

Smart Acoustic Sensor Array (SASA) System for Real-time Sound Processing Applications

Marcos Turqueti, Erdal Oruklu and Jafar Saniie

Abstract:

This chapter describes the design and implementation of a Smart Acoustic Sensor Array (SASA) System consists of a 52 microphone MEMS array embedded in an FPGA (Field Programmable Gate Array) platform with real-time data acquisition, signal processing, and network communication capabilities. The SASA system is evaluated using several case studies in order to demonstrate the versatility and scalability of the sensor array platform for real-time sound analysis applications. These case studies include sound mapping and source localization.

Keywords: Acoustic MEMS array, sound localization, field-programmable gate arrays, smart sensors

Smart Acoustic Sensor Array (SASA) System for Real-time Sound Processing Applications

Marcos Turqueti, Erdal Oruklu and Jafar Saniie

1.1 Introduction

Designing Smart Acoustic Sensor Array (SASA) system presents significant challenges due to the number of sensors, processing speed, and complexity of the targeted applications. This chapter describes the design and implementation of a 52 microphone MEMS array, embedded in an FPGA (Field Programmable Gate Array) platform with real-time processing capabilities. In this type of system, characteristics such as speed, scalability, and real-time signal processing are paramount and furthermore, highly advanced data acquisition and processing modules are necessary. Sound processing applications such as sound source mapping, source separation and localization have several important characteristics that the system must meet such as; array spatial resolution, low reverberation, and real-time data acquisition and processing.

The acoustic MEMS array system presented in this study is designed to meet these challenges. The design takes into account mechanical factors such as array geometry, sensor disposition, and microphone reverberation. In addition, the design of electronics system addresses electronic noise, power decoupling, cross-talk, and connectivity. The system integrates the acoustic array with real-time data acquisition, signal processing, and network communication capabilities. FPGA based smart sensor nodes are examined with several case studies demonstrating the versatility and scalability of the sensor array platform for real-time applications. These case studies include sound mapping and source localization.

Review Copy

Microphone arrays are capable of providing spatial information for incoming acoustic waves. Arrays can capture key information that would be impossible to acquire with single microphones, however, this extra information comes with a price in the increased system complexity. The design of microphone arrays faces several challenges such as number of detectors, array geometry, reverberation, interference issues and signal processing (Buck et al. 2006). These factors are crucial to the construction of a reliable and effective microphone array system.

Microphone arrays present additional challenges for signal processing methods since large numbers of detecting elements and sensors generate large amounts of data to be processed. Furthermore, applications such as sound tracking and sound source localization, require complex algorithms to properly process the raw data (Benesty et al. 2008). Challenging environments such as multiple sound sources moving with background noise are especially difficult to deal with when real-time processing is required. It is also important for such a system to be easily scalable as demonstrated by (Weinstein et al. 2004) and (Benesty et al. 2008). The performance of a microphone array typically increases linearly with the size of the array.

The system presented in this project is designed to provide an acoustic data acquisition system that is flexible and expandable, with powerful real-time signal processing capability in order to be interfaced with a wide variety of sensor arrays. The data acquisition and processing architecture presented in this work is called CAPTAN (Compact And Programmable daTa Acquisition Node) (Turqueti et al. 2008) and the acoustic array embedded on the system is called

Review Copy

AMA (Acoustic MEMS Array). The combination of AMA and CAPTAN is called SASA (Smart Acoustic Sensor Array) system.

The CAPTAN architecture is a distributed data acquisition and processing system that can be employed in a number of different applications ranging from single sensor interface to multi-sensor arrays data acquisition, and high performance parallel computing (Rivera et al. 2008).

This architecture has unique features of being highly expandable, interchangeable, adaptable and with a high computation power inherent to its design. The AMA array was designed to conform to the CAPTAN architecture and take advantage of its architecture. It is an acoustic array that employs sound or ultrasound sensors distributed in two dimensions.

1.2 Fundamentals of acoustic sensor arrays and applications

There is currently a significant amount of research and numerous applications which use sound or ultrasound for communications, detection and analysis. Some of the research topics are multi-party telecommunications, hands-free acoustic human-machine interfaces, computer games, dictation systems, hearing-aids, medical diagnostics, structural failure analysis of buildings or bridges, and mechanical failure analysis of machines such as vehicle or aircraft, and robotic vision, navigation and automation (Alghassi 2008, Eckert et al. 2011, Kim et al. 2011, Llata et al. 2002, Nishitani et al. 2005, Kunin et al. 2010, Kunin et al. 2011). In practice, there are a large number of issues encountered in the real world environment which make the realistic applications of this theory significantly more difficult (Shaw 2002, Brandstein and Ward 2001, Rabinkin et al. 1996). This includes ambient sound and electrical noise, presence of wideband

Review Copy

non-stationary source signals, the presence of reverberation echoes, the high frequency sound sources which require higher speed systems, and fluctuation of ambient temperature and humidity which affect the speed at which sound waves propagate.

The geometry of the array can play an important part in the formulation of the processing algorithms (Benesty et al. 2008). Different applications require different geometries in order to achieve optimum performance. In applications such as source tracking, the array geometry is very important to determine the performance of the system (Nakadai et al. 2002). Regular geometries where sensors are evenly spaced are preferred in order to simplify the algorithm development. Linear arrays are usually applied in medical ultrasonography, planar arrays are often used in sound source localization, and three dimensional spherical arrays are most frequently used in sophisticated SONAR applications. In other applications such as source separation, the geometry of the array is not as important as the transducer characteristics such as dynamic range and transducer aperture. The size of an array is usually determined by what frequency the array will operate and what kind of spatial resolution the application using the acoustic array requires (Weinstein et al. 2004).

A data acquisition system is necessary to condition, process, store and display the signals received by the array. Data acquisition for an acoustic array can be very challenging, but most of the issues are co-related with the type of sensors, number of sensors and array geometry. Also, the application plays an important role in defining the needs of the array for signal processing or real-time operation requirements.

Review Copy

Beamforming is a technique used in transducer arrays for signal beam field and directivity for transmission or reception. The spatial directivity is achieved by the use of interference patterns to change the angular directionality of the array. When used for transmission of signals the transmitting signal will be steered, where the amplitude and phase of its individual elements will be controlled to act as one through patterns of constructive and destructive interference. The resulting wavefront will have the energy of the array concentrated on the desired direction. There are two categories of beamforming; static and adaptive (Hodgkiss 1980, Campbell 1999). Static beamforming involves using a fixed set of parameters for the transducer array. Adaptive beamforming on other hand can adapt the parameters of the array in accordance with changes in the application environment. Adaptive beamforming, although computationally demanding, can perform better than static beamforming in noise rejection.

1.3 Design and implementation of smart acoustic MEMS array

The design process for the Acoustic MEMS Array (AMA) involves deciding the best topological distribution of sensors, number of microphones, inter-microphone spacing and microphone type. In order to avoid spatial aliasing and have a good acoustic aperture, the inter-microphone distance is determined to be 10.0 mm center to center. This spacing makes it possible to obtain relevant phase information of incoming acoustic sound waves, increasing the array sensitivity and allowing spatial sampling of frequencies up to 17 kHz without aliasing. The inter-microphone array spacing is approximated by dividing the sound speed by the inter-microphone distance and then dividing the result by two in order to satisfy the Nyquist–Shannon sampling theorem.

Review Copy

The number of microphones is a compromise between the size of the board, the electronics needed to deal with the massive amount of data, and the need to achieve a signal-to-noise ratio of at least 20dB. The AMA board (see Figure 1.1) consists of 52 MEMS microphones distributed in an octagonal layout having 8 columns and 8 rows with two central transmitter elements. The central elements can be used for calibration purposes or sonar applications. When used as a calibration element, the central element emits a set of mono-component frequencies. Then, these signals are captured by the microphones for calibrating the response of the array taking into account the spatial distribution of microphones. When the central element is used for active sonar applications, a series of pre-programmed pulses are emitted and then, when the emitted waves encountering obstacles they bounce back toward the array, giving information about the distance and position of obstacles.

The MEMS microphones are the fundamental components of the array, and their composition must offer high sensitivity and low reverberation noise. The overall sensitivity of the array increases monotonically with the number of sensors. The MEMS microphone chosen for this array (SPM0208) has a sensitivity of 1V/Pa at 1 kHz (Knowles 2006). Figure 1.2 shows the selected MEMS microphone. These microphones are omni-directional and when combined into an array they provide a highly versatile acoustic aperture. The selected MEMS microphone frequency response is essentially flat from 1 to 8 kHz, and it has a low frequency limit of 100 Hz and a high frequency limit of 25 kHz. The microphones are glued with silver epoxy to the copper pads in the board in order to shield the microphone response against acoustic noise and reverberation.

Review Copy

In support of the microphones, the AMA board also provides analog-to-digital converters that interface to the CAPTAN board through the four board to board vertical bus connectors (Turqueti et al. 2008). The readout system is based on the CAPTAN architecture and its implementation is presented in the next section. Physically the system can be separated into three parts namely the Acoustic MEMS Array (AMA), Node Processing and Control Board (NPCB), and the Gigabit Ethernet Link (GEL) board. The NPCB is the backbone board that contains the FPGA which contains the system's configware. The GEL board controls Ethernet communications. The NPCB and GEL boards are part of the CAPTAN system. The AMA board is the hardware which contains the microphones, amplifiers and analog-to-digital converters (ADC). Figure 1.3 shows the three hardware components that make up the smart acoustic sensor array (SASA) system.

The board containing the MEMS array is responsible for the data acquisition. Each of the 52 MEMS microphones is equipped with its front end signal conditioning circuit. This circuit is presented in Figure 1.4a and Figure 1.4b. It has the objective of amplifying and filtering the microphone signal. The first stage of the amplifier gives a 20 dB gain and is embedded within the microphone case as indicated by the dashed area in Figure 1.4b. The second stage of the amplifier gives further 20 dB of gain where the gain can be adjusted by the feedback resistor. The amplifiers also provide a second order highpass filter with low cutoff frequency of 400 Hz.

This board supports two different commercially available MEMs microphones from Akustica, the SPM208 with a dynamic range varying from 100 Hz to 12 kHz and the SPM204 with a dynamic range varying from 10 kHz to 65 kHz, where these microphones can be intermixed on

Review Copy

the array or used as the solo type. The two different sensors allow sound analysis and ultrasound ranging applications. SPM208 frequency response is shown in Figure 1.5.

