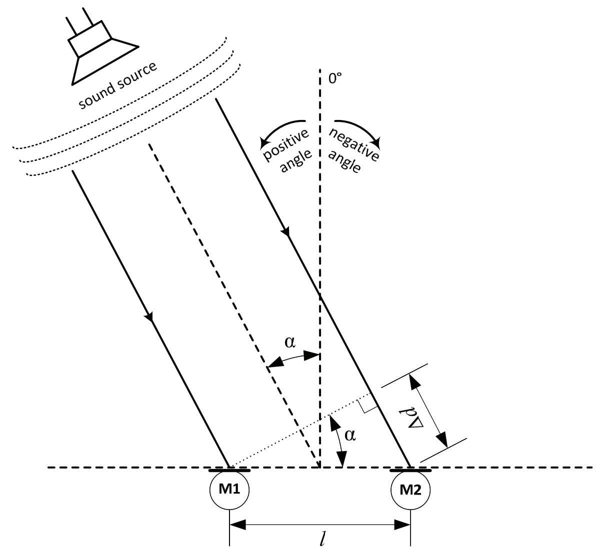

Fig. 1. The time delay of sound signal propagation between microphones

The distance difference between signals arriving to both microphones is derived as [14]:

$$\Delta d = l \cdot \sin(\alpha), \tag{1}$$

where Δd denotes the distance difference and l is the distance between microphones.

Assuming that the sound speed is constant and equal to c, ($c \approx 343.8$ m/s at 20 °C), the time delay Δt corresponding to the distance difference is:

$$\Delta t = \frac{l \cdot \sin(\alpha)}{c} \tag{2}$$

The angle of a sound source is then:

$$\alpha = \sin^{-1}\left(\frac{\Delta t \cdot c}{l}\right) \tag{3}$$

The maximum achievable time delay of a sound arrival between both channels depends on a distance between microphones. For example, the delay of approximately 582 μs can be observed for distances at least 20 cm and a sound speed of $c \approx 343.8$ m/s at 20 °C.

The delay calculation is the most important stage of an angle determination. A binaural hearing is based on Jefress model [6]. One of the possible implementations of this model use the cross-correlation as a similarity measure of two signals:

$$r_{x_L,x_R}(\tau) = \sum_{k=0}^{N} x_L(k) \cdot x_R(k-\tau), \quad (4)$$

where x_L and x_R are signals obtained respectively from left and right microphone.

The time delay can be derived as an argument for which cross-correlation function reaches its maximum:

$$\Delta t = \arg\max\ r_{x_L,x_R}(\tau). \quad (5)$$

The computational complexity of the cross-correlation algorithm based on the relation (4) is proportional to N^2. There are several algorithms with lower complexity like GCC-PHAT [8], implemented in hardware [7]. However, the requirement is a greater frame size, due to the need of Fast Fourier Transform calculation.

It should be mentioned that for digital signal sampled at 44100 Hz, the maximum delay of 582 µs corresponds approximately to about 25 samples.

3. Hardware implementation

One of the additional goals of a solution proposed in this work was the implementation of the sound localization system as a reusable and scalable hardware module. The design is based on the VHDL language. As a base test platform the CoolRunner II Starter Board containing Xilinx XC2C256 chip was selected [3]. The use of the above-mentioned platform requires dedicated software tools. In this case it was the Xilinx ISE 14.4. Apart from other features of the CPLDs, the practical reason for hardware platform choice was the total time of design implementation processes with special emphasis on the time of the synthesis level.

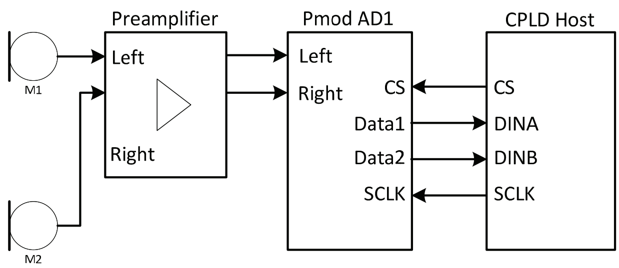

Fig. 2. The scheme of a sound source localization system

The selected board itself does not contain any sound acquisition devices. In order to feed an audio signal to the host module an external analog-to-digital converter (ADC) has to be used. The overall scheme of a proposed system is presented in Fig. 2.

