

## **SPEAKER IDENTIFICATION MENGGUNAKAN METODE MEL FREQUENCY CEPSTRUM COEFFICIENTS (MFCC) DAN DYNAMIC TIME WARPING (DTW)**

**Speaker Identification Using Mel Frequency Cepstrum Coefficients  
(MFCC) And Dynamic Time Warping (DTW) Method**

**Sri Agustini Yumiati<sup>1</sup>, S. Mariyanto Al Sasongko<sup>2</sup>, L. A Syamsul Irfan Akbar<sup>3</sup>**

### **ABSTRAK**

Suara merupakan alat komunikasi yang paling dasar yang digunakan oleh manusia. Pengenal penutur merupakan proses mengenali identitas dari seorang penutur dengan membandingkan fitur-fitur suara yang di-input-kan dengan semua fitur-fitur dari setiap penutur yang ada dalam database. Dalam identifikasi penutur terdiri dari dua tahap yaitu tahap pengekstrakan ciri dan tahap pengenalan pola. Untuk tahap pengekstrakan ciri digunakan metode mel frequency cepstrum coefficient (MFCC). Pada tahap pengenalan pola digunakan metode dynamic time warping (DTW) dimana membandingkan antara data uji dengan data referensi pada database dengan cara menghitung warping path. Hasil yang didapat dari penelitian membuktikan bahwa penggunaan nilai threshold akan menentukan keberhasilan sistem pengenalan penutur. Semakin rendah nilai threshold akan menyebabkan penurunan pada persentase tingkat pengenalannya. Ini terbukti dari pengenalan penutur dengan volume suara berbeda persentase terbesar terdapat pada nilai threshold 5,4 sebesar 94% dan nilai threshold terendah yaitu 0,9 sebesar 41%. Sedangkan untuk pengujian penutur diluar database hasil yang didapat adalah persentase tingkat pengenalan sebesar 79%. Ini dipengaruhi oleh besarnya nilai threshold. Nilai threshold antara 3,6 sampai 4,5 memiliki tingkat terima dan ditolak paling baik.

**Kata Kunci** : MFCC, DTW, Pengenalan Penutur, Nilai threshold

### **ABSTRACT**

The sound is the most basic means of communication used by humans. Identification of speakers is the process of recognizing the identity of a speaker by comparing the features-the inputted voice with all the features of each speaker in the database. In the speaker identification consists of two phases: feature extraction phase characteristics and pattern recognition. For the characteristic extraction phase used mel frequency cepstrum method coefficient (MFCC). At this stage of pattern recognition used method of dynamic time warping (DTW) which compares the test data with the reference data in the database by calculating warping path. The result from the research show that using the threshold value will determine the success of the speaker recognition system. The lower the threshold value will cause a decrease in the percentage level of recognition. This is evident from the introduction of different speakers with the largest percentage of the volume contained in the 5.4 threshold value of 94% and the lowest was 0.9 threshold value of 41%. As for testing the speakers outside the database results obtained recognition rate is the percentage of 79%. It is influenced by the magnitude of the threshold value. Threshold value between 3.6 to 4.5 have high levels of the most well received and rejected.

**Keywords** : MFCC, DTW, Speaker Identification, Threshold Value

### **PENDAHULUAN**

Suara merupakan alat komunikasi yang paling dasar yang digunakan oleh manusia. Dewasa ini ketika teknologi semakin berkembang pesat, manusia tidak hanya berinteraksi dengan manusia saja, tetapi seringkali juga harus berinteraksi dengan perangkat-perangkat teknologi seperti komputer. Namun komputer tidak dapat

mengerti bahasa yang kita ucapkan, sehingga diperlukan sebuah sistem pengenalan penutur (*speaker identification*).

Pengenal penutur merupakan proses mengenali identitas dari seorang penutur dengan membandingkan fitur-fitur suara yang di-input-kan dengan semua fitur-fitur dari setiap penutur yang ada dalam database.

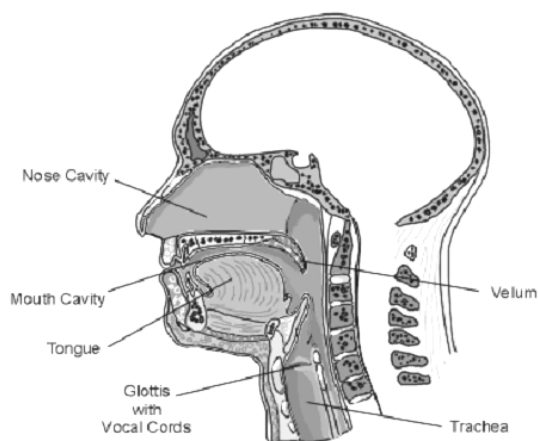
<sup>1</sup> Jurusan Teknik Elektro Fakultas Teknik Universitas Mataram, Nusa Tenggara Barat, Indonesia  
Email : [agustini.yumiati@gmail.com](mailto:agustini.yumiati@gmail.com)<sup>1</sup>, [mariyantosas@gmail.com](mailto:mariyantosas@gmail.com)<sup>2</sup>, [laluirfan@gmail.com](mailto:laluirfan@gmail.com)<sup>3</sup>

Dalam tugas akhir ini akan dirancang simulasi identifikasi penutur. Sistem akan mengidentifikasi (mengenal) suara penutur dari 12 orang penutur yang terdiri dari 6 suara pria dan 6 suara wanita dengan umur rata-rata 22-24 tahun dengan menggunakan software matrix laboratory (Matlab).

**Pengenalan Suara Pada Manusia.** Pada sistem pengenalan suara oleh manusia terdapat tiga organ penting yang saling berhubungan yaitu telinga yang berperan sebagai transduser dengan menerima sinyal masukan suara dan mengubahnya menjadi sinyal syaraf, jaringan syaraf yang berfungsi mentransmisikan sinyal ke otak, dan otak yang akan mengklasifikasi dan mengidentifikasi informasi yang terkandung dalam sinyal masukan.

**Proses Produksi Suara.** Proses produksi suara pada manusia dapat dibagi menjadi tiga buah proses fisiologis, yaitu: pembentukan aliran udara dari paru-paru, perubahan aliran udara dari paru-paru menjadi suara, baik *voiced* maupun *unvoiced* yang dikenal dengan istilah *phonation*, dan artikulasi yaitu proses modulasi/ pengaturan suara menjadi bunyi yang spesifik.

Organ tubuh yang terlibat pada proses produksi suara adalah : paru-paru, tenggorokan (*trachea*), laring (*larynx*), faring (*pharynx*), pita suara (*vocal cord*), rongga mulut (*oral cavity*), rongga hidung (*nasal cavity*), lidah (*tongue*), dan bibir (*lips*), seperti dapat dilihat pada gambar 1.



Gambar 1 Organ wicara manusia (Munawar, 2010)

Faktor Penyebab Kesalahan Proses Verifikasi dan Identifikasi adalah sebagai berikut

1. Kesalahan dalam pengucapan (*misspoken*) dan pembacaan (*misread*) frasa.
2. Keadaan emosional yang ekstrim (misalnya stres).
3. Pergantian penempatan mikropon
4. Kekurangan atau ketidak-konsistenan akustik dari ruