Spoken Isolated Word Recognition of Punjabi Language Using Dynamic Time Warp Technique

Ravinder Kumar¹ and Mohanjit Singh²

¹ Department of Computer Science & Engineering,
Thapar Univeristy, Patiala – 147004, India
ravinder@thapar.edu

² Tech., Shaw Communications,
4950 47 St NE, Calgary, Canada
mohanjitsd@gmail.com

Abstract. This research work is to develop a speech recognition system for speaker dependent, real time, isolated words of Punjabi language. The methods used for speech recognition have since been developed and improved with increasing accuracy and efficiency leading to a better human machine interface. In this work, I have developed a speech recognition system, which has a medium size dictionary of isolated words of Punjabi language. The study involved the detailed learning of the various phases of the signal modeling process like preprocessing and feature extraction as well as the study of multimedia API (Application Programming Interface) implemented in Windows 98/95 or above. Visual C++ has been used to program sound blaster using MCI (Media Control Interface) commands. In this system the input speech can be captured with the help of microphone. I have used MCI commands and record speech. The sampling frequency is 16 kHz, sample size is 8 bits, and mono channels. The Vector Quantization and Dynamic Time Warping (DTW) have been used for the recognition system and some modifications have been proposed to noise detection, word detection algorithms. In this work, vector quantization codebook of size 256 is used. This size selection is based on the experimental results. The experiments were performed with different size of the codebook (8, 16, 32, 64, 128, and 256). In DTW, there are two modes: one is training mode and other is testing mode. In training mode the database of the features (LPC Coefficients or LPC derived coefficients) of the training data is created. In testing mode, the test pattern (features of the test token) is compared with each reference pattern using dynamic time warp alignment that simultaneously provides a distance score associated with the alignment. The distance scores for all the reference patterns are sent to a decision rule, which gives the word with least distance as recognized word. Symmetrical DTW algorithm is used in the implementation of this work. The system with small isolated word vocabulary on Punjabi language gives 94.0% accuracy. System can recognize 20 – 24 words per minute of interactive nature with recording time 3 – 2.5 seconds respectively.

Keywords: dynamic time warp (DTW), linear predictive coding (LPC), Punjabi language, vector quantization (VQ), application programming interface (API), media control interface (MCI).