

Aplikasi Pengenalan Ucapan Sebagai Pengatur Mobil Dengan Pengendali Jarak Jauh

Ajub Ajulian Z.
Achmad Hidayatno
Muhammad Widyanto Tri Saksono

Abstract: Growth in Digital signal processing technology gives positives influences in human life. One of the branch of science that gives significant influence is digital speech processing. It can be expand into some applications that make human life easier. Digital speech processing is appropriate to speech recognition.

Speech recognition is used to arrange the movement of remote control car. The remote control car will move according to our speech. This final project is closely related with speech recognition. The LPC (Linear Predictive Coding) method will extract the speech signal features and HMM (Hidden Markov Model) to modelling the speech signal are used. It is done by comparing model from extracting feature that is available in HMM modelling. Models will be used in the speech recognition process, if the models have highest level of conformity.

The experiment has been done in two conditions, i.e ideal condition in room with low noise level and unideal condition in room with noise. Result of the experiment from the whole sistem performace at ideal condition is 97,71% for people that have been inputed in the database, and 95,42% for people which have not been inputed in the database. Result of the experiment of database say it regulary is 97,14%. Result of the experiment of database at unideal condition is 54,28% for inputing word at noisy area, result of the experiment of database for inputing word at high frequency from siren voice is 98,57% and result of the experiment of inputing similar word with database is 97,6%

Key Words : speech recognition, LPC, HMM..

Kemajuan teknologi dalam bidang Pengolahan Sinyal Digital (*Digital Signal Processing*) telah membawa dampak positif dalam kehidupan manusia. Salah satu disiplin ilmu dalam pengolahan sinyal digital yang memberikan dampak yang cukup besar ialah bidang Pengolahan Suara Digital. Pengolahan suara digital dapat dikembangkan berbagai aplikasi yang dapat mempermudah kehidupan manusia. Salah satu aplikasi yang dapat dibuat ialah aplikasi untuk pengatur mobil dengan pengendali jarak jauh. Dalam proses pengatur mobil dengan

pengendali jarak jauh, sistem pengolahan suara digital dapat digunakan untuk mengenali suara masukan dan memberikan suatu keputusan tentang arti dari ucapan tersebut. Proses ini sering disebut dengan proses pengenalan ucapan.

Selama ini, jika ingin menjalankan mobil pengendali jarak jauh perlu dilakukan penekanan suatu tombol. Hal ini menjadi tidak efektif jika ingin menjalankan mobil pengendali jarak jauh masih dengan sistem manual. Tentu saja hal tersebut akan membuang waktu dan tenaga untuk menjalankan mobil pengendali jarak jauh

Ajub Ajulian Z., Achmad Hidayatno (ajub_achmad@elektro.ft.undip.ac.id), adalah dosen di Jurusan Teknik Elektro Fakultas Teknik Universitas Diponegoro (Undip) Semarang Jl. Prof. Sudharto, S.H. Tembalang Semarang 50275

Muhammad Widyanto Tri Saksono adalah mahasiswa di Jurusan Teknik Elektro Fakultas Teknik Universitas Diponegoro (Undip) Semarang Jl. Prof. Sudharto, S.H. Tembalang Semarang 50275

tersebut. Oleh karena itu perlu dikembangkan suatu sistem yang dapat menjalankan mobil pengendali jarak jauh melalui ucapan, yang dapat dijalankan dari jarak jauh.

Dalam Tugas Akhir ini dibuat sebuah aplikasi untuk pengatur mobil dengan pengendali jarak jauh dengan menggunakan masukan berupa suara dan menghasilkan keluaran yang dapat digunakan sebagai pengaktif mobil dengan pengendali jarak jauh

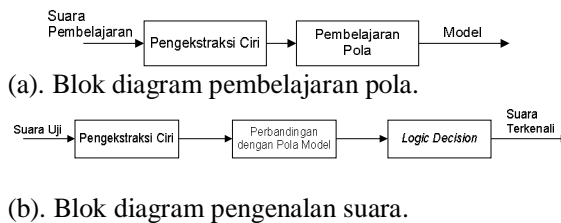
Tujuan

Tujuan dari Penelitian ini ialah mengatur mobil dengan pengendali jarak jauh melalui ucapan.

