

VISUALISASI PENGENALAN UCAPAN VOKAL BAHASA INDONESIA DENGAN METODE LPC-DTW

Syaiful Rachman (L2F001644)
Jurusan Teknik Elektro, Fakultas Teknik Universitas Diponegoro
Semarang, Indonesia
Ipoelteknik2001@yahoo.com

Abstrak- Usaha manusia untuk memaksimalkan fungsi PC sebagai alat untuk mempermudah pekerjaan manusia terus dikembangkan. Salah satunya menciptakan PC yang mampu berinteraksi secara langsung dengan manusia melalui ucapan, menggantikan fungsi keyboard dan mouse pada PC. Berdasarkan hal ini berkembanglah studi tentang pengenalan suara digital.

Untuk dapat mengenali kata-kata dalam lingkup yang luas perlu penelitian terhadap elemen penyusun kata yaitu vokal. Visualisasi terhadap pengenalan ucapan berguna bagi seseorang yang ingin mempelajari dan memahami proses pengenalan ucapan dengan metode tertentu. Untuk itu pada tugas akhir ini divisualisasikan proses pengenalan ucapan vokal dalam Bahasa Indonesia yakni /a/, /i/, /u/, /e/, dan /o/ menggunakan metode LPC-DTW. Visualisasi yang dilakukan adalah visualisasi hasil ekstraksi ciri dari metode LPC berupa nilai koefisien-koefisien cepstral LPC, dan juga visualisasi pengenalan pola menggunakan DTW melalui suatu matrik yang dinamakan matrik distorsi.

Ada beberapa hasil yang diperoleh dari penelitian yang dilakukan. Koefisien-koefisien cepstral LPC dapat digunakan untuk mewakili sinyal ucapan dalam proses pengenalan pola DTW. Panjang frame yang digunakan dalam pembingkaiannya suatu sinyal ucapan mempengaruhi jumlah frame yang dihasilkan dari analisis LPC. Pencocokan waktu yang digunakan oleh metode DTW untuk mengatasi adanya variasi yang terjadi pengucapan suatu vokal dapat digambarkan melalui jalur terbaik. Jalur terbaik dari vokal yang dikenali berbentuk diagonal dan akan berupa garis diagonal sempurna apabila 2 buah pola yang dicocokkan identik.

Kata kunci: ucapan, pengenalan, DTW, visualisasi

I. Pendahuluan

1.1 Latar Belakang

Teknologi pengenalan ucapan dikembangkan untuk mewujudkan keinginan manusia dalam memaksimalkan fungsi PC sebagai alat yang mampu mempermudah pekerjaan manusia di segala aspek. Hal yang hendak dicapai adalah menciptakan PC yang mampu berinteraksi dengan manusia secara langsung menggunakan bahasa manusia sehari-hari sesuai tata bahasa yang berlaku. Studi tentang pengenalan ucapan sudah dilakukan selama bertahun-tahun untuk mencapai sukses yang ideal, tetapi hal tersebut belum juga dapat terpenuhi sampai saat ini. Masih perlu dilakukan penelitian dan peningkatan lebih lanjut terhadap metode pengenalan yang sudah ada.

Untuk dapat mengenali kata-kata dalam lingkup yang luas maka perlu penelitian terhadap penyusun kata yakni vokal. Dalam Tata Bahasa Indonesia terdapat lima buah vokal yakni: /a/, /i/, /u/, /e/, /o/. Kelimanya memiliki cara pengucapan berbeda yang menghasilkan ciri sinyal yang berbeda pula. Untuk itu pada tugas akhir ini divisualisasikan proses pengenalan terhadap kelima buah vokal ini, dengan metode LPC (*Linear Predictive Coding*) sebagai pengekstraksi ciri dan DTW (*Dynamic Time Warping*) sebagai pengenalan pola. Visualisasi proses pengenalan ini bermanfaat untuk mempermudah bagi orang yang ingin mempelajari dan memahami proses pengenalan ucapan dengan menggunakan metode LPC dan DTW.

1.2 Tujuan

Tugas akhir ini bertujuan untuk melakukan visualisasi terhadap proses pengenalan ucapan vokal

yang ada dalam Tata Bahasa Indonesia yakni: /a/, /i/, /u/, /e/, /o/, dengan menggunakan metode LPC dan DTW.