After the analog signal is conditioned, it is processed by the channel ADC. Every single channel has its own ADC, an ADC121S101 (National 2010). This is a serial 12-bit A to D converter with a maximum sampling rate of 1 M samples/second. The system presented in this work was set to work at 36 K samples/second but it can simply be adjusted in the configware to any sampling rate as desired up to 1 M samples/second. After the ADC digitizes the signal, it generates a serial bit stream at 432 K bits/second. This bit stream is continuously generated by all 52 ADCs creating an overall data rate of 22.5 M bits/second transferred to the NPCB board through the four vertical bus connectors.

1.4 System implementation of AMA and CAPTAN

The smart acoustic sensor array (SASA) system (see Figure 1.3) developed in this project is a novel combination of PC/FPGA based data acquisition system (CAPTAN) with an embedded MEMS based microphone array (AMA). The NPCB board, as part of the CAPTAN, contains a VIRTEX-4 XC4VFX12 FPGA where all data from the array is stored and processed. The FPGA is connected to a 32MB EPROM that contains a specially designed configware for dealing with the data coming from the array. This configware is automatically loaded to the FPGA every time the system is powered up. It is important to mention that the configware plays a central part on the system architecture and it is divided in three distinct modules: 1) Acquisition Control Module, 2) Signal Processing Module, and 3) Ethernet Communication Module.

Review Copy

The Acquisition Control Module is the block that contains the SPI (Serial Peripheral Interface) interface to communicate with the ADCs, this block is responsible for controlling the ADCs by programming its registers; it is also responsible for receiving and formatting the data sent by the ADCs. Data coming from the Acquisition Control Module is then sent to the Signal Processing Block. The processing block is application dependent and usually contains DSP modules but can be as simple as a buffer. From this block the data is forwarded to the Ethernet block. The Ethernet Communication Block is part of the CAPTAN architecture and is a specially designed configware that formats data to UDP protocol (Turqueti et al. 2008) and sends it to the GEL board. This block does all network communication using UDP, and is a full duplex system able to transmit and receive data at Gigabit rate (see Figure 1.6). Data are transmitted out of the FPGA through eight lines at a 125MHz clock. These transmission lines connect to the GEL card which electrically formats the data to conform to Ethernet physical layer using a PHY (physical layer converter). Once on the Ethernet, the information can flow directly to a computer or it can go to a network. Each GEL has a unique MAC (Media Access Control) address and an IP (Internet Protocol) address that can be configured by the user.

An integral part of the SASA system is the data acquisition and network manager software. After the information is broadcast to the Ethernet, a computer connected to the network is capable of retrieving the data using the CAPTAN data acquisition software. This software contains a custom class created for the SASA system, which is able to interface with the array, programming it by sending commands through the Ethernet. Commands sent by the software are interpreted by the NPCB and forwarded to the array. The software is also capable of processing and displaying the data broadcasted by the SASA in real-time due to the wide link bandwidth.

Review Copy

Due to the scalability of the SASA system, there is the possibility of networking multiple microphone array boards and creating a highly versatile smart acoustic sensing system.

1.5 SASA system operation

The SASA system operation is controlled by a GUI interface software, specifically designed for the sound sensor array based on the CAPTAN architecture. This software has the capability of programming the hardware of the microphone array with the following parameters; sampling rate, ADC sampling rate, and number of sensor elements to be read. At system level, the GUI can select the board IP's with which the user wishes to establish connection. Once the GUI programs the array or arrays, the system goes into data acquisition mode. In data acquisition mode, the boards are continuously sending data to the computer and the computer pipes this data to a file on the hard disk. At the same time, a second thread of the software provides visualization of the raw data on the screen if the system is set for raw data mode. If the system is set instead to processed data mode, the computer simply dumps the data in a file for subsequent analysis.

There are three modes of operation that the system can run: raw mode, processed mode and mix mode. The raw mode simply sends the following information to the computer in a 64-bit word: board ID, microphone number and 12 bit acquired data. The processed mode, on the other hand, is dependent on the configware designed by the user. It can vary from a simple signal averaging to FFT's and more complex analysis as desired. The mix mode transmits both raw and processed data in real-time. The Ethernet communication bandwidth is 800 Mbps which by far exceeds the acquired raw data bandwidth of 22 Mbps for a 36 K samples/seconds. The

Review Copy

difference between the communication and raw data bandwidths allows the simultaneous transmission of raw and processed data in real-time. All network communications are transparent to the user and are managed by the CAPTAN system. The user must set the IP switches on the GEL board before any network communication can take place.

A case study to demonstrate the system operating in raw mode is illustrated by Figure 1.7. In this test, the sound source is an omnidirectional microphone (1 cm diameter) emitting a single tone tuned at 4 kHz and located at 20 cm away from the microphone array. In the right side of Figure 1.7, the sound source is placed to the left of the center of the array board, while in the left side of Figure 1.7, the sound source is placed to the right side of the center of the array board. It is also important to observe that for this test the array is not calibrated and therefore, numeric results are not absolute but relative. Nevertheless, it is possible to observe that even with the uncalibrated raw data mode, important information about the sound intensity distribution can be obtained by the SASA system.

1.6 SASA system calibration

It is necessary to calibrate the SASA system in order to conduct reliable experiments. The first step of the calibration is to calibrate the individual gain of each channel on the array. The test stand used for this end is shown in **Error! Reference source not found.**Figure 1.8 and is composed of the array itself, a CDMG13008L sound wave generator (Cui 2006) and the clamp holder. The scale in this figure is in μPa .

Review Copy

The next step for the calibration consists of acquiring data with the array, positioning the sound source at 1 cm from the target microphone and acquiring one second of data. This process is repeated 52 times for the entire sensor array. The result of this process is presented in Figure 1.9 and provides the mask for which all the data will be modified so that the array responds homogeneously when excited. This also provides an absolute calibration of the array since at 1 cm the sound pressure was set to 5 μPa . In this case, as shown in **Error! Reference source not found.**Figure 1.9, we have normalized the scale set from 1 to 0 where 1 corresponds to 5 μPa . The next calibration is based on a chirp waveform ranging from 100 Hz to 8 kHz to extract the frequency response of every sensor. As expected, the frequency response is homogeneous throughout the array. Figure 1.10 shows the frequency mean, minimum and maximum responses of all 52 acoustic sensors.

1.7 Sensor array for time-of-flight measurements

To create a controlled environment for acoustic experimentation and Time-of-Flight (TOF) measurements, a 52"x52"x27" anechoic chamber was designed and built. The key features of the chamber are its ability to isolate the experiment inside the chamber from outside noise and to absorb sound inside the chamber to prevent multiple reflections (reverberation). It should be noted that while the noise from outside of the chamber could be either reflected back outside or absorbed by the chamber, sound inside the chamber has to be absorbed by the surfaces of the chamber to prevent reflections. The material used for sound absorption was a high density 2" thick polyester-based polyurethane convoluted foam which is specifically designed for sound absorption. Furthermore, a sensor array test stand was designed and built in order to conduct a

Review Copy

wide variety of acoustic and ultrasound experiments. Figure 1.11 shows the outside view of the assembled anechoic chamber and a picture of the sensor array test stand.

A time-of-flight estimation experiment was performed to determine the direction of and distance of the sound source with respect to the microphones. The TOF was calculated based on the recorded phase delay and the distance between the receiving microphones. The first set of experiments involved the CAPTAN based microphone array data acquisition system and a transmitting sound source. The microphone array data acquisition system was mounted on the vise base, the transmitting sound source was also mounted on a vise base. The experiment was carried out in a laboratory room with various random objects around the area of experimentation resulting in a high noise and a highly reflective environment. The parameters varied in this experiment were the distance to the sound source, frequency of the sound source, the pairs of microphones used, and the angle between the sound source and the microphones. The overall geometry of the experiment is shown in Figure 1.12.

For the first test, the distance between the receivers and transmitter was two feet and for the second test the distance was 5 inches. The sound source signal used was a continuous sine wave of frequencies 1 kHz and 2 kHz, generated by an arbitrary waveform generator. The distance between the inner set of receivers, as shown in Figure 1.13, was 10.0 mm, and the distance between the outer set of receivers, was 70.0 mm. The upper frequency was limited to 2 kHz to allow all microphones of the microphone array system to be used without aliasing. The maximum source signal frequency which could be used can be obtained from the spatial sampling theorem. This theorem states that for a given maximum temporal frequency in the

Review Copy

source signal, there is a minimum spatial sampling, i.e., there is a maximum distance between the receivers used in the acquisition system. Specifically, this maximum distance is given by $\delta \leq c / (2 * f_{max}) = \lambda_{min} / 2$, where δ is the maximum distance between the receivers, c is the speed of sound, f_{max} is the maximum frequency of the source signal, and λ_{min} is minimum wavelength of the source signal.

A sample of the overall data collected by the MEMs array is shown in Figure 1.13. In this figure, each window represents the data collected by one of the MEM microphones. The sine wave pattern of the data collected by two of the microphones is compared in Figure 1.14. Here, the phase difference between the data collected by the two MEM microphones can be clearly seen. The sets of microphones used for experimentation were the outer set and the central inner set (see Figure 1.13). For each set, the delay was obtained from each pair of opposing microphones; the results from each pair were then averaged to obtain the delay measurement for the set. Figure 1.15 shows a results graph comparing the expected results versus those collected at 2 feet and at 5 inches for the 2 kHz source signal. It can be seen from these results that phase based TOF in a highly noisy and reverberant environment does not produce reliable results at larger distances. This can be seen from the mismatch between the collected data and the expected results, and also from the fact of the time delays increase and decrease, instead of just increasing when the angle between the sound source and receiving microphones increases. The phase based TOF estimation works better at close distances where the power of the original signal is high compared to the power of the reflection based noise.

Review Copy

In order to reduce the effect of reflection and ambient noise, a second set of experiments was performed inside the anechoic chamber. These experiments also used the CAPTAN based microphone array data acquisition system and a transmitting speaker. The experimental setup was the same as in the first set of experiments except for the use of the anechoic chamber.

Figure 1.16 shows a results graph comparing the expected results versus those collected at 2 feet and at 5 inches for the 2 kHz source signal. From these results, it can be seen that performing phase based sound source TOF inside an acoustic chamber, which absorbs sound and thus reduces reflections, produces some improvement over performing sound source TOF estimation in a general room environment. Furthermore, the frequency of the source signal and the distance between the receiving microphones have an effect on the accuracy of the results.

