

ABSTRAK

Seiring perkembangan teknologi informasi, proses belajar - mengajar semakin banyak menggunakan media alternatif. Salah satu media pembelajaran alternatif yang digunakan adalah video. Untuk mempermudah pemahaman, biasanya video pembelajaran dilengkapi dengan *caption* atau teks keterangan tentang apa yang dibicarakan oleh pembicara. Akan lebih menghemat waktu dan energi apabila *caption* dihasilkan secara otomatis berdasarkan apa yang diucapkan pembicara. Oleh karena itu pada penelitian kali ini akan dibangun sistem *Automatic Video Captioning* menggunakan teknologi *Speech Recognition*.

Sinyal suara dari video masukan diekstrak dan diproses dengan sistem *speech recognition* untuk menghasilkan teks yang sesuai. Pada penelitian ini sistem *speech recognition* dibangun dengan *Linear Predictive Coding* untuk ekstraksi ciri dan *Hidden Markov Model* untuk pencocokan ciri. Teks yang dihasilkan dari sistem *speech recognition* tersebut kemudian digunakan sebagai *caption* dari video masukan.

Pengujian sistem dilakukan dengan mengubah-ubah jumlah data latih dan parameter HMM yaitu jumlah *state* dan jumlah *cluster* untuk mencari parameter dengan akurasi paling optimal. Dari hasil pengujian didapatkan akurasi tertinggi sebesar 75,50% pada jumlah *state* 6, jumlah *cluster* 256 dan data latih sebesar 90 untuk setiap suku kata dalam *database*.

Kata kunci: *Automatic Video Captioning*, *Speech recognition*, *Linear Predictive Coding* (LPC), *Hidden Markov Model* (HMM)