

PENERAPAN METODE STATISTIK DAN AVERAGE ENERGY UNTUK MENGUJI TINGKAT KEMIRIPAN PADA IDENTIFIKASI SUARA

Erdoria Kristina⁽¹⁾
@yahoo.com

Jong Jek Siang⁽²⁾
j_j_siang@yahoo.com

Gunawan Santosa⁽³⁾
gunawan@ukdw.ac.id

Abstraksi

Sistem pengenalan suara saat ini sudah dapat mengenal suara dengan ketepatan yang cukup untuk perkataan tertentu. Sistem ini tentunya lebih cepat dan mudah dibandingkan dengan sistem pengenalan lainnya hanya saja masih dibutuhkan beberapa cara yang efektif agar suara manusia dapat dikenali dengan baik, sehingga dilakukan penelitian agar memiliki hasil yang sesuai.

Penelitian sistem identifikasi suara dilakukan uji sampel untuk mengidentifikasi suara orang yang berbeda. Penelitian pertama-tama dilakukan dengan pengambilan 10 sampel suara yang memiliki format wav lalu disimpan ke dalam database. Setelah itu kita dapat menginputkan suara yang ingin diidentifikasi, bisa menggunakan metode statistik maupun metode average energy untuk mendapatkan ciri. Setelah itu sampel suara tersebut dapat dibandingkan dengan menggunakan jarak euclidian yang menghasilkan nilai prosentase error antara suara-suara yang dibandingkan. Semakin kecil prosentase errornya berarti suara tersebut semakin mirip. Pada penelitian ini didapatkan hasil perbandingan antara average energy dan metode statistik. Pada metode statistik didapatkan hasil lebih baik daripada average energy.

Kata kunci: *speaker identification, statistik, average energy*

1. Pendahuluan

Sistem pengenalan suara saat ini sudah dapat mengenal suara dengan ketepatan yang cukup untuk perkataan tertentu. Sistem ini tentunya lebih cepat dan mudah, hanya saja masih dibutuhkan beberapa cara yang efektif agar suara manusia dapat dikenali dengan baik, sehingga memiliki hasil yang sesuai.

Untuk itu dilakukan penerapan metode statistik dan average energy agar dapat dilakukan pengujian identifikasi suara. Dalam penerapan ini dapat diketahui bagaimana

¹ Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Duta Wacana

² Teknik Informatika, Fakultas Teknologi Informasi Universitas Kristen Duta Wacana

³ Teknik Informatika, Fakultas Teknologi Informasi, Universitas Kristen Duta Wacana

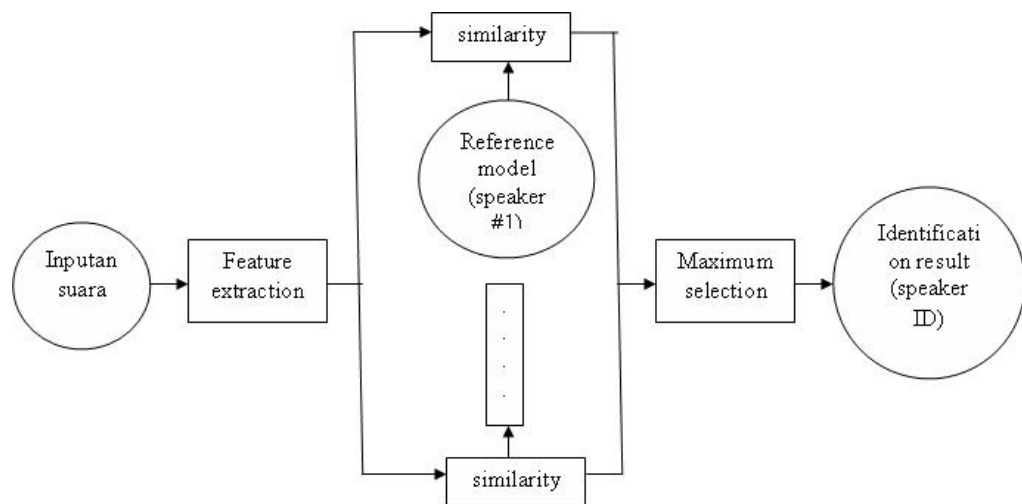
suara manusia diekstraksi dan bagaimana rumusan statistik maupun *average energy* dapat digunakan untuk menjadi bahan perbandingan dalam menentukan ciri suara.

