



**IMPLEMENTATION OF A HARDWARE SYSTEM TO
ASSIST ILLEGIBLE
USING A HIDDEN MARKOV MODEL BASED
SPEAKER INDEPENDENT, CONTINUOUS SPEECH
RECOGNITION SYSTEM FOR SINHALA LANGUAGE**

Submitted by
W.G.D.M. Samankula
(FGS/06/02/16/2012/01)



A thesis submitted to the Faculty of Graduate Studies, University of Kelaniya
in fulfillment of the requirements for the degree of
Master of Philosophy in Computer Science

| | |
|------------------|------|
| ප්‍රවේශ අංකය: | 1307 |
| විවෘත අංකය: | |



May 2016

Abstract

In this thesis, a speaker independent speech recognition system was built to recognize the continuous Sinhala speech sentences using the toolkit, HTK-3.4.1 based on the statistical approach, Hidden Markov Model (HMM). Three hundred sentences were considered for recording. Data recordings were done with 50 males and 50 females and testing was performed by 10 speakers who had and had not participated for the training. The recognized sequence of words are the commands to automate home appliances such as light, television and radio etc., to help people with differently-able to operate equipment.

The different feature extraction methods such as Mel Frequency Cepstral Coefficient (MFCC), Perceptual Linear Prediction (PLP), Linear Predictive Coding (LPC), Bark Frequency Cepstral Coefficients (BFCC), Linear Prediction Reflection Coefficients (LPREFC), LPC Cepstral Coefficients (LPCEPSTRA), log mel-filter bank channel outputs (FBANK) and linear mel-filter bank channel outputs (MELSPEC) were used with different number of feature parameters varied between 4 to 12 by adding log energy coefficients, and their first, second and third derivatives in order to find the optimal number of parameters for each method. The context-independent and context-dependent acoustic models: word-internal and cross-word triphones and tied state triphones were developed. Decision tree state clustering method was applied for creating tied state triphones and the optimal threshold values for the outlier threshold (RO) and the threshold controlling clustering termination (TB) were determined to create the phonetic decision tree in order to get the optimal result. Finally, tied state triphone based multiple mixture models were applied with 2 mixture, 4 mixture, 8 mixture, 16 mixture and 32 mixture systems. The comparison of above mentioned approaches is discussed in detail.

The speech recognition system was physically implemented to provide access from a PC or laptop, based on Arduino UNO board (ATmega328 microcontroller). The identified command is transferred to the Arduino UNO board through serial communication and then a signal is transmitted using Radio Frequency (RF) to operate an electrical home appliance using a wireless transceiver module.

Keywords: Sinhala speech recognition, Hidden Markov Model, Feature Extraction, Operate electrical appliances, Microcontroller.