Signal Processing in Practice

Assignment# 3 (Total Points = 100)
Due date: 11:59PM, February 05, 2024

Using spectrogram as representation for speech sound

1 Dataset

Link: https://tinyurl.com/ycyy3w53

The shared dataset contains three zip files 'clean_speech.zip', 'noisy_speech.zip' and 'rir.zip'.

The 'clean_speech.zip' contains 2 sec segments of sustained phoneme utterances recorded from 30 subjects. The phonemes include three vowels - /a/, /i/, /u/ and three fricatives - /s/, /sh/, /f/. Multiple utterances of a phoneme are recorded by a subject. The utterances are in .wav format. The files are named as: SUB< 3 digit subject ID >_< phoneme >_< utterance ID >.wav. The data is divided into 'train' and 'test' segments.

The 'noisy_speech.zip' contains noisy versions of the files present in 'clean_speech.zip'. For each clean speech file, one among white noise, pink noise, babble noise and high-frequency radio channel noise is selected randomly and added to the clean speech at 5dB SNR.

The 'rir.zip' contains 12 different room impulse responses.

2 Problem Statement

1. Build a 3-class vowel classifier to perform /a/ vs. /i/ vs. /u/ classification.

   (a) Use spectrogram of clean speech.

   (b) Use spectrogram of noisy speech.

   (c) Randomly select $x\%$ of rows of clean spectrogram and replace with uniformly distributed random noise for $x = 1, 5, 10, 20, 50$. Use the modified spectrogram.

   (d) Randomly select $x\%$ of columns of clean spectrogram and replace with uniformly distributed random noise for $x = 1, 5, 10, 20, 50$. Use the modified spectrogram.

   (e) Time-scale the clean speech with scale factor 1.2 and 0.8. Use the spectrogram of time-scaled speech.

   (f) Add reverberation of three types to the clean speech and use the revereberated spectrogram. Take any three RIR from the shared dataset.

   In each of the above cases, train your classifier with the train segments and report the classification accuracy on the test segments. Also report the confusion matrix in each of the above cases.

2. Repeat for 3-class consonant classification - /s/ vs. /sh/ vs. /f/.

Write a report providing your explanation and interpretation of the classification results in all the above cases.