

Abstrak

Speech to Text adalah salah satu aplikasi dari *speech recognition* yang mengolah sinyal bicara (*voice*) sehingga menghasilkan output berupa informasi yang terkandung dalam suara tersebut yang direpresentasikan dalam bentuk teks. *Speech to Text* berperan sangat penting dalam dunia elektronik, karena sistem *Speech to Text* menjadi jembatan komunikasi antara mesin dengan manusia.

Beberapa tahun belakangan ini metode Mel frequency cepstral coefficient (MFCC) telah banyak diteliti dan terbukti keberhasilannya dalam proses ekstraksi ciri sinyal suara pada sistem *speech to text*, namun metode ini menjadi kurang robust ketika harus menangani sinyal yang mengandung noise. Karena itulah pada tugas akhir ini akan diujikan penggunaan metode ekstraksi ciri MFCC yang digabung dengan Gaussian Mixture Models (GMM), untuk memproses kondisi sinyal mengandung noise.

Prinsip kerja ekstraksi ciri adalah dengan mengkonversi sinyal suara ke dalam beberapa parameter, dimana ada sebagian informasi tidak berguna yang dibuang tanpa menghilangkan arti sesungguhnya dari sinyal suara tersebut. Hasil outputan dari ekstraksi ciri ini menjadi masukan pada proses pengenalan pola. Metode yang digunakan pada pengenalan pola adalah metode hidden markov model (HMM). Prinsip kerja sistem pengenalan suara ini adalah dengan membandingkan informasi suara yang ada pada referensi model dengan informasi ucapan yang menjadi masukan sistem tersebut. Dari hasil penelitian menggunakan metode MFCC-GMM dapat meningkatkan tingkat kebenaran sebesar 4.54% yakni berasal dari MFCC+GMM 10, sedangkan untuk panjang frame yang sesuai dengan sistem ini adalah 240 dengan jarak antar frame sebesar 80.

Kata Kunci : *sinyal suara, ekstraksi ciri, MFCC, GMM, noise*