

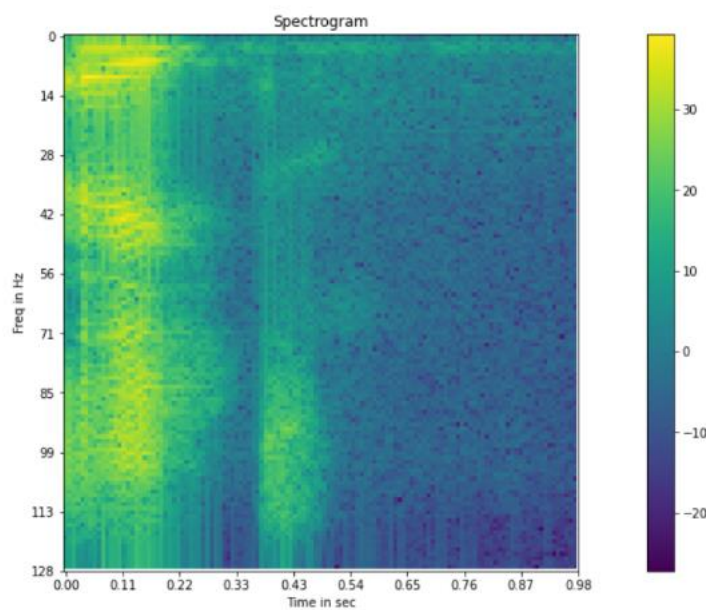
## MCA ASSIGNMENT-2 REPORT

### Spectrogram

It is used to analyse properties of signals over time. The basic idea behind spectrogram is:

- Colour code the magnitude to get a plot of 2 variables received from the DFT; dark represents small magnitude while brighter colour represents larger magnitude.
- Use logarithmic to compress larger values
- Plot spectral slices one after another to obtain image like picture

### **Spectrogram of an audio file**



After receiving a spectrogram of each file, the following features were extracted for more deeper analysis.

- Zero Crossing Rate
- Spectral Energy
- Spectral Entropy

SVM shows an accuracy of around 10%.

### MFCC

MFCC are used in speech recognition. Main steps are:

- Frame signal into short frames
- Apply mel filterbank
- Take logarithmic of all filterbank energy
- Take DCT

- Keep DCT coefficients 2 to 13

Along with the extracted features, SVM shows accuracy of around 31%.

### Noise Augmentation

3 noise files were taken. Out of these, one file was randomly selected and added in each file after multiplying by a factor of 0.005.

Using SVM, the class was predicted on the testing set. Below are the results.

Accuracy with noise augmented audio files:

- Spectrogram : 10%
- MFCC : 40%

### ANALYSIS

Spectrogram uses a linear frequency scaling, so each frequency bin is spaced the equal number of Hertz apart. The mel-frequency scale on the other hand, is a quasi-logarithmic spacing roughly resembling the resolution of the human auditory system. Since, MFCC features are more closely related to biological features, they work better with speech recognition. These files had human audio speaking out various digits. Thus, accuracy and precision were both better in case of MFCC.