

LOCAL DTW COEFFICIENTS AND PITCH FEATURE FOR BACK-PROPAGATION NN DIGITS RECOGNITION

Rubita Sudirman¹, Sh-Hussain Salleh¹, Shaharuddin Salleh²

¹Faculty of Electrical Engineering, ²Faculty of Science

Universiti Teknologi Malaysia

81310 Skudai, Johore

MALAYSIA

ABSTRACT

This paper presents a method to extract existing speech features in dynamic time warping path which originally was derived from LPC. This extracted feature coefficients represent as an input for neural network back-propagation. The coefficients are normalized with respect to the reference pattern according to the average number of frames over the samples recorded. This is due to neural network (NN) limitation where a fixed amount of input nodes are needed for every input class. The new feature processing used the famous frame matching technique, which is Dynamic Time Warping (DTW) to fix the input size to a fix number of input vectors. The LPC features vectors are aligned between the source frames to the template using our DTW frame fixing (DTW-FF) algorithm. By doing frame fixing, the source and template frames are adjusted so that they have the same number of frames. The speech recognition is performed using the back-propagation neural network (BPNN) algorithm to enhance the recognition performance. The results compare DTW using LPC coefficients to BPNN with DTW-FF coefficients. Added pitch feature investigate the improvement made to the previous experiment using different number of hidden neurons.

KEY WORDS

dynamic time warping, normalization, linear predictive coding, pitch feature, back-propagation neural network