1.3 Pembatasan Masalah

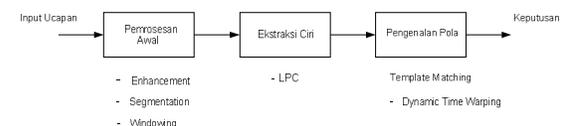
Tugas akhir ini memiliki batasan-batasan antara lain :

1. Pengenalan ucapan bersifat *speaker dependent*. Pengenalan dapat dilakukan dengan baik terhadap orang yang telah memiliki referensi.
2. Merupakan pengenalan vokal /a/, /i/, /u/, /e/, /o/.
3. Metode ekstraksi ciri yang digunakan dalam tugas akhir ini adalah LPC.
4. Metode pengenalan pola yang dipakai untuk tugas akhir ini adalah DTW.

II. Landasan Teori

2.1 Pengenalan Suara Digital

Proses pengenalan ucapan pada program komputer terdiri atas beberapa tahap. Gambar 1 memperlihatkan gambaran secara umum tahapan-tahapan pengenalan suara.



Gambar 1 Tahapan dalam pengenalan ucapan.

Tahap pertama dari pengenalan ucapan adalah pemrosesan awal yang di dalamnya terdapat 3 proses yakni peningkatan kualitas sinyal (*enhancement*), pembingkaiannya (*segmentation*), dan penjendelaan

(*windowing*). Tahap selanjutnya adalah ekstraksi ciri yang bisa dilakukan menggunakan metode LPC. Hasil dari ekstraksi ciri dengan LPC ini berupa nilai koefisien-koefisien *cepstral* LPC yang merupakan ciri khusus dari sinyal suara. Setelah ciri dari sinyal suara didapat, dilakukan pengenalan pola yang bisa dilakukan dengan metode DTW. Metode DTW melakukan perhitungan jarak antara seluruh pola referensi dengan pola uji yang ingin dikenali untuk mendapatkan jarak-jarak pencocokan keseluruhan.

2.2 Analisis LPC^[6]

Suatu sinyal ucapan yang masuk disegmentasi atau dibingkai dengan panjang bingkai (*frame*) tertentu. Terhadap sinyal ucapan terbingkai dilakukan penjendelaan yang dapat dilakukan menggunakan jendela hamming (*hamming window*), yang memiliki persamaan:

$$w(n) = 0,54 - 0,46 \cos\left(\frac{2\pi n}{N-1}\right), \quad 0 \leq n \leq N-1$$

maka akan didapat sinyal ucapan terbingkai yang telah dijendelakan menggunakan jendela *hamming* dengan urutan:

$$s_n(m) = s(n+m) \cdot w(n), \quad 0 \leq m \leq N-1$$

Analisis LPC dilakukan menggunakan metode autokorelasi, yang memiliki bentuk persamaan fungsi autokorelasi:

$$R_n(i-k) = \sum_{m=0}^{N-1-(i-k)} s_n(m) s_n(m+i-k)$$

Jika fungsi autokorelasi bersifat simetris $R_n(k)=R_n(-k)$, sehingga persamaan LPC dapat dinotasikan sebagai:

$$\sum_{k=1}^p R_n(|i-k|) \hat{a}_k = R_n(i), \quad 1 \leq i \leq p$$

Nilai koefisien-koefisien LPC didapat dengan menyelesaikan persamaan matrik berikut:

$$\begin{bmatrix} R_n(0) & R_n(1) & R_n(2) & \dots & R_n(p-1) \\ R_n(1) & R_n(0) & R_n(1) & \dots & R_n(p-2) \\ R_n(2) & R_n(1) & R_n(0) & \dots & R_n(p-3) \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ R_n(p-1) & R_n(p-2) & R_n(p-3) & \dots & R_n(0) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \dots \\ \dots \\ \alpha_p \end{bmatrix} = \begin{bmatrix} R_n(1) \\ R_n(2) \\ R_n(3) \\ \dots \\ \dots \\ R_n(p) \end{bmatrix}$$

