Development of an affordable ultrasonic recording array

Zachary Levy

Ultrasound is useful for a broad range of scientific, engineering, and educational purposes, such as for studying bioacoustics. The study aimed to develop a low cost ultrasonic array capable of acting as a recording and analysis device for frequencies between 20Hz and 85kHz. While such devices are available commercially, they are typically expensive and therefore inaccessible for many use cases. Ultrasonic sensors were selected and arranged to optimize the antenna pattern of the device, which consists of MEMS microphones, a custom circuit board, and a microcontroller. A software program was written in Python to derive the frequency spectrum of the ultrasonic signal and allow for conventional beamforming. The device was tested by measuring various frequencies of sound between 400Hz and 40kHz. The device was able to detect both audible and ultrasonic sounds, but an issue with data transmission resulted in inaccuracies in the recordings.

Keywords: *ultrasound, recording device, array*

Sound composed of frequencies above the maximum of the human hearing range at approximately 20KHz is referred to as ultrasound. Ultrasound has numerous applications in technology such as for distance sensors and medical imaging (Carovac et al., 2011). A variety of animals can hear and produce ultrasound. Of particular interest are bats, certain species of which are capable of hearing frequencies as high as 150 KHz and utilizing ultrasound in order to echolocate (Koay, Heffner, Bitter, & Heffner, 2003). Therefore, devices which are capable of recording and analyzing ultrasonic signals are useful for both scientific and engineering purposes, and might also be useful for education relating to ultrasound. Although such devices are widely available commercially, they are often expensive and specialized which limits access to the technology.

Sound and ultrasound can be detected using a type of sensor known as a transducer. The design of a transducer determines if it can transmit signals, receive signals, or both as well as its sensitivity to different frequencies and signals from different directions, known as the antenna pattern (McCowan, 2001).

The electrical signal outputted by a transducer is a voltage that varies over time. This signal is analog, meaning that it is continuous and can occupy any voltage in a range of values. In order for the signal to be processed by a computer, it must be quantized and converted to a digital signal. This is done by passing the signal through an analog-to-digital converter (ADC). A measurement of the voltage of the signal, known as a sample, is taken over increments in time. Each sample of the signal can only occupy a discrete value in a range of values. The frequency of the samples, the number of samples of the signal taken each second, is known as the sampling rate (Smith, 1999).

In order for a signal to be properly sampled, there can't be any components in the signal with a frequency greater than half of the sampling rate. Otherwise, these signals will be aliased, meaning that they will appear as lower frequency signals that were not present in the original signal (Smith, 1999). Because the frequency of ultrasound is higher than the frequency of audible sound, the sampling rate must also be higher to avoid aliasing.

Frequency components in a given range of frequencies can be removed from a signal using a filter. Filters can act on analog or digital signals, but the performance of digital filters is typically superior. There are two types of digital filters: finite impulse response (FIR) filters and infinite impulse response (IIR) filters, which are implemented using convolution and recursion, respectively (Smith, 1999).

Recording and processing ultrasonic signals presents numerous challenges. The absorption of ultrasound in air is significant, and is particularly high at higher frequencies which limits the range of a recording device (Bass et al., 1995). Background noise can also make observations difficult, and many common items such as fluorescent lights and fans can produce ultrasound (Holcomb et al., 2015). To overcome these challenges, it is useful to have a device that is both sensitive and highly directional. This can be achieved by arranging multiple transducers into an array, a group of sensors functioning as a single device (McCowan, 2001). A variety of array designs and geometries exist such as linear, grid, and circular arrays.

In a simple linear array, the transducers are arranged along a line at regular intervals and the signals from each transducer are summed together. Signals arriving from perpendicular to the array will reach each element at the same time and reinforce each other, resulting in a signal with a larger amplitude, while signals arriving from other directions may not, resulting in a smaller amplitude.

A variety of techniques can be used to modify the shape of an array's beam, the region the array is most sensitive to, or steer it in a particular direction without mechanical movement of the array. The simplest of these techniques is conventional delay and sum beamforming. This consists of adding a time delay to the signal from each transducer, effectively changing when the wave arrives at each element (McCowan, 2001).

In this study, an array design was combined with beamforming techniques to create a recording device that is useful for a wide range of purposes. Although the device could be mass manufactured, the focus of this study was to design it so that interested users could build the device themselves.

Methods

Array Design

Micro-electromechanical system (MEMS) microphones were selected for the device due to their small size which allows for arrays with small spacings between microphones. The specific microphones used were TDK InvenSense ICS-41352, an omnidirectional bottom port microphone (InvenSense, 2016). This model was chosen due to its availability, high sampling rate, and wideband frequency response. Since the microphones are digital, additional analog circuitry was not required to process the output of the microphones, simplifying the device.

To select a geometry of the array, a program was written in Python to plot antenna patterns for a given array geometry. Several different geometries were tested including linear arrays, circular arrays, and grid-shaped arrays. The array was chosen to have 25 microphones arranged in a square grid with a spacing of 5mm between adjacent microphones due to the design having high directivity and because of the simplicity of the geometry. The antenna pattern of the chosen geometry is shown for 20kHz, 40kHz and 60kHz in Appendix A. The total number of microphones was limited to 25 because of constraints created by the microcontroller that limit the maximum bitrate of the device.