Since the surrounding surfaces inside the anechoic chamber absorb most of the sound, the reflections causing the distortions in this set of experiments came from the microphone array itself. In order to reduce the effect of the reflection, an alternate physical setup and acquisition system was used for a third set of experiments. The physical setup consisted of a vise with generic 60° beam angle microphone attached to each of the two arms of the vise through foam with nothing in between the vise arms. The distance between the microphones was increased to 6.3". To comply with the spatial sampling theorem, the frequency of the source signals was changed to 700 Hz with the greater distance between the receivers. Figure 1.17 shows a graph comparing the expected results versus those collected at 2 feet for the 700 Hz source signal. It can be observed that by increasing the distance between the microphones and removing any reflective surfaces from in between the microphones significantly improves the accuracy of phase based TOF measurements even at larger distances. For the 700 Hz sound signal the

Review Copy

measured delays follow the correct pattern but the delays are a reduced version of the expected values. This bias in estimation is again attributed to the reflection within the environment.

To further reduce the effects of reflection on the TOF, another set of experiments was performed. Here, the physical setup and the test parameters were the same as for the third set of experiments except for the type of sound source signal. This time instead of using a continuous sine wave for the source signal only a 20 cycle sine wave was transmitted. Then, only the first cycle of the received pulsed sine wave is used for the phase measurement. This is expected to further reduce errors due to reflections since the first sine cycle arrives before undesirable wave reflections, and thus the phase information should be immune to distortion. Figure 1.18 shows a graph comparing the expected results versus those collected at 2 feet for the 700 Hz source signal. From these results it can be observed that phase based TOF measurement which uses the phase information from only the first wave pulse provides accurate results even at larger distances and independent of the frequency of the source signal.

1.8 3D sound source localization

2D localization of the sound source can be performed with 3 receivers (Benesty et al. 2008) using only the TDOA (time difference of arrival) information. The geometry for 2D localization using three receivers arranged in a line is shown in Figure 1.19. As shown in this figure, the location of the transmitter in a plane can be obtained by measuring the time difference of arrival between receivers 1 and 2, and also receive 1 and 3.

Review Copy

3D localization can be decomposed into two problems of 2D localizations as shown in Figure 1.20. Here, the transmitter is labeled Tx , and the receivers are labeled Rx_1 through Rx_5 . Three receivers in each plane are used in the same way that they were used for 2D localization, with both results using the same coordinate system centered at receiver 5. Geometrically, this is the same as finding two circles or semi-circles, the intersection of which is the location of the sound source.

An experiment consisting of six 40 kHz ultrasonic sensors, five of which acted as receivers and one of which acted as a transmitter, was performed for 3D source localization. The transmitter used a 20 cycle sine wave pulse train. The sensor array test stand as was used to hold the transmitter and receivers. The geometry and dimensions of the setup are shown in Figure 1.21. The receivers were arranged in a plus sign pattern. The results of this set of experiments are shown in Table 1. As presented in the table, this experiment for 3D sound source localization produces accurate results and the collected values are within a few percent of each other.

1.9 SASA system for mapping of the heart sound

The SASA system for source separation was tested to monitor and localize the heart sound. Figure 1.22 shows the system collecting data on the chest of the subject. The system was configured with eight microphones, with the sampling rate of 36K samples/second for a total digitizing time of 11 seconds. Each heart beat in a healthy adult human is composed of basically two distinct sounds. The first heart sound (S1) is caused by the Mitral (M) and Tricuspid (T) atrioventricular valves and the second heart sound (S2) is caused by the Aortic (A) and

Review Copy

Pulmonary (P) semilunar valves. The localization of these heart valves on the human chest is illustrated in Figure 1.23 (Bates 2005).

The results from the digitization are shown in Figure 1.24, although 400K samples were captured, only 200K are displayed on this picture for clarity. It is important to observe that this time the signals on the microphones are very different from each other. Since the array is very close to the distributed sound source due to the position of the microphones (see Figure 1.25), they will have very different localized signals. Figure 1.26 displays sound localization information collected during the heart experiment. The heart is here imaged using the sound intensity measured at each microphone, while a subsequent interpolation of the data was done and illustrated on the same picture. It is possible to observe two very distinct patterns for S1 and S2 on the image.

1.10 Conclusion

This work describes the design, development and capabilities of a scalable SASA system that can be used to enhance sound processing and acquisition for a diverse set of applications. The system uses MEMS microphone arrays associated with the CAPTAN scalable architecture in order to deliver a powerful real-time signal processing and data acquisition platform. This research demonstrates that integration of key technologies such as MEMS sensors, high performance FPGA devices, and Gigabit Ethernet can produce a very compact, network enabled, versatile acoustic array with very high performance applicable to a broad range of applications. The SASA platform can benefit many areas of acoustic signal processing, specifically multiple source separation, mapping and localization. Other possible applications include digital cardiac

Review Copy

auscultation, structural health monitoring using acoustic emission, robotic auditory systems, voice based man-machine interfacing and ultrasound beacon based object tracking.

1.11 References

Alghassi, H., Eye Array Sound Source Localization, Doctoral dissertation, University of British Columbia, Vancouver, (2008).

Bates, B., A Guide to Physical Examination and History Taking, Lippincott Williams & Wilkins, 9th ed., (2005).

Benesty J., Chen, J., Huang, Y., Microphone Array Signal Processing, Springer, Berlin, 1st ed., Vol. 1, (2008).

Brandstein, M., and Ward, D. (Eds.), Microphone Arrays Signal Processing Techniques and Applications, Berlin Heidelberg: Springer-Verlag. (2001).

Buck, M., Haulick, T., and Pfleiderer H., "Self-Calibrating Microphone Arrays for Speech Signal Acquisition: A Systematic Approach", Applied Speech and Audio Processing, Vol. 86, No.6, pp. 1230-1238, (2006).

Campbell, D.K., Adaptive Beamforming Using a Microphone Array for Hands-Free Telephony, Master's Thesis, Bradley Department of Electrical and Computer Engineering, (1999).

CUI Inc, "CDMG13008L-02 Micro Dynamic Speaker", 2006.

Eckert, J, German, R., and Dressler, F., "An Indoor Localization Framework for Four-Rotor Flying Robots Using Low-Power Sensor Nodes", IEEE Transactions on Instrumentation and Measurement, 60 (2), pp. 336-344, (2011).

Hodgkiss, W. "Dynamic Beamforming of a Random Acoustic Array", IEEE Acoustics, Speech and Signal Processing, IEEE International Conference on ICASSP '80, Vol. 5, pp. 311- 314, (1980).

Kim, S. J., and Kim, B. K., "Accurate Hybrid Global Self-Localization Algorithm for Indoor Mobile Robots with Two-Dimensional Isotropic Ultrasonic Receivers", IEEE Transactions on Instrumentation and Measurement, pp.(99), pp.1-14, 0 (2011).

Knowles acoustics, "SPM204 and SPM208 Datasheet Rev C", 2006.

Kunin, V., Cardoso, B., Saniie, J., and Oruklu, E., "Acoustic Sensor Array for Sonic Imaging in Air", Proceedings of the IEEE Ultrasonics Symposium, pp. 1833-1836, (2010).

Kunin, V., Jia, W., Turqueti, M., Saniie, J., and Oruklu, E., "3D Direction of Arrival Estimation and Localization using Ultrasonic Sensors in an Anechoic Chamber", Proceedings of the IEEE Ultrasonics Symposium, pp. 756-759, (2011).

Llata, J. R., Sarabia, E. G., and Oria, J. P., "Three-Dimensional Robotic Vision Using Ultrasonic Sensors. Journal of Intelligent and Robotic Systems", 33 (3), pp. 267-284, Berlin Heidelberg: Springer-Verlag. (2002).

Review Copy

Nakadai, K., Okuno, H.G., Kitano H., "Real-Time Sound Source Localization and Separation for Robot Audition", IEEE International Conference on Spoken Language Processing, pp. 193-196, (2002).

National Semiconductor, "ADC121S101 Datasheet", 2010.

Nishitani A., Nishida Y., and Mizoguch H., "Omnidirectional Ultrasonic Location Sensor", IEEE sensors conference, October 2005-November 2005, Irvine, CA., (pp. 4). (2005).

Rabinkin, V., D., Renomeron, J., R., French, C., J., and Flanagan, L., J., " Estimation of Wavefront Arrival Delay Using the Cross-Power Spectrum Phase Technique", 132nd Meeting of the Acoustic Society of America, Honolulu, HI. (1996).

Rivera, R., Turqueti, M., and Prosser, A., "A Software Solution for the Control, Acquisition, and Storage of CAPTAN Network Topologies", IEEE Nuclear Science Symposium Conference Record, pp. 805-808, (2008).

Turqueti, M., Rivera, R., Prosser, A., Andresen, J. and Chramowicz, J., "CAPTAN: A Hardware Architecture for Integrated Data Acquisition, Control and Analysis for Detector Development," IEEE Nuclear Science Symposium Conference Record, pp. 3546-3552, (2008).

Shaw, B., "Source Localization and Beamforming", IEEE Signal Processing Magazine, 19 (2), (2002).

Weinstein E., Steele, K., Agarwal, A., Glass, J., "A 1020-Node Modular Microphone Array and Beamformer for Intelligent Computing Spaces", MIT/LCS Technical Memo, MIT-LCS-TM-642, (2004).

