Gesture Controlled Audio System

ECE445 Design Document Check

By

Kehang Chang (kehangc2)

Ruofan Chen (ruofanc2)

Ruohua Li (ruohual2)

1. Introduction

I will build a Gesture Controlled Audio System for people who would like to enjoy the convenience of controlling their speakers remotely with gestures. My customer has a problem that when they are cooking or working, they lack a robust and convenient way to control their speakers in potentially noisy environments robustly. My product solves my customer's problem by providing a solution allowing them to control their speakers with simple gestures.

2. High-level Requirements

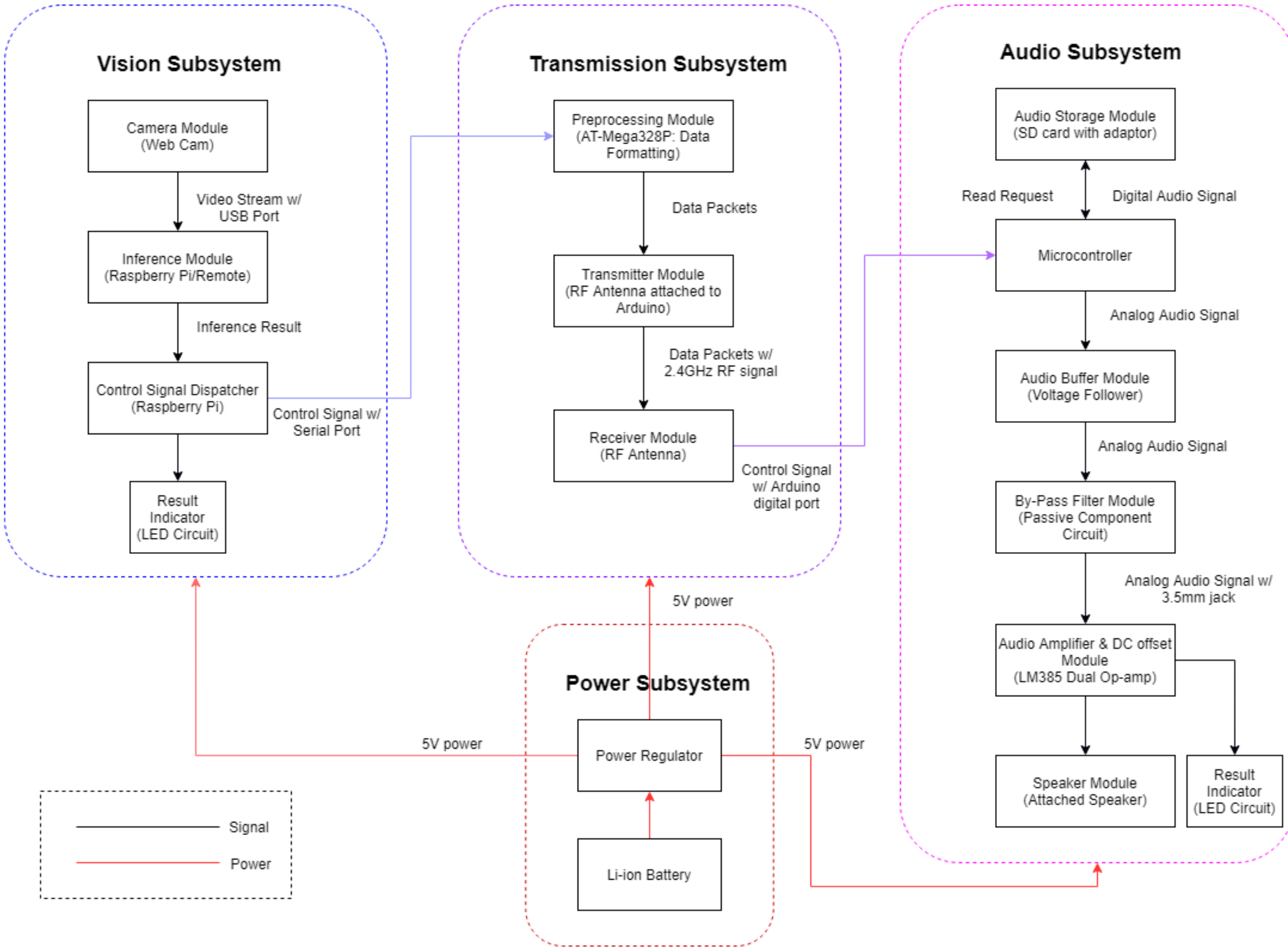
- Visual Characteristics

  The vision subsystem must show its ability to recognize human gesture with precision and recall > 80% to accurately determine gestures. The gestures so far including play/pause, next track, previous track. We might let users create their own gesture profiles as well.
- Audio Characteristics