The ready-made Pmod AD1 [4] peripheral module meets the requirements of the system. The peripheral module contains two Analog Devices AD7476A ADCs enabling two independent channel analog signal acquisitions with a resolution of 12 bits per each channel. As input devices we used condenser microphones (Behringer ECM8000) connected to SM ProAudio EP84 microphone preamplifier.

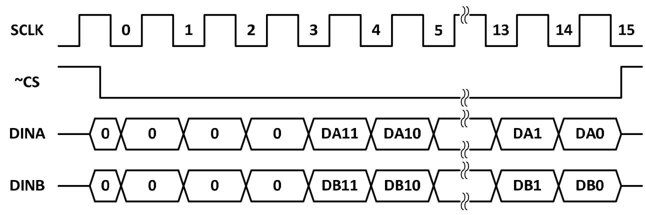

Fig. 3. Serial interface timing

From the host side, ADCs are controlled by a serial interface. Both modules are enabled by the same ~CS signal (Fig. 3), and both input data are read in parallel. Therefore there is no latency between both signals. The host is clocked at 8 MHz. Next, the system clock is divided, and SCLK and ~CS signals are generated. The A/D conversion is performed at every falling edge of ~CS, which aligns with the sampling frequency. The frequency was set to 21.4 kHz for the test purpose. As noted above, the value of a sampling rate at the given maximum time delay between channels is 12 samples. For the given sampling frequency the angle resolution is equal approximately to 7.2°.

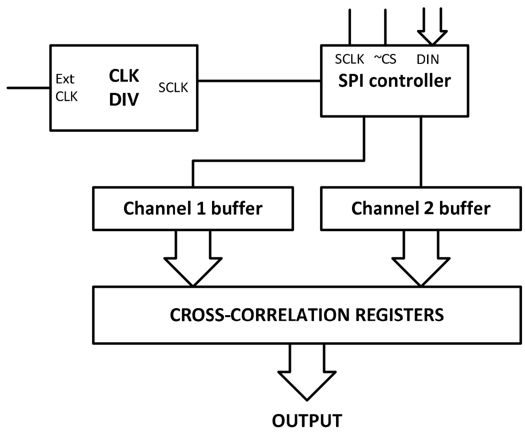

Fig. 4. The block diagram of the host unit

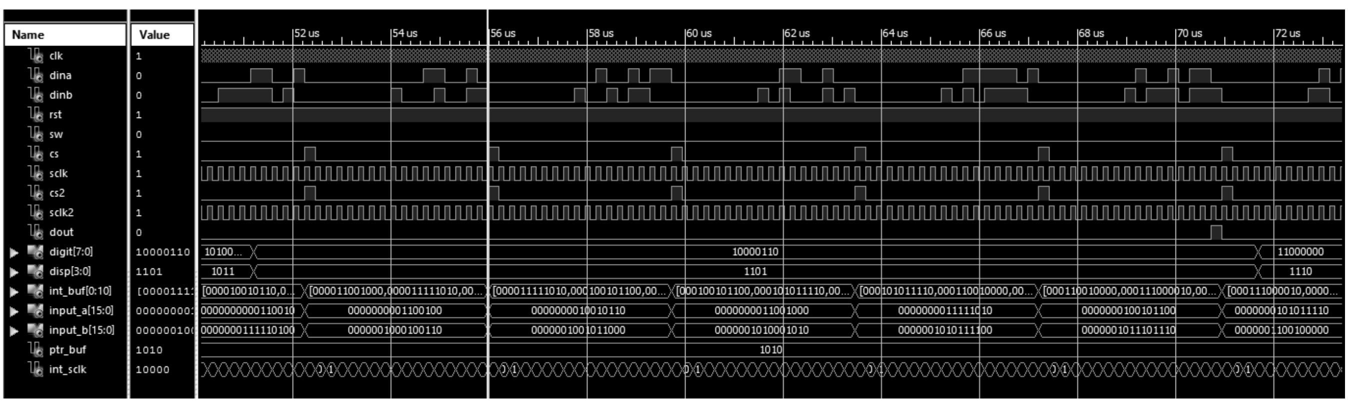

Fig. 5 The part of waveform simulation of the VHDL module of the sound localization system

In Fig. 4 the simplified block diagram of the designed host module is shown. Except the internal buffers for audio data, the main part of the module is a cross-correlation register block using the recursive form of the cross-correlation function. The system also supports the embedded seven-segment display used as a "local logger" for results presentation (not shown in figure).

The result of the main VHDL module simulation is presented in Fig. 5. Only the subset of signals is visible including display control and external data for each channel. The input digital audio signals in the waveform set are generated by VHDL test bench module.