Electrical System

A custom printed circuit board (PCB) was designed for the device (Fig 1, 2). The board contains the MEMS microphones, 25 0.1μF capacitors, and a 30 pin terminal (Appendix B). The terminal and circuit board provide each microphone with a connection to a clock signal, ground, an input voltage, and a data line. The select pad on each microphone is connected to ground in order to configure the microphones in the right data channel. Holes are also present under the microphones to allow for sound to pass through the board and into the sensor. The capacitors connect the input voltage to ground to prevent noise and are placed close to each microphone. The capacitors and microphones are surface mount devices (SMD) and were attached using reflow soldering. The terminal is through-hole and was attached using through-hole soldering.

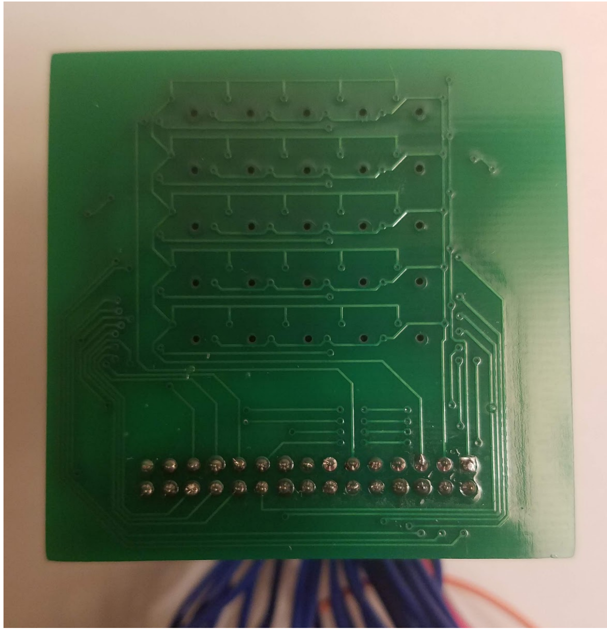


Figure 1. A picture of the front of the assembled circuit board. The small holes arranged in a grid are the ports for the microphones.

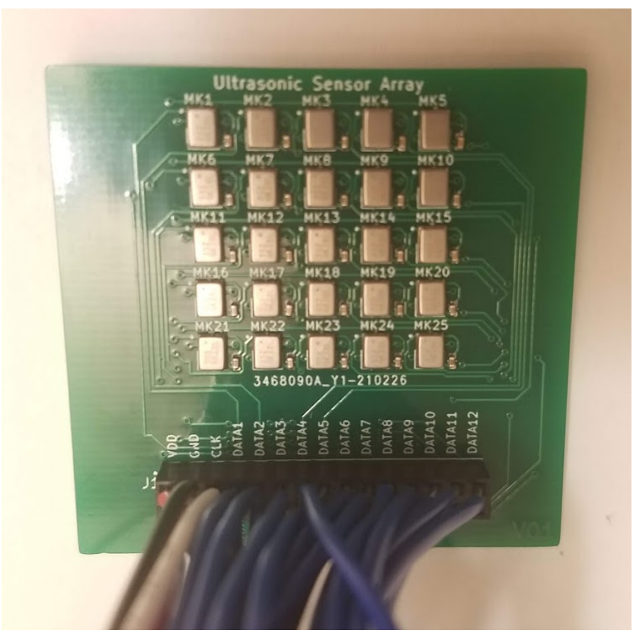


Figure 2. A picture of the back of the fully assembled circuit board. The microphones, capacitors, and the terminal are visible.

A teensy 4.1 microcontroller was also used as part of the electrical system. The input voltage and ground on the PCB are connected to the 3.3 volt and ground pins on the microcontroller, respectively. The clock signal is generated with a pulse width modulation (PWM) signal, and the data channels are connected to the microcontroller pins in an order that was determined to optimize the placement of the corresponding bits in the general-purpose input/output (GPIO) registers, minimizing the amount of time required to read the data from the sensors. The pin generating the clock signal was also connected to another microcontroller pin in series with a 330Ω resistor so that the signal can be read and used for timing. An external 10kΩ pull-down resistor was connected to each pin of the microcontroller that was used for reading sensor or clock signal data to help stabilize the digital input.