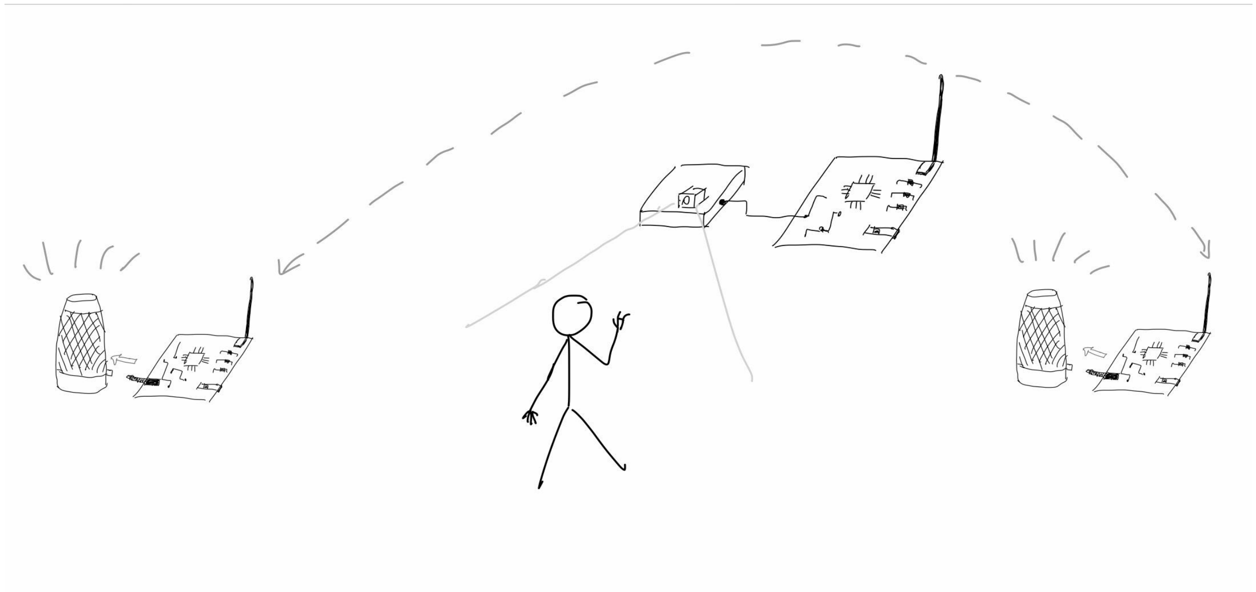
  The audio subsystem should be able to distribute and play music with 8-bit resolution. Noise out of [20Hz, 20KHz] range should be filtered properly such that the noise/authentic signal power ratio < 25%..
- Transmission Characteristics

  The transmission subsystem should be able to deliver correct control signal and audio signal with lagging < 1s. To be more precise, the system should complete its entire processing pipeline within 1 second: detect gesture, decode gesture, execute audio control.

3. Block Diagram



4. Physical Design (if applicable)



5. Requirements & Verification Tables

5.1 Vision Subsystem

5.1.1 Gesture Inference Module

Requirement	Verification
Detect a valid gesture (pause, resume, next song) based on images captured in a 30fps-capable surface-mounted camera	Do 100 gestures with 50 valid gestures and 50 invalid gestures, and according to Monte-Carlo method if we can obtain 40/50 correct response for valid gestures as well as at least 80% of valid gesture detected is actually valid, then the model is running as expected.

5.2 Transmission Subsystem

5.2.1 Transmitter/Receiver Module

Requirement	Verification
Transmit/Receive a data stream in 100 Bps for the world wide 2.4 -2.5 GHz ISM band.	Transmit a data packet with size of 100 Byte within one second, with distance no shorter than 5 meters.

5.3 Audio Subsystem

5.3.1 By-pass filter Module

Requirement	Verification
Filter out noises out of range of 40kHz ~ 100kHz	Play a piece of audio signal with frequencies span from 20kHz to 200kHz then detect if the output signal has successfully filtered out frequencies within 40kHz and beyond 100kHz.

