

ABSTRAK

Informasi sinyal suara mempunyai karakteristik yang unik. Berbagai penelitian dalam pengolahan sinyal suara pun telah banyak dikembangkan. Salah satu hasil penelitian tersebut adalah aplikasi *speech-to-text*. Aplikasi ini merupakan cabang dari aplikasi pengenalan suara yang mengolah sinyal suara, mengenalinya, dan mengubahnya menjadi representasi tekstual. Biasanya sistem *speech-to-text* menggunakan fonem sebagai unit pengenalannya. Namun masih terdapat kesulitan dalam mengenali kelas-kelas fonem tertentu yang mempunyai durasi yang cukup singkat. Oleh karena itu, dikembangkan suatu metode pengenalan suara dengan menggunakan database *diphone* (kombinasi fonem) sebagai pengganti unit pengenalannya.

Pada tugas akhir ini penulis mencoba mengeksplorasi penggunaan *diphone* sebagai database referensi dalam sistem pengenalan suara. *Diphone* lebih mudah dikenali karena memberikan informasi akustik yang lebih banyak daripada fonem. Metode klasifikasi yang digunakan dalam tugas akhir ini adalah *hybrid Hidden Markov Model* dan Algoritma Genetika (HMMGA). Ekstraksi ciri sinyal suara menggunakan tiga metode yang berbeda, yaitu: LPC (*Linier Predictive Coding*), MFCC (*Mel Frequency Cepstral Coefficient*) dan HFCC (*Human Factor Cepstral Coefficient*).

Pengujian menggunakan tiga metode ekstraksi ciri sinyal suara pada sistem HMM untuk pengenalan kata menghasilkan tingkat akurasi terbaik 85.56 % (LPC), 57.78 % (MFCC), dan 66.67 % (HFCC), dan untuk pengenalan *diphone* pada kata menghasilkan tingkat akurasi terbaik 92.22 % (LPC), 76.67 % (MFCC), dan 80.00 % (HFCC). Sedangkan pada pengujian sistem *hybrid HMMGA* terjadi peningkatan antara 3-15 % dari sistem HMM terburuk.

Kata Kunci: pengenalan suara, *speech-to-text*, *hybrid Hidden Markov Model* dan Algoritma Genetika (HMMGA), *Linier Predictive Coding* (LPC), *Mel Frequency Cepstral Coefficient* (MFCC), *Human Factor Cepstral Coefficient* (HFCC).