Tab. 1. Summary of used resources

Platform	Macrocells Used/Max	Pterms Used/Max	Registers Used/Max	Pins Used/Max	Function Block Inputs Used/Max
XC2C256	**150/256 (59%)**	**355/896 (40%)**	**134/256 (53%)**	**22/118 (19%)**	**248/640 (39%)**

Table 1 shows the summary of resources used in designed module based on the ISE summary report tool. The relative low usage of external ports indicates that the design can be extended to incorporate more microphones. The high number of macrocells and registers used comes, in part, from the code version containing, from the essential module point of view, redundant internal registers.

4. Conclusion

In this paper the hardware implementation of binaural sound source localization system is described. Although the VHDL code of the system is dedicated to the specific hardware platform, it can be adapted, as assumed, to other ones. The module was written in behavioral style and functionally simulated. However, the implemented system requires more comprehensive tests in physical environment. The advantages of reprogrammable devices and use of HDLs allow to flexible modify hardware parameters of the system. This causes the implementation more sophisticated localization algorithms, too. The future work will include the expanding of number of microphones, which can be completed by using similar Pmods without any additional effort.

5. References

[1] Calmes L.: Biologically Inspired Binaural Sound Source Localization and Tracking for Mobile Robots. PhD thesis, RWTH Aachen University, 2009.

[2] Calmes L.: The sound Localizer Software, http://www.laurent calmes.lu/soundloc_software.html, (access: November 2016).

[3] Digilent website: CoolRunner-II Starter Board Reference Manual. https://reference.digilentinc.com/_media/coolrunner-ii:coolrunner-ii_rm.pdf, (access: November 2016).

[4] Digilent website: PmodAD1 Reference Manual. https://reference.digilentinc.com/_media/reference/pmod/pmodad1/pmodad1_rm.pdf, (access: November 2016).

[5] Fisher B. J., Seidl A. H.: Resolution of interaural time differences in the avian sound localization circuit – a modeling study. Frontiers in Computational Neuroscience, Vol. 8, Article 99, 2014.

[6] Jeffress L. A.: A place theory of sound localization. Journal of Comparative & Physiological Psychology, 41(1): 35–39, 1948.

[7] Jin J., Jin S., Lee S., Kim H. S., Choi J. S., Kim M., Jeon J. W.: Real-time Sound Localization Using Generalized Cross Correlation Based on 0.13 μm CMOS Process. Journal of Semiconductor Technology and Science, Vol.14, no. 2, April, 2014.

[8] Knapp C. H., Carter G. C.: The generalized correlation method for estimation of time delay. IEEE Transactions on Acoustics, Speech, and Signal Processing, 24(4): 320–327, August 1976.

[9] Kruczkowski P., Mąka T.: LibLaura: A Library for Binaural Sound Source Localization. Measurement Automation Monitoring, vol. 62, no. 12, 2016

[10] Lunati V., Manhès J.; Danès P.: A versatile System-on-a-Programmable-Chip for array processing and binaural robot audition. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Portugal, 2012.

[11] Meraoubi H., Boudraa B.: Fixed sound source localization in reverberant environments using a multimicrophone set. Société Française d'Acoustique. Acoustics 2012, Nantes, France, April 2012.

[12] Moore B. C. J.: An Introduction to the Psychology of Hearing. BRILL, 6th ed., 2013.

[13] ST website: Getting started with osxAcousticSL real-time sound source localization software expansion for STM32Cube. User manual. http://www.st.com/resource/en/user_manual/dm00239859.pdf, (access: November 2016).

[14] Strumiłło P. (ed.): Advances in Source Localization. Intech, 2011, http://www.intechopen.com/books/advances-in-sound-localization

Received: 12.09.2016 *Paper reviewed* *Accepted: 02.11.2016*

Mirosław ŁAZORYSZCZAK, PhD

Mirosław Łazoryszczak received the PhD degree in computer science from the Szczecin University of Technology in 2005. He works as an assistant professor at the West Pomeranian University of Technology Szczecin, Faculty of Computer Science and Information Technology. His research interests concerns embedded systems and applications of programmable devices.

e-mail: mlazoryszczak@wi.zut.edu.pl

Cezary WERNIK, BSc

Cezary Wernik received the BSc degree in computer science from the Szczecin University of Technology in 2016. His research interests concerns embedded systems programming.

e-mail: cwernik@wi.zut.edu.pl