DASAR TEORI

Pengenalan Suara

Pengenalan suara merupakan salah satu upaya agar suara dapat dikenali atau diidentifikasi sehingga dapat dimanfaatkan. Pengenalan suara dapat dibedakan ke dalam tiga bentuk pendekatan, yaitu pendekatan akustik-fonetik (*the acoustic-phonetic approach*), pendekatan kecerdasan buatan (*the artificial intelligence approach*), dan pendekatan pengenalan-pola (*the pattern recognition approach*). Blok diagram pengenalan pola pada pengenalan suara ditunjukkan pada Gambar 1



Gambar 1 Blok diagram pembelajaran pola dan pengenalan suara.

Berikut ini merupakan penjelasan dari masing-masing blok:

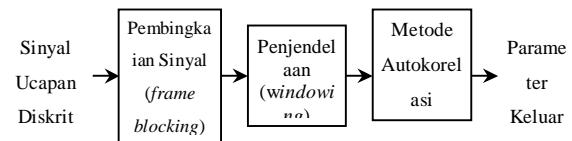
1. **Pengekstraksi ciri.**
Merupakan proses mendapatkan sederetan besaran pada bagian sinyal masukan untuk menetapkan pola pembelajaran atau pola uji.
2. **Pembelajaran Pola**
Satu atau lebih pola pembelajaran yang berhubungan dengan bunyi ucapan dari kelas yang sama, digunakan untuk membuat pola representatif dari ciri-ciri kelas tersebut. Hasilnya yang biasa disebut dengan pola referensi, dapat menjadi

sebuah model yang mempunyai karakteristik bentuk statistik dari ciri-ciri pola referensi.

3. **Perbandingan dengan Pola Model**
Pola uji yang akan dikenali dibandingkan dengan setiap kelas pola referensi. Kesamaan besaran antara pola uji dengan setiap pola referensi akan dihitung.
4. **Logic Decision**
Menentukan kelas pola referensi mana yang paling cocok untuk pola uji berdasarkan klasifikasi pola.

Ekstraksi Ciri

Ciri sinyal ucapan sangat berguna pada sistem pengenalan suara. Salah satu metode yang digunakan untuk proses ekstraksi ciri adalah LPC. Analisis prediksi linear adalah suatu cara yang digunakan untuk mendapatkan sebuah pendekatan sinyal suara. Peramalan linear secara khusus merupakan metode yang cocok dalam pengolahan sinyal suara. Metode ini dapat juga diterapkan dalam pengenalan kata. Tujuan dari digunakannya metode ini adalah untuk mencari nilai koefisien LPC dari suatu sinyal. Gambar 2 menunjukkan blok diagram analisis LPC.



Gambar 2 Blok diagram analisis LPC

Prinsip dasar dari ekstraksi ciri sinyal dengan menggunakan LPC adalah bahwa contoh sinyal ucapan $s(n)$ pada waktu ke- n dapat diperkirakan sebagai kombinasi linear p sampel sinyal ucapan sebelumnya yaitu :

$$s(n) \approx a_1s(n - 1) + a_2s(n - 2) + \dots + a_p s(n - p)$$

Pemodelan Dengan HMM (Hidden Markov Model)

HMM berkembang dengan cepat pada akhir tahun 1960 dan awal tahun 1970 karena pemodelan ini sangat kaya akan struktur matematika dan bisa digunakan untuk beragam aplikasi. HMM merupakan salah satu bentuk model Markov dengan observasi merupakan sebuah fungsi probabilitas dari *state* (keadaan) dengan model yang dihasilkan adalah sebuah proses stokastik. Proses stokastik tidak dapat

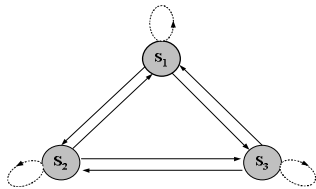
diobservasi langsung tetapi dapat diobservasi hanya melalui rangkaian proses stokastik yang menghasilkan runtun observasi.