5.4 Power Subsystem

5.4.1 Power Regulator Module

Requirement	Verification
Output a 4.95~5.05V voltage with a maximum current of 50mA when a 7.2~14V input voltage is applied	A. Connect a 5 pack Li-ion battery pack to the power regulator module B. Probe the power line coming out of power regulator module to measure the supply voltage

6. Tolerance Analysis

Trade off between responding robustness and inference time (overall responding time)

$f(I_{t1},..., I_{tn}) = f(I_{t1}) AND ... AND f(I_{tn})$, where n denotes the number of frames the model need to do inference on to generate a control signal and $f(.)$ denotes inference model which produces boolean value indicating whether the input frame I_t constitutes a valid gesture.

$FP(f(I_{t1},..., I_{tn})) = FP(f(I_{t1}))... FP(f(I_{tn}))$, where FP denotes false positive rate.

If we respond on fewer frames to decide what the next command is going to be, i.e., n is relatively small, we can respond faster to each user input, but this strategy will also potentially increase the false positive rate of the model since users, from time to time, could accidentally pose valid gestures temporarily when they do not mean to trigger a control signal.

7. Safety & Ethics

7.1 Ethics

We understand the importance of ethical and safety concern during the process of designing this product and put “safety, health, and welfare of the public” from #1 of the IEEE Code of Ethics at the forefront of our design thinking [1]. Since this product contains a camera module which will capture image input of users at real time, it’s important to protect the privacy of end users. We are planning to provide users the full disclosures about how we handle the user inputs and give users the ability to choose if captured images should be stored temporarily or erased immediately. Based on the Consumer Data Privacy and Security Act, we would “directly obtain the individual’s consent” before we start to collect user information [2]. We will also maintain a security program to maintain “security, confidentiality, and integrity of personal data” from malicious usages [2].

7.2 Electric Hazards

The audio system needs to process the signal from its receiver module and output the signal into the speaker. There could be a potential electric hazard caused by a short circuit. The audio system is powered by an atmega328P chip which will source a 5V voltage. The max transmission operating current is 115mA and max receiving operating current is 45mA for a typical RF communication module [3]. An electric current as low as 30 mA current could potentially induce a ventricular fibrillation [4]. Thus special circuitry has been implemented to cut down the power immediately once a short circuit has been detected.

7.3 Fire Hazards

The RF communication module and its peripheral signal processing units could be affected by surging of current due to a short circuit. The RF communication module has a total power dissipation of 60mW [3]. When power is not cut off promptly after a short circuit, the temperature of the PCB board will potentially induce a fire. Thus, we have taken the power regulation of the whole module into consideration during the PCB design of our product.

8. Citations

[1] “IEEE Code of Ethics.” *IEEE*, www.ieee.org/about/corporate/governance/p7-8.html.

[2] Moran, Jerry. “Text - S.3456 - 116th Congress (2019-2020): Consumer Data Privacy and Security Act of 2020.” *Congress.gov*, 12 Mar. 2020, www.congress.gov/bill/116th-congress/senate-bill/3456/text#toc-ida8f663638e07477a8a47c52bb9e5f876.

[3] “Single Chip 2.4 GHz Transceiver.” *Sparkfun*, www.sparkfun.com/datasheets/Components/nRF24L01_prelim_prod_spec_1_2.pdf.

[4] “Electrical Injury.” *Wikipedia*, Wikimedia Foundation, 15 Jan. 2021, en.wikipedia.org/wiki/Electrical_injury.

[5] Martin, Taylor. “34 Alexa Tips and Tricks for Beginners.” *CNET*, www.cnet.com/how-to/amazon-echo-alexa-tips/.

[6] *All-New Echo (4th Gen) | With Premium Sound, Smart Home Hub, and Alexa | Charcoal*. www.amazon.com/All-New-Echo-4th-Gen/dp/B07XKF5RM3/ref=asc_df_B07XKF5RM3.

[7] “Buy HomePod.” *Apple*, www.apple.com/shop/buy-homepod/homepod?afid=p238%7CsE8EkhnRj-dc_mtid_1870765e38482_pcrid_487811458778_pgrid_54140720810_&cid=aos-us-kwgo---slid-SUuEsuxd--product-.

[8] “Full-Range 8’ Speaker Pioneer Type B20FU20-51FW.” *Parts Express*, www.parts-express.com/GRS-8FR-8-Full-Range-8-Speaker-Pioneer-Type-B20FU20-51FW-292-430?gclid=Cj0KCQiAvbiBBhD-ARIsAGM48bxiUtTOtZXyLCI_CmjwfI6am72oGwLDXQFNWUhIsYm-aY3hDzTdWGAaAsEpEALw_wcB.