CS304 Project1 Report

Yiwei Liang, Yuchen Song

January 2024

1 Introduction

This project includes three parts and we combined the second part and the third part together. For the first part, we are asked to do actual speech capture. For the second part, we are asked to build a program with hit-to-talk and automatically when the end of speech is detected. For the third part, the recorded sound is required to be saved in ms wav or raw pcm format.

2 Part I

2.1 Overview

In this part, we used the numpy and pyaudio libraries to record and read the data from the buffer, and then used the Ipython library to display the audio.

2.2 Code explaination

Initialization Define CHUNK size, audio resolution, number of channels, sampling rate, recording time, and the expected number of chunks.

Setup pyaudio Initialize pyaudio and open a stream to capture audio.

Data process The list of frames is converted into a NumPy array, and then the hexadecimal data is converted to decimal using np.frombuffer.

Display audio The captured speech signals are displayed using the ipd.Audio function from the IPython package.

2.3 Result

This part contains the code that can record given time audio and can display after the record.

3 Part II

3.1 Code

Hit-to-talk The hit-to-talk function waits for the 's' key to be pressed, then waits for 1 second before printing "Please talk."

Energy Calculation The EnergyPerSampleInDecibel function calculates the energy of audio frames in decibels.

Initial Background Energy The find-ini-background function calculates the initial background energy.

Endpointing Classification The classifyFrame function classifies frames as speech or non-speech based on energy levels and background.

Recording Loop The record function continuously captures audio frames, classifies them, and accumulates frames until a threshold of non-speech duration is reached. It then saves the recorded speech as a WAV file.

Start Recording System The start function calls the hit-to-talk function and then starts the recording process using the record function.

3.2 Result

The result of this code will save a wav file that records starts when people hit button s and stops at people stop talking.