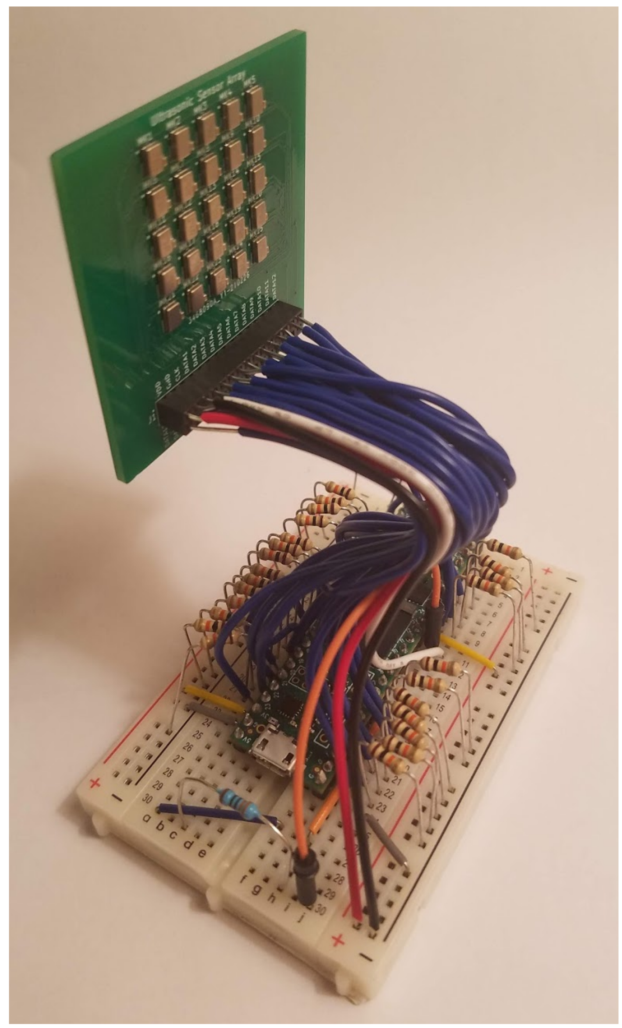


Figure 3. A picture of the fully assembled recording device.

An image of the fully assembled recording device is shown in fig 3. The total cost of all components and supplies used to create the recording device was approximately 90 USD.

Software Design

A program for the microcontroller was written in Arduino to generate the clock signal, read data from the sensors, and transmit the data over USB to a computer for storage and processing. The ICS-41352 microphones are capable of a sample rate of up to 4.8MHz (InvenSense, 2016), but a sample rate of only 3MHz was used in this study because of limitations relating to the speed of the code. This rate was set by setting the frequency of the clock signal to the desired sample rate.

The computer program was written in Python and functions to read data over USB, apply the beamforming algorithm, filter the data, save the data, and generate graphs representing the data. The raw microphone is first split into signals for each microphone, and then conventional beamforming is performed by adding all the signals together. The capability to steer the beam by applying delays to each signal was not added.

To convert the pulse density modulation (PDM) signal outputted by the microphones to a pulse code modulation signal (PCM), the high frequency noise introduced by the delta-sigma modulator must be removed (kite, 2012). A variety of digital filters were designed for different configurations of the system, but an order 358 50kHz low-pass FIR filter was used for all testing done for this study. The filter has a passband between 0Hz and 50kHz, a transition band between 50kHz and 65kHz, and a stopband between 65kHz and the maximum frequency of 1500kHz. The program can also downsample the signal to reduce the amount of data and increase processing speed. The optimal downsampling factor was calculated to be roughly 20 for this configuration. However, the signal was not downsampled during testing so that the amount of data received could be maximized, increasing the clarity of any features in the signal. The program then applies a discrete Fourier transform (DFT) to determine the frequency spectrum of the signal. Finally, the signal is saved as both an audio and text file.

While the microphones and therefore the device are capable of detecting sounds between approximately 20Hz and 85 kHz, the configuration used in this study only allows the device to detect sounds between 20Hz and 50kHz (InvenSense, 2016). This range was chosen because of the difficulty associated with accurately generating sounds above this range and the increased processing speed that results from less data being needed to describe the signal.

Testing

The device was tested at ultrasonic and audible frequencies. For all trials, the sound source was placed 0.5m directly in front of the array to ensure that it was in the far field and was pointed directly at the array. For audible frequencies, two trials were done. One trial used a 400Hz sine wave, and another used a 4kHz sine wave. A standard computer and speaker were used to play the sine waves.

For ultrasonic frequencies, an ultrasonic transmitter was powered by a 40kHz square wave. Since a square wave is composed of all odd multiples of the base frequency, all other frequencies in the signal will be removed because they are above the pass and transition bands of the filter. This means that the 40kHz square wave can be treated as a sine wave for this experiment.

A recording was also made where no sound was generated, although some background noise was present. This trial was done as a control to ensure that any sound the device detected was actually present. For each trial, the sounds were recorded using the device for 5 seconds. A sample of the time series and a frequency spectrum of the entire recording were generated for each recording. Samples of speech were also recorded to test more complex sound waves.

Results

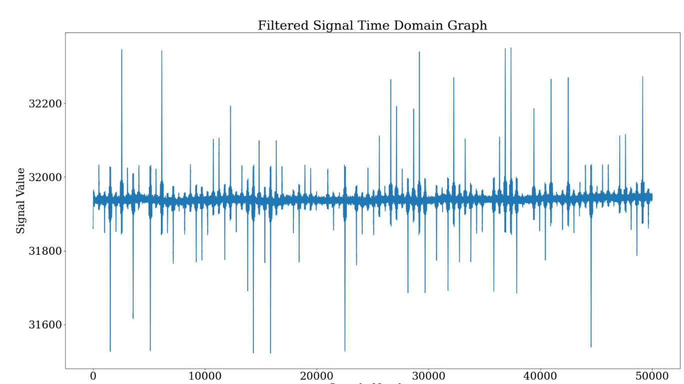


Figure 4. A sample of the time series for a recording with no sound played.

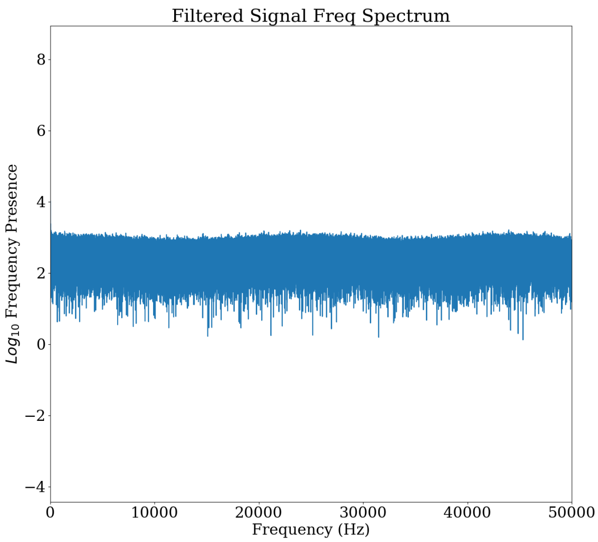


Figure 5. The frequency spectrum of a recording with no sound played.