Review Copy

Table 1. Experimental results for measured 3D distances

| From Receiver | Collected (inches) | Expected (inches) | Error (inches) |
|---------------|--------------------|-------------------|----------------|
| 1 | 37.921 | 37.26 | 0.661 |
| 2 | 37.378 | 36.71 | 0.668 |
| 3 | 34.807 | 34.04 | 0.767 |
| 4 | 35.213 | 34.61 | 0.603 |
| 5 | 36.567 | 35.53 | 1.037 |

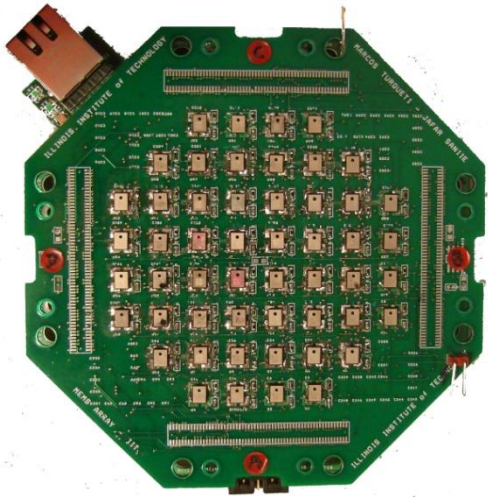


Figure 1.1. Top view of the AMA board hosting the MEMS elements

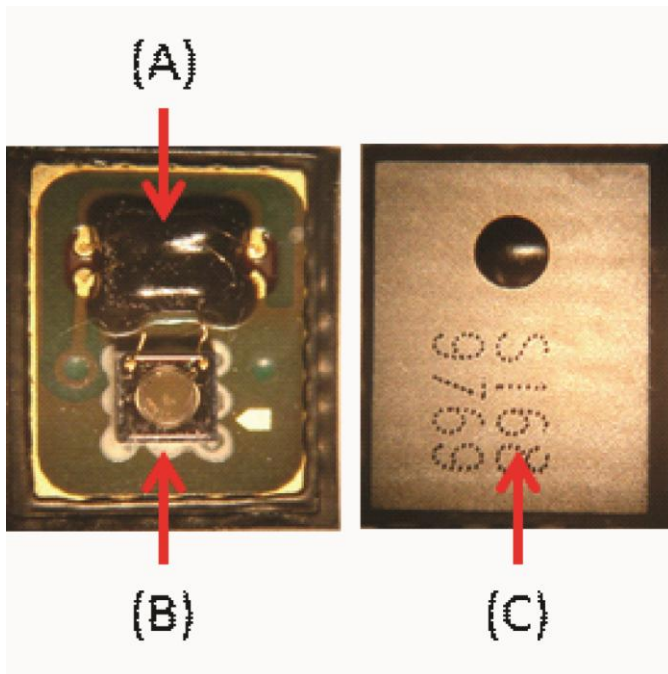


Figure 1.2. MEMS microphone, (A) amplifier, (B) microphone and (C) aluminum cover

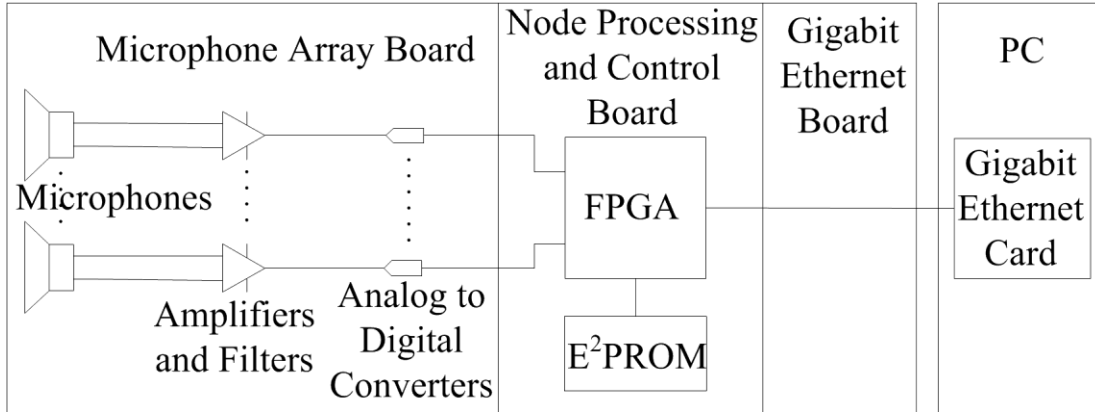


Figure 1.3. Functional block diagram of the smart acoustic sensor array (SASA) system

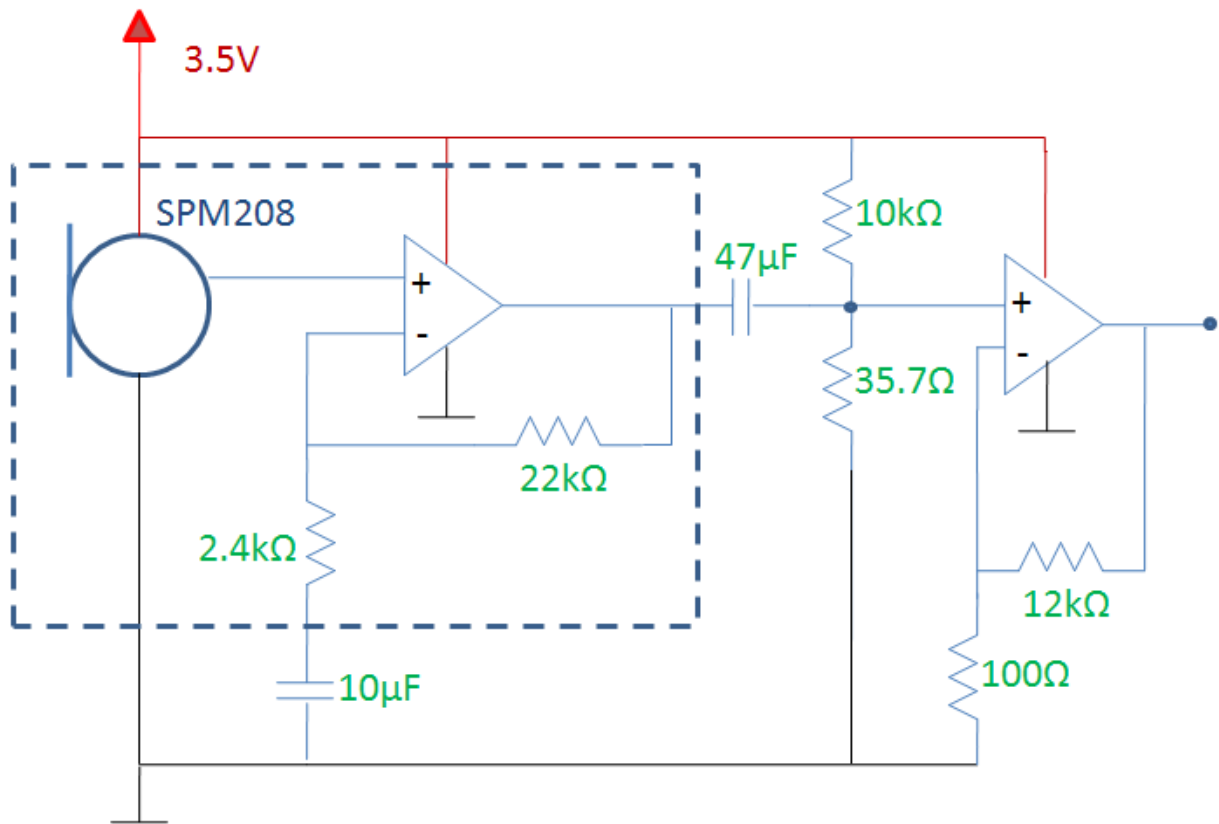


Figure 1.4a. Frontend electronics of the AMA array.

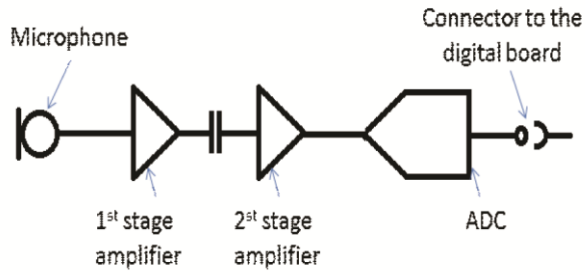


Figure 1.4b. Two-stage amplifier circuit for each microphone

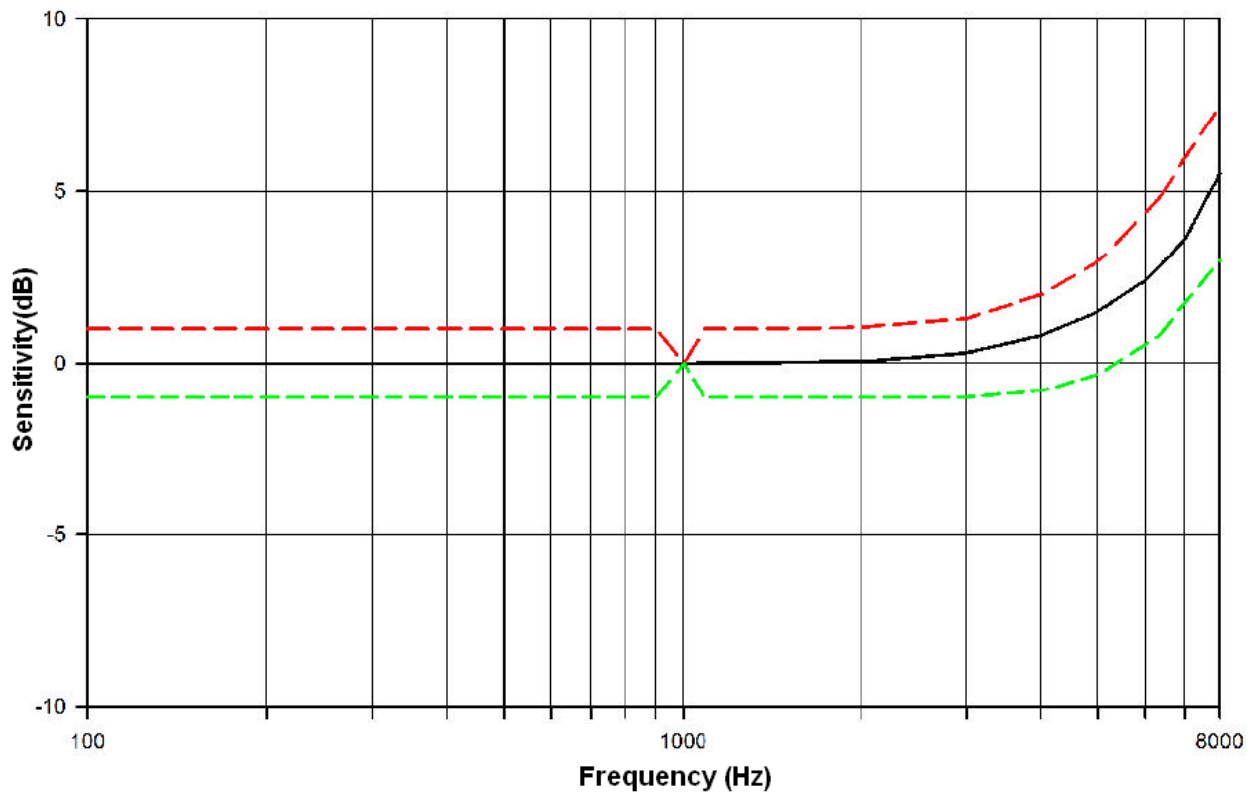


Figure 1.5. SPM208 frequency response. The lines on the top and on the bottom are the error margins.

