

KAMUS DIGITAL DENGAR TULIS UCAP BAHASA INDONESIA

Ricky*, Wahyudi, ST, MT **, Achmad Hidayatno, ST, MT **
Jurusan Teknik Elektro, Fakultas Teknik, Universitas Diponegoro
Jl. Prof. Sudharto, Tembalang, Semarang, Indonesia

Abstrak – Setiap kebudayaan besar di dunia bangga akan kamus bahasanya. Pada kenyataannya, kamus itu tidak hanya menjadi lambang kebanggaan suatu bangsa, tetapi juga mempunyai fungsi dan manfaat praktis. Bangsa Indonesia juga boleh berbangga karena memiliki Kamus Besar Bahasa Indonesia. Salah satu cara pelestarian perbendaharaan kata, yaitu dengan melakukan digitalisasi kamus Bahasa Indonesia. Pada Tugas Akhir ini ditambahkan kemampuan kamus ini, yaitu sistem pengenalan suara yang berfungsi mengubah atau mengenali suatu ucapan (bahasa lisan) menjadi teks (bahasa tulisan) dan sistem pembentuk (sintesis) ucapan yang berfungsi yang mengubah teks (bahasa tulisan) menjadi ucapan (bahasa lisan).

Metode pengenalan suara yang digunakan adalah kombinasi metode LPC (Linear Predictive Coding) untuk pengekstraksian suara masukan dan HMM (Hidden Markov Model) untuk pemodelan suara. Pengenalan suara dilakukan dengan membandingkan nilai yang dihasilkan dari ekstraksi ciri dengan model HMM yang telah tersedia. Jika ditemukan model yang cocok, maka model itulah yang digunakan dalam pengisian kata. Proses pembentukan ucapan digunakan bantuan MBROLA speech engine.

Hasil pengujian basisdata suara responden yang memiliki basisdata dengan mengucapkan lema yang terdapat pada basisdata tanpa derau adalah 84,62 %, sedangkan hasil pengujian dengan derau adalah 66,67 %, dan hasil pengujian responden yang memiliki basisdata dengan mengucapkan lema yang tidak terdapat pada basisdata suara adalah 93,75 %. Hasil pengujian basisdata suara untuk responden yang tidak memiliki basisdata suara adalah 75,99%. Hasil pengujian basisdata kata dengan memasukkan kata yang terdapat pada basisdata kata, yang tidak terdapat pada basisdata kata, dan hasil pengujian dengan memasukkan huruf atau suku kata adalah 100 %.

Kata Kunci : LPC, HMM, MBROLA speech engine.

1. Pendahuluan

Secara historis, Bahasa Indonesia merupakan salah satu dialek temporal dari Bahasa Melayu, yang strukturnya maupun khazanahnya sebagian besar masih sama atau mirip dengan dialek-dialek temporal terdahulu, seperti Bahasa Melayu Klasik dan Bahasa Melayu Kuno. Secara sisiologis, dapat dikatakan bahwa Bahasa Indonesia baru dianggap "lahir" atau diterima eksistensinya pada tanggal 28 Oktober 1928. Secara Yuridis, baru tanggal 18 Agustus 1945 Bahasa Indonesia secara resmi diakui adanya.

1.1 Latar Belakang

Pada perkembangannya, ahli linguistik selalu mencari solusi agar perbendaharaan kata dalam Bahasa Indonesia tidak tersisih oleh bahasa-bahasa dari negara lain. Sejalan dengan pemikiran tersebut, mereka mulai melakukan kodifikasi bahasa yang menjadi bagian dari pembakuan bahasa tersebut atau yang lebih dikenal sebagai proses penyusunan kamus.

Sekarang ini, pemikiran akan pelestarian perbendaharaan kata mulai dimasuki oleh bidang ilmu lain, misalnya para insinyur telekomunikasi telah mengembangkan pemikiran untuk mengubah teks menjadi ucapan dan ucapan menjadi teks. Sub-sistem teks menjadi ucapan berfungsi untuk mengubah teks (bahasa tulisan) menjadi ucapan (bahasa lisan), sedangkan sub-sistem ucapan menjadi teks adalah

kebalikan dari fungsi sistem sebelumnya, yaitu sistem yang mengubah atau mengenali suatu ucapan (bahasa lisan) menjadi teks (bahasa tulisan).

Pada era digital, pembangkitan dan pengenalan ucapan dilakukan menggunakan berbagai algoritma pengolahan sinyal digital yang dijalankan dengan perangkat lunak. Berkembangnya teknologi digital menyebabkan perkembangan sistem tulis-ucap atau TTS (*Text to Speech*) dan *Speech Recognition* melahirkan beberapa alternatif baru untuk menjalankan bagian pembangkit ucapan untuk mencari pendekatan yang menghasilkan ucapan yang lebih alami.

1.2 Tujuan

Tugas akhir ini bertujuan untuk membuat program kamus digital yang dapat mengucapkan setiap kata beserta arti kata dan yang dapat mengenali masukan berupa suara diucapkan dalam Bahasa Indonesia.

1.3 Pembatasan Masalah

Untuk mengoptimalkan kinerja program, maka dilakukan hal-hal sebagai berikut.

1. Data masukan berupa suara yang merupakan sebuah lema dari berbagai pengguna.
2. Format lema yang diucapkan hanya berupa sebuah kata.