Tipe HMM

HMM dibagi menjadi dua tipe dasar yaitu HMM *ergodic* dan HMM Kiri-Kanan

1. HMM *ergodic*

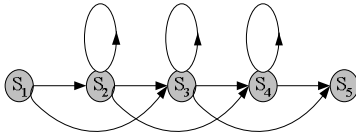
Pada HMM model *ergodic*, perpindahan *state* yang satu ke *state* yang lain semuanya memungkinkan, hal ini ditunjukkan pada Gambar 3.



Gambar 3 HMM model *ergodic*.

2. HMM Kiri-Kanan

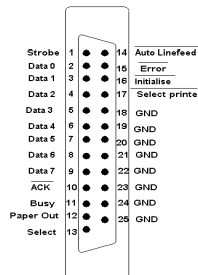
Pada HMM Kiri-Kanan, perpindahan *state* hanya dapat berpindah dari kiri ke kanan saja tidak dapat mundur ke belakang, hal ini ditunjukkan pada Gambar 4.



Gambar 4 HMM model kiri-kanan.

Konfigurasi Port Paralel

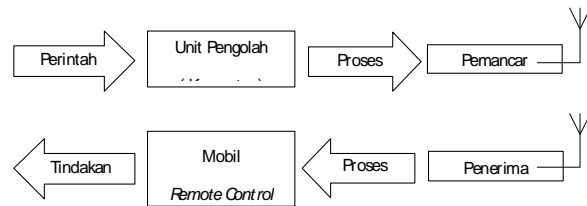
Port paralel adalah *port* yang pada umumnya digunakan sebagai *port* penghubung dengan printer. *Port* paralel juga dapat digunakan untuk keperluan lain, karena *port* paralel dapat digunakan untuk data *in* dan data *out*. Saat *port* paralel mengirimkan data keluaran melalui 8 buah pin-nya, maka pin ini akan mempunyai tegangan 5 volt, tergantung data biner yang dikirim. Gambar 5 menunjukkan konfigurasi *port* paralel DB-25 *female* yang dapat ditemukan di belakang PC.



Gambar 5 Konfigurasi *Port* Paralel DB-25 *female*.

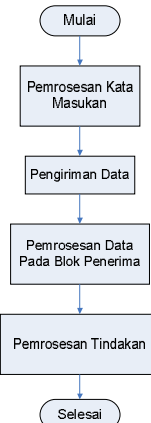
PERANCANGAN DAN IMPLEMENTASI SISTEM

Sistem yang dibuat merupakan suatu sistem yang mampu mengendalikan mobil *remote control* dengan ucapan. Melalui sistem ini, mobil tersebut dapat dikendalikan dengan menyebutkan perintah maju, mundur, kiri, kanan dan mati. Gambar 6 merupakan diagram blok dari sistem mobil pengendali jarak jauh dengan perintah suara.



Gambar 6 Diagram blok sistem pengaktif Mobil Pengendali Jarak Jauh Dengan Perintah Suara.

Perintah dari *port* paralel komputer akan diteruskan ke blok pemancar untuk selanjutnya ditransmisikan ke blok penerima. Blok penerima berfungsi untuk mengolah data yang diterima untuk kemudian diteruskan ke blok alat. Blok alat berfungsi sebagai saklar untuk menghidupkan atau mematikan peralatan elektronik yang terhubung pada blok alat. Blok pemancar dan penerima menggunakan pemancar dan penerima dari mobil *remote control* yang ada di pasaran. Diagram alir dari sistem pengenalan ucapan sebagai pengatur mobil dengan pengendali jarak jauh dapat dilihat pada Gambar 7.