2.3 Konversi Parameter LPC ke Koefisien Cepstral^[7]

Koefisien-koefisien *cepstral* LPC adalah parameter yang penting dalam pengenalan ucapan. Koefisien-koefisien *cepstral* LPC dapat diperoleh langsung dari koefisien-koefisien LPC, melalui persamaan:

$$c_m = a_m + \sum_{k=1}^{m-1} \left(\frac{k}{m}\right) c_k a_{m-k}, \quad 1 \leq m \leq p$$

$$c_m = \sum_{k=m-p}^{m-1} \left(\frac{k}{m}\right) c_k a_{m-k}, \quad m > p$$

Koefisien *cepstral* LPC merupakan koefisien-koefisien dari representasi transformasi fourier spektrum magnitude logaritmik. Koefisien-koefisien *cepstral* LPC terbukti memiliki sifat yang lebih *robust* dibandingkan koefisien-koefisien LPC pada pengenalan ucapan. Apabila Q menyatakan jumlah koefisien *cepstral* LPC, maka biasanya representasi dengan nilai $Q > p$, yaitu $Q \sim (3/2)p$ koefisien, sering digunakan pada pengenalan ucapan.

2.4 Pengenalan Pola DTW^[8]

Dynamic Time Warping adalah metode pengenalan pola yang didasarkan pada pencocokan pola sinyal (*template matching*) antara semua *template* referensi yang ada dengan suara uji masukan. Pencocokan dilakukan menggunakan pendekatan untuk mencari jarak yang terkecil. *Template* referensi adalah pola-pola yang dijadikan acuan (referensi) untuk pencocokan.

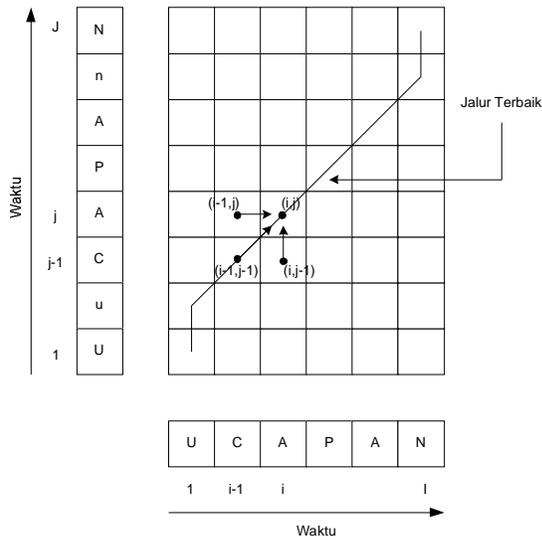
Perhitungan untuk mencari jarak antara 2 vektor ciri diperlukan dalam pengenalan pola DTW. Besarnya jarak antara dua vektor ciri ini yang disebut sebagai jarak lokal (*local distance*). Jarak lokal antara sebuah vektor ciri dari suatu *template* referensi x dengan sebuah vektor ciri dari suara uji y yang masuk dirumuskan dengan sebagai berikut :

$$d(x, y) = (x - y)^2$$

Ucapan adalah proses yang bergantung terhadap waktu. Oleh karena itu pengucapan suatu kata yang sama dapat memiliki durasi waktu yang berbeda. Selain itu suatu kata dengan durasi yang sama dapat memiliki komposisi yang berbeda. Oleh karena itu diperlukan proses pencocokan waktu (*time-alignment*) untuk mengatasi hal itu.

Contoh *time-alignment* dapat dilihat pada Gambar 2, yang menggunakan matrik dalam satuan waktu. Apabila dalam suatu proses pengenalan menggunakan nilai koefisien *cepstral* LPC untuk melakukan perbandingan, maka koefisien-koefisien *cepstral* LPC pada *frame* yang ada digunakan untuk mewakili satuan waktu dari matrik tersebut. Matrik pada Gambar 2.5 digunakan untuk memvisualisasikan secara sederhana bagaimana proses pencocokan waktu dikerjakan. Posisi suara uji berada di sebelah kiri matrik, dengan waktu yang menuju ke atas, sedangkan *template* referensi pada bagian bawah matrik, dengan waktu yang menuju ke kanan. Pada ilustrasi tersebut, suara uji UuCAPAnN adalah salah satu bentuk variasi yang ada yang akan dicocokkan dengan *template* referensi UCAPAN. Cara pendekatan sederhananya adalah untuk huruf yang paling mendekati maka itu dijadikan acuan huruf yang

dipilih, yang juga berarti memiliki nilai jarak yang terkecil bila dibandingkan nilai jarak huruf yang lain.



Gambar 2 Proses pencocokan waktu DTW.

Contohnya adalah “n” pada UuCAPAnN paling dekat adalah dengan “N” dari semua huruf yang ada pada UCAPAN, begitu juga dengan “u” dari suara uji akan memiliki jarak yang terkecil dengan “U” dari *template* referensi. Pendekatan seperti ini akan dilakukan untuk semua *template* referensi yang ada pada sistem. Pencocokan waktu diperlihatkan oleh jalur terbaik yang terbentuk pada matrik.

Matrik yang digambarkan oleh Gambar 2 dinamakan matrik distorsi. Matrik distorsi digunakan untuk menggambarkan jarak-jarak antara seluruh koefisien *cepstral* LPC dari *template* referensi dengan seluruh koefisien *cepstral* LPC dari suara uji, beserta jalur terbaik yang terbentuk hasil pengenalan pola DTW. Satuan waktu dari matrik distorsi diwakili oleh urutan koefisien *cepstral* LPC.

Dynamic Time Warping bekerja berdasarkan pemrograman dinamis (*dynamic programming*) untuk melakukan pencocokan pola, dengan jalan membandingkan jarak, yaitu jarak pada posisi diagonal bawah (i-1,j-1), samping kiri (i-1,j) dan bawah (i,j-1) jarak manakah yang paling kecil jarak itulah yang dipilih. Berdasarkan Gambar 2, dapat dihitung jarak global (D) yang didapat dari jarak lokal yang ada (d). Apabila d(i,j) adalah jarak lokal pada posisi (i,j) pada matrik, maka jarak global D(i,j) pada posisi (i,j) adalah:

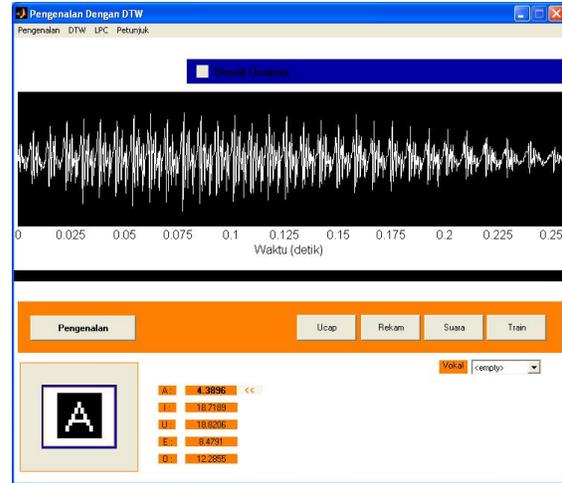
$$D(i, j) = \min[D(i-1, j-1), D(i-1, j), D(i, j-1)] + d(i, j)$$

dengan kondisi awal (inisialisasi) $D(1,1) = d(1,1)$. Nilai jarak global yang terakhir $D(I,J)$ pada posisi (I,J) pada matrik distorsi, menyatakan jarak pencocokan keseluruhan untuk pencocokan sebuah *template* referensi dengan suara uji. Masukkan ucapan vokal akan dikenali sebagai sebuah vokal sesuai *template* referensi yang memiliki jarak pencocokan keseluruhan yang

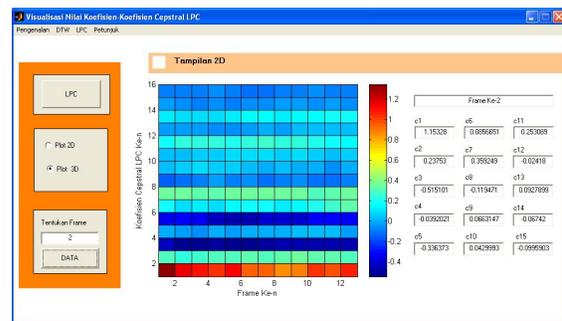
terkecil, atau nilai $D(I,J)$ terkecil bila dibandingkan nilai $D(I,J)$ dari pencocokan dengan *template* referensi lainnya.

III. Tampilan Program

Program visualisasi pengenalan vokal Bahasa Indonesia dengan metode LPC-DTW memiliki 3 mode utama tampilan yakni: mode pengenalan, mode visualisasi nilai koefisien *cepstral* LPC, dan mode visualisasi pengenalan pola dengan DTW. Tampilan mode pengenalan, mode visualisasi nilai koefisien *cepstral* LPC, dan mode visualisasi DTW dapat dilihat pada Gambar 3-Gambar 5.



