

Abstrak

Sistem pengenalan suara membutuhkan model akustik dan model bahasa agar dapat bekerja. Model akustik berasal dari file suara yang kemudian diubah menjadi hmm oleh HTK. Model bahasa statistik menggunakan teks corpus Bahasa Indonesia yang kemudian diubah menjadi model bahasa trigram, untuk model bahasa *syntactic*, digunakan Julius untuk pembentukan grammar. Setelah model akustik dan model bahasa statistik dan *syntactic* didapat, digunakanlah Julius untuk sistem pengenalan suara.

Pada penggabungan model bahasa dengan model akustik 16khz, didapat kebenaran kalimat tertinggi yaitu 82%, kebenaran kata 94.7%, dan akurasi kata 94.7%, dengan akurasi kata model bahasa statistik 93.43% dan akurasi kata pada model bahasa *syntactic* 82.63%.

Kata kunci: sistem pengenalan suara kosakata besar, julius, htk, model bahasa.