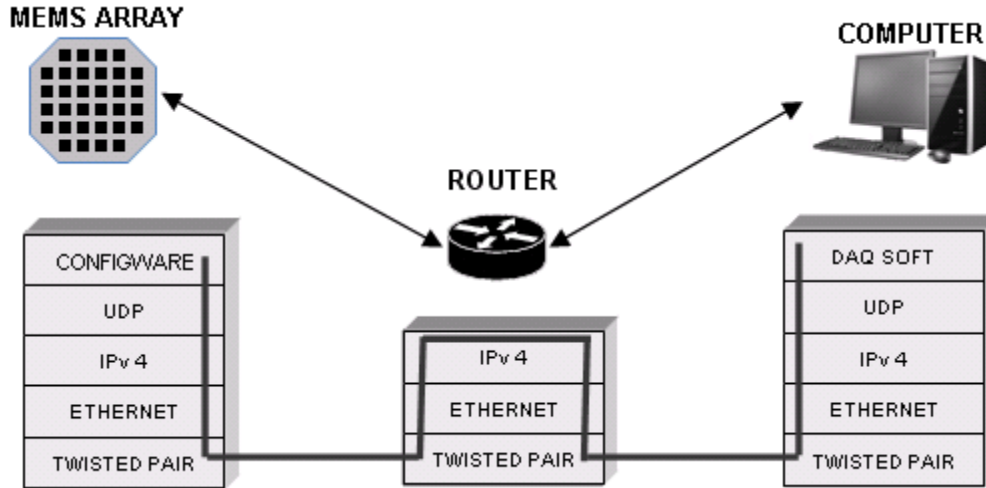


Figure 1.6. Data path and framing overview representing several layers of data encapsulation that allows the data flow through the Ethernet.

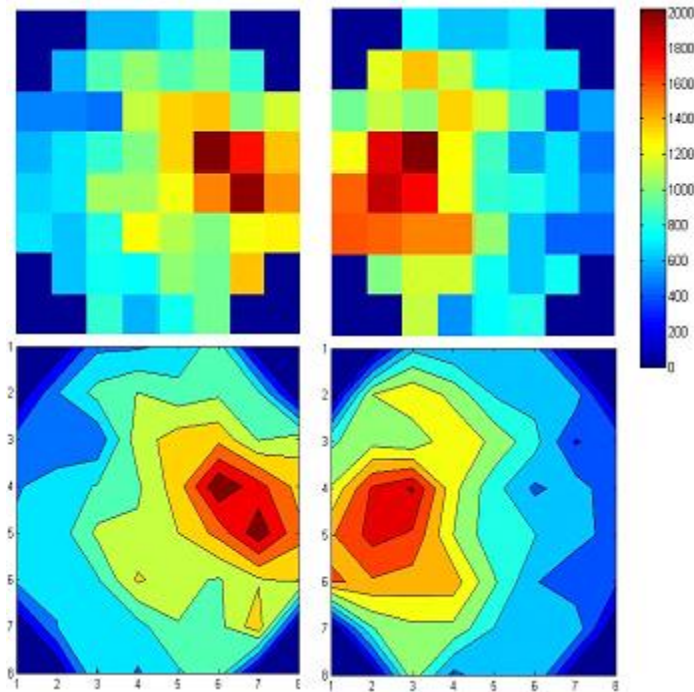


Figure 1.7. A 2D display of the system acquiring data in two different situations. Top 2D images are the raw data. Bottom 2D images are the subsequent interpolation of the raw data.

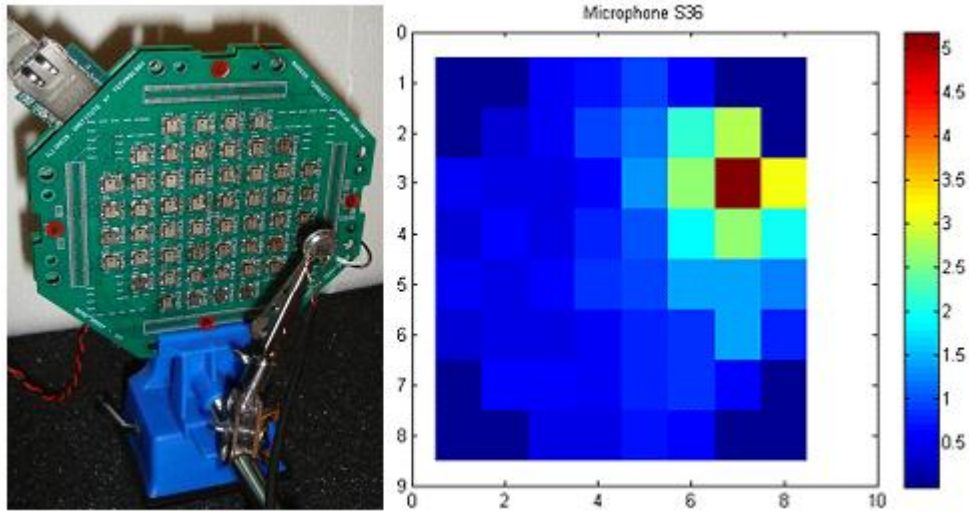


Figure 1.8. On the left the test setup; on the right the mapping of the array responding to the stimulus of the microphone.

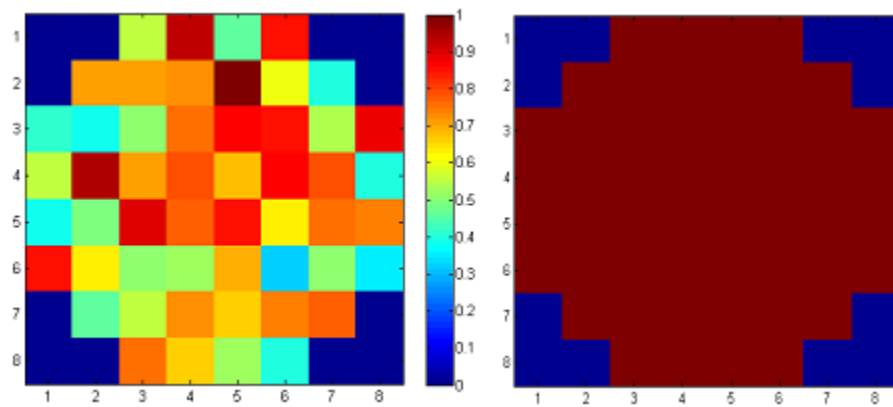


Figure 1.9. On the left the gain map of the array; on the right the array digitally equalized.

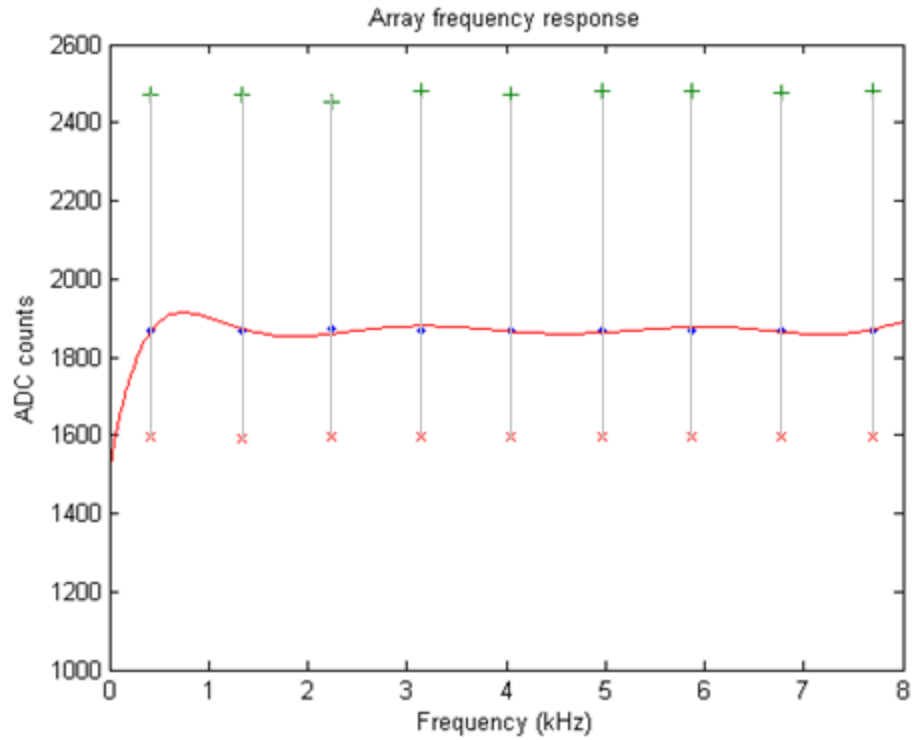


Figure 1.10. Frequency response of the array with the mean represented by the solid line, the maximum and minimum gain are represented by the bars.