For the trial with no sounds played, a low amplitude low frequency wave was visible in the time series (Fig 4) along with spikes varying in size spaced approximately 500 samples apart. In the frequency spectrum (Fig 5), there was some low frequency noise and almost no higher frequencies.

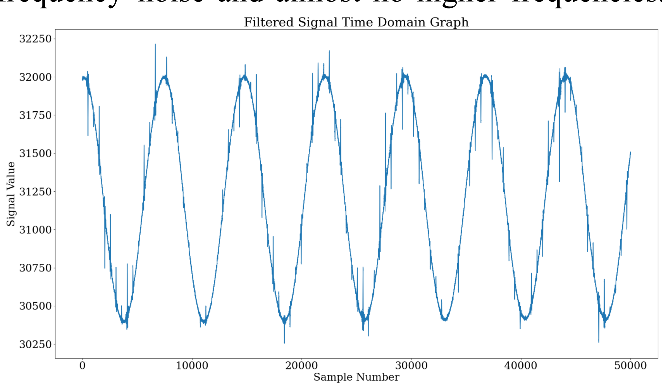


Figure 6. A sample of the time series for a recording with a 400Hz sine wave.

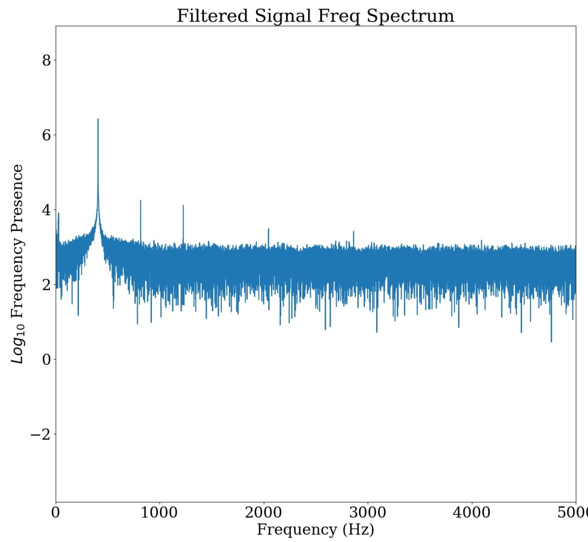


Figure 7. The frequency spectrum of a recording with a 400Hz sine wave.

For the first audible sound trial, a 400Hz sine wave was clearly visible in both the frequency spectrum and time series (Fig 6, 7). There were also spikes spaced approximately 500 samples apart in the time series and spikes at several other frequencies in the frequency spectrum, although they are much weaker than the 400Hz wave.

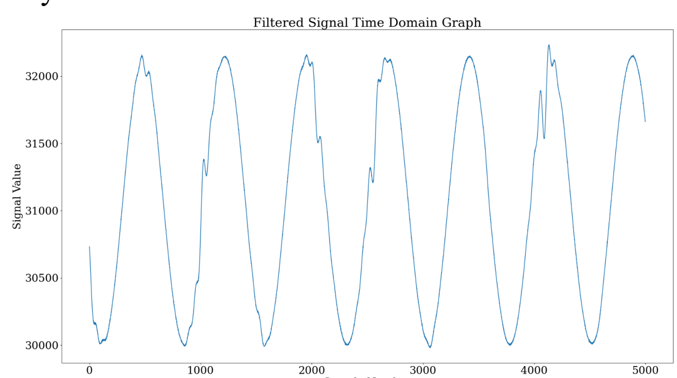


Figure 8. A sample of the time series for a recording with a 4kHz sine wave.

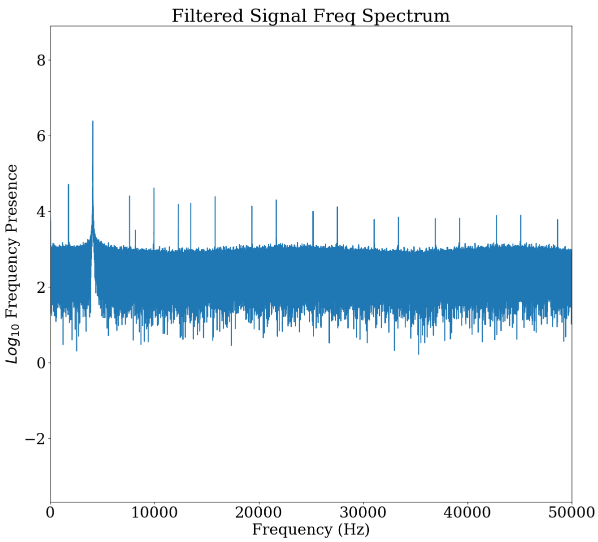


Figure 9. The frequency spectrum of a recording with a 4kHz sine wave.

