

Abstract

Speech to Text is an application of speech recognition that process speech signal (voice) so as to produce output in the form of the information contained in these voices are represented in the form of text. Speech to Text plays an important role in the world of electronics, because the system Speech to Text is to bridge the communication between machines with humans.

In recent years the method of Mel frequency cepstral coefficient (MFCC) has been widely researched and proven success in the feature extraction process voice signals at speech to text system, but this method becomes less robust when it comes to dealing with signals that contain noise. So on this final project will be tested using MFCC feature extraction method combined with Gaussian Mixture Models (GMM), to process the signals contain noise.

The working principle feature extraction is to convert voice signals into multiple parameters, where there is some useless information is discarded without losing the true meaning of the voice signal. Output of feature extraction is to be input in the process of pattern recognition. The methods used in pattern recognition is the method of hidden Markov model (HMM). The working principle of this speech recognition system is to compare the sound information available on the reference model with the input speech information system. From the research results using MFCC-GMM method can increase the rate of 4.54% truth that is derived from the MFCC + GMM 10. As for the length of the frame in accordance with this system is 240 by the distance between frames is 80.

Keyword : *noise, feature extraction, MFCC, GMM*