Gambar 7 Diagram alir Sistem Pengenalan Ucapan Sebagai Pengaktif Peralatan Elektronik.

Pengenalan Ucapan

Pembuatan sistem pengenalan ucapan terdiri dari proses pengolahan basisdata, proses ekstraksi ciri, dan proses pemodelan dan pelatihan. Ketiga proses tersebut sering disebut dengan proses pembelajaran. Proses pengenalan ucapan merupakan proses perbandingan antara suara uji dengan model suara yang didapat dari proses pembelajaran.

Pembentukan Basisdata

Basisdata berisi perintah-perintah yang diperlukan untuk menjalankan program ini, seperti kata hidup, mati, satu, dua, tiga, dan empat. Dalam pembuatan basisdata utama digunakan program bantu Cool Edit Pro 2. Basisdata yang diambil memiliki frekuensi cuplik 8000 Hz, 16 bit, mono, yang diperoleh dari lima orang responden yang masing-masing tiga orang berjenis kelamin pria dan dua orang yang berjenis kelamin wanita dengan rentang usia 22-23 tahun. Setelah dilakukan perekaman, suara dari setiap responden akan dipotong-potong dengan durasi satu detik untuk setiap kata.

Dalam basisdata terdapat 300 potong kata yang akan diproses, yang terdiri dari pengucapan kata cepat, pelan, kanan, kiri, mundur, dan mati (tiap kata diucapkan sepuluh kali oleh lima responden).

Ekstraksi Ciri Dengan Metode LPC

Proses pengekstraksian ciri dapat dilakukan dengan menggunakan senarai program sebagai berikut:

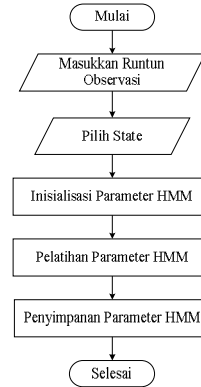
```
s = wavread( 'ABC.wav' );
X = PAR_LPC( s, p, N, M );
```

Dengan:

- s = matrik berkas suara
- p = orde LPC
- N = jumlah *sample* tiap *frame*
- M = jarak antara *frame* yang berurutan

Pemodelan Dengan Menggunakan HMM

Parameter HMM didapatkan melalui lima tahap, yaitu memasukkan runtun observasi hasil dari proses ekstraksi ciri, memilih *state*, inialisasi parameter HMM, pelatihan HMM, pelatihan HMM digunakan untuk mendapatkan parameter yang lebih baik, dan penyimpanan parameter.



Gambar 8 Bagan pemodelan parameter HMM

Proses Antarmuka Sistem

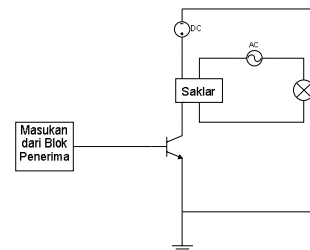
Proses antarmuka sistem menggunakan *port* paralel, tetapi sebelum bisa memakainya terlebih dahulu *port* paralel tersebut perlu didaftarkan. Untuk mendaftarkan *port* paralel yang sudah tersedia di belakang panel komputer, menggunakan instruksi seperti di bawah ini :

```

out =
daqregister('parallel');
dio =
digitalio('parallel', 'LPT1');
lines =
addline(dio, 0:7, 'out');
putvalue(dio, [1 0 0 0 0 0 0 0]);
    
```

Rangkaian Blok Alat

Rangkaian blok alat memerlukan komponen semikonduktor berupa transistor yang digunakan sebagai saklar. Saat transistor mencapai keadaan jenuh, arus akan mengalir dari basis ke kolektor. Sebaliknya saat transistor mencapai keadaan *cut-off* arus mengalir dari basis ke *emitter*. Gambar 9 menunjukkan rangkaian skematik sistem yang dibuat sebagai pengaktif peralatan elektronik.



