ABSTRACT

On children dubbers, often problems happen in sound recording such as hard to find a good mood for children, numbers of children who are limited to be dubbers, and others that are still common discrepancy to the expected results. A technique to solve this problem is Voice Conversion.

The most important part in voice conversion is how to model the target signal to be emulated by the source signal. In previous project, Gaussian Mixture Model (GMM) is used as signal modeling method. In that project, GMM modeling can be used to implement conversion function and converted speech signal sound natural. Therefore, this project will be use Hidden Markov Model (HMM) because modeling method for dynamic speech signal is required.

The most optimal HMM parameters is using 4-state from logarithm sum of probability. The highest increase of cepstral RMSE before conversion and after conversion is equal to 32,35% and an average 25,84% which obtained from 400 samples input signals. MOS testing have an average value of 2,505 in terms of similarities and have an average value of 2,805 in terms of quality which obtained from 30 respondents.

Key Words: **Cepstral Analysis, HMM, Voice Conversion, RMSE**