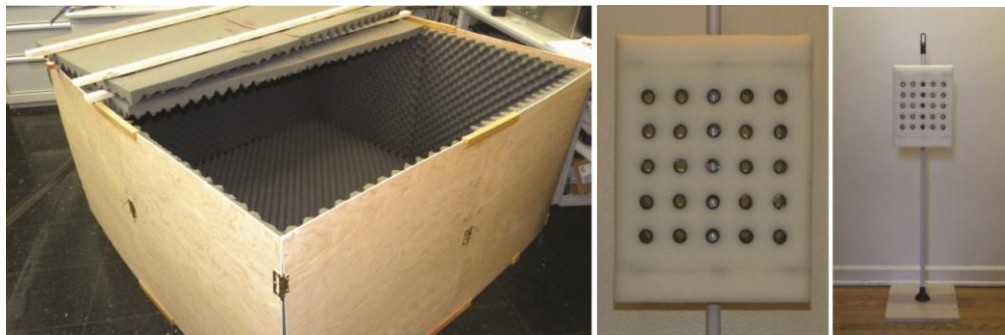


Figure 1.11. Anechoic chamber and sensor array test stand

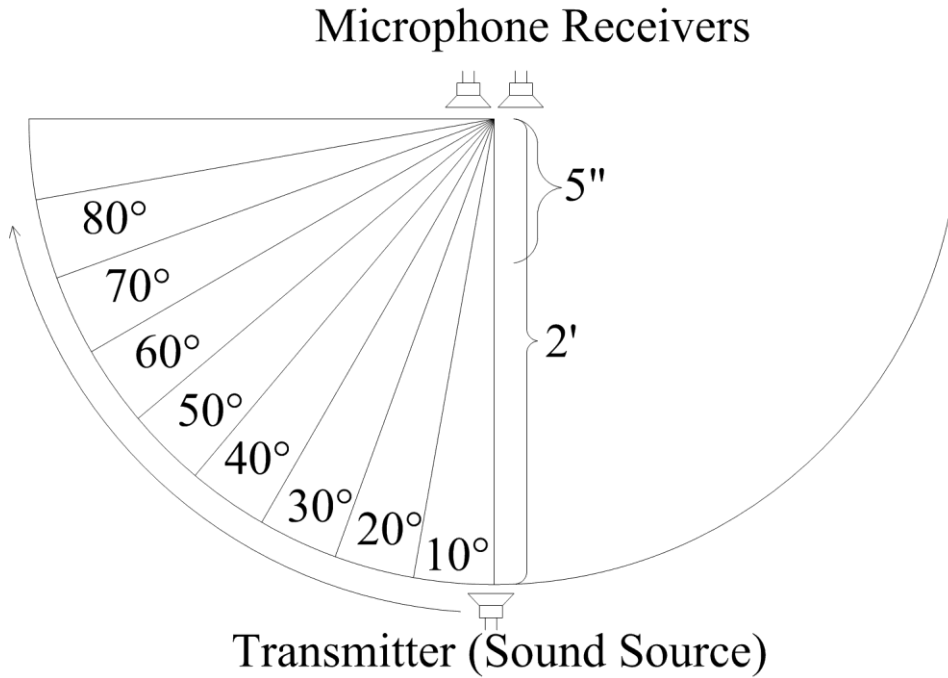


Figure 1.12. Sound source TOF estimation setup

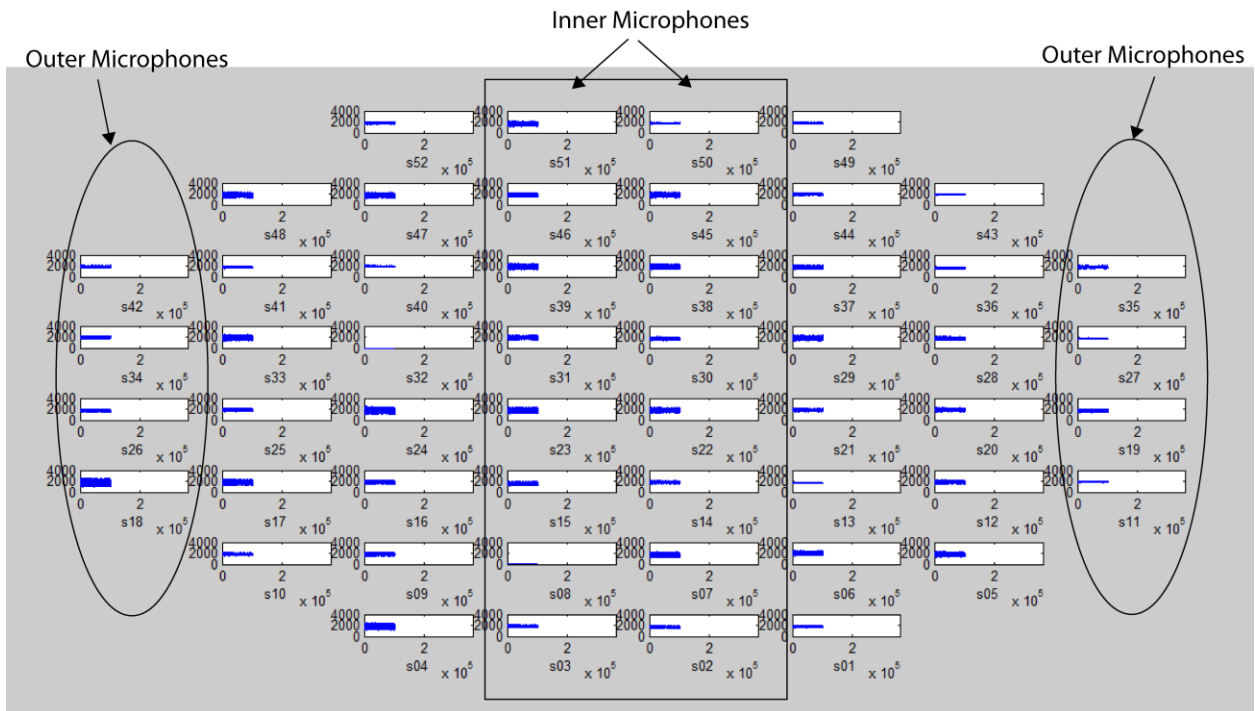


Figure 1.13. Sample data collected (blue patterns) by MEMS microphone array

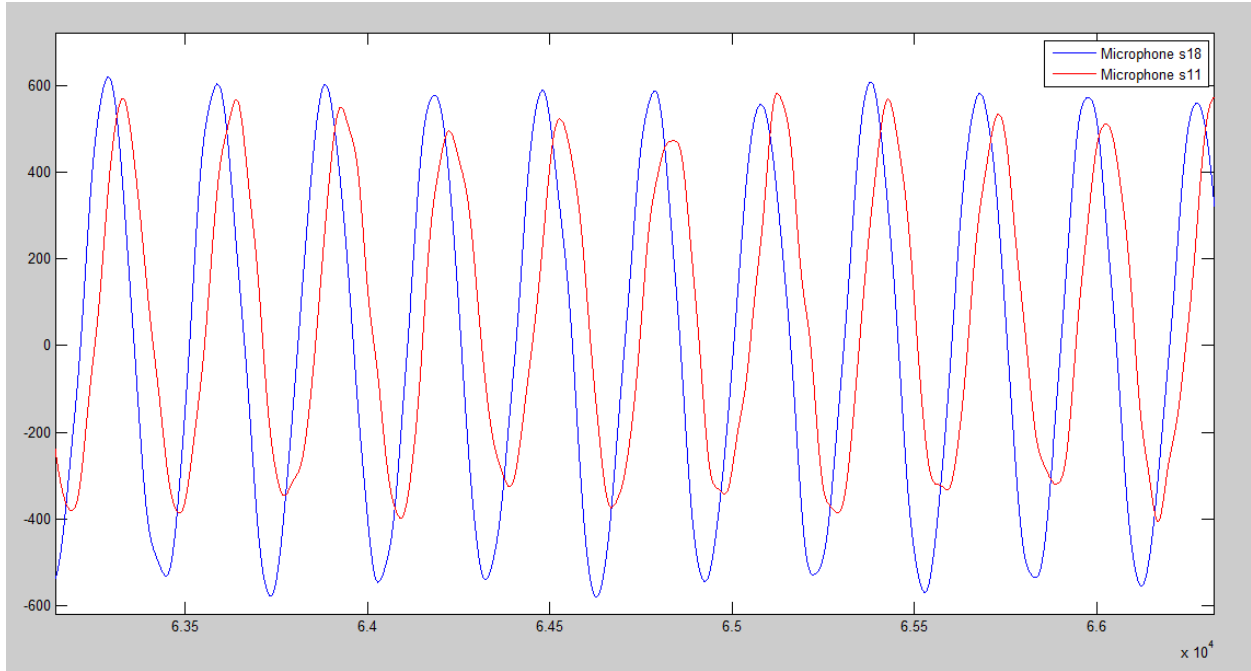


Figure 1.14. Data collected by two MEMS microphones

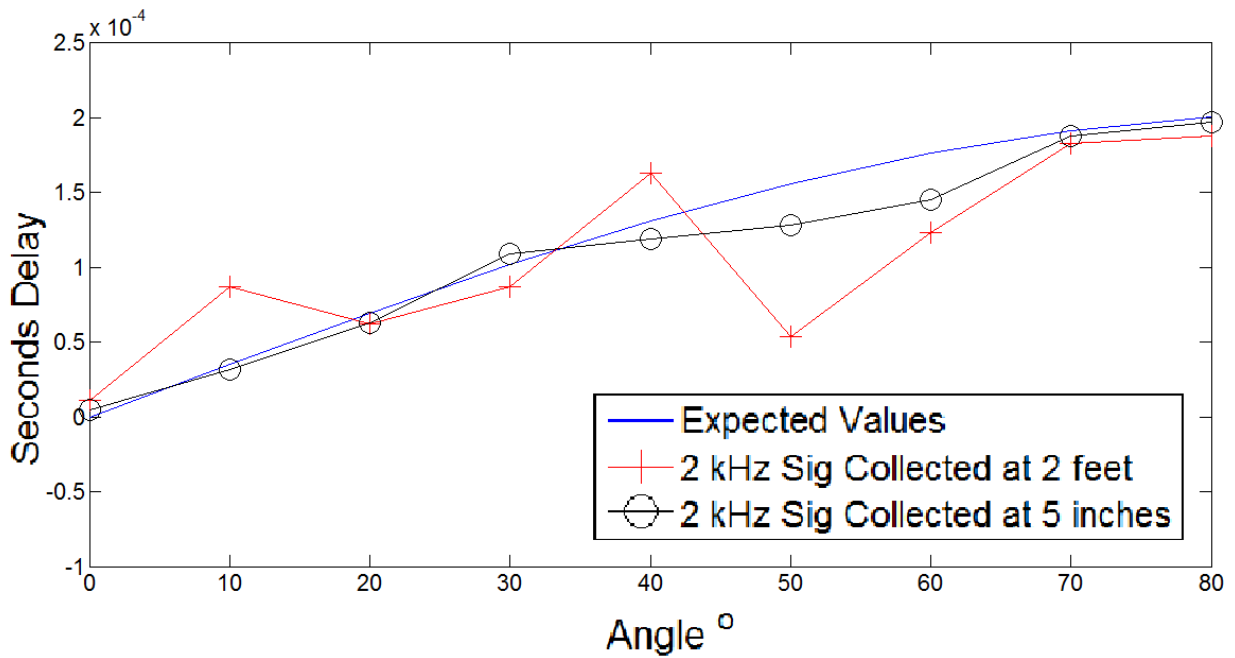


Figure 1.15. TOF estimation experiment in a room environment for 2 kHz sound source at 2' and at 5"