Gambar 9 Rangkaian skematik perancangan piranti.

PENGUJIAN

Pengujian dilakukan dengan pengujian program pengenalan ucapan.

Pengujian Program Utama

Pada tampilan awal terdapat beberapa pilihan menu, salah satunya adalah menu Program Utama. Menu program utama merupakan menu untuk menampilkan jendela program utama. Untuk membuka jendela program utama maka pada tampilan jendela awal dipilih menu Program Utama, setelah menu program utama dipilih, maka akan muncul jendela program utama yang ditunjukkan pada Gambar 12.



Gambar 12 Tampilan Menu Program Utama.

Pada tampilan program utama terdapat tombol Mulai, tombol ini berfungsi untuk menjalankan program pengenalan ucapan. Bila tombol ini ditekan maka program akan berjalan. Bila pada saat program berjalan terdapat sinyal suara, maka sinyal suara tersebut akan diproses dengan dibandingkan dengan model parameter yang tersedia.

Pengujian Program Pengenalan Ucapan

Pengujian Dalam Kondisi Ideal

Pengujian dalam kondisi ideal dilakukan pada kondisi ruangan yang memiliki derau rendah. Pengujian Basisdata dilakukan dengan cara mengucapkan kata-kata "cepat", "pelan", "kanan", "kiri", "mundur", "mati" oleh 10 responden dengan masing-masing perekaman sebanyak 10 kali. Pada pengujian kinerja sistem, dilakukan dengan cara menjalankan sistem secara keseluruhan. Untuk menghitung persentase keberhasilan digunakan persamaan:

$$(\%) \text{ Pengenalan} = \frac{\text{Jumlah ucapan total yang dikenali}}{\text{Jumlah ucapan total}} \times 100 \%$$

Tabel 1 Hasil pengujian responden yang memasukkan basisdata.

Responden	Jumlah		Keberhasilan
	Masukan	Berhasil	
I	60	54	90%
II	60	60	100%
III	60	50	83,33%
IV	60	52	86,67%
V	60	60	100%

Tabel 2 Hasil pengujian responden yang belum memasukkan basisdata.

Responden	Jumlah		Keberhasilan
	Masukan	Berhasil	
VI	60	60	100%
VII	60	47	78,33%
VIII	60	54	90%
IX	60	53	88,3%
X	60	51	85%

Tabel 3 Hasil pengujian kinerja sistem.

Kata Masukan	Jumlah Pengujian	Keberhasilan
Cepat	10	100%
Pelan	10	80%
Kanan	10	80%
Kiri	10	100%
Mundur	10	80%
Mati	10	100%

Pengujian Dalam Kondisi Tidak Ideal

Untuk selanjutnya dilakukan pengujian dengan kondisi tidak ideal. Pengujian dengan kondisi tidak ideal meliputi:

1. Masukan merupakan kata yang tidak terdapat dalam basisdata. Kata sebagai masukan : "Belok", "Rem", "Gas", "Belakang", "Berhenti", "Depan".
2. Pengujian dengan kondisi ruang pengujian yang berderau. Derau yang digunakan dalam pengujian ini dihasilkan dari pembangkitan sinyal derau dengan menggunakan *software* bantu Cool Edit Pro.

Tabel 4 Hasil pengujian basisdata dengan pengucapan kata di luar basisdata.

Kata Masukan	Jumlah Pengujian	Keberhasilan
Belok	10	70%
Rem	10	70%
Gas	10	70%
Belakang	10	70%
Berhenti	10	80%
Depan	10	80%

Tabel 5 Hasil pengujian kinerja sistem dengan pengucapan kata di luar basisdata.

Kata Masukan	Jumlah Pengujian	Keberhasilan
Belok	10	80
Rem	10	80
Gas	10	80
Belakang	10	80
Berhenti	10	70
Depan	10	70

Tabel 6 Hasil pengujian basisdata dengan kondisi lingkungan berderau.

Kata Masukan	Jumlah Pengujian	Keberhasilan
Cepat	10	50%
Pelan	10	60%
Kanan	10	60%
Kiri	10	50%
Mundur	10	50%
Mati	10	60%

Tabel 7 Hasil pengujian kinerja sistem dengan kondisi lingkungan berderau

Kata Masukan	Jumlah Pengujian	Keberhasilan
Cepat	10	50%
Pelan	10	60%
Kanan	10	60%
Kiri	10	60%
Mundur	10	60%
Mati	10	70%

Dapat dilihat pada Tabel 3 dan Tabel 7 kinerja sistem mengalami penurunan yang sangat jauh bila dibandingkan dengan pengujian dalam kondisi ideal. Penurunan ini dapat dipengaruhi oleh:

1. Pelafalan kata masukan.
Pengucapan ketika pengujian sangat berpengaruh dalam penentuan keberhasilan proses pengenalan. Untuk menghasilkan pengenalan kata masukan sesuai dengan yang diinginkan, perlu diperhatikan kata-kata masukan yang diucapkan.
2. Derau
Derau yang ditambahkan dalam proses pengujian sangat berpengaruh dalam proses pengenalan. Ketika derau ditambahkan pada saat pengucapan kata masukan, derau akan merubah nilai-nilai amplitudo sinyal masukan dan merusak sinyal suara masukan tersebut. Perubahan ini mengakibatkan

kinerja proses pengenalan mengalami penurunan.

KESIMPULAN

Dari hasil penelitian dan pembahasan dapat disimpulkan bahwa :

1. Hasil pengujian basisdata perintah pada kondisi ideal adalah 92% untuk responden yang sudah memasukkan basisdata dan 88,3% untuk responden yang belum memasukkan basisdata.
2. Hasil pengujian kinerja sistem secara keseluruhan pada kondisi ideal adalah 90%.
3. Hasil pengujian basisdata perintah pada kondisi tidak ideal adalah 73,33% untuk pengucapan kata di luar basisdata dan 55% untuk pengucapan kata masukan pada lingkungan berderau.
4. Hasil pengujian kinerja sistem pada kondisi tidak ideal adalah 76,66% untuk pengucapan kata diluar basisdata dan 60% untuk pengucapan masukan pada kondisi lingkungan berderau.
5. *Amplitude* derau yang melebihi ambang tetap terproses dan memberikan sebuah keluaran, tetapi hasil dari pemrosesan derau akan memberikan keluaran yang tidak dapat diprediksi.
7. Lafal pengucapan kata masukan harus jelas dan sesuai dengan pemenggalan kata.

DAFTAR RUJUKAN

- Barmawi, M., Tjia, M., *Elektronika Terpadu*, Jakarta: Erlangga, 1997.
- Edyanto, J., *MATLAB Bahasa Komputasi Teknis*, Yogyakarta: Andi, 2000.
- Hapsari, J.P., *Aplikasi Pengenalan Suara dalam Pengaksesan Sistem Informasi Akademik*, Skripsi S-1, Universitas Diponegoro, Semarang, 2007.
- Rabiner, L., Biing-Hwang Juang, *Fundamentals Of Speech Recognition*, New Jersey: Prentice Hall, 1993.
- Rabiner, L., *Tutorial of Hidden Markov Models and Selected Application in Speech Recognition*.
<http://www.cs.ubc.ca/~murphyk/Bayes/rabiner.pdf>. Februari 1989.
- Upperman, G., *Linear Predictive Coding In Voice Conversion*.
<http://cnx.org/content/m12473/latest/>. Desember 2004.