System Requirements Specification

for

Audio Surveillance System

Version 1.0 approved

Prepared by Zachary Tauscher, Jacob Attia, Caleb Leeb, Jaclyn Welch, and Jorge Garcia

Embry Riddle Aeronautical University Senior Capstone Design I

1 February 2022

Table of Contents

Revision History

Name	Date	Reason For Changes	Version
Initial Draft	9/28/21	Proposal of Draft Document	1
Second Draft	10/26/21	V2 Submission	2

Third Draft	11/30/21	V3 Submission - Final for Semester	3
Fourth Draft	02/04/22`	V4 Submission - First for Spring Semester	4

1. Introduction

This product is designed to collect and process sound signals via a microphone array and send that data to a secure computer. This computer will utilize machine learning algorithms to identify and classify the sounds recorded by this microphone array and ascertain the position of the sound relative to the microphone array. The secure computer will provide a GUI for security personnel to view the source location of the sound and the class assigned to that sound. The goal of this product is to provide cheap surveillance without the need for invasive camera systems.

1.1 Purpose

The purpose of this document is to identify the requirements of a system which provides positional and identifying information on people within a room based solely on the audio data received by several arrays of microphones placed within the room. There are two experimental designs which are proposed by this document, each with their own unique requirements, which will be specified in this document below. However, for the most part these requirements and basic statements of intent are shared between the two designs. This product will encompass the scope of single-room surveillance and will be built to function in environments with common sources of background noise.

1.2 Document Conventions

For future reference within this document, all mentions of the first experimental product design, multiple arrays placed at disjointed points in a room connected via wireless connectivity modules, will be referred to as design “A”. The second design, consisting of a ring of microphone arrays to be modularly placed at a central point within a room, will be referred to as design “B”.

1.3 Intended Audience and Reading Suggestions

This document is intended to be read and reviewed by the product owner (Dr. Wang), the customer (Dr. Ilhan Akbas), and to be contributed to and continually modified by all members of the development team: Jacob Attia, Caleb Leeb, Zachary Tauscher, Jorge Garcia and Jaclyn Welch.

For those readers unfamiliar with general concepts of machine learning, audio processing, or triangulation, reviewing the included documents for reference before identifying the scope of this project may be helpful. For readers with a sufficient background in these concepts who feel comfortable with an integrated approach of these subjects to further surveillance, the following section regarding the product scope may be of more relevance.

To any reader of this document whose intention is to continue development in the pursuit of future research, both the product perspective (2.1) and our analysis models used (Appendix B) are of particular interest.

1.4 Product Scope

The product is not intended to cover more than a single room, and it's scope is not to encompass more than positional information in more than two dimensions, as a more labor-intensive, intrusive, and indiscreet method of setting up a room would have to be employed to simulate three dimensions, robbing the product of any viable use case.

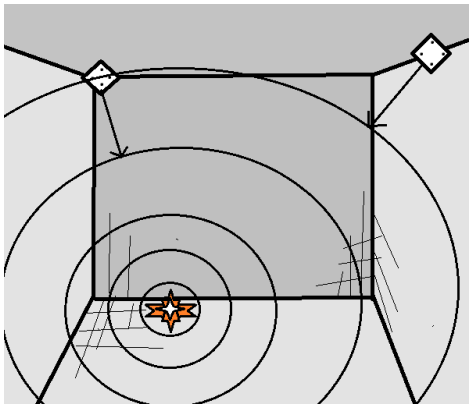
1.5 References

The following link: https://github.com/karatekidcaleb/CS490-Option-2-Audio/ will take the reader to the GitHub repository in which reside all source code and build-relevant working directories. As our product's software is not currently maintained or protected by any license, redistributors and those intent on modification can expect to follow guidelines relevant to the GNU Lesser General Public License 3.0. For those interested in some literature relevant to the formation of this product's research and a cursory literature review, please refer to the references to prior research in this AoC detailed later in 2.1.

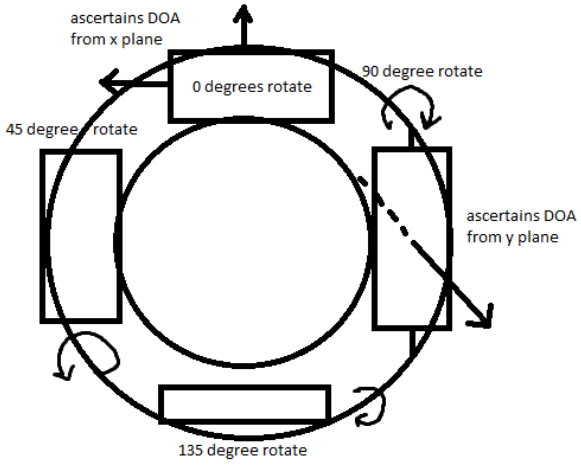
2. Overall Description

2.1 Product Perspective

Design A



Design B



The proposed product exists within the field of intelligence, surveillance, and reconnaissance (ISR). This product provides a service unoccupied currently by any product on the market, and whose closest equivalent in enterprise solutions is non-modular, and in the price range over over 500 USD (See Appendix C.a). In cases where visual obstructions to the target of surveillance exist, the proposed system bypasses any need for visual line of sight, or any visual cues whatsoever.

