

## ABSTRAK

Pada *dubber* anak-anak, sering terjadi masalah dalam proses perekaman suara. Masalah yang sering terjadi umumnya adalah sulit menemukan *mood* yang baik bagi anak-anak, jumlah anak-anak yang terbatas untuk mengisi suara dan lain sebagainya sehingga masih sering terjadi ketidaksesuaian terhadap hasil yang diharapkan. Salah satu teknik untuk mengatasi masalah ini adalah dengan *Voice Conversion*.

Bagian terpenting pada proses *Voice Conversion* adalah bagaimana memodelkan sinyal target agar bisa ditiru oleh sinyal sumber. Pada penelitian sebelumnya digunakan metode pemodelan *Gaussian Mixture Model* (GMM) [10]. Pada penelitian tersebut pemodelan GMM bisa digunakan untuk mengimplementasikan fungsi konversi dan sinyal hasil konversi yang dihasilkan terdengar natural. Oleh karena itu, dalam penelitian ini digunakan metode *Hidden Markov Model* (HMM) karena dibutuhkan suatu metode pemodelan untuk sinyal yang memiliki karakteristik dinamis.

Didapatkan parameter HMM yang paling optimal adalah dengan menggunakan 4 *state* dari logaritma peluang pemodelan yang dihasilkan. Peningkatan hasil RMSE cepstral dari sinyal sebelum dikonversi dan sinyal setelah dikonversi yang tertinggi yaitu sebesar 32,35% dan rata-rata sebesar 25,84% yang didapat dari 400 data sinyal input. Pengujian MOS memiliki nilai rata-rata 2,51 dari segi kemiripan sedangkan dari segi kualitas memiliki nilai rata-rata 2,81 yang diperoleh dari 30 responden.

**Kata Kunci:** *Cepstral Analysis*, HMM, *Voice Conversion*, RMSE