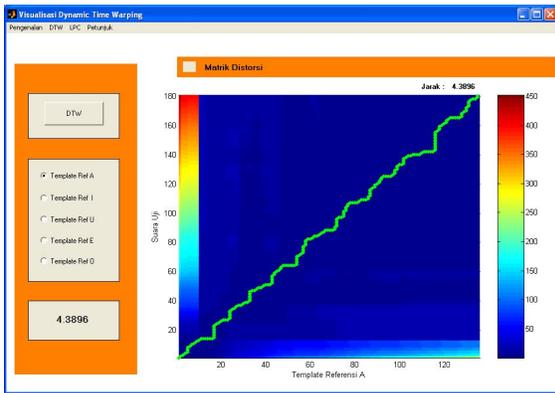
Gambar 3 Mode pengenalan.



Gambar 4 Mode visualisasi nilai koefisien *cepstral* LPC.

Mode pengenalan yang diperlihatkan Gambar 3 digunakan untuk melakukan uji langsung terhadap vokal yang ingin dikenali. Untuk dapat mengenali suatu vokal, maka pemakai terlebih dahulu harus memiliki *template* referensi untuk semua vokal. Sifat pengenalan ucapan yang digunakan pada tugas akhir ini adalah *speaker dependent*, pengenalan dapat dilakukan dengan baik terhadap orang yang telah memiliki *template* referensi. *Template* referensi dapat dibuat langsung pada mode pengenalan. Pemakai dapat langsung merekam ucapan untuk dijadikan *template* referensi dengan menekan **tombol ucap**, diikuti **tombol rekam**. Pelatihan *template* referensi dapat dilakukan dengan menekan **tombol ucap** diikuti **tombol train**. Penjelasan tentang cara menggunakan

mode pengenalan secara lebih detail dapat dilihat pada **menu petunjuk** yang ada pada mode pengenalan. Mode visualisasi nilai koefisien *cepstral* LPC yang diperlihatkan Gambar 4 digunakan untuk menampilkan visualisasi terhadap nilai-nilai koefisien *cepstral* LPC yang didapat dari ekstraksi ciri dengan metode LPC. Sinyal ucapan yang ditampilkan nilai koefisien *cepstral* LPC-nya berasal dari mode pengenalan. **Tombol LPC** digunakan untuk menampilkan visualisasi nilai koefisien-koefisien *cepstral* LPC. Pemakai dapat memilih tampilan visualisasi yang diinginkan baik bentuk visualisasi 2 dimensi maupun 3 dimensi menggunakan **radio button (pilihan tampilan)** yang ada pada sisi sebelah kiri. **Tombol data** digunakan untuk menampilkan nilai koefisien *cepstral* LPC pada *frame* tertentu. Sebelumnya pemakai terlebih dahulu harus mengisi *frame* mana yang ingin dilihat nilai koefisien *cepstral* LPC-nya pada **edit text (tentukan frame)**. Data nilai koefisien *cepstral* LPC ditampilkan pada sisi sebelah kanan dari jendela mode visualisasi koefisien LPC.



Gambar 5 Mode visualisasi DTW.

Mode visualisasi pengenalan pola DTW yang diperlihatkan Gambar 5 digunakan untuk melihat matrik distorsi beserta jalur terbaiknya yang dihasilkan dari proses pengenalan pola menggunakan DTW dari mode pengenalan. **Tombol DTW** digunakan untuk melihat visualisasi matrik distorsi beserta jalur terbaik yang dibentuk dari proses pencocokan suara uji dengan *template* referensi tertentu yang dapat dipilih melalui **radio button (pilihan template referensi)**.

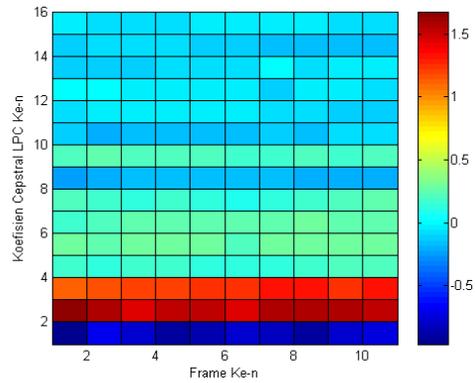
IV. Visualisasi

Visualisasi pengenalan ucapan vokal Bahasa Indonesia menggunakan metode LPC-DTW terdiri atas 2 bagian, yakni visualisasi nilai koefisien *cepstral* LPC dan visualisasi DTW. Visualisasi nilai koefisien *cepstral* LPC menjelaskan tentang visualisasi nilai koefisien-koefisien *cepstral* LPC hasil proses ekstraksi ciri menggunakan LPC. Visualisasi DTW menjelaskan tentang visualisasi proses DTW menggunakan sebuah matrik yang dinamakan matrik distorsi.

