ABSTRACT

Speech emotion recognition (SER) is the ability to identify the type of emotion from human speech, intending to create a natural interaction between humans and computers or commonly known as human-computer interaction (HCI). To increase human satisfaction while using digital devices such as computers and smartphones, the development of HCI is very much needed. Thus, the research entitled "Emotion Recognition through Speech Signals using Linear Predictive Coding (LPC) with Hidden Markov Model (HMM) and K-Nearest Neighbor (KNN) Classification" is important to do.

This study uses the LPC feature extraction method using orders 1, 8, and 16 for the classification methods used are HMM and KNN. The feature extraction process with LPC is carried out without frame blocking and windowing processes. In each classification method, the number of state parameters for HMM will be changed and the number of neighbors for KNN will be changed. The dataset used is the primary dataset. The dataset retrieval process was carried out centrally in a soundproof room with the help of 10 actors (5 men and 5 women). Each actor will say the word "Dia pacar saya" 10 times for each emotion class (angry, disappointed, sad, and happy). Thus, the dataset obtained is 400 voice samples. Of the 400 voice samples, 80 were selected as the best samples for the development of the SER system in this study.

Based on the system testing process carried out, the LPC parameter of order 16 and KNN classification with neighbors 1, 3, and 5 resulted in the best system performance, with a value of 96.88% and a computation time of 0.51 seconds. Meanwhile, when using HMM the resulting accuracy is 62.50% and the computation time is 17.38 seconds with a total of 10 states. Thus, it can be concluded that the best system performance in identifying the type of emotion occurs when using LPC characteristics of order 16 with classification KNN.

Keywords: *Speech Emotion Recognition, Linear Predictive Coding, Hidden Markov Model, K-Nearest Neighbor.*