APLIKASI VOICE COMMAND PADA WINDOWS XP DENGAN METODA HIDDEN MARKOV MODEL (HMM)

Ahmad Yani[1], Iwan Iwut Tritoasmoro[2], Hertog Nugroho Pd.[3]

[1]Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom

---

Abstrak
Hidden Markov Model adalah suatu rantai Markov dimana keluaran atau fungsi peluang yang menggambarkan simbol keluaran berhubungan dengan state dan transisi antar state.

Pada tugas akhir ini telah dirancang dan direalisasikan suatu sistem yang dapat mengidentifikasikan suara manusia. Masukan berupa suara manusia yang berisi perintah-perintah dan keluaran eksekusi perintah pada windows. Sistem identifikasi suara ini terdiri dari pre-processing, ekstraksi ciri menggunakan Mel Frequency Cepstral Coefficient (MFCC) dan pengklasifikasian suara, dengan metode pengklasifikasiannya adalah Hidden Markov Model (HMM).

Adapun hasil pengujian dapat membedakan antar perintah dan dapat melakukan eksekusi aplikasi windows. Sistem bekerja dengan akurasi terbaik sebesar 73% dengan kondisi S/N sebesar 33 dB dan pada S/N 20 dB akurasi turun menjadi 42%.

Kata Kunci : hidden markov model, HMM, MFCC, speech recognition

---

Abstract
Hidden Markov Model is a Markov chain where the output or function that describes the chances of the output symbols associated with the state and interstate transition.

At the end of this final project has been designed and realized a system that can identify the human voice. Input is a human voice which contains the voice commands and the output is an execution of windows application. Voice identification system consists of pre-processing, feature extraction uses Mel Frequency Cepstral Coefficient (MFCC) and classification of voice, with the method of classification is Hidden Markov Model (HMM).

The results of the test can distinguish between the command with the best accuracy by 73% with the condition S/N of 33 dB and with S/N 20 dB, accuracy dropped to 42%.

Keywords : hidden markov model, HMM, MFCC, speech recognition

---

BAB I
PENDAHULUAN

1.1 Latar Belakang

Suara merupakan suatu bentuk lazim dari sebuah komunikasi antar manusia dan saat ini berkembang menjadi komunikasi antara manusia dan sebuah perangkat mesin. Suara mewakili suatu komponen penting dari komunikasi *digital* dan mendasari perkembangan teknologi telekomunikasi.

Perkembangan terakhir menciptakan kemungkinan bahwa suara dapat digunakan dalam sistem *security*. Dalam pengenalan suara, tujuan yang ingin dicapai adalah menggunakan *sample* dari suara untuk menentukan identitas seseorang yang mengeluarkan suara tersebut diantara sejumlah orang yang berbicara. Teknik ini memungkinkan penggunaan suara seseorang untuk memverifikasi identitas dan mengendalikan akses terhadap suatu layanan tertentu misalnya panggilan suara (*voice dialing*), telepon *banking*, akses *database*, akses informasi, bahkan akses ke suatu tempat/ruangan tertentu. Tidak sampai disini, perkembangan ini akan terus berlanjut dan suara pasti akan tetap menjadi pusat penelitian sebagai kunci dari sistem komunikasi *universal*.

Teknologi *speech recognition* adalah teknologi yang menggunakan peralatan dengan sumber masukannya adalah suara, seperti mikrofon untuk menginterpretasikan suara manusia untuk transkripsi atau sebagai metode alternatif interaksi dengan komputer. Dengan menggunakan teknologi *speech recognition* memungkinkan pengguna untuk berbicara secara langsung dan cepat serta efisien daripada harus mengetikkan suatu perintah dengan menggunakan *keyboard* dan *mouse*.

Oleh karena itu dengan memanfaatkan metoda *Hidden Markov Model* (HMM) dikembangkan suatu perangkat lunak *voice command* berbasis *speech recognition* yang berfungsi untuk menjalankan file-file *executable* pada sistem operasi *Windows*.

1.2 Tujuan

Adapun tujuan yang ingin dicapai yakni:

1. Merancang dan mebuat program aplikasi *voice command* pada *windows* dengan menggunakan pendekatan metoda *Hidden Markov Model* (HMM).
2. Mengaplikasikan program untuk menjalankan file-file *executable* pada sistem operasi *Windows*.
3. Melakukan analisa kinerja sistem terhadap background noise.

1.3 Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Masukan perintah suara dibatasi hanya satu orang pembicara. Dalam tugas akhir ini suara penulis dipakai sebagai sumber suara.
2. Perintah-perintah yang dicoba hanya terdiri dari 10 perintah sederhana pada *Windows*, baik berupa kata maupun frasa.
3. Masukan dilakukan ketika pembicara tidak mengalami gangguan suara baik karena sakit ataupun gangguan lainnya.
4. Bahasa yang digunakan merupakan bahasa inggris.

1.4 Metodologi Penulisan

Metodologi penyelesaian masalah yang digunakan adalah sebagai berikut :

1. *Study* literatur

   Pada tahap ini, penulis mencoba mencari literatur yang terkait mengenai *Hidden Markov Model* (HMM). Berdasakan itu, penulis mencoba menentukan spesifikasi teknis yang lebih rinci.

2. Perancangan

   Dengan spesifikasi yang telah ditentukan maka pada tahap ini, penulis mencoba merancang algoritma umum dari sistem kemudian barulah mencoba mencari skema rinci dari masing-masing blok.

3. Realisasi

   Rekayasa dan perancangan sistem berdasaarkan algoritma yang telah dipelajari dan menyesuaikan dengan bahasa pemrograman yang digunakan

4. Pengujian

   Setelah realisasi selesai penulis mencoba melakukan pengujian dengan masukan kata yang bervariasi.

5. Analisa dan Evaluasi

   Tahap ini diperlukan untuk mengevaluasi kinerja dan kehandalan perangkat lunak dalam menerjemahkan perintah.

6. Perbaikan dan Penyempurnaan

   Bila terdapat beberapa kesalahan yang masih dapat diperbaiki, maka pada kesempatan ini penulis berusaha untuk memperbaikinya dan menyempurnakannya.

7. Prototipe

   Tahap ini diperlukan untuk merampungkan dan mengemas sistem kedalam bentuk yang lebih representatif.

8. Penyusunan Laporan

   Pada tahap ini dilakukan penyusunan laporan hasil penelitian yang telah dilakukan dan membuat kesimpulan dari hasil penelitian tersebut.

1.5 Sistematika Penulisan

Untuk memudahkan dalam penyusunan laporan proyek akhir ini, penulis menggunakan sistematika penulisan sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi penjelasan tentang informasi dan identifikasi masalah mulai dari latar belakang masalah, tujuan pembuatan proyek akhir. Disini juga dikemukakan metodologi, batasan masalah dan sistematika penulisan yang digunakan dalam penyusunan laporan proyek akhir ini.

BAB II LANDASAN TEORI

Bab ini menjelaskan secara ringkas dan jelas tentang dasar teori penunjang yang berhubungan dengan perancangan dan realisasi proyek akhir misalnya teori tentang *Hidden Markov Model*.

BAB III PERANCANGAN DAN REALISASI

Pada bab ini dipaparkan mengenai tahap-tahap proses perancangan sistem.

BAB IV PENGUJIAN DAN ANALISA

Bab ini menjelaskan secara ringkas mengenai hasil pengujian kinerja perangkat lunak dan analisa sistem.

BAB V PENUTUP

Bab ini membahas mengenai hasil akhir analisa yang didapat dalam bentuk kesimpulan serta saran-saran penulis yang didasarkan pada perancangan dan realisasi proyek akhir ini, yang mungkin diperlukan untuk pengembangan dan penelitian untuk hasil yang lebih baik.

Telkom University

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil analisis terhadap pengujian yang dilakukan pada aplikasi *voice command* pada *windows xp* dengan metoda *hidden markov model*, maka dapat ditarik kesimpulan sebagai berikut :

1. Telah dirancang dan direalisasikan suatu Aplikasi *Voice Command* pada *Windows XP* menggunakan *Mel Frequency Cepstral Coefficient* dan *Hidden Markov Model*.
2. Proses perekaman suara dapat mempengaruhi performansi sistem. Jika sinyal yang terekam tercampur *noise* maka sistem akan sulit mengekstrak ciri dari sinyal suara sehingga deteksi menjadi salah dan secara keseluruhan dapat menurunkan tingkat akurasi sistem.
3. Pada tugas akhir ini performa sistem mempunyai akurasi terbaik sebesar 73%, yaitu ketika S/N bernilai 34 dB. Data yang diujikan sebanyak 100 data uji tiap percobaan dan dengan data latihnya sebanyak 100 data latih.
4. Dari pengujian sistem terhadap *noise*, didapat bahwa akurasi sistem berbanding terbalik dengan *noise background*. Semakin tinggi tingkat *noise* maka tingkat akurasi sistem pun akan semakin menurun.
5. Pada saat S/N 14 dB sistem sudah tidak dapat melakukan pemrosesan suara dengan baik. Hal ini dikarenakan pada pemrosesan awal (*pre processing*), sistem sudah tidak mampu membedakan antara sinyal *voice* dan *unvoice*

5.2 Saran

Pengembangan yang dapat dilakukan pada tugas akhir ini antara lain :

1. Perekaman suara, khususnya untuk database training sebaiknya di dalam ruangan yang bebas noise
2. Pengembangan tugas akhir ini dapat dicoba dengan membuat aplikasi *voice command bahasa Indonesia* dengan menggunakan bahasa pemrograman yang mampu memproses data lebih cepat.
3. Untuk diaplikasikan secara *real-time* dapat ditambahkan metode *Speaker Adaptation*, *Pitch Convertion* dan *Gender Normalization.*

4. *Continuous Hidden Markov Model* (CHMM) dapat dijadikan seebagai pengganti *Discrete Hidden Markov Model* (DHMM) untuk pengenalan sinyal kontinu (kalimat)
5. Penggunaan metode lain untuk pengklasifikasian seperti GA, JST atau yang lainnya.

DAFTAR PUSTAKA

1. Away, Gunaidi Abdia. (2006). "***The shortcut of MATLAB***", INFORMATIKA Bandung

2. B. Kura, Vijay. (2003). "***Novel Pitch Detection Algorithm with Application to Speech Coding***" , B.Tech., Jawaharlal Institute of Technological University.

3. Kanungo, T.,Mount, D. Netanyahu, N., dan Piatko, C. D. (2002). "***An efficient k-means clustering algorithm: Analysis and implementation***".

4. Prahallad, Kishore. "***Speech Technology: A Practical Introduction Topic: Spectrogram, Cepstrum and Mel-Frequency Analysis***"

5. Rabiner, L. and Juang, B. (1993). "***Fundamentals of Speech Recognition***", Prentice-Hall.

6. Shannon, Ben, dan Kuldip K. (2000). "***A Comparative Study of Filter Bank Spacing for Speech Recognition***"