

MULTI-ACCENT SPEECH RECOGNITION FOR TAMIL-ENGLISH MIXED LANGUAGE SYSTEM

Ragavan Mathusagar

A dissertation submitted in partial fulfilment of the requirement for
MSc Advanced Software Engineering degree in Computer Science

Department of Computing

Informatics Institute of Technology, Sri Lanka

in collaboration with

University of Westminster, UK

2021

Abstract

Speech recognition has been a hot topic as the intelligent era has developed. Even though numerous automated speech recognition (ASR) programs have been available, a significant number of them do not support Tamil with full features. And when it comes to practical usability, just supporting only Tamil language words will not be an option since the English words are often used in the day-to-day conversation. But still considering a mixed language scenario will not totally satisfy the scenario without considering the different accent in language. So that the research has been done by considering both Tamil-English mixed language and accented speech.

Data has been collected from two group of people with different Tamil accent (Sri Lankan Tamil , Indian Tamil) and the ASR models have been created using the open-sourcetool “Kaldi”. “Mel Frequency Cepstral Coefficient”(MFCC) has been used for feature extracting and Hidden Markov Model (HMM) and Gaussian Mixture Model (GMM) based monophone and triphone models were created and tested. From the triphone model, the best model has been selected and used for the hybrid model creation by replacing the GMM with neural network.

Model accuracy has been compared based on the WER and SER value for each model and also benchmarked with the previous systems. The results showed accuracy improvement for the hybrid model compared to triphone models.

Key words : Automatic speech recognition, Neural network , Acoustic models, Kaldi, Mixed language ASR