2. Landasan Teori

2.1. Speaker Identification

Speaker identification adalah proses untuk mencari dan mendapatkan identitas dari seorang pengguna dengan membandingkan pola suara yang diinputkan dengan semua pola suara yang ada di dalam database. Proses ini melakukan perbandingan one-to-many (1:N).

Pembicara yang merupakan bagian dari pengenalan pembicara dapat dibagi ke dalam metode teks bebas dan teks tertentu. Pada sistem teks bebas, model pembicara meng-capture karakteristik ucapan seseorang melalui sinyal ucapan dengan mengabaikan apa yang diucapkannya, dalam artian kata-kata yang diucapkan sembarang (bebas). Sebaliknya pada sistem teks tertentu, pengenalan identitas pembicaranya didasarkan pada ucapan seseorang dengan kata-kata yang spesifik atau telah disepakati, seperti *password*, card numbers, kode PIN dan sebagainya.



Gambar 1. Bagan Pengenalan Suara *Speaker Identification*

2.2. Metode Statistik

Metode statistik merupakan prosedur-prosedur yang digunakan dalam penyajian data yang meliputi pengumpulan, pengorganisasian, peringkasan, dan penyajian data. Selain penyajian digunakan juga untuk penafsiran data yang meliputi pendugaan, pengujian dugaan dan penarikan kesimpulan.

Untuk menerapkan metode statistik dibutuhkan beberapa ukuran statistik dalam berbagai analisis data sampel. Ukuran Statistik digolongkan menjadi 2 kelompok, yaitu ukuran pemusatan dan ukuran penyebaran. Ukuran pemusatan merupakan ukuran lokasi

pusat atau ukuran yang mempunyai kecenderungan memusat, sedangkan ukuran penyebaran merupakan ukuran yang mengukur keragaman antarpengamatan. Beberapa ukuran pemusatan sederhana yang biasa digunakan yaitu mean, median, dan modus. Sedangkan beberapa ukuran penyebaran yaitu standar deviasi dan simpangan.

Dibawah ini merupakan beberapa ukuran yang dipakai dalam sistem, yaitu:

1. Rata-rata (*Mean*)

Rata-rata merupakan sebuah ukuran yang khas yang mewakili suatu himpunan data. Rata-rata dapat dihitung menggunakan rumus sebagai berikut:

$$\text{Mean } (\bar{x}) = \frac{\sum X_i}{n} \quad [2.1]$$

Dengan, X_i = nilai data ke i

n = banyak data

2. Simpangan Rata-rata (SR)

Simpangan rata-rata merupakan jumlah nilai mutlak dari selisih semua nilai dengan nilai rata-rata dibagi banyaknya data

$$\text{SR} = \frac{\sum (X_i - \bar{X})}{n} \quad [2.2]$$

Dengan, X_i = nilai data ke $-i$

\bar{X} = nilai rata-rata

n = banyaknya data

3. Standar Deviasi dan Variansi

Salah satu teknik statistik yg digunakan untuk menjelaskan homogenitas kelompok. Varians merupakan jumlah kuadrat semua deviasi nilai-nilai individual thd rata-rata kelompok. Sedangkan akar dari varians disebut dengan standar deviasi atau simpangan baku.

$$\text{Variansi} = \frac{\sum (X_i - \bar{X})^2}{n} \quad \text{Standar deviasi} = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}} \quad [2.3]$$

