Acoustic Motion Tracking

Team 35 - Sean Nachnani and Hojin Chun

ECE 445 Project Proposal - Spring 2018

TA: Yuchen He

1 Introduction

1.1 Objective

People are always striving for a better life and are trying to make each day more comfortable than the last. As technology advances further, and AI comes more into focus, we try to satisfy the comfort needs by building autonomous cars, creating smart cities, and building robots that can eventually replace humans. Even now, we have products like Alexa, Siri, Bixbi, and many more that can understand a person's vocal commands and get data or perform tasks for the user, without them having to lift a finger.

Our objective is to build a device that use sound rather than video as a means of motion recognition. Current devices mentioned above are limited to only using natural language processing to interpret a user's need. We want to expand upon this further and allow devices to perform commands using simple gestures. Not only our device will make people's life more comfortable with this feature, but also will help people with speech impediments or certain accents, allowing them to still use these devices.

Our implementation will have 4-input microphone array that allows for at least a 48 Khz sample rate, and use a speaker that can reproduce sounds up to 24 kHz. We will use a microcontroller to act as the bridge between all the hardware components and the software running on external device. Our device will be designed to be plugged into the a regular power outlet.

Our current software implementation involves pulsing a pseudo-random wave, and calculating the distance of the hand by measuring the peaks of the Channel Impulse responses received from the microphone array. If we are able to finish the project ahead of schedule, we plan to go further and implement FMCW (Frequency Modulated Continuous Waveform) radar as a basis for this approach. This will allow us to also take advantage of Doppler effects and calculate velocity, further increasing the accuracy of the device.

1.2 Background

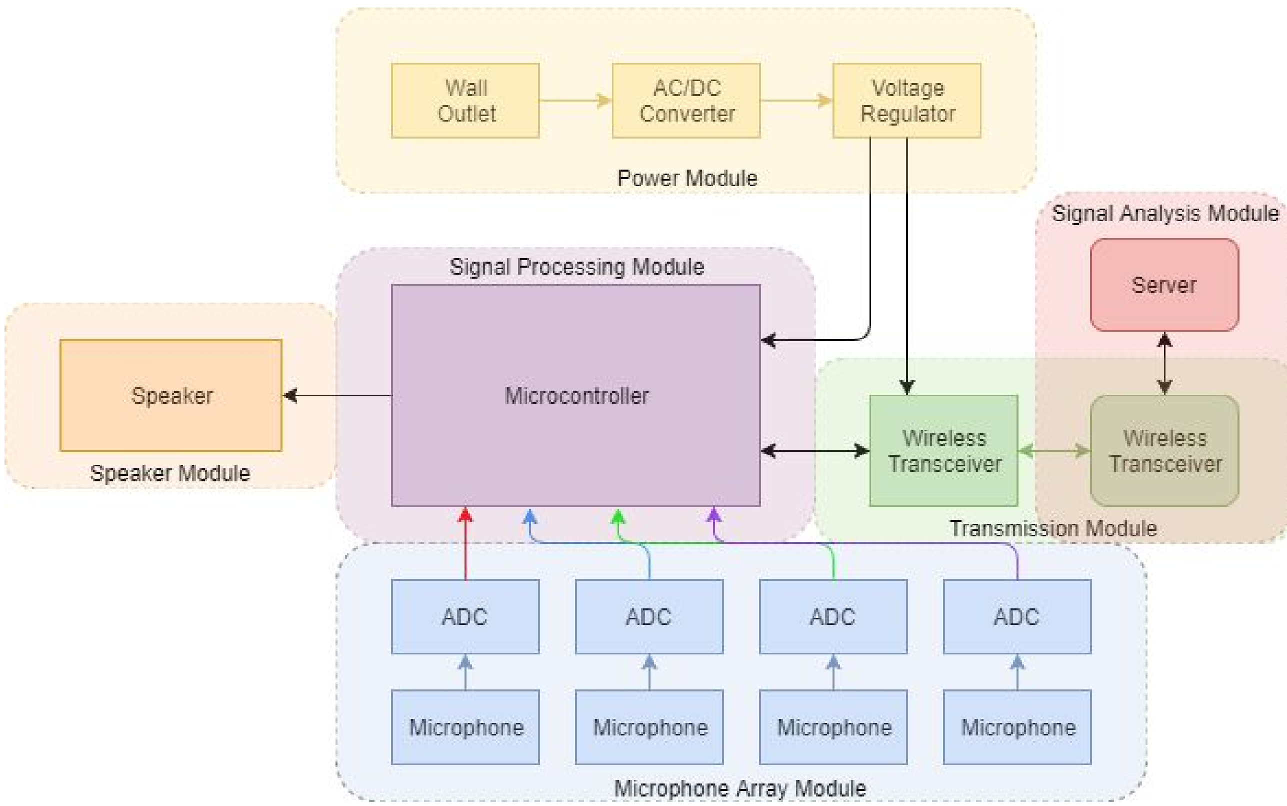
Motion tracking has always been widely researched, but the majority of it centered around video, or radar. Now, with the increasing number of mobile devices that contain microphones and speakers, research has been drawn towards the acoustic region. A lot of this research has been been focused around VR[1] and smartphones[2], but we intend to implement this with the smart assistants that are becoming ever more prevalent. Natural Language Processing has come very far and allowed these devices to understand human speech. However, they still struggle when dealing with speech impediments and accents. Therefore, implementing gesture control for these devices will help solve these issues. Moreover, this project will open the door to other motion recognition uses for the smart devices, such as the ability to act as a motion alarm if the user isn't home, and also as possible forms of authentication to counter acoustic DDoS attacks[3].

1.3 High-Level Requirements List

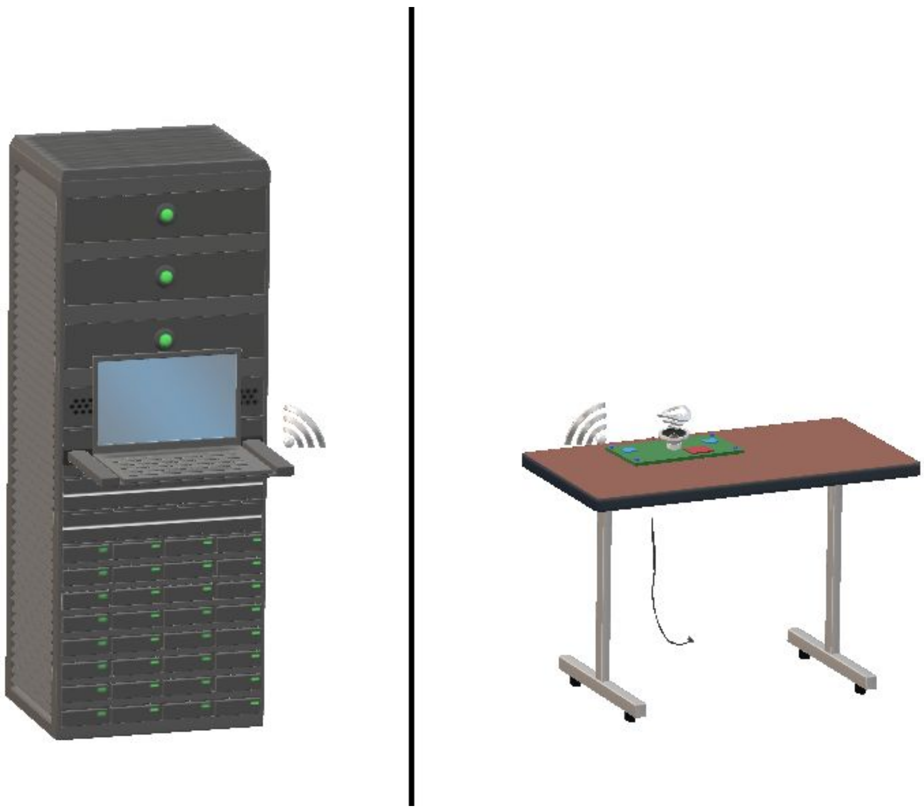
- The Device must be able to wirelessly transmit the recorded signal without loss to another device that can process it.
- The microphone array must be able to record frequencies above 20khz to allow for the speaker to transmit in the inaudible range
- The Device must be able to accurately measure the distance of the hand performing the gestures to within 3 cm, to allow for proper identification of the motion.

2 Design

2.1 Block Diagram



2.2 Physical Design



The Physical design above provides a 3-D model of the system. The device itself in this model is located on the table (on the right panel) and is wirelessly transmitting to the server (on the left panel). The speaker for the device will be central to the 4 microphone array. The device will be designed such that it can be placed on any flat surface in a room and be connected to an outlet (not show).

2.3 Functional Overview

2.3.1 Microcontroller

The microcontroller is used to sample from the four microphone and ADC pairs while simultaneously transmitting the sound pulses for the speaker to play. It stores the data for the full listening duration, and then transmits it to the wireless transceiver to send off to the server for signal analysis. We plan on using one of the TI MSP430 family of mcu's, for the low power consumption, the 12-bit DAC, 512kb of onboard memory, and both SPI and I^2C capabilities.

