

Abstrak

Pengenalan suara dapat dilakukan dengan cara merepresentasikan sinyal suara ke dalam suatu bentuk yakni representasi dalam bentuk gelombang atau representasi dalam bentuk parameter. Dalam Tugas Akhir ini, dilakukan pengenalan sinyal suara dalam bentuk parameter, dimana sinyal suara diekstraksi vektor cirinya untuk mengambil informasi sinyal dan kemudian dibuat model pembentuknya untuk masing-masing kata yang ingin dikenali.

Setiap sinyal memiliki ciri khasnya masing-masing. Hal ini dapat dilihat jika kita melakukan pembesaran gelombang sinyal suara sebesar 10 ms sampai 30 ms. Sinyal suara yang sangat singkat tersebut kemudian disimpan ke dalam frame-frame dan dicari koefisien cepstralnya (vektor ciri) menggunakan *Linear Predictive Coding* (LPC).

Selanjutnya dilakukan kuantisasi pada tiap vektor ciri yang diperoleh dengan menggunakan metode *k-means cluster* sehingga menghasilkan simbol observasi. Tiap urutan simbol observasi membentuk sekuen *state* yang berbeda. Setiap kata yang ingin dikenali dimodelkan dengan arsitektur HMM sehingga menghasilkan model kata. Untuk proses pengenalannya diterapkan evaluasi pada tiap-tiap model. Model yang paling mendekati adalah model kata yang dicari. Hasil menunjukkan pemilihan ukuran *codebook* yang tepat dapat memberikan keakuratan data di atas 90%, sedangkan pemilihan jumlah *state* tidak terlalu berpengaruh terhadap keakuratan data.

Kata Kunci: *speech-to-text*, *Linear Predictive Coding* (LPC), *Hidden Markov Model* (HMM), ukuran *codebook*, jumlah *state*