

Speech Verification and Spoofing Project

The project has 4 separate systems to be developed

1. Automatic Speaker Verification
2. Speech Recognition : Speech to Text
3. Speech Spoofing (Synthetic Voice)
4. Noise Removal System (Digital Signal Processing)

=>

For Speaker Verification, there are two different variants : Text Dependent & Text Independent.
There are 2 main stages in Speaker Verification

1. Feature Extraction
2. Model Building & Testing

There is also a Normalisation Step before Model Building to make the system independent of Input Device used.

Technology available for Feature Extraction

1. **Mel Frequency Cepstral Coefficients (MFCC)**
2. Perceptual Linear Prediction (PLP)
3. Linear Predictive Codes (LPC)

MFCC or PLP is a good choice for Feature Extraction. We can also use Praat Software Package to help with the Feature Extraction.

There are 3 feasible approaches for Model Building

1. **Deep Learning (ANN)**
2. Hidden Markov Models (HMM)
3. Gaussian Mixture Model (GMM)

Most Suitable Technology is Deep Learning as it is more effective and easier to implement. We can probably include elements from HMM or GMM and combine it with our Neural Network Model.

=>

For Speech to Text,

Deep Learning has a clear advantage over other models like GMM.
There are stages in this also

1. Feature Extraction
2. Finding the Most Probable Word

Technology for finding Most Probable word

1. GMM - HMM based
2. Deep Neural Network - HMM based
3. **Deep Belief Network. (DBN)**

Out of these, DBN method was found to be most effective by a Research Paper and DBNs can be implemented using TensorFlow.

=>

Implementation Platforms

1. Feature Extraction can be implemented on Python / MATLAB
2. Noise Removal Algorithms can be prototyped on MATLAB and implemented in Python.
3. DBN, DNNs can be implemented using TensorFlow in Python
4. HMMs and GMMs do not have many stable packages.

=>

Flow of Work