Dengan, X_i = nilai data ke $-i$

\bar{X} = nilai rata-rata

n = banyaknya data

2.3. Average Energy

File audio diwakili oleh sejumlah fitur yang diekstrak dari file tersebut dan disimpan dalam database sistem. Fitur ini digunakan untuk pencocokan antara file audio

query dan file audio dalam database baik secara individual maupun simultan. Fitur ini akan diekstraksi dengan menggunakan energi rata-rata (*Average Energy*) untuk diperoleh ciri *Average Energy*. Ciri *Average Energy* adalah ciri yang digunakan untuk menyatakan nyaring atau tidaknya sinyal audio. Ciri ini diperoleh dengan rumus:

$$E = \frac{\sum_{n=0}^{N-1} x(n)^2}{N} \quad [2.4]$$

E adalah energi rata-rata dari potongan audio, $x(n)$ menyatakan nilai dari *sample* ke n sedangkan N merupakan jumlah total *sample* pada potongan audio.

2.4 Jarak Euclidian

Euclidian adalah pengukuran jarak garis lurus (*straight line*) antara titik X (X_1, X_2, \dots, X_n) dan titik Y (Y_1, Y_2, \dots, Y_n). Jarak Euclidian dapat dianggap sebagai jarak yang paling pendek antar dua poin, maka dari itu dalam tugas akhir ini digunakan fungsi jarak Euclidian yang akan menjadi hasil akhir berupa data minimum. Secara matematis Jarak Euclidian dapat dituliskan di dalam persamaan berikut :

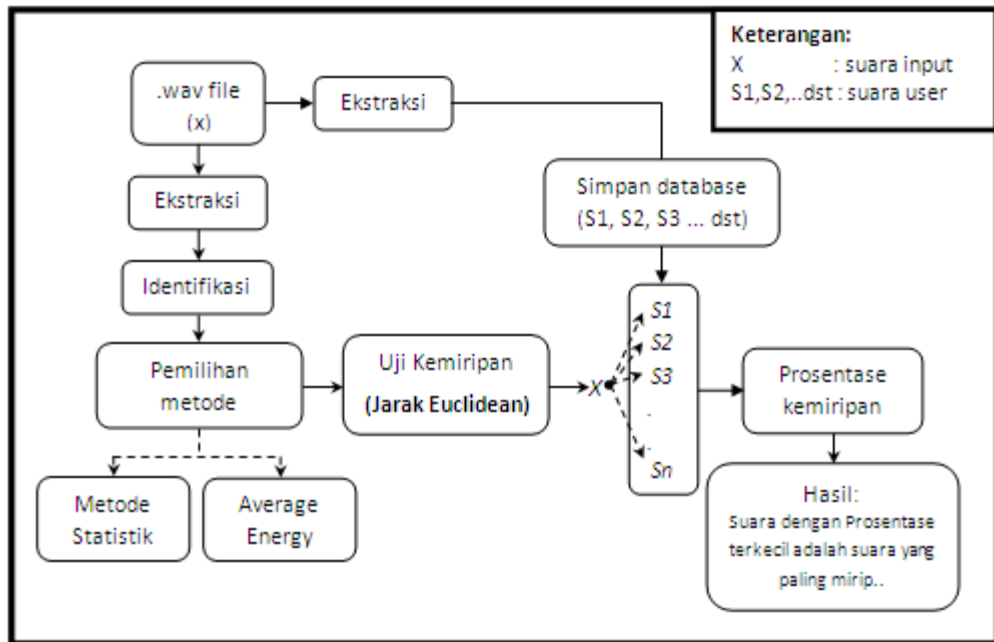
$$d(i,j) = \sqrt{|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2} \quad [2.5]$$

Keterangan :

$d(i,j)$	=	nilai jarak
x_i	=	nilai nilai pada fitur 1
x_j	=	nilai nilai pada fitur 2

3. Analisis dan Perancangan Sistem

Sistem ini dimulai dengan membuka file audio dengan format wav lalu dilakukan ekstraksi dan menyimpan dokumen file hasil ekstraksi tersebut ke dalam database (S1, S2, S3, ... dst) atau menggunakan ekstraksi untuk melakukan identifikasi suara dengan dua macam pilihan metode, yaitu dengan memilih metode statistik atau average energy. Setelah itu dilakukan uji kemiripan dengan jarak euclidian, yaitu dengan membandingkan setiap suara dalam database dengan suara input yang ditunjukkan dengan prosentase error. Prosentase error terkecil merupakan suara termirip.