4.1. Visualisasi Nilai Koefisien Cepstral LPC

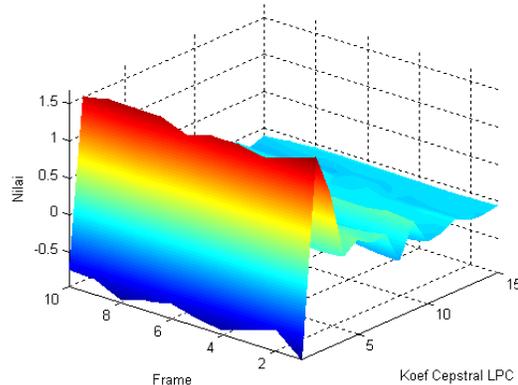
Visualisasi nilai koefisien-koefisien *cepstral* LPC ucapan vokal secara 2 dimensi menggunakan panjang

frame 20 ms atau 160 data cuplikan, dan orde-10 untuk analisis LPC ditunjukkan oleh Gambar 6.



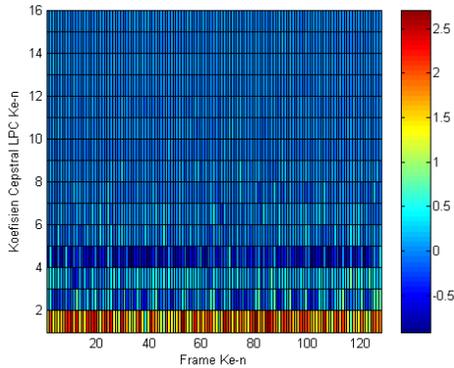
Gambar 6 Visualisasi 2D koefisien LPC.

Visualisasi nilai koefisien *cepstral* LPC secara 2 dimensi digambarkan menggunakan kotak-kotak warna untuk menampilkan nilai koefisien *cepstral* LPC. Sumbu yang ada pada tampilan, yaitu sumbu *y* menggambarkan nilai koefisien *cepstral* LPC ke-*n* sedangkan sumbu *x* mewakili *frame* ke-*n*. Nilai koefisien *cepstral* LPC diwakili oleh sebuah warna pada kotak warna visualisasi 2D. Tingkatan nilai koefisien *cepstral* LPC tersebut dapat dilihat berdasarkan tingkatan nilai warnanya di sebelah kanan dari tampilan 2D. Berdasarkan tingkatan nilai warna, warna merah menunjukkan nilai yang tinggi dan warna biru menunjukkan nilai yang rendah. Nilai koefisien-koefisien *cepstral* LPC yang digambarkan melalui kotak-kotak warna disajikan secara bertingkat untuk setiap *frame*, mulai kotak warna paling bawah menggambarkan nilai koefisien *cepstral* LPC pertama (c_1) hingga kotak warna paling atas menggambarkan nilai koefisien *cepstral* LPC terakhir (c_{15}). Terdapat 15 buah kotak-kotak warna yang disusun bertingkat pada visualisasi nilai koefisien *cepstral* LPC secara 2 dimensi, karena jumlah koefisien *cepstral* yang didapat dari ekstraksi ciri dengan LPC sebanyak $(3/2) \times$ orde LPC yang digunakan, dengan orde 10 maka didapatkan 15 koefisien-koefisien *cepstral* LPC. Tampilan 3 dimensi membawa tampilan 2 dimensi koefisien LPC ke bentuk 3 dimensi ditunjukkan pada Gambar 7.



Gambar 7 Visualisasi 3D koefisien LPC.

Gambar 8 memperlihatkan visualisasi nilai koefisien *cepstral* LPC dengan pembingkai sinyal ucapan menggunakan panjang bingkai (*frame*) 2 ms atau 16 data cuplikan.



Gambar 8 Pembingkai dengan *frame* pendek.