The results were similar for the 4kHz sine wave, with the signal clearly present in both domains (Fig 8, 9). However, the spikes were significantly more intense in the frequency spectrum.

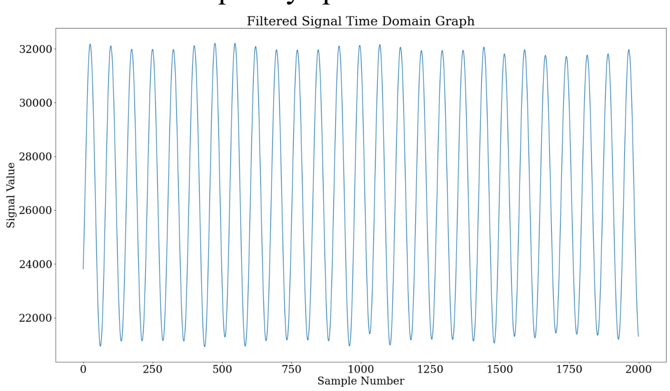


Figure 10. A sample of the time series for a recording with a 40kHz sine wave.

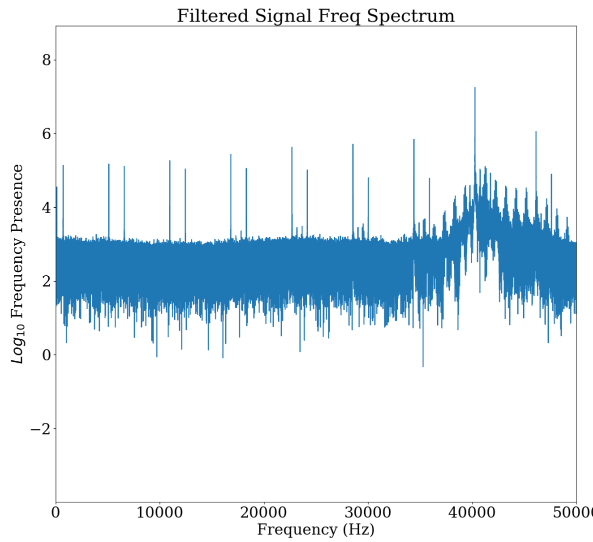


Figure 11. The frequency spectrum of a recording with a 40kHz sine wave.

The results were also similar for the ultrasonic trial, with a signal slightly above 40kHz strongly present in both the time series and frequency spectrum (Fig 10, 11). Strong spikes at other frequencies were present in the frequency spectrum.

The speech recordings were clearly intelligible, although there was significant noise.

Discussion

The low amplitude waveform present in the trial with no sounds played was likely a result of background noise, such as the noise generated by the computer used to record the data.

The sine waves present in all the other trials matched the incoming sound and expected data, but the spikes in the time series and the other frequencies present were unexpected. One possible explanation for the spikes is that they were a result of an issue relating to data transmission. While the program for the microcontroller was being written, an issue was encountered where the device would transmit approximately 2000 bytes (equivalent to 500 samples) at an appropriate speed, take much longer to transmit the next few bytes, and then repeat the cycle. This would cause some of the samples to be missed during the period of slow transmission because the device would skip reading data from the microphones while it was transmitting the data. Since some of the samples were effectively deleted, the density of ones and zeros in the PDM signal was changed, changing the signal. When the signal was passed through the filter, the effect was spread out over multiple samples.

Depending on which samples were deleted, the resulting signal could be similar or very different from the actual microphones measurements. This means that the resulting spike could be large, small, or nonexistent. This is consistent with the observed variation in spike sizes. When the DFT is performed, the spikes in the time domain would be converted into the spikes in the frequency domain.

One possible reason the spikes have a relatively high amplitude in the 40kHz trial is that a square wave was used. The deletion of samples likely has different effects depending on the exact shape of the waveform.

Another factor that may create spikes in the frequency domain is that the speakers and transducers used may generate significant amounts of sound at frequencies that are different then the desired frequency.

One significant limitation of the device is that the frequency response of the MEMS microphones is far from flat (InvenSense, 2016). This means that the output signal is distorted because some frequencies in a recording will appear stronger than in the original signal, while others will appear weaker. Another limitation of the device is the recording range. Although no precise testing was performed to determine the sensitivity and range of the device, during testing it was found that the source of the sound would have to be close to the device for it to be audible in the recordings.

To improve the quality of the device, the spikes should be investigated and the issue with data transmission should be addressed. Alternatively, it may be possible to search through the recording, identify spikes, and then replace the spikes with a predicted waveform.

Testing could be done to measure the antenna pattern at various frequencies to determine if it matches the calculated antenna pattern. More array geometries could be evaluated to find a more optimal configuration for the device.

The experiments performed in this study could be repeated in an environment that minimizes acoustic reflections off surfaces to provide more accurate measurements. Calibrated speakers and transmitters could also be used so that the amplitude of the recorded waves could be determined, allowing the device to measure the intensity of sound.

The device could be improved by integrating more of the electronics, such as the pull-down resistors, wires, and microcontrollers into the PCB to make the device more compact and simplify assembly. The software could be modified to offer a live spectrogram, allow for faster data processing, and allow for steering and other beamforming algorithms to be used.

Additionally, some parts of the assembly, particularly reflow soldering, can be difficult and time consuming due to having to place a large number of components precisely. It may be possible to modify the design to allow for easier assembly.

