Hardware Implementation of a Hidden Markov Model Based, Speaker Independent, Continuous, Sinhala Speech Recognition System

W.G.D.M. Samankula[1], N.G.J. Dias[2]

A speaker independent speech recognition system is built to recognize the continuous Sinhala speech sentences using the toolkit, HTK 3.4.1 based on the statistical approach, Hidden Markov Model (HMM). Mel Frequency Cepstral Coefficient (MFCC), Perceptual Linear Prediction (PLP) and Linear Predictive Coding (LPC) are considered as the feature extraction methods. The recognition performance is considered for number of feature parameters varied from 4 to 12, by adding energy coefficients, first and second derivatives of each coefficient, in order to find the optimal number of parameters for each feature extraction method. Three hundred Sinhala sentences were considered for recording in order to create the phonetically balanced dictionary. Data recordings were done with 50 males and 50 females and testing was performed by 25 speakers who had participated 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., and this can help people with motor disabilities to operate equipment.

The speech recognition system was physically implemented to provide access from a PC or a laptop, based on Arduino UNO board (ATmega328 microcontroller). Arduino comes with a simple integrated development environment (IDE) and allows the programmer to write programs for Arduino in C language. The identified command is transferred to the Arduino UNO board through serial communication and the signal is transmitted using Radio Frequency (RF) to operate electrical home appliances from anywhere up to 150 meters using wireless transceiver modules (C1101) with operating frequency 433MHz. Software was developed to operate more than 18 home appliances, but in hardware implementation, only four are tested. Four Arduino UNO boards are used to implement the light and fan on/off control and the door and curtain angle control. On/off control is operated using relays to switch on and switch off. The door and curtain angle control are constructed by motor with the MOSFET transistors (IRFZ44N). Since a high recognition rate of 85% was achieved for MFCC with 7 feature parameters and adding energy coefficients, first and second derivatives in the software analysis of the previous studies, the same model was used to implement the hardware. A different grammar file is created in the language model of the software to achieve high recognition rate, by considering words and phrases that are only needed to operate the hardware.

Key words: Sinhala speech recognition, Hidden Markov Model, Operate electrical appliances, Microcontroller

1 Department of Statistics & Computer Science, University of Kelaniya, Sri Lanka. dileepamadhu@gmail.com

2 Department of Statistics & Computer Science, University of Kelaniya, Sri Lanka. ngjdias@kln.ac.lk