

ABSTRAK

Dubbing adalah sebuah istilah untuk pengisian suara pada video. Pada proses perekaman suara untuk *dubbing* dalam pembuatan film, terdapat beberapa masalah pada *dubber* (pengisi suara) anak-anak untuk tokoh yang diperankannya, seperti sulit menemukan *mood* yang baik bagi anak-anak untuk diajak dalam proses rekaman, jumlah anak-anak yang terbatas untuk mengisi suara, dan lain-lain menyebabkan waktu pengerjaan film tersebut menjadi lebih lama. Untuk itu diperlukan sebuah sistem yang dapat meniru suara seseorang (konversi suara) dengan tingkat kemiripan yang tinggi.

Pada tugas akhir ini dirancang sebuah sistem konversi suara untuk memodelkan sinyal suara target anak-anak agar bisa ditiru oleh sinyal masukan yang baru. Suara target didapatkan dari rekaman suara anak-anak yang berumur 5 s.d. 15 tahun dan suara peniru dapat berasal dari suara orang dewasa. Ekstraksi ciri yang digunakan adalah *Linear Predictive Coding (LPC)* dan metode pemodelan yang digunakan adalah *Hidden Markov Model (HMM)*. HMM merupakan model statistika yang digunakan untuk aplikasi di bidang pengenalan pola temporal, salah satunya adalah analisis sintesis suara. Dalam proses pembuatan sistem ini, bagian yang disintesis ialah frekuensi dasar (F0) dan spektral suara.

Hasil dari konversi suara ini, yaitu orde LPC terbaik yang digunakan adalah orde 19, *state* terbaik yang digunakan dalam pemodelan HMM adalah *state* ke-5. RMSE F0 dari pria dewasa ke anak-anak setelah konversi mengalami peningkatan sebesar 57,7%, sedangkan RMSE F0 dari wanita dewasa ke anak-anak setelah konversi mengalami peningkatan sebesar 15,29%. RMSE Ceptral setelah konversi mengalami peningkatan sebesar 33,25%. Dari segi kemiripan, pengujian MOS untuk HMM memiliki nilai rata-rata 2,64 dan dari segi kualitas, pengujian MOS untuk HMM memiliki nilai rata-rata 3,23.

Kata kunci: Ceptral, *Dubbing*, HMM, LPC, MOS, RMSE, Spektral, *Voice Conversion*