At a price of $90, the recording device meets the criteria of being affordable. However, an issue with the current method of transmitting data to a computer for analysis generates unwanted noise in any recordings. Despite this issue, the device is still capable of detecting sound across the measured frequency, although the introduction of frequencies that aren't part of the original sound limits its use as a recording device, especially because the frequencies introduced are not easily predictable.

References

Bass, H. E., Sutherland, L. C., Zuckerwar, A. J., Blackstock, D. T., & Hester, D. M. (1995). Atmospheric absorption of sound: Further developments. The Journal of the Acoustical Society of America, 97(1), 680–683. https://doi.org/10.1121/1.412989

Carovac A, Smajlovic F, Junuzovic D. Application of ultrasound in medicine. Acta Inform Med. 2011;19(3):168-171. doi:10.5455/aim.2011.19.168-171

Holcomb, M.R., Schneider, S.N., & Briggs, J.F. (2015). 100 kHz Bandwidth Ultrasonic Recording System. Instrumentation Science & Technology, 43, 214 - 221

InvenSense. (2016, August 30). ICS-41352 data sheet.

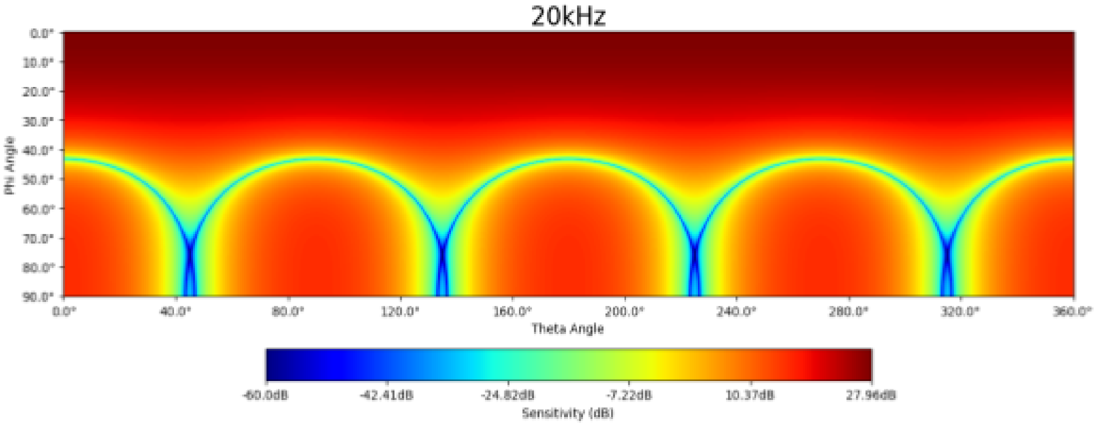
Kite, T. (2012). Understanding PDM Digital Audio. Audio Precision.

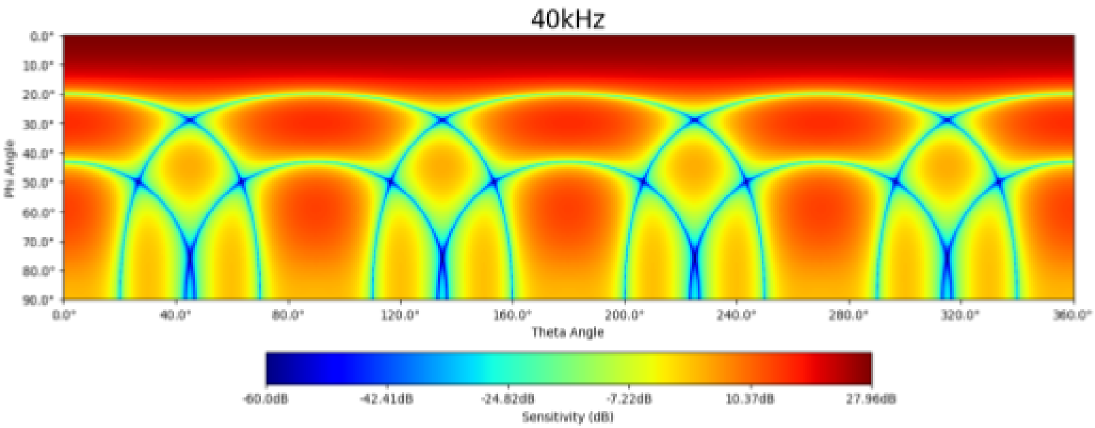
Koay, G., Heffner, R. S., Bitter, K. S., & Heffner, H. E. (2003). Hearing in American leaf-nosed bats. II: Carollia perspicillata. Hearing Research, 178(1-2), 27-34. doi:10.1016/s0378-5955(03)00025-x