2.2 Product Functions

2.2.1 The product must be able to discern/label what types of sound are coming in (such as footsteps, normal conversation)

2.2.3 The product must be able to determine the location of a sound source in a room and display them on the user interface.

2.2.4 The product must distinguish between different kinds of sound and display them with the differentiation.

2.2.5 The product must transmit data from the microphone array to the host computer.

2.3 User Classes and Characteristics

2.3.1: Security Guard in Command
2.3.2: Administrator of Audio Surveillance System
2.3.3: Technical Support (IT)

- Security Guard
- Administrator of System
- IT Support

2.4 Operating Environment

2.4.1 User Interface

The designed interface runs on a secure host computer in communication with the Raspberry Pi microphone array and utilizes the processed data to display to the user the location and identity of multiple sound sources. This display will be in the form of a continually updating two-dimensional window with text labels and colored polygons on the screen to display associated characteristics of a sound source.

2.4.2 Data Collection

A microphone array of 4 microphones is linked to a Raspberry Pi and is used to collect audio files and time delay of arrival data. This data is processed using Python to turn the audio .wav file into a spectrogram for use in machine learning and package the time delay of arrival to be used for locational calculation. It is then transmitted to the secure host computer.

2.4.3 Data Processing

The data processing environment takes in the collected data received from the Raspberry Pi microphone array. It uses the spectrogram of the recorded audio as an input to Python machine learning libraries, then classifies what the audio is of. It separately uses the time delay of arrival data to calculate locational coordinates to be projected on a map of the room.

2.5 Design and Implementation Constraints

1. Design Constraints:
    - System shall weight under 1 pound
    - System shall be able to process data and display results without continual user request
    - System must maintain a low physical profile when mounted, not exceeding 4 inches vertically
    - System must be able to support computational requirements of Machine Learning algorithms, whether on the Raspberry Pi subsystem or otherwise
2. Implementation Constraints
    - The wireless components of the system must maintain their Bluetooth connection solely with the intended associated component
    - The raw audio received by the microphone arrays must be transformed into a usable format which can be parsed for telemetry, class, and multiplicity by the GUI display

2.6 Assumptions and Dependencies

2.7.1 The product will operate in a rectangular room

2.7.2 The product will be mounted in the center of the room

2.7.3 The user has a secure network to transmit security data

2.7.4 The user has a secure computer

3. External Interface Requirements

3.1 User Interfaces

3.2 Hardware Interfaces

The Audi Surveillance System consists of multiple components working together to collect, transmit and compute data. The hardware consists of a host computer acting as a security terminal and a Raspberry Pi linked to a microphone array to collect and prepare data. The host computer links with the Raspberry Pi and collects the processed data to be presented via the User Interface.

3.2.1 The data collection Hardware will consist of a Raspberry Pi and a ReSpeaker Microphone Array V2 and a host computer.

3.2.2 The host computer shall link with the Raspberry Pi and Microphone Array

3.2.3 The Microphone Array and Raspberry Pi configuration will send data back to the host computer containing location data of the audio source and classification data of the type of audio

3.3 Software Interfaces

3.4 Communications Interfaces

3.4.1 The array configuration will collect audio files, .WAV, and directional data in degrees from the microphones.

3.4.2 Communication between the Raspberry Pi and microphone array shall be handled by USB

4. System Features

5. Other Nonfunctional Requirements

Section 5 details the current known requirements and regulations associated with this project as of its initial release.

5.1 Performance Requirements

5.2 Safety Requirements

5.2.1 The system shall be securely mounted in the room as to not fall and cause potential injury

5.2.2 The system will not be used in unapproved rooms or locations

5.3 Security Requirements

5.3.1 As the system will be recording audio the files and data must not be contained on the device itself but rather sent to a host system.

5.3.2 As the system will be left in a room the hardware must not allow for tampering and should be completely contained.

5.3.3 No other device shall be able to receive the transmitted audio

5.3.4 The operator of the host computer shall have the correct clearance to access the system

Appendix A: Glossary

Appendix B: Analysis Models

Appendix C: External Links

(A) ETS SMI-9 8 Zone Microphone Audio Surveillance Kit (a1securitycameras.com)
https://www.a1securitycameras.com/ets-smi-9-8-zone-audio-surveillance-kit.html