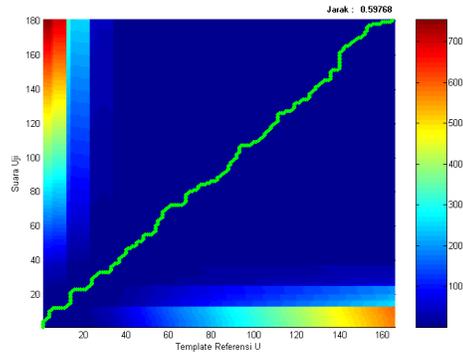
Apabila sinyal ucapan dibingkai menggunakan panjang bingkai yang lebih pendek maka akan didapat jumlah *frame* sinyal ucapan yang lebih banyak.

4.2. Visualisasi DTW

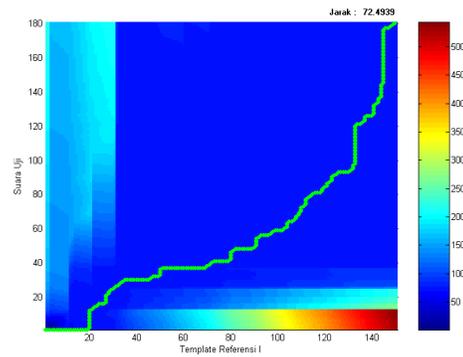
Proses pengenalan pola menggunakan DTW melakukan proses perhitungan jarak antara dua buah pola yang dibandingkan. Jarak yang dihitung adalah jarak antara nilai koefisien *cepstral* LPC dari *template* referensi dengan nilai koefisien *cepstral* LPC dari suara uji yang masuk. Kumpulan seluruh jarak-jarak ini digambarkan melalui sebuah matrik yang dinamakan matrik distorsi. Matrik distorsi diperlihatkan oleh Gambar 9-Gambar 11. Sumbu *x* pada matrik distorsi mewakili seluruh koefisien *cepstral* LPC dari *template* referensi, sedangkan sumbu *y* matrik distorsi mewakili seluruh koefisien *cepstral* LPC dari suara uji. Pada tampilan matrik distorsi ini juga diperlihatkan jalur terbaik yang digambarkan dengan warna hijau. Jalur terbaik menggambarkan proses pencocokan waktu DTW, yang digunakan untuk mengatasi adanya variasi yang terjadi pada pengucapan suatu vokal yang ingin dikenali. Jalur terbaik berisi jarak-jarak terkecil bila dibandingkan jarak lainnya. Jalur terbaik dibentuk dengan cara penelusuran kembali matrik distorsi mulai baris terakhir, kolom terakhir hingga baris pertama, kolom pertama. Penelusuran ini membandingkan jarak-jarak yang ada pada posisi diagonal bawah (*i-1,j-1*), posisi bawah (*i,j-1*), posisi kiri (*i-1,j*) dari posisi acuan (*i,j*). Posisi yang dipilih sebagai pembentuk jalur terbaik adalah posisi yang memiliki nilai jarak yang terkecil.

Pengamatan matrik-matrik distorsi hasil pengenalan pola DTW dilakukan dengan memperhatikan bentuk jalur terbaiknya. Jarak pencocokan keseluruhan paling kecil ditunjukkan dengan jalur terbaik suatu matrik distorsi yang berbentuk diagonal. Hal ini dapat dilihat pada tampilan matrik distorsi dari Gambar 9 dan Gambar 10. Gambar 9 memperlihatkan pencocokan antara suara uji dengan *template* referensi U, didapatkan jarak pencocokan

keseluruhan sebesar 0,59768. Jalur terbaik pada matrik distorsi berbentuk diagonal.



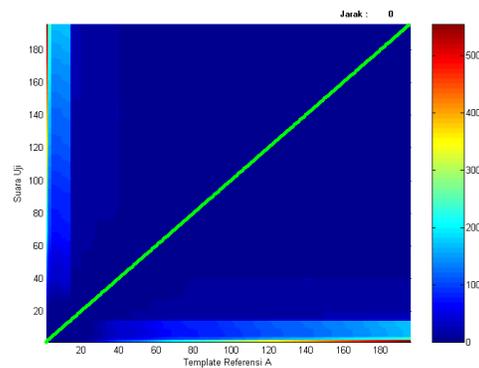
Gambar 9 Pencocokan dengan *template* referensi U.