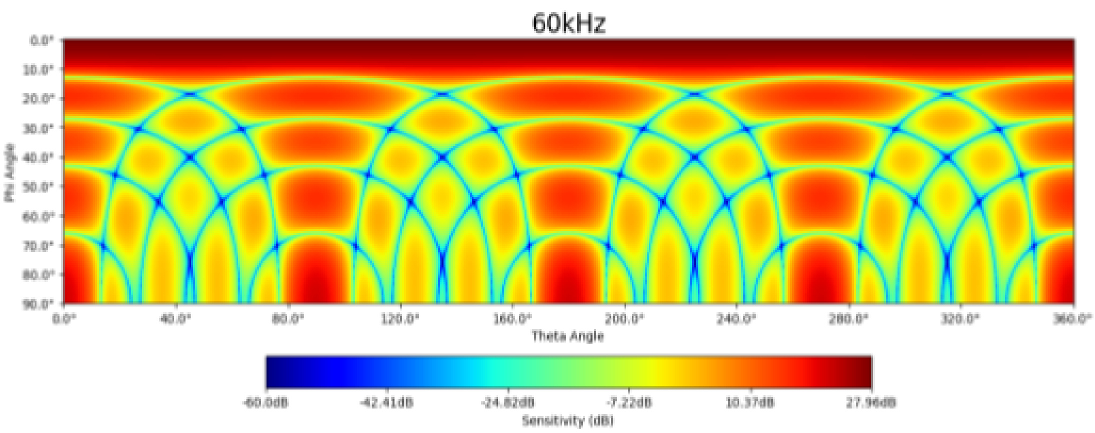
McCowan, I. (2001). Microphone Arrays : A Tutorial.

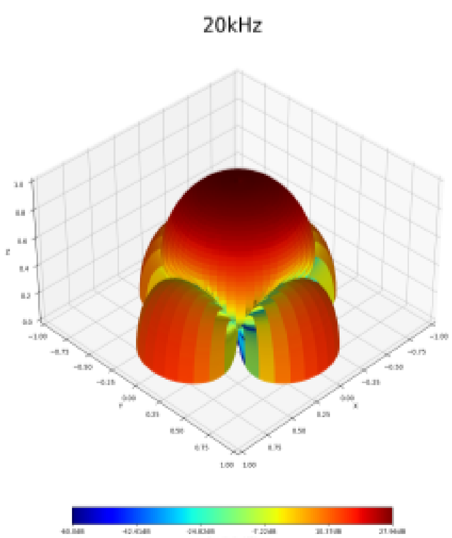
Smith, S. W. (1999). The scientist and engineer's guide to digital signal processing (2nd ed.). California Technical Publishing.

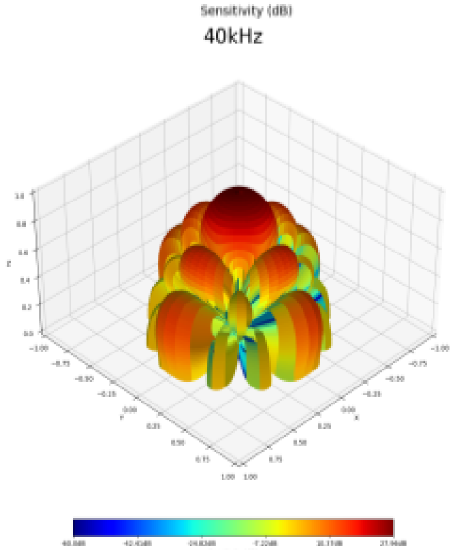
Appendix A

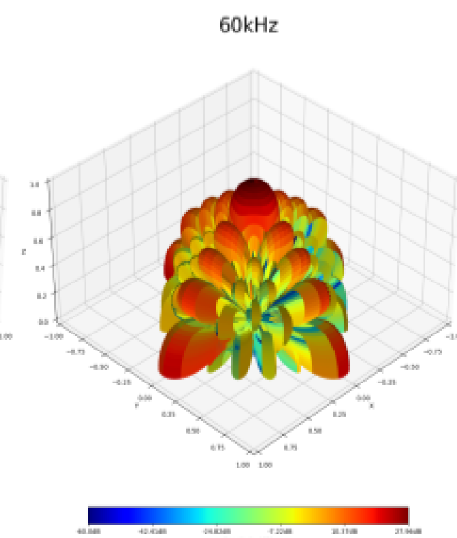

20kHz
Phi Angle
0.0° 10.0° 20.0° 30.0° 40.0° 50.0° 60.0° 70.0° 80.0° 90.0°
Theta Angle
0.0° 40.0° 80.0° 120.0° 160.0° 200.0° 240.0° 280.0° 320.0° 360.0°
-60.0dB -42.41dB -24.82dB -7.22dB 10.37dB 27.96dB
Sensitivity (dB)


40kHz
Phi Angle
0.0° 10.0° 20.0° 30.0° 40.0° 50.0° 60.0° 70.0° 80.0° 90.0°
Theta Angle
0.0° 40.0° 80.0° 120.0° 160.0° 200.0° 240.0° 280.0° 320.0° 360.0°
-60.0dB -42.41dB -24.82dB -7.22dB 10.37dB 27.96dB
Sensitivity (dB)


60kHz
Phi Angle
0.0° 10.0° 20.0° 30.0° 40.0° 50.0° 60.0° 70.0° 80.0° 90.0°
Theta Angle
0.0° 40.0° 80.0° 120.0° 160.0° 200.0° 240.0° 280.0° 320.0° 360.0°
-60.0dB -42.41dB -24.82dB -7.22dB 10.37dB 27.96dB
Sensitivity (dB)