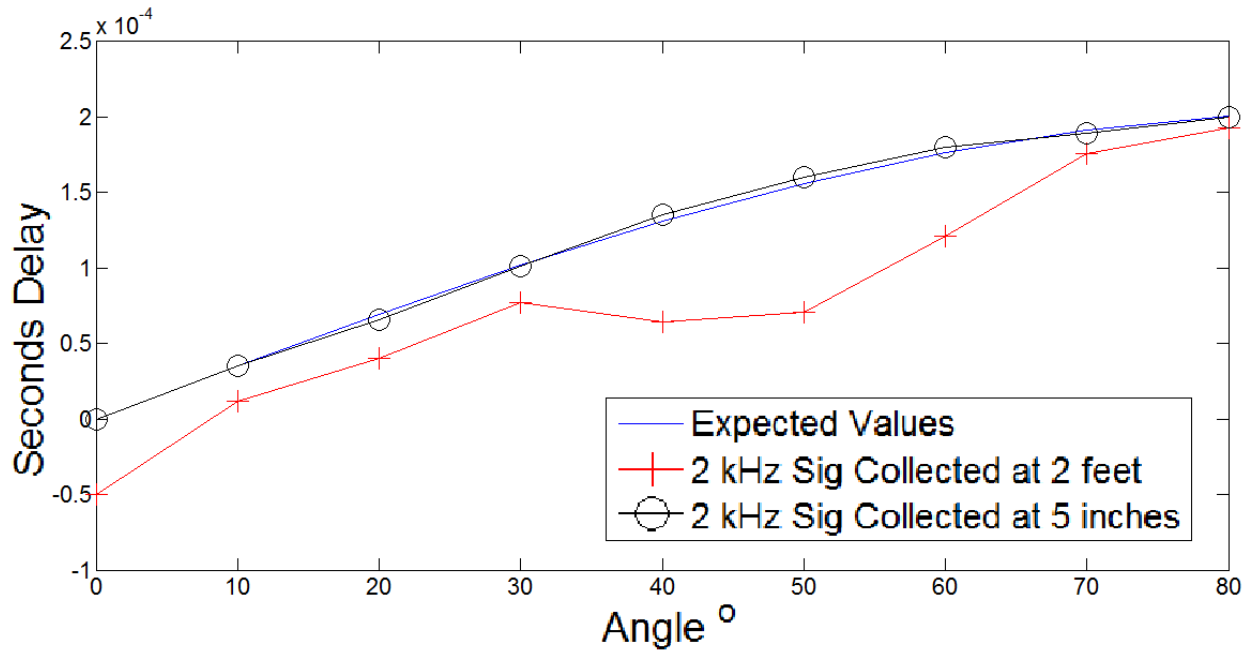


Figure 1.16. TOF estimation experiment in an anechoic Chamber for 2 kHz sound source at 2' and at 5''

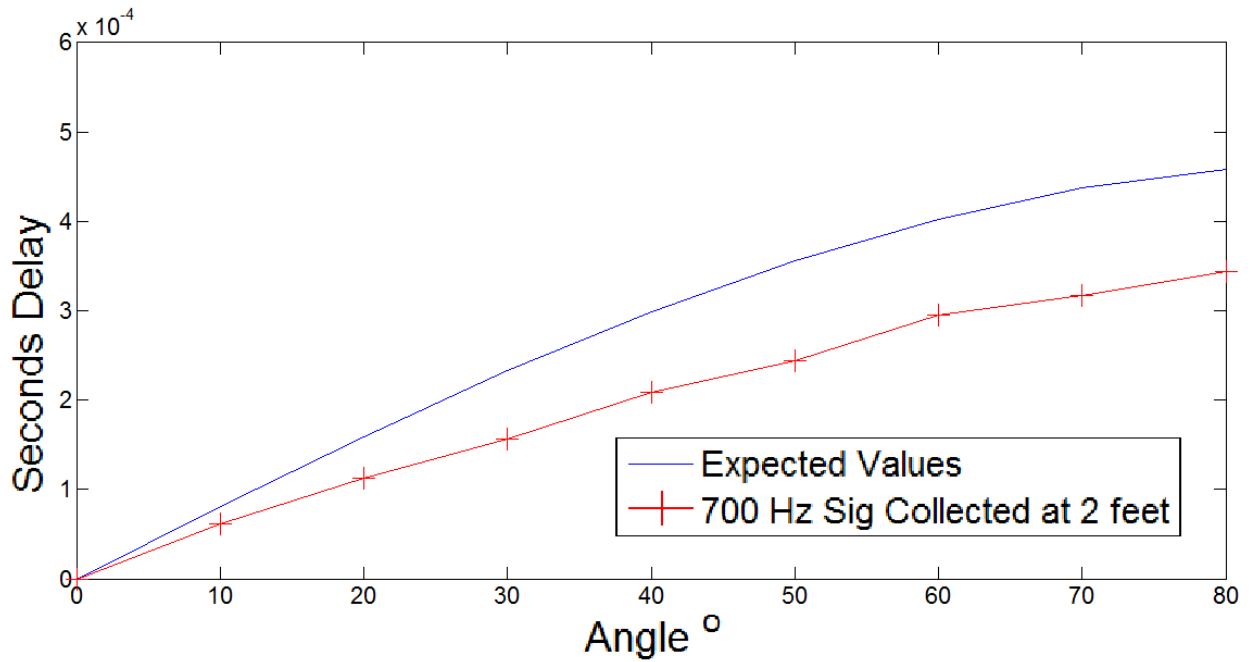


Figure 1.17. TOF estimation experiment in an anechoic chamber with sensor array test stand using 700 Hz sound source at 2-feet away from the microphone.

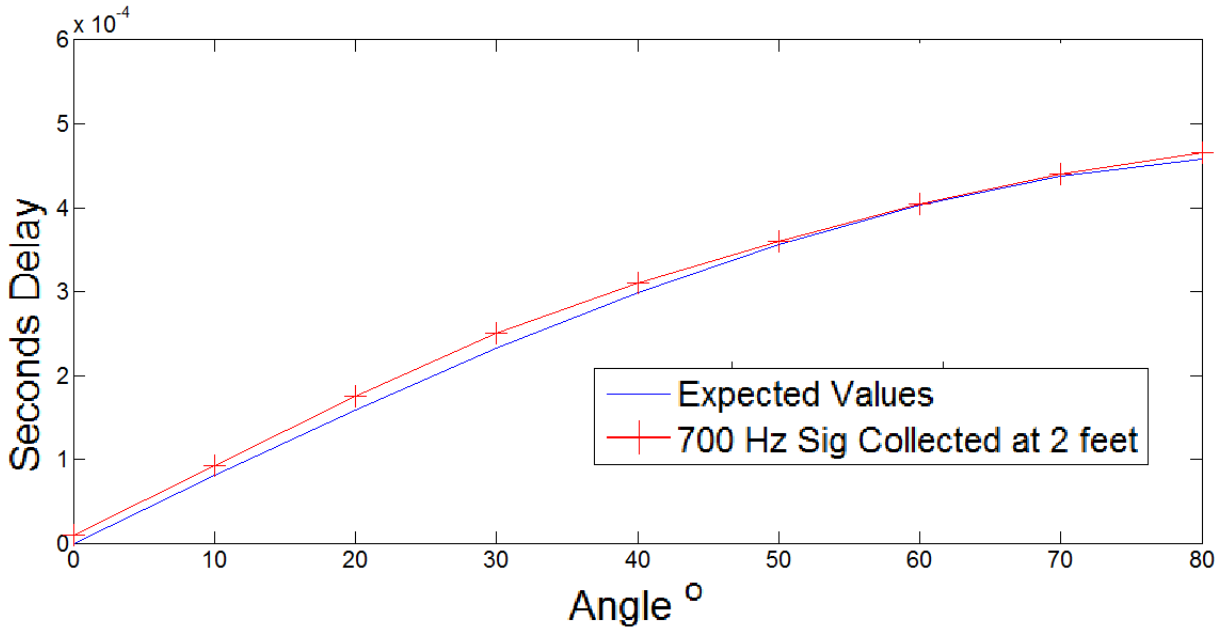


Figure 1.18. TOF estimation experiment in an anechoic chamber with sensor array test stand using a 20 Cycle 700 Hz sine wave sound source at 2-feet away from the microphone.