Gambar 10 Pencocokan dengan *template* referensi I

Gambar 10 memperlihatkan pencocokan antara suara uji dengan *template* referensi I, didapatkan jarak pencocokan keseluruhan sebesar 72,4339. Jalur terbaik pada matrik distorsi tidak berbentuk diagonal. Suara uji yang masuk akan dicocokkan dengan *template* referensi /a/, /i/, /u/, /e/, /o/. Suara uji akan dikenali sebagai vokal dengan *template* referensi yang memiliki nilai jarak pencocokan keseluruhan terkecil. Nilai jarak pencocokan keseluruhan yang ditunjukkan Gambar 9 lebih kecil dibandingkan Gambar 10, maka suara uji akan dikenali sebagai vokal /u/.

Apabila 2 buah pola identik dicocokkan akan didapat matrik distorsi seperti Gambar 11.



Gambar 11 Pencocokan 2 pola yang identik.

Jalur terbaik yang terbentuk dari pencocokan 2 pola yang identik berupa garis diagonal sempurna, dan jarak pencocokan keseluruhannya nol.

V. Kesimpulan

Berdasarkan hasil pengujian dan analisis yang dilakukan pada tugas akhir ini didapatkan beberapa kesimpulan:

1. Koefisien-koefisien *cepstral* LPC dapat digunakan untuk mewakili sinyal ucapan pada proses pengenalan pola DTW.
2. Panjang *frame* yang digunakan dalam pembungkaiian suatu sinyal ucapan mempengaruhi jumlah *frame* yang dihasilkan dari analisis LPC.
3. Semakin panjang sinyal ucapan, semakin banyak jumlah *frame* yang dihasilkan pada analisis LPC.
4. Matrik distorsi menggambarkan jarak-jarak antara nilai koefisien-koefisien *cepstral* LPC *template* referensi dengan suara ucapan uji.
5. Pencocokan waktu yang digunakan oleh metode DTW untuk mengatasi adanya variasi yang terjadi pada pengucapan suatu vokal dapat digambarkan melalui jalur terbaik.
6. Jalur terbaik dari vokal yang dikenali berbentuk diagonal dan akan berupa garis diagonal sempurna apabila 2 pola yang dicocokkan identik.

VI. Saran

Visualisasi terhadap metode pengenalan ucapan berguna untuk memperjelas bagaimana teknik pengenalan tersebut bekerja. Berdasarkan hal ini visualisasi menggunakan metode pengenalan lainnya perlu dilakukan.

Referensi

- [1] Furui, S., *Digital Speech Processing, Synthesis, and Recognition*, Marcel Dekker, Inc., New York, 1989.
- [2] Gold, B., and N. Morgan, *Speech and Audio Signal Processing : Processing and Perception of Speech and Music*, John Wiley & Sons, Inc., New York, 1999.
- [3] Hansel, D., B Littlefield, J Edyanto, *Matlab Bahasa Komputasi Teknis*, Andi, Yogyakarta, 2001.
- [4] Nouza J., "Visualization Software for Speech Processing", SpeechLab, Dept. of Electronics and Signal Processing, Technical University of Liberec, Halkova 5, 461 17 Liberec, Czech Republic.1997.
- [5] Proakis, J., G., Manolakis D., G., *Pemrosesan Sinyal Digital: Prinsip, Algoritma, dan Aplikasi*, Edisi Bahasa Indonesia, PT Prehallindo, Jakarta, 1997.
- [6] Rabiner, L., Biing-Hwang Juang. *Fundamentals Of Speech Recognition*, New Jersey: Prentice Hall, 1993.
- [7] ---, *Cepstrum*, http://www.owl.net.rice.edu/~elec532/PROJECTS_98/speech/cepstrum/cepstrum.html, Juli 2006.

- [8] ---, *Dynamic Time Warping*, <http://www.cnel.ufl.edu/~kkale/dtw.html>, Januari 2006.

Biodata penulis



Syaiful Rachman, lahir di Jakarta, tanggal 2 Agustus 1983, menjalani pendidikan TK di TK Aisyiah Jkt tahun 1988-1989, pendidikan SD di SDN 04 pagi Jkt tahun 1989-1995, pendidikan SLTP di SMPN 139 Jkt tahun 1995-1998, pendidikan SMU di SMUN 12 Jkt tahun 1998-1999, SMUN 5 Smg tahun 1999-2001. Saat ini penulis sedang menyelesaikan pendidikan Strata Satu Teknik di Jurusan Teknik Elektro, Fakultas Teknik Universitas Diponegoro, Semarang, Indonesia.

Menyetujui,

Pembimbing I

Pembimbing II

Wahyudi, ST, MT
NIP. 132 086 662

Achmad Hidayatno, ST, MT
NIP. 132 137 933