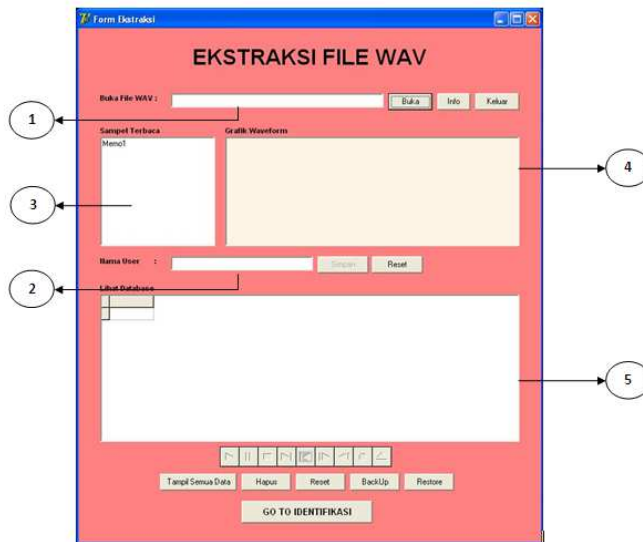


4. Hasil Implementasi

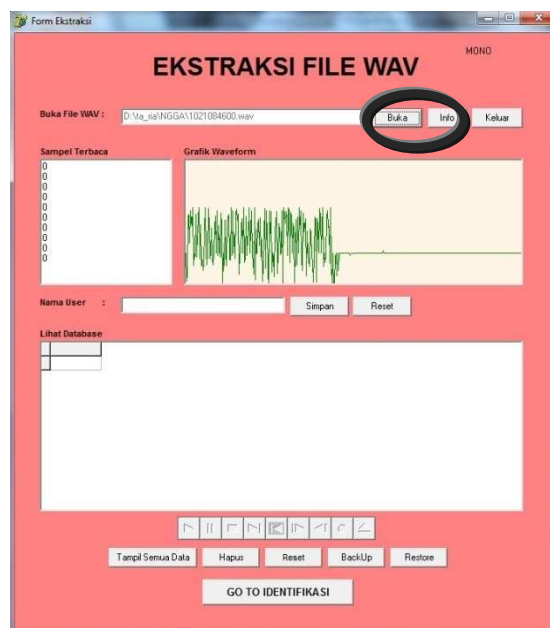
4.1. Form Ekstraksi

Form Ekstraksi merupakan form yang berfungsi untuk mengubah sinyal analog wave menjadi data sampel yang nantinya dapat disimpan ke dalam database access . Dalam form ini akan ditampilkan data sampel, grafik waveform dan media untuk menyimpan suara dengan memberikan nama usernya. Selain itu, dapat digunakan untuk menampilkan seluruh data yang tersimpan dalam database, membuat *BackUp* dan *Restore*.

Hasil dari program ini berdasarkan rancangan form ekstraksi adalah sebagai berikut:

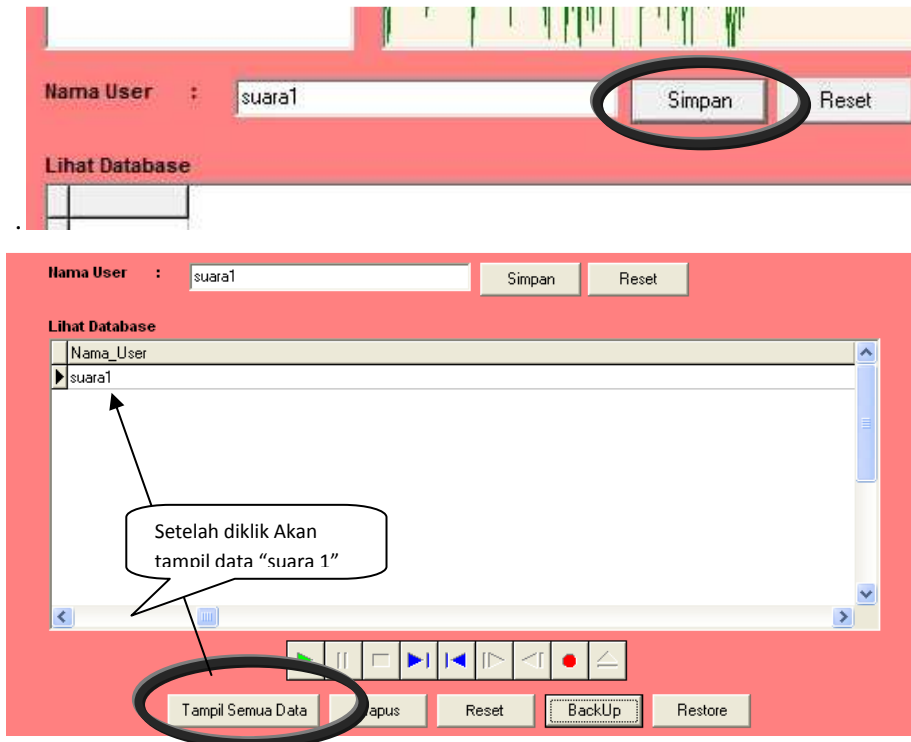


Pada form ekstraksi terdapat dua buah *textbox* (no 1 dan 2), dua buah *memo* (no 3 dan 4), dan sebuah *combobox* (no 5). Pada *textbox* no 1 akan membuka masukan berupa sebuah file berekstensi wav yang nantinya akan diekstraksi menjadi sampel-sampel yang akan ditampilkan pada *memo* no 3 dan menampilkan gambar grafik waveform pada *memo* no 4 dengan isi tampilan form seperti gambar di bawah ini:



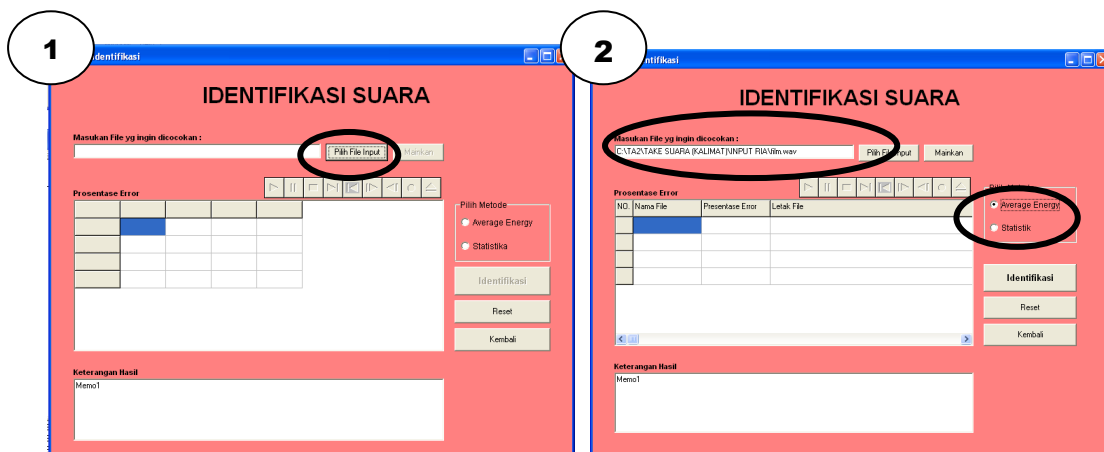
Selanjutnya jika file tersebut sudah diekstraksi lalu dapat disimpan ke dalam database “dbSuara.mdb” dengan memberikan nama user pada *textbox* no 2 dan klik *button Simpan*. Lalu akan ada keterangan yang muncul melalui *messagebox* bahwa dokumen telah