20kHz

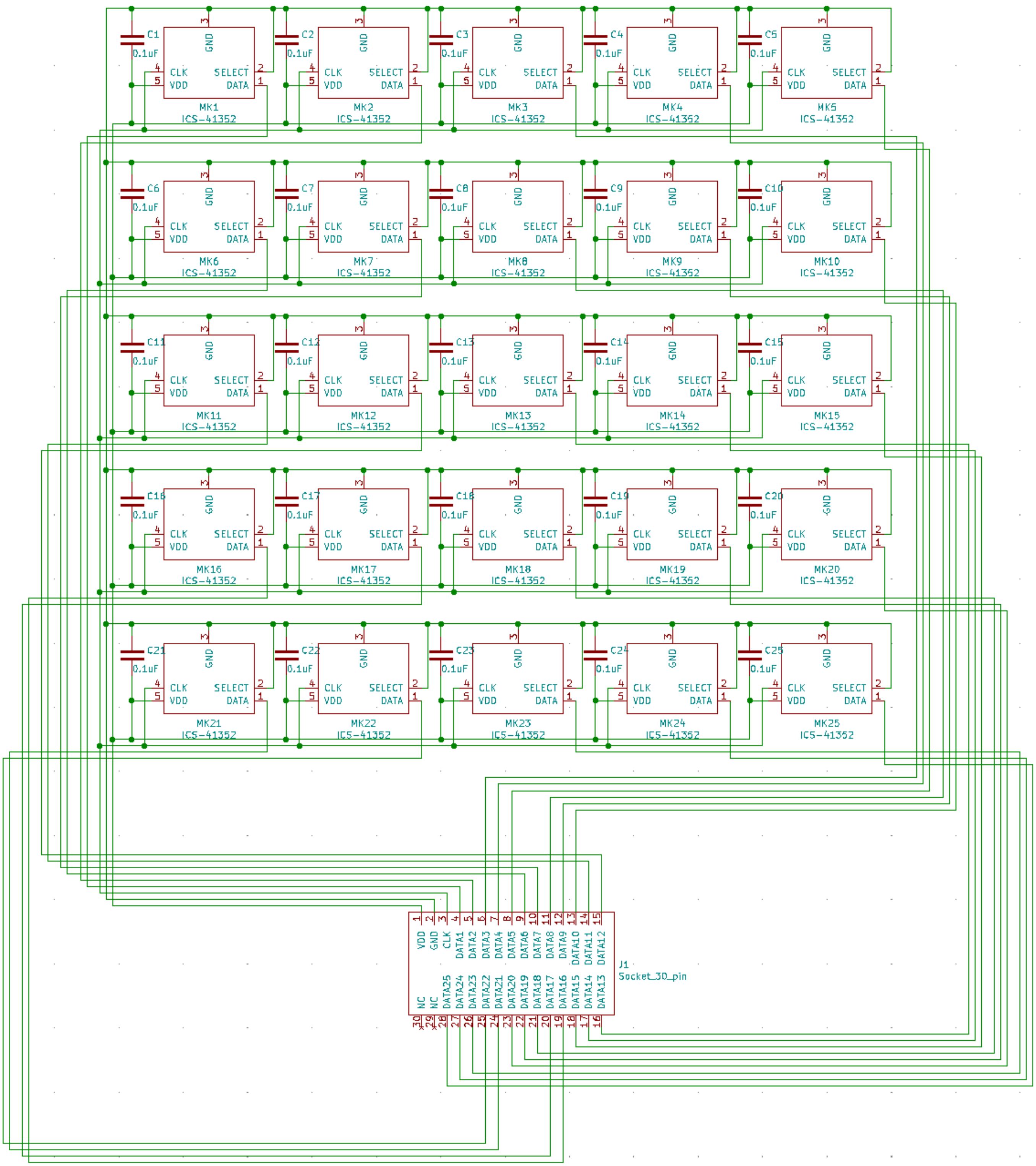

40kHz


60kHz

Appendix A. Antenna patterns for a 25 element grid-shaped array with 5 mm spacing between elements at various frequencies. The plots show sensitivity as a function of angle in two and three dimensions.

Appendix B

Appendix B. A schematic of the custom PCB designed for the device containing microphones, capacitors, and a terminal.


C1 0.1uF
C2 0.1uF
C3 0.1uF
C4 0.1uF
C5 0.1uF
C6 0.1uF
C7 0.1uF
C8 0.1uF
C9 0.1uF
C10 0.1uF
C11 0.1uF
C12 0.1uF
C13 0.1uF
C14 0.1uF
C15 0.1uF
C16 0.1uF
C17 0.1uF
C18 0.1uF
C19 0.1uF
C20 0.1uF
C21 0.1uF
C22 0.1uF
C23 0.1uF
C24 0.1uF
C25 0.1uF
GND
CLK
SELECT
VDD
DATA
MK1 ICS-41352
MK2 ICS-41352
MK3 ICS-41352
MK4 ICS-41352
MK5 ICS-41352
MK6 ICS-41352
MK7 ICS-41352
MK8 ICS-41352
MK9 ICS-41352
MK10 ICS-41352
MK11 ICS-41352
MK12 ICS-41352
MK13 ICS-41352
MK14 ICS-41352
MK15 ICS-41352
MK16 ICS-41352
MK17 ICS-41352
MK18 ICS-41352
MK19 ICS-41352
MK20 ICS-41352
MK21 ICS-41352
MK22 ICS-41352
MK23 ICS-41352
MK24 ICS-41352
MK25 ICS-41352
VDD GND CLK DATA1 DATA2 DATA3 DATA4 DATA5 DATA6 DATA7 DATA8 DATA9 DATA10 DATA11 DATA12
NC NC DATA25 DATA24 DATA23 DATA22 DATA21 DATA20 DATA19 DATA18 DATA17 DATA16 DATA15 DATA14 DATA13
J1
Socket_30_pin