Requirement 1: The controller must support at least 12-bit DAC

Requirement 2: The controller must support SPI and I^2C

Requirement 3: The controller must have enough onboard memory (at least 512kb) to be able to store the entire recording before sending it to the transmitter

2.3.2 Microphone Array

We have chosen to use 4 microphones to create the array. Our initial plan is to use the Invensense ICS-40730 analog MEMS microphone. This microphone has a −32 dBV Differential Sensitivity, −38 dBV Single Ended Sensitivity, and a package size of 4.72 × 3.76 × 3.5 mm. With the addition of an ADC that has a sample rate above 48khz, we can then use sounds outside the audible range. We are considering using two TI PCM1804 chips, since they have a resolution of 24bits, both SPI and I^2C, two analog inputs, and a max sampling rate of 192khz. This will also allow us to use only 2 ADC chips, reducing overall area cost.

Requirement 1: A sample rate of at least 48khz to allow for transmitting in the inaudible range

Requirement 2: The ADC must have SPI or I^2C capabilities

2.3.3 Speaker

The speaker needs to be able to transmit frequencies up to 24khz. For this device we are planning on using the XS-GTF1027 speaker by Adafruit which has a frequency response of 60Hz - 24kHz.

Requirement 1: The speaker must be able to play frequencies up to at least 24khz

2.3.4 Wireless Transmission

For our project the gesture recognition occurs on the server side, allowing us to implement it in software. To transmit and receive data from the microcontroller, we will be using a wifi module. By choosing to transmit over wifi, we will enable the device to be located anywhere, not just within a certain range of the server.

Requirement 1: The wireless transmitter must be able to transmit over 2.4Ghz channel and support MAC protocol

2.3.5 Power

We will be using a wall outlet to power the device. We will need an AC to DC adapter that can convert the 120V at 60hz to DC voltage that is within the range of the devices components. Since the different parts of the device have different voltage requirements, we will also be using a voltage regulator to allow each component to receive the correct power specifications.

Requirement 1: The adapter must be able to handle converting from 120V, 60hz to DC voltage.

Requirement 2: The voltage regulator must be able to supply voltages in ranges of: 1.8V-3.6V for the microcontroller, 5V for the ADCs, and voltage requirements for the wireless transmitter

2.3.6 Server

We will be using a laptop to function as the server for our project. It will be able to receive the transmission from the device over wireless transmission and compute the signal analysis to identify the gesture performed. It will then transmit the identified gesture back to the device, allowing it to perform the action.

Requirement 1: The server must be able to receive and transmit a wireless transmission

Requirement 2: The server must be able to process the signal and return a correctly identified gesture

2.4 <u>Risk Analysis</u>

The microcontroller is the component of the device that poses the greatest risk to successfully completing the project. This part is the centerpiece of the device, it receives the signals from all four microphone ADC pairs simultaneously, while also controlling the speaker playing the sound pulses. On top of that, it has to properly store the data and send it to the wireless transmitter. It is the most critical piece, and needs to synchronize all the other components together properly. If this microcontroller fails to do so, the device cannot function.

3 Ethics and Safety

For our project, there are few safety concerns. As mentioned above, our device will be plugged into a wall outlet. To be able to do that, we will have to have a AC/DC converter. The danger comes when we are dealing with wall outlet voltage at 120V AC and converting it to 24V DC. We will need to make sure the wall outlet contains a ground using one hand method and we will also need to make sure that AC/DC conversion is off limit from the user so they never have to come into contact with high voltages.

When dealing with high voltages, the concern of large current comes along with it. With large current, it can also dissipate heat. So when we are dealing with high voltages and high current, we will have to careful to the heat and we will make sure that the user will never be exposed to the excessive heat.

We are responsible for all decision we make for the design of our device and it is our responsibility to disclose any issues that might be dangerous to the user per Section 1 of the IEEE code of Ethics. We believed that if our device is properly and well designed, we will lessen these hazards to create a enjoyable experience for the user.

References

[1] W. Mao, J. He, and L. Qiu, “CAT: High-Precision Acoustic Motion Tracking,” Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking - MobiCom 16, 2016.

[2] R. Nandakumar, V. Iyer, D. Tan, and S. Gollakota, “FingerIO,” Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI 16, 2016.

[3] N. Roy, H. Hassanieh, and R. R. Choudhury, “BackDoor,” Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services - MobiSys 17, 2017

[4] IEEE (2017). IEEE Code of Ethics. [Online]. Available: http://www.ieee.org/about/corporate/governance/p7-8.html [Accessed: 7- Feb- 2018].