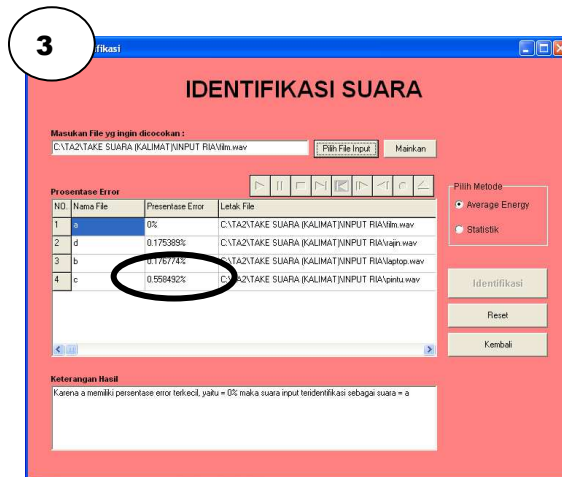
tersimpan, yaitu dokumen simpangan, rata-rata, standar deviasi dan average energy. Seperti yang terlihat pada tampilan gambar berikut:



4.2. Form Identifikasi Suara

Form identifikasi suara berfungsi untuk mengidentifikasi suara antara suara yang diinputkan dengan suara-suara yang ada di dalam database sehingga suara input dapat dikenali berdasarkan tingkat kemiripannya. Berikut ini merupakan hasil dari program berdasarkan rancangan form ekstraksi beserta penggunaannya:





Penjelasan:

- 1 Tampilan awal form identifikasi suara, penggunaan pertama diawali dengan memilih file input dengan menekan *button* 'Pilih File Input'.
- 2 File input ditampilkan pada *textbox*. Setelah itu, pilih metode yang ingin digunakan, dengan memilih salah satu *radiobutton* 'Statistik' atau 'Average Energy'.
- 3 Lalu tekan *button* Identifikasi hingga hasil akhir keluar berupa urutan suara-suara yang teridentifikasi oleh suara yang diinputkan mulai dari yang termirip hingga tidak terlalu mirip berdasarkan prosentase errornya mulai dari yang terkecil hingga besar.

4.3. Hasil Pengujian dengan Membandingkan Kemiripan Suara dengan User Berbeda

No.	Nama User	Prosentase Error		
		Tempo Lambat	Tempo Cepat	Kekerasan
1	Ria	0,86%	1,49%	0,72%
2	Adhe	1,93%	1,13%	0,79%
3	Christine	0,04%	1,75%	0,06%
4	Eva	2,04%	1,04%	2,31%
5	Mitha	1,53%	1,33%	2,38%
6	Corry	2,39%	3,03%	1,15%
7	Eris	0,98%	2,29%	0,64%
8	Femmy	2,58%	2,6%	2,21%
9	Ully	2,14%	1,58%	0,07%
10	Ana	2,51%	4,55%	4,73%

Dari hasil prosentase menurut tempo suara cepat dan tempo suara lambat, didapatkan suara RIA input teridentifikasi sebagai suara RIA yang ada pada database, dengan orang yang sama dengan hasil perolehan prosentase error terkecil untuk tempo lambat 1,87% dan tempo cepat 3,32%. Sedangkan suara RIA dalam database menurut kekerasan suaranya tidak teridentifikasi sebagai suara Adhe dengan hasil prosentase error terkecil 6,27% dan suara RIA menempati posisi kedua suara dengan prosentase terkecil 8,7%.

No.	Nama User	Prosentase Error		
		Tempo Lambat	Tempo Cepat	Kekerasan
1	Ria	0,86%	1,49%	0,72%
2	Adhe	1,93%	1,13%	0,79%
3	Christine	0,04%	1,75%	0,06%
4	Eva	2,04%	1,04%	2,31%
5	Mitha	1,53%	1,33%	2,38%
6	Corry	2,39%	3,03%	1,15%
7	Eris	0,98%	2,29%	0,64%
8	Femmy	2,58%	2,6%	2,21%
9	Ully	2,14%	1,58%	0,07%
10	Ana	2,51%	4,55%	4,73%

Dari hasil prosentase menurut tempo suara cepat, tempo suara lambat, dan kekerasan suara didapatkan bahwa suara RIA input tidak teridentifikasi sebagai suara RIA yang ada pada database, melainkan suara user lain yang ada dalam database dengan hasil perolehan prosentase error terkecil untuk tempo lambat 0,04% untuk suara Christine, tempo suara cepat 1,04% untuk suara Eva dan kekerasan suara 0,06% untuk suara Christine. Jadi tidak dikenali sebagai suara Ria.

Dari kedua metode ini dapat disimpulkan bahwa menggunakan metode statistik bisa mengidentifikasi suara seseorang lebih baik daripada average energy. Semakin banyak rumusan statistik lain untuk menghitung sample, lebih akurat hasil yang diperoleh untuk mengidentifikasi suara. Untuk average energy bisa diperoleh hasil lebih baik jika menggunakan pemotongan suara dengan durasi.

5. Kesimpulan

Berdasarkan percobaan yang telah dilakukan dan hasil analisa dari bab-bab sebelumnya maka dapat disimpulkan bahwa :

- 1) Metode Statistik dan average energy merupakan metode yang cukup baik digunakan untuk mengenali fitur suara seseorang.
- 2) Pada percobaan identifikasi suara, metode statistik memberikan hasil prosentase lebih baik daripada *average energi*.
- 3) Setelah melakukan pengujian dengan satu suara dengan cara pengambilan suara yang berbeda-beda maka dapat dilihat bahwa identifikasi suara dapat dipengaruhi oleh volume suara, tempo suara, dan tinggi rendahnya suara.

Daftar Pustaka

- , http://id.wikipedia.org/wiki/Pengenalan_ucapan (diakses 20 November 2010).
- Arkhiansyah, Yuni. (2007) "Implementasi Ciri Energi Rata-rata, Cacah Perpotongan pada Titik, dan Rasio Periode Keheningan pada Pencarian File Audio" diakses 19 Oktober 2010
<http://e-riset.darmajaya.ac.id/jurnal-ik/wp-content/uploads/2009/10/7-yuni.pdf>
- Dajan, Anto (1986). Pengantar Metode Statistik, Jilid I. Jakarta: LP3ES.
- Kurniawan, Harry & Taufiq Hidayat. (2010) "Perancangan Program Pengenalan Wajah Menggunakan Fungsi Jarak Metode Euclidean Pada Matlab, diakses 29 Desember 2010
"http://journal.uui.ac.id/index.php/Snati/article/viewFile/885/839
- Leung. "Audio Indexing and Retrieval". (2004) , diakses 15 Agustus 2010
<http://www.cs.cityu.edu.hk/~howard/Teaching/IT6902/AudioIndexingAndRetrieval.pdf>.
- Mustofa, Ali (2007) "Sistem Pengenalan Penutur dengan Metode *Mel-frequency Wrapping*", diakses 30 November 2009, <http://puslit2.petra.ac.id/ejournal/index.php/elk/article/viewFile/16704/16696>
- Soejoed, Zanzawi (1986). Metode Statistika I. Jakarta: Karunika.
- Spiegel, Murray R (1994). Statistika, Edisi Kedua; alih bahasa, I Nyoman Susila dan Ellen Gunawan. Jakarta: Erlangga.
- Supranto, J (2007). Statistik untuk Pemimpin Berwawasan Global, Edisi Kedua. Jakarta: Salemba Empat.
- Suyanto, M (2003). Multimedia, alat untuk keunggulan bersaing. Yogyakarta: Andi.
- Wilson, Scott (2003) "WAVE PCM soundfile format", diakses 19 Oktober 2010
<https://ccrma.stanford.edu/courses/422/projects/WaveFormat>