

Project Title: Speech Recognition for Chhattisgarhi.

Work Place / Department: Department of Electrical Engineering, NIT Raipur.

Sanction letter No: 910/CCOST/MRP/2013 dated on 20-08-2013.

SUMMARY

Automatic speech recognition (ASR) is a feature based data driven technology, requires a relatively large amount of labelled data to train acoustic models with several machine learning paradigms. Basically, feature extraction in front end and pattern recognition at back ends are used in ASR. In this work, Mel frequency cepstral coefficient (MFCC) is used for feature extraction as it mimics the functioning and working of human auditory system very efficiently. Based on data collection and data processing ASR may be implemented with any of isolated, connected or continuous mode of operation. In *isolated speech recognition*, information falling between two silence zones of a word is used to extract speech features. In *continuous speech recognition* system, it is difficult to find several word boundaries present in a sentence. Different utterances of same word may change in different sentences due to co-articulation effects and position of word relative to other words present in a sentence. In given work, we have implemented three machine learning paradigms (ANN, multiclass SVM and HMM) to recognize isolated chhattisgarhi words as well as continuous chhattisgarhi speech under speaker dependent (SD) and speaker independent (SI). Those machine learning techniques have implemented on self-recorded speech data in various geographical locations of Chhattisgarh state of India, which consisting of **19000 utterances** of **95** common but acoustically challenging **words** from **200 speakers** and **2640 sentences** having **125 words** from **220 speakers**. In this work, speaker variability analysis has also done by analyzing the speech in two categories intra-speaker variability and inter-speaker variability. Isolated word recognition rate by using ANN, SVM and HMM are **99.84%**, **94.25%**, **76.45%**, respectively. Also, for the continuous speech recognition, rate is **90.88%**, **87.17%** and **81.25%** for ANN, SVM and HMM, respectively. It is found by experimental results that **ANN provides better accuracy** for isolated word as well as continuous speech recognition.