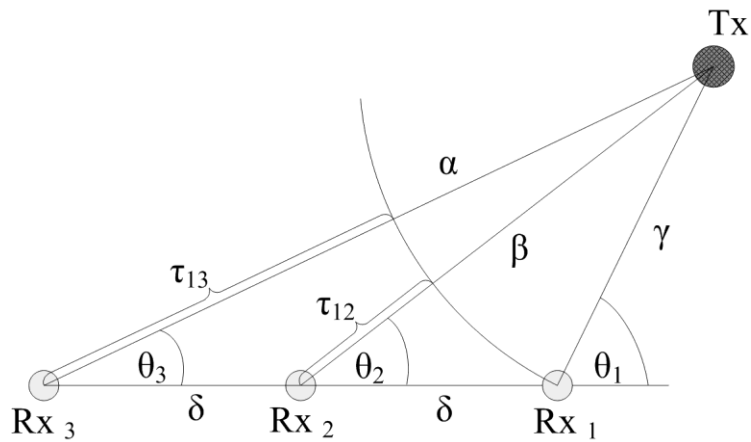


Figure 1.19. 2D localization geometry

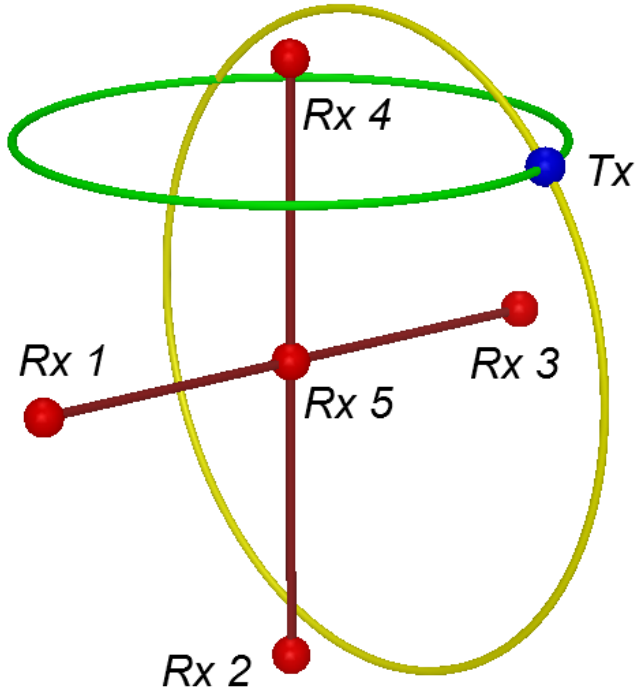


Figure 1.20. 3D localization geometry

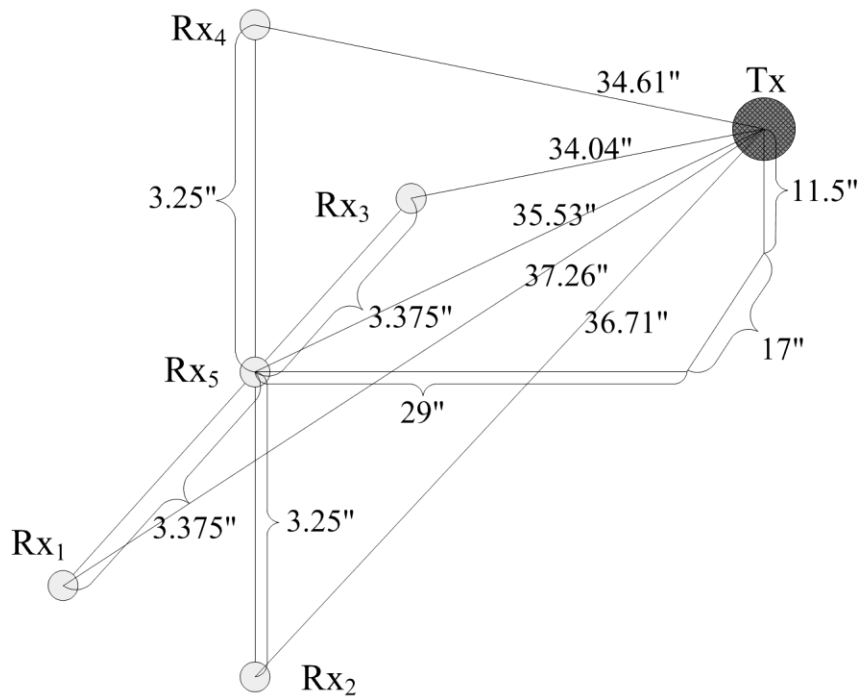


Figure 1.21. 3D Distances between the receivers and the transmitter

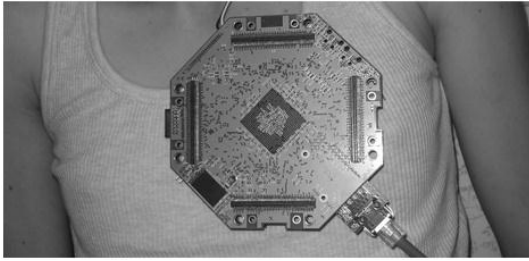


Figure 1.22. System collecting data on the subject's heart.

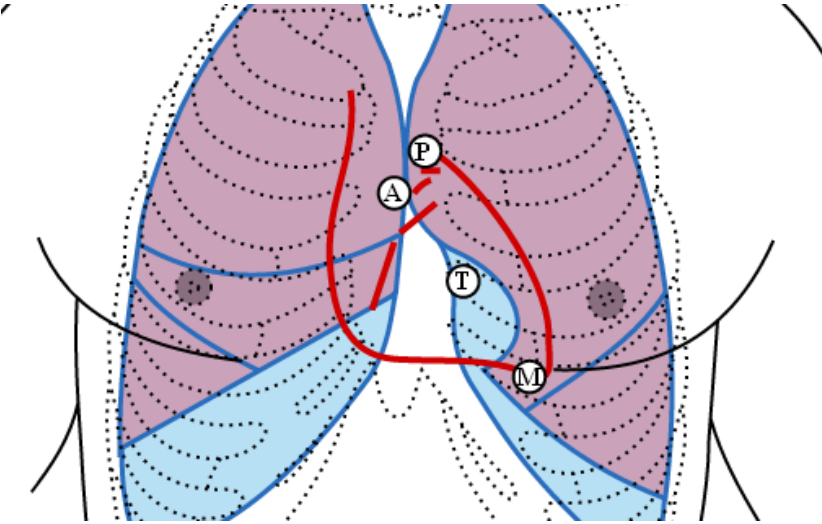


Figure 1.23. Localization of heart valves.

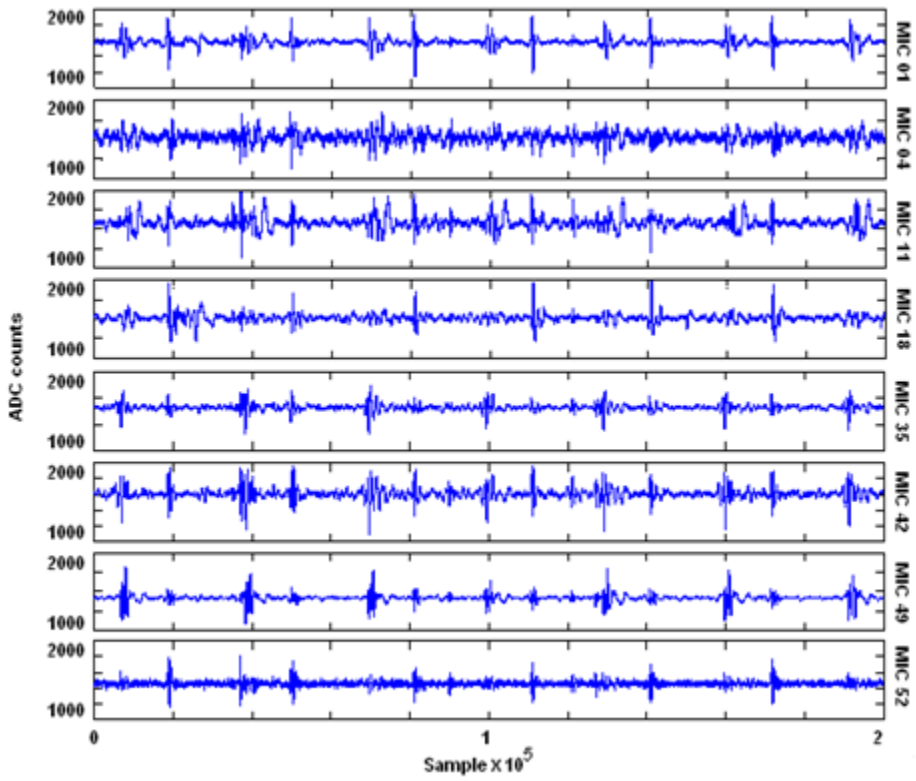


Figure 1.24. From the top to the bottom data acquired by microphones 1,4,11,18,35,42,49 and 52.

| | | | | | | | |
|----|----|----|----|----|----|----|----|
| | | 52 | 51 | 50 | 49 | | |
| | 48 | 47 | 46 | 45 | 44 | 43 | |
| 42 | 41 | 40 | 39 | 38 | 37 | 36 | 35 |
| 34 | 33 | 32 | 31 | 30 | 29 | 28 | 27 |
| 26 | 25 | 24 | 23 | 22 | 21 | 20 | 19 |
| 18 | 17 | 16 | 15 | 14 | 13 | 12 | 11 |
| | 10 | 9 | 8 | 7 | 6 | 5 | |
| | | 4 | 3 | 2 | 1 | | |

Figure 1.25. Numbering convention for the microphone array

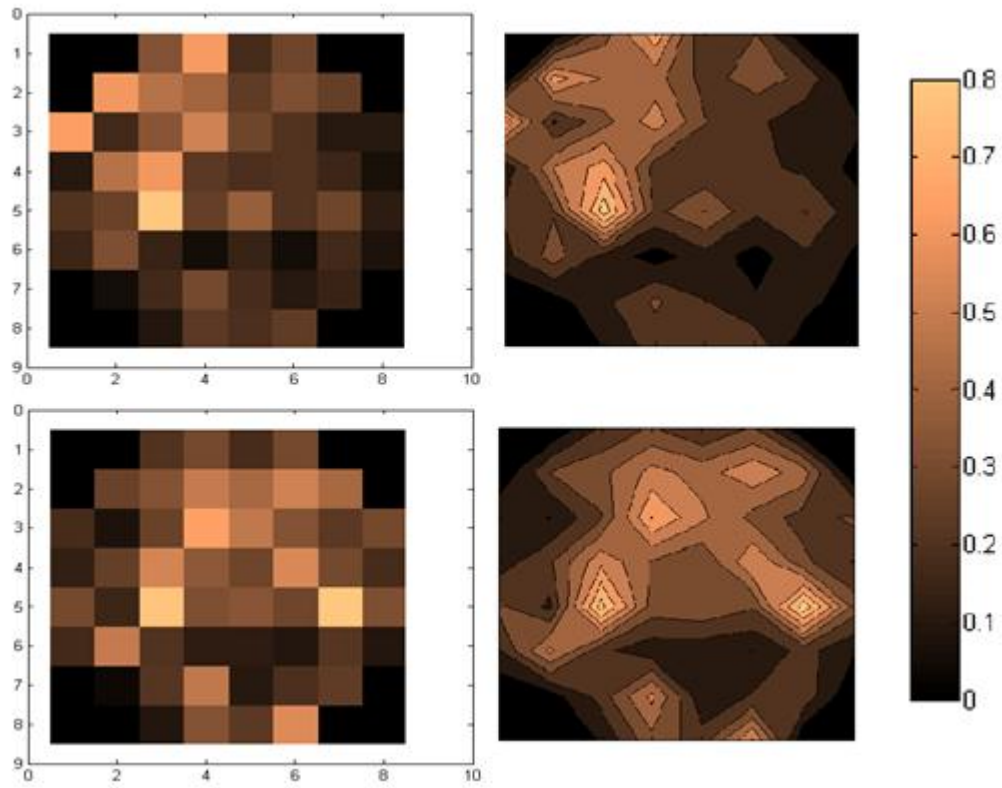


Figure 1.26. Sound imaging of the heart. On the top the sound image captured when the heart beat was at the S1 stage and on the bottom when it was at S2 stage. The 2D images on the right side are the subsequent interpolation of the raw 2D images on the left side.

Review Copy