* Mahasiswa Teknik Elektro UNDIP

** Dosen Teknik Elektro UNDIP

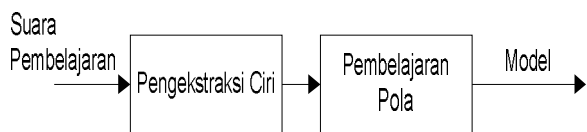
3. Pengujian basisdata suara digunakan sebanyak 106 lema yang diambil dari Kamus Besar Bahasa Indonesia (KBBI).
4. Pengucapan lema yang tidak sesuai dengan waktu yang ditentukan tidak akan direspon oleh program.
5. Basisdata informasi dibuat menggunakan Microsoft Office Access 2003 yang berisi lema beserta artinya (maksimum memiliki lima arti), gabungan kata beserta artinya (maksimum memiliki lima gabungan kata beserta artinya), sublema beserta artinya (maksimum memiliki lima sublema beserta artinya).
6. Pembuatan program dilakukan dengan menggunakan bahasa pemrograman Visual Basic 6 yang digunakan untuk tampilan program dan pengambilan basisdata kata dan Matlab 7.1 yang digunakan untuk proses pengenalan ucapan.
7. Digunakan LPC (*Linear Predictive Coding*) untuk mengekstraksi ciri suara dan HMM (*Hidden Markov Model*) untuk pengenalan pola.
8. Digunakan IndoTTS dalam sistem tulis-ucap.
9. Pengujian basisdata suara untuk responden yang memiliki basisdata dan tidak memiliki basisdata.
10. Pengujian basisdata kata dilakukan dengan mengetikkan lema yang ada, tidak ada pada basisdata, dan mengetikkan huruf atau sukukata.

2. Landasan Teori

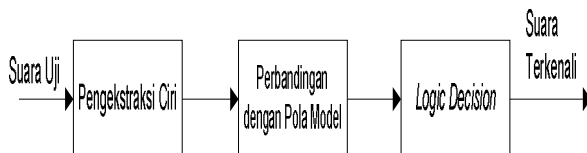
2.1 Pengenalan Suara Digital^{[7] [8] [12]}

Pengenalan suara merupakan salah satu upaya agar suara dapat dikenali atau diidentifikasi sehingga dapat dimanfaatkan. Salah satu pendekatan dalam pengenalan suara ialah pengenalan-pola (*the pattern recognition approach*).

Pendekatan pengenalan pola terdiri dari dua langkah yaitu pembelajaran pola suara dan pengenalan pola melalui perbandingan pola. Tahap perbandingan pola adalah tahap bagi ucapan yang akan dikenali, dibandingkan polanya dengan setiap kemungkinan pola yang telah dipelajari dalam fase pembelajaran, untuk kemudian diklasifikasi dengan pola terbaik yang cocok. Blok diagram pengenalan pola pada pengenalan suara ditunjukkan pada Gambar 1.



(a). Blok diagram pembelajaran pola.



(b). Blok diagram pengenalan suara.

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. Untuk sinyal suara, ciri-ciri besaran biasanya merupakan keluaran dari beberapa bentuk teknik analisis spektrum seperti *filter bank analyzer*, LPC atau DFT (*Discrete Fourier Transform*).

2. Pembelajaran Pola

Satu atau lebih pola pembelajaran yang berhubungan dengan bunyi suara 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.

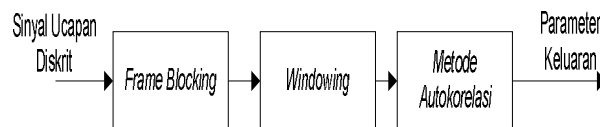
Pengenalan suara secara umum dapat dibagi menjadi tiga tahap, yaitu : ekstraksi ciri, pemodelan, dan pengenalan. Ekstraksi ciri adalah upaya untuk memperoleh ciri dari sinyal suara. Salah satu metode yang dapat digunakan untuk proses ekstraksi ciri adalah LPC. Setelah didapatkan ciri, kemudian dilakukan pemodelan. Sinyal suara dapat dikarakteristikan sebagai variabel proses acak, sehingga untuk pemodelan ini dapat dilakukan dengan pemodelan statistik yaitu HMM. Dari pemodelan akan didapatkan parameter yang selanjutnya digunakan dalam proses pengenalan.

2.2 Ekstraksi Ciri dengan Metode LPC (*Linear Predictive Coding*)^{[7] [8] [12]}

Ciri sinyal ucapan sangat berguna pada sistem pengenalan suara. Salah satu metode yang digunakan untuk proses ekstraksi ciri adalah *Linear Predictive Coding* (LPC). Prinsip dasar dari pemodelan sinyal dengan menggunakan LPC adalah bahwa contoh sinyal ucapan $s(n)$ pada waktu ke- n dapat diperkirakan sebagai kombinasi linier p sampel sinyal ucapan sebelumnya yaitu

$$s(n) \approx a_1 s(n-1) + a_2 s(n-2) + \dots + a_p s(n-p)$$

dengan koefisien a_1, a_2, \dots, a_p diasumsikan bernilai konstan pada *frame* analisis ucapan. Prosedur untuk mendapatkan koefisien LPC diperlihatkan pada Gambar 2.

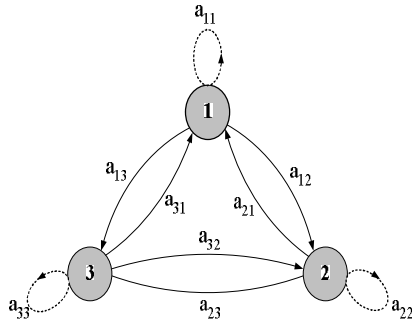


Gambar 2 Blok diagram analisis LPC.

2.3 Pemodelan dengan HMM (*Hidden Markov Model*) [7][12]

Algoritma HMM didasari oleh model matematik yang dikenal dengan rantai Markov. Rantai Markov secara umum ditunjukkan pada Gambar 3. Beberapa hal yang dapat dijelaskan tentang rantai Markov, yaitu:

- ▲ Transisi keadaan dari suatu keadaan tergantung pada keadaan sebelumnya.
 $P[q_t = j | q_{t-1} = i, q_{t-2} = k, \dots] = P[q_t = j | q_{t-1} = i]$
- ▲ Transisi keadaan bebas terhadap waktu.
 $a_{ij} = P[q_t = j | q_{t-1} = i]$



Gambar 3 Rantai Markov.

Elemen yang terdapat pada HMM yaitu :

1. N , jumlah keadaan (*state*) dalam model.
2. M , jumlah simbol observasi yang berbeda tiap keadaan.
3. Distribusi keadaan transisi $A = \{a_{ij}\}$ dengan
 $a_{ij} = P[q_{t+1} = j | q_t = i], \quad 1 \leq i, j \leq N$
4. Distribusi probabilitas simbol observasi, $B = \{b_j(k)\}$ dengan
 $b_j(k) = P(o_t = v_k | q_t = j), 1 \leq j \leq N, 1 \leq k \leq M$
5. Distribusi keadaan awal $\pi = \{\pi_i\}$
 $\pi_i = P[q_1 = i] \quad 1 \leq i \leq N$

2.4 Tulis-Ucap^{[1][2][3][4]}

Pensintesis ucapan adalah suatu sistem yang dapat mengubah deretan kata-kata sebagai masukan menjadi ucapan sebagai keluaran. Sistem pensintesis ucapan juga biasa disebut dengan sistem tulis-ucap. Sistem ini dapat digunakan untuk berbagai aplikasi, sebagai contoh sistem informasi tagihan telepon atau sistem informasi lainnya yang diucapkan secara lisan. Metode sintesis ucapan memungkinkan mesin dapat melewati perintah atau informasi kepada pengguna lewat "ucapan". Proses ini melibatkan pemecahan kata menjadi fonem, menganalisis penanganan khusus dari teks seperti angka, jumlah mata uang, perubahan nada suara, dan juga pemberian tanda-tanda baca.

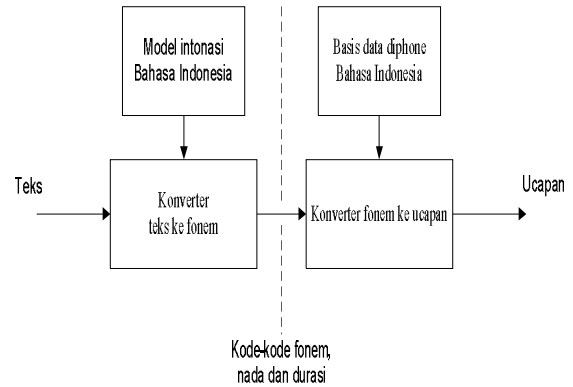
Pesan atau informasi yang dikirimkan lewat ucapan, memiliki kelebihan dalam pentransmisian pesan atau informasi, antara lain:

1. Pengguna dapat dengan mudah memahami pesan atau informasi tanpa perlu intensitas konsentrasi

tinggi. Hal ini berbeda dengan proses membaca. Selain harus mengartikan makna per kalimat, pembaca juga harus berkonsentrasi dengan posisi baris kalimat yang dibaca.

2. Pesan atau informasi dapat diterima saat pengguna sedang terlibat dengan aktivitas lain, misalnya saat berjalan, menangani atau sedang melihat objek lain.

Bagan sub sistem pensistesa ucapan dapat dilihat pada Gambar 4.



Gambar 4 Sub sistem pensintesis ucapan.

Sistem pensintesis ucapan pada prinsipnya terdiri dari dua subsistem dasar, yaitu:

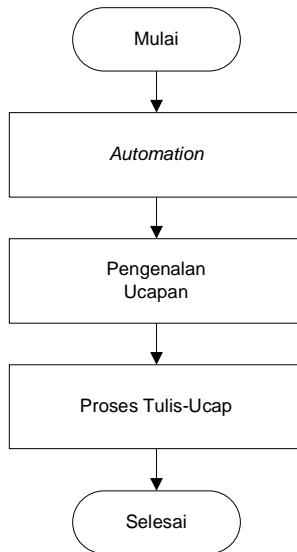
1. Subsistem konverter teks ke fonem.
 Subsistem konverter teks ke fonem berfungsi untuk mengolah kalimat masukan dalam suatu bahasa tertentu yang berbentuk teks menjadi urutan kode-kode bunyi yang direpresentasikan dengan kode fonem, durasi serta nadanya. Kode-kode fonem adalah kode yang merepresentasikan unit bunyi yang ingin diucapkan. Pengucapan kata atau kalimat pada prinsipnya adalah urutan bunyi atau secara simbolik adalah urutan kode fonem.
2. Subsistem konverter fonem ke ucapan.
 Bagian konverter fonem ke ucapan akan menerima masukan kode-kode fonem serta nada dan durasi yang telah dihasilkan oleh bagian sebelumnya. Berdasarkan kode-kode tersebut bagian ini akan menghasilkan bunyi atau sinyal ucapan yang sesuai dengan kalimat yang ingin diucapkan. Ada beberapa teknik yang dapat digunakan untuk implementasi bagian ini. Salah satu teknik yang digunakan adalah penyambungan diphone. Pada sistem yang menggunakan teknik penyambungan diphone, sistem harus didukung oleh suatu basis data diphone yang berisi rekaman segmen-segmen ucapan yang berupa diphone. Ucapan dari suatu bahasa dibentuk dari satu set bunyi mungkin berbeda untuk setiap bahasa, oleh karena itu setiap bahasa harus dilengkapi dengan basis data diphone yang berbeda. Setiap fonem dilengkapi dengan informasi durasi dan nada. Informasi durasi diperlukan untuk menentukan berapa lama suatu fonem diucapkan, sedangkan informasi nada diperlukan untuk menentukan tinggi rendahnya nada pengucapan suatu fonem. Durasi

dan nada bersama-sama akan membentuk intonasi suatu ucapan

Setiap bahasa memiliki aturan cara pembacaan dan cara pengucapan teks yang sangat spesifik. Hal ini menyebabkan implementasi unit konverter teks ke fonem menjadi sangat spesifik terhadap suatu bahasa.

3. Perancangan dan Implementasi Sistem

Dalam perancangan sistem digunakan dua bahasa pemrograman yaitu, Matlab 7.1 dan Visual Basic 6. Kedua bahasa pemrograman tersebut berturut-turut berperan dalam proses pengenalan ucapan dan proses pensintesis ucapan. Secara garis besar, jalannya program dapat dilihat pada Gambar 5.



Gambar 5 Diagram alir jalannya program.

Proses diawali dengan *Automation* yang berfungsi untuk menghubungkan sistem yang telah dibuat pada Matlab 6.5 dengan sistem yang dibuat pada Visual Basic 6. Proses berikutnya yaitu melakukan pengenalan ucapan yang diikuti dengan proses tulis-ucap.

3.1 Automation

Matlab 7.1 dan Visual Basic 6 merupakan dua bahasa pemrograman yang saling terpisah satu sama lain, sehingga dibutuhkan sebuah metode untuk menghubungkan kedua bahasa pemrograman tersebut. Metode yang digunakan dalam Tugas Akhir ini ialah metode *Automation*. *Automation* merupakan protokol COM (*Component Object Model*) yang memungkinkan suatu aplikasi yang bertindak sebagai *client* untuk mengatur jalannya aplikasi lain yang bertindak sebagai *server*. Pada perancangan Tugas Akhir ini, Visual Basic 6 bertindak sebagai *client* sedangkan Matlab 7.1 bertindak sebagai *server*. Setelah hubungan *client-server* terbentuk, maka dapat dijalankan rutin-rutin perhitungan selanjutnya. Bertindak sebagai *client*, Visual Basic 6 harus mengenali aplikasi Matlab sebagai sebuah objek. Senarai program yang digunakan ialah sebagai berikut.

```

Dim matlab as object
Setmatlab=createObject("Matlab.application")
  
```

Senarai program diatas digunakan untuk deklarasi variabel matlab berupa objek sekaligus mengaktifkannya. Nama "Matlab.application" merupakan nama aplikasi matlab yang khusus digunakan dalam proses *Automation*. Setelah dilakukan pengaktifan aplikasi matlab, pemanggilan rutin-rutin pada *server* dapat dilakukan, contoh pemanggilan rutin ialah sebagai berikut.

```
Call matlab.execute("awal")
```

Pada contoh Senarai program diatas dilakukan pemanggilan M-File dengan nama "awal".

3.2 Ekstraksi Ciri

Proses pengestraksian ciri merupakan proses untuk mendapatkan parameter-parameter sinyal suara. Parameter-parameter inilah yang nantinya digunakan untuk membedakan satu kata dengan kata yang lain. Parameter yang dibuat akan disimpan dalam sebuah berkas dengan ekstensi .mat (ParMMN.mat), dengan MM menunjukkan state, dan N menunjukkan nomor urut kata. Selanjutnya pada saat proses pengenalan dilakukan, berkas (ParMMN.mat) inilah yang akan diproses. Proses pengestraksian ciri dapat dilakukan dengan menggunakan senarai program sebagai berikut.

```

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

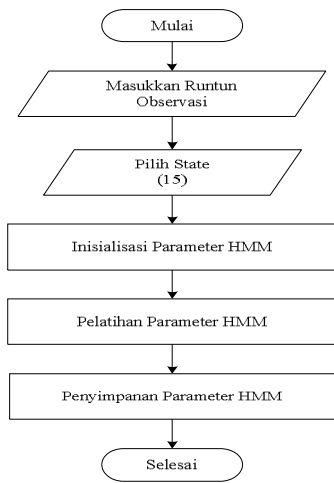
dengan:

s = matrik berkas suara
 p = orde LPC, dengan nilai 10
 N = jumlah *sample* tiap *frame*, dengan nilai 700
 M = jarak antara *frame* yang berurutan, bernilai 100

Dengan kombinasi variabel masukan, maka akan dihasilkan keluaran berupa matrik X dengan MxN, M menunjukkan banyaknya *frame*, sedangkan N menunjukkan bahwa setiap baris terdiri dari koefisien cepstral dan turunan koefisien cepstral terhadap waktu. Hasil dari vektor ciri ini sangat berguna untuk proses pemodelan, pelatihan dan pengenalan.

3.3 Pemodelan

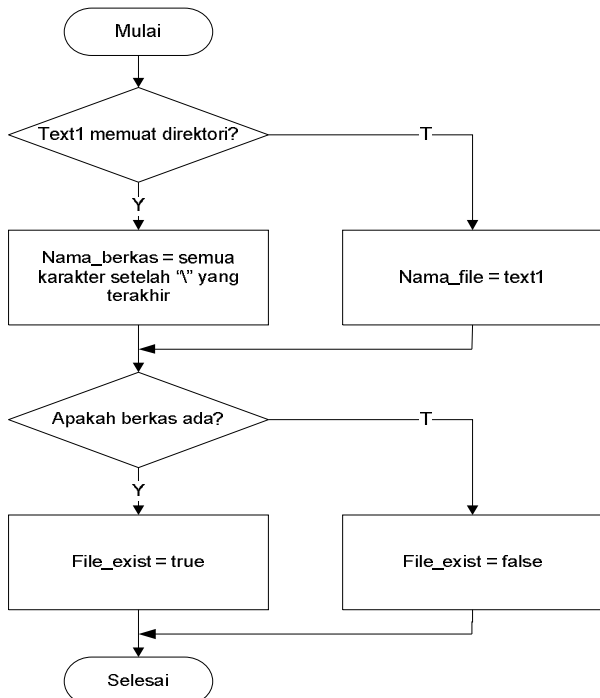
Untuk mendapatkan parameter HMM melalui lima tahap, yaitu : memasukkan runtun *observasi* hasil dari proses ekstraksi ciri, memilih *state*, inialisasi parameter HMM, pelatihan HMM, pelatihan parameter HMM dengan tujuan untuk mendapatkan parameter yang lebih baik dan penyimpanan. *State* untuk pemodelan parameter HMM dapat dipilih dari *state* 15. Proses mendapatkan parameter HMM ditunjukkan pada Gambar 6.



Gambar 6 Bagan pemodelan parameter HMM.

3.4 Pendeteksian Keluaran Server

Proses pengenalan ucapan memberikan hasil akhir berupa variabel dengan nama "maksim". Variabel maksim merupakan sebuah *cell array* (matriks cell). *Cell array* diubah menjadi *char array* (matriks karakter) dengan tujuan agar dapat dibandingkan dengan matrik inisialisasi. Proses yang terakhir yaitu pembentukan keluaran berupa berkas dengan nama pengguna.txt, yang nantinya digunakan sebagai masukan bagi sistem pensintesis ucapan. Gambar 7 menunjukkan diagram alir dari rutin pendeteksi keluaran server.



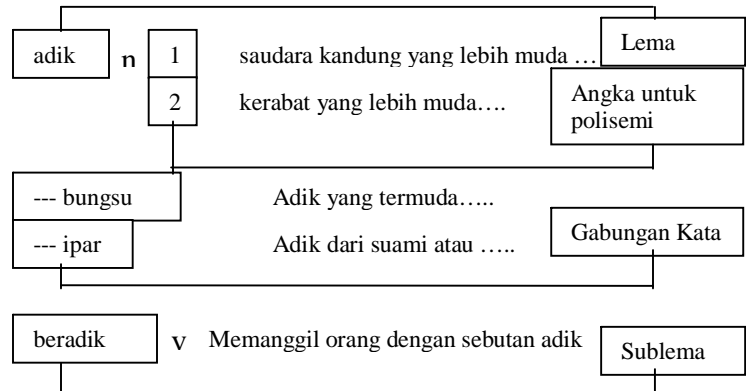
Gambar 7 Diagram alir rutin pendeteksi keberadaan berkas.

Jika terdeteksi berkas yang diinginkan, maka MBROLA akan membaca informasi yang sesuai dengan isi dari *file* pengguna.txt yang terdeteksi.

3.5 Pengolahan basisdata

Basisdata kamus digital ini dibuat dengan menggunakan Microsoft Office Access 2003. Basisdata ini berisi informasi tentang lema, lima arti lema, lima gabungan kata dan artinya, lima sublema dan artinya. Pada rancangan basisdata *field* lema dijadikan *primary key*, ini bertujuan agar tidak terjadi duplikasi data.

Data yang terdapat pada kamus digital ini diambil dari Kamus Besar Bahasa Indonesia (KBBI). Data yang dapat ditambahkan pada kamus digital ini dibatasi pada lema, sublema, dan gabungan kata. Lemma adalah kata dasar, sublema adalah kata dasar yang diberi imbuhan, dan gabungan kata adalah kelompok kata yang merupakan frasa. Masing-masing lema dibatasi hanya memiliki maksimum lima arti (polisemi). Polisemi menunjukkan satu kata memiliki lebih dari satu arti. Sublema dibatasi hanya sampai lima sublema beserta artinya. Gabungan kata dibatasi hanya sampai lima gabungan kata beserta artinya, secara singkat dapat dilihat pada Gambar 8.



Gambar 8 Lemma, gabungan kata, dan sublema.

4. Pengujian

4.1 Jalannya Program

Program utama dijalankan dari kamus.exe. Saat dijalankan, program akan menuju jendela utama. Pada jendela utama terdapat 2 pilihan yaitu mulai, dan keluar. Tampilan jendela utama dapat dilihat pada Gambar 9.



Gambar 9 Jendela utama.

Jika tombol masuk ditekan maka akan muncul jendela pilihan metode untuk mencari kata yang dikehendaki seperti terlihat pada Gambar 10.



Gambar 10 Jendela pilihan metode.

Jendela pilihan metode berfungsi untuk mencari lema yang dikehendaki. Terdapat dua pilihan yaitu metode suara dan metode ketik.



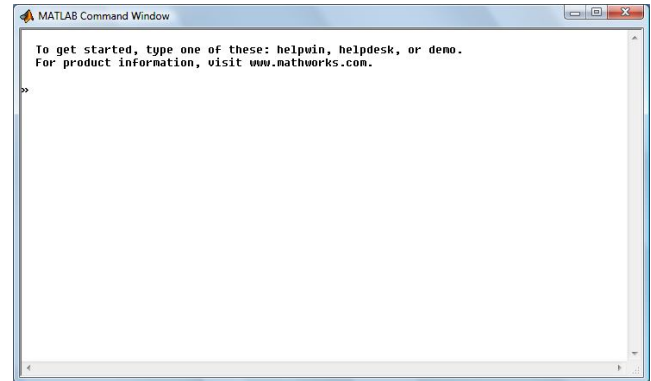
Gambar 11 Jendela metode ketik kata (ketik).



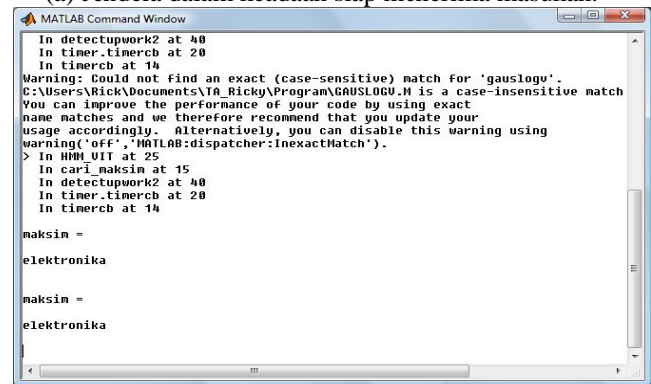
Gambar 12 Jendela metode ucap kata (suara).

Pada pilihan metode suara, munculnya jendela metode suara diikuti dengan terdengarnya sebuah tanda yang menginformasikan kepada pemakai bahwa program siap menerima masukan. Jendela metode ucap kata dapat dilihat pada Gambar 11. Jendela lain yang akan tampil jika metode suara yang dipilih adalah jendela perintah matlab (*Matlab Command Window*). Jendela perintah

Matlab dapat dilihat pada Gambar 13. Pada pilihan metode ketik tidak diikuti dengan jendela perintah Matlab.



(a) Jendela dalam keadaan siap menerima masukan.



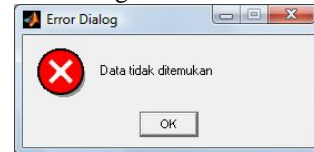
(b) Jendela dalam keadaan telah menerima masukan.

Gambar 13 *Matlab Command Window*.

Jika lema yang dicari tidak terdapat pada basisdata kata dan suara, maka akan muncul kotak dialog seperti yang diperlihatkan pada Gambar 14(a) dan Gambar 14(b).



(a) Kotak dialog tidak ada basisdata kata.



(b) Kotak dialog tidak ada basisdata suara.
Gambar 14 kotak dialog tidak ada basisdata.

4.2 Kehandalan Sistem

4.2.1 Pengujian basisdata suara

➤ Pengujian responden yang memiliki basisdata

- Pengujian lema yang ada pada basisdata

Pada pengujian ini, lema terucap sebagai masukan diucapkan oleh empat responden yang telah memasukkan basisdata suara. Tiap-tiap responden memasukkan basisdata suara masing-masing sebanyak seratus enam kata yang sama untuk semua responden.

Pengujian dilakukan dengan mengucapkan lema yang telah dimasukkan sebanyak sepuluh kali untuk tiap lema dengan memakai basisdata suara responden masing-masing.

o **Pengujian tanpa derau**

Pengujian dilakukan dengan mengucapkan lema pada kondisi ruangan derau rendah. Kondisi derau rendah adalah kondisi dimana variasi dari nilai *amplitude* yang dihasilkan oleh matrik maksim tidak berbeda jauh untuk kata yang sama pada waktu yang berbeda. Hasil pengujian ini ditampilkan pada Tabel 1.

Tabel 1. Hasil pengujian basisdata suara responden yang memiliki basisdata suara tanpa derau.

Responden	kata yang dikenali	Persentase
1	903	85,19%
2	921	86,89%
3	891	84,06%
4	873	82,36%

Hasil rata-rata pengujian dari keseluruhan responden yang memiliki basisdata suara tanpa derau menunjukkan nilai 84,62%.

o **Pengujian dengan derau**

Pengujian dilakukan dengan mengucapkan lema pada kondisi ruangan berderau. Kondisi berderau adalah kondisi dimana variasi dari nilai *amplitude* yang dihasilkan oleh matrik maksim berbeda jauh untuk kata yang sama pada waktu yang berbeda. Derau yang digunakan dalam pengujian ini dihasilkan dari pembangkitan sinyal derau dengan menggunakan *software* bantu Cool Edit Pro. Hasil pengujian ini ditamilkan pada Tabel 2.

Tabel 2 Hasil pengujian basisdata suara responden Yang memiliki basisdata suara dengan derau.

Responden	kata yang dikenali	Persentase
1	721	68,02%
2	745	70,28%
3	669	63,11%
4	692	65,28%

Hasil rata-rata pengujian dari keseluruhan responden yang memiliki basisdata suara dengan derau menunjukkan nilai 66,67 %.

• **Pengujian lema yang tidak ada pada basisdata**

Pada pengujian ini, lema terucap sebagai masukan diucapkan oleh empat responden yang telah memasukkan basisdata suara. Tiap-tiap responden menguji dua puluh dua lema yang sama untuk semua responden dan lema yang diuji tidak terdapat pada basisdata suara sebelumnya. Pengujian dilakukan dengan mengucapkan sepuluh kali untuk tiap lema dengan memakai basisdata suara masing-masing responden. Hasil Pengujian ini ditampilkan pada Tabel 3.

Tabel 3 Hasil pengujian basisdata suara dengan masukan lema di luar basisdata.

Responden	kata yang tidak dikenali	Persentase
1	208	94,46%
2	206	93,64%
3	200	90,91%
4	211	95,91%

Hasil rata-rata pengujian dengan memasukkan lema terucap yang tidak terdapat pada basisdata suara adalah 93,75 %.

➤ **Pengujian terhadap responden yang tidak memiliki basisdata**

Pada pengujian ini, lema terucap sebagai masukan diucapkan oleh empat responden yang belum memasukkan basisdata suara. Pengujian dilakukan dengan mengucapkan lema sebanyak seratus enam lema dengan lema yang diucapkan sama seperti yang ada pada basisdata suara sebelumnya. Tiap-tiap responden mengucapkan sepuluh kali untuk tiap lema dengan memakai basisdata suara responden pertama. Hasil pengujian ini ditampilkan pada Tabel 4.

