

The Struggle

By: Akram Kassay

Technical Overview

- Sampler
 - Identified speech segments in live audio stream
 - Uses the Rabiner/Sambur algorithm
- Trainer
 - Produces HMM's from training data input via the EM algorithm
 - Generated .csv results for different state configurations
- Recognizer
 - Linkage between sampler and HMM's
 - Tracks recognition state to only check against appropriate HMM's

Problems...SO MANY Problems

- Too much computation in audio callback
- Performance degradation due to live plotting
- Loss of numerical precision in EM calculation
- Signal scaling
- Determination of states per phrase
- Performance inconsistencies between desktop and laptop

Resolutions...Happy Ending Right?

- Too much computation in audio callback
 - Audio callback bottlenecked by phrase detection, put upper limit on window to check
- Performance degradation due to live plotting
 - Use FuncAnimation and configure code to remove selective plots
- Loss of numerical precision in EM calculation
 - LOG, LOG EVERYTHING
- Signal scaling
 - Sounddevice from -1 to 1 and reading from .wav is -65536 to 65536
- Determination of states per phrase
 - Brute force run different configurations and save results to .csv
- Performance inconsistencies between desktop and laptop
 - Bluetooth sucks, laptop mic sucks, life is hard

Demo?