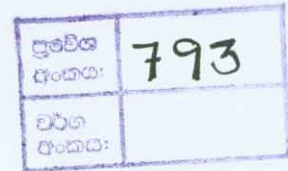# Speaker Dependent Speech Recognition on a Selected Set of Sinhala Words

The Thesis Submitted for the Degree of Master of Philosophy in Computer Science

by

P.G.N.PRIYADARSHANI

(FGS/06/02/08/2008/03)

Department of Statistics and Computer Science

University of Kelaniya

July 2012

ABSTRACT

The thesis is concerned about the Automatic Sinhala Speech Recognition using three distinct pattern recognition approaches: Dynamic Time Warping (DTW), Vector Quantization (VQ), and Genetic Algorithm (GA). First two approaches were speaker dependent and the third approach was speaker independent. The objective of the study was to fulfill the gap currently exists in the field of Sinhala Speech Recognition in Sri Lanka.

The speech recognition process consisted of several phases. Initially, the speech corpus was built for a vocabulary with thousand and fifty five frequently used different Sinhala words. Five speakers were participated in the construction of the corpus of recordings. Next, the cepstral analysis was carried out in order to represent the speech signals in parametric form for further processing. In the feature extraction phase, the speech could be represented with Mel Frequency Cepstral Coefficients in conjunction with their first and second order derivatives. Finally, three pattern recognition approaches were used separately to recognize the Sinhala word utterances. Considerable success was achieved for each individual approach but with significant differences of recognition rates. Maximum recognition rates of VQ and DTW were 94% and 98% respectively. A comparison of these two approaches is discussed in detail. GA achieved 96% overall accuracy.

Keywords- ASR; VQ; DTW; GA; PLP; LPCC; MFCC; hamming window; FFT; DCT; Euclidean distance