Tabel 4 Hasil pengujian basisdata suara responden yang tidak memiliki basisdata suara.

Responden	kata yang dikenali	Persentase
1	823	77,64%
2	809	76,32%
3	799	75,38%
4	791	74,62%

Hasil rata-rata pengujian dari keseluruhan responden yang tidak memiliki basisdata suara menunjukkan nilai 75,99%.

4.2.2 Pengujian Basisdata Kata

➤ **Pengujian lema yang terdapat pada basisdata**

Pada pengujian basisdata kata, empat reponden diberi kesempatan mengetik tiga puluh lema tertulis secara bebas masing-masing sepuluh kali untuk tiap lema yang terdapat pada basisdata. Hasil pengujian ini ditampilkan pada Tabel 5.

Tabel 5 Hasil pengujian basisdata kata dengan memasukkan kata yang terdapat pada basisdata kata.

Responden	Jumlah kata benar	Persentase
1	300	100%
2	300	100%
3	300	100%
4	300	100%

Hasil pengujian basisdata kata dengan memasukkan kata yang terdapat pada basisdata kata menunjukkan bahwa semua kata yang terdapat pada basisdata kata dikenali oleh program.

➤ Pengujian lema yang tidak terdapa pada basisdata

Pada pengujian basisdata kata, empat reponden diberi kesempatan mengetik tiga puluh lema tertulis secara bebas masing-masing sepuluh kali untuk tiap lema yang tidak terdapat pada basisdata kata. Hasil pengujian ditampilkan pada Tabel 6.

Tabel 6 Hasil pengujian basisdata kata dengan memasukkan kata yang tidak terdapat pada basisdata kata.

Responden	Jumlah kata salah	Persentase
1	300	100%
2	300	100%
3	300	100%
4	300	100%

Hasil pengujian basisdata kata dengan memasukkan kata yang tidak terdapat pada basisdata kata menunjukkan bahwa semua kata yang terdapat pada basisdata kata tidak dikenali oleh program.

➤ Pengujian huruf atau sukukata

Pada pengujian basisdata kata, empat reponden diberi kesempatan mengetik tiga puluh huruf atau sukukata secara bebas masing-masing sepuluh kali yang terdapat pada basisdata

Tabel 7 Hasil pengujian basisdata kata dengan memasukkan huruf atau suku kata.

Responden	Jumlah pengujian benar	Persentase
1	300	100%
2	300	100%
3	300	100%
4	300	100%

Hasil pengujian basisdata kata dengan memasukkan huruf atau suku kata sembarangan menunjukkan bahwa semua kata yang tampil mengandung huruf atau suku kata yang dimasukkan.

4.2 Analisis

Berdasarkan hasil analisis data pengujian, maka dalam pengenalan pengucap menggunakan analisis LPC dan HMM ini perlu diperhatikan hal-hal sebagai berikut.

1. Kondisi Lingkungan

Kondisi ruangan saat perekaman tidak kedap suara sehingga memungkinkan adanya suara-suara lain yang ikut terekam, sehingga menghasilkan ekstrasi ciri yang tidak mencirikan sinyal asli. Hal ini diatasi dengan melakukan perekaman di ruangan yang memiliki nilai derau yang minimal.

2. Kondisi suara responden

Setiap sinyal suara yang diucapkan oleh seseorang (pria atau wanita) selalu memiliki karakteristik yang berbeda, baik itu panjang-pendek, keras-pelan dan

lain lain. Hal ini dapat diatasi dengan mengusahakan pengucapan sesuai kondisi normalnya.

3. Letak mikrofon.

Perekaman dilakukan secara bertahap sehingga dalam peletakan mikrofon tidak sama. Jarak dan sudut mikrofon sangat mempengaruhi pada sinyal suara yang dihasilkan. Hal ini diatasi dengan mengatur jarak mikrofon dan sudut mikrofon yang sebisa mungkin sama.

4. Kondisi mikrofon.

Perekaman sinyal ucapan harus dilakukan saat mikrofon tidak rusak. Keadaan rusak disini dimaksudkan keadaan mikrofon tidak menambahkan derau. Mikrofon sering jatuh dan kabel yang terkelupas dapat mempengaruhi proses pengenalan dan pengambilan ciri sinyal ucapan.

5. Cara perekaman sinyal suara.

Perekaman sinyal suara yang tepat mempengaruhi ketepatan pengenalan. Cara perekaman yang baik adalah ucapan diucapkan sewajar mungkin dan tidak dibuat-buat, suara diucapkan tidak terlalu keras dan tidak terlalu lemah.

6. Lama waktu perekaman basisdata.

Penentuan lama waktu perekaman yang digunakan sangat berpengaruh, karena untuk waktu ± 1 detik banyak kata-kata pendek yang dapat terwakili, tetapi menjadi masalah pada kata-kata yang panjang misalnya, radioaktifitas. Hal sebaliknya terjadi jika penggunaan waktu yang lebih dari ± 1 detik.

7. Penggunaan basisdata

Basisdata berpengaruh pada akurasi pengenalan, semakin banyak basisdata yang digunakan maka ciri karakteristik sinyal suara yang diperoleh juga akan semakin beragam, sehingga peluang pengenalan sinyal yang diujikan akan semakin kecil. Basisdata yang digunakan sebaiknya direkam tidak pada suatu waktu, sehingga ragam cara pengucapan yang dimiliki seseorang akan terwakili, maka peluang pengenalan akan semakin besar.

5. Penutup

5.1. Kesimpulan

Dari hasil penelitian dan pembahasan dapat disimpulkan bahwa :

1. Metode LPC menghasilkan ciri suara yang berbeda untuk setiap sinyal suara yang berbeda, sedangkan metode HMM menghasilkan nilai maksim yang berbeda untuk setiap lema yang diperoleh dari keluaran metode LPC.
2. Nilai *amplitude* yang dihasilkan oleh matriks maksim tergantung pada lema dan waktu pengucapan.
3. Hasil pengujian basisdata suara responden yang memiliki basisdata dengan mengucapkan lema yang terdapat pada basisdata tanpa derau adalah 84,62 %, sedangkan hasil pengujian dengan derau adalah 66,67 %, dan hasil pengujian

- responden yang memiliki basidata dengan mengucapkan lema yang tidak terdapat pada basidata suara adalah 93,75 %.
4. Hasil pengujian basidata suara untuk responden yang tidak memiliki basidata suara adalah 75,99 %.
 5. Pemberian derau diatas nilai $1.7e+003$ menghasilkan keluaran yang salah.
 6. Hasil pengujian basidata kata dengan memasukkan kata yang terdapat pada basidata kata adalah 100 %, sedangkan hasil pengujian dengan memasukkan lema yang tidak terdapat pada basidata kata adalah 100 %, dan hasil pengujian dengan memasukkan huruf atau suku kata adalah 100 %.

5.2 Saran

Saran yang dapat diberikan sehubungan dengan pelaksanaan penelitian ini adalah :

1. Sistem basis data dapat dikembangkan untuk memuat semua yang terdapat pada Kamus Besar Bahasa Indonesia (KBBI), antara lain: jenis kata, majas, pribahasa, dan lain-lain.
2. Sistem pengenalan ucapan dapat dikembangkan menjadi sistem yang tidak peka terhadap lingkungan derau tinggi, dan gaya ucapan semua pengguna pada berbagai macam kondisi.
3. Disediakan alat perekaman khusus yang dapat mengurangi derau lingkungan agar tingkat kealiamihan kata terucap sebagai masukan semakin tinggi.

Daftar Pustaka

- [1] Akhmad, Arman Arry., *Konversi dari Teks ke Ucapan*. <http://www.indotts.com>. September 2004.
- [2] Akhmad, Arman Arry., *Perkembangan Teknologi TTS Dari Masa ke Masa*. <http://www.indotts.com>. September 2004.
- [3] Akhmad, Arman Arry., *Proses Pembentukan dan Karakteristik Sinyal Ucapan*. <http://www.indotts.com>. September 2004.
- [4] Arrman Arry Akhmad., *Teknologi Pemrosesan Bahasa Alami sebagai Teknologi Kunci untuk Meningkatkan Cara Interaksi antara Manusia dengan Mesin*, <http://www.indotts.com>. Agustus 2004.
- [5] Firdaus., *Visual Basic 6.0 untuk orang awam*, Maxikom, Palembang, 2006.
- [6] Firdaus., *Pemrograman Database dengan Visual Basic 6.0 untuk orang awam*, Maxikom, Palembang, 2005.
- [7] Hapsari J.P., *Aplikasi Pengenalan Suara dalam Pengaksesan Sistem Informasi Akademik*, Skripsi S-1, Universitas Diponegoro, Semarang, 2007.

- [8] Ifeachor, Emmanuel, Barrie W. Jervis., *Digital Signal Processing*, Addison-Wesley Publishing Company, 1993.
- [9] Kridalaksana, Harimurti., *Pembentukan Kata dalam Bahasa Indonesia*, PT. Gramedia Pustaka Utama, Jakarta, 2007.
- [10] Madcoms., *Microsoft Access 2003*, Penerbit Andi, Yogyakarta, 2003.
- [11] Martinet., Andre, *Ilmu bahasa: Pengantar*, Kanisius, Yogyakarta, 1987.
- [12] Rabiner, L., Biing-Hwang Juang, *Fundamentals Of Speech Recognition*, New Jersey: Prentice Hall, 1993.
- [13] Sugiharto, Aris., *Pemrograman GUI dengan Matlab*, Andi, Yogyakarta, 2006.
- [14] -----, *Kamus Besar Bahasa Indonesia*, Edisi ketiga, Balai Pustaka, Jakarta, 2003.
- [15] -----, <http://www.planet-source-code.com>

Biodata Penulis



Ricky (L2F003535)

Dilahirkan di Jakarta, 5 juni 1984. Saat ini sedang menyelesaikan Tugas Akhir pada program S1 di Universitas Diponegoro jurusan teknik elektro konsentrasi Elektronika Telekomunikasi.

Menyetujui,

Dosen Pembimbing I,

Dosen Pembimbing II,

Wahyudi, ST, MT
NIP. 132 086 662

Achmad Hidayatno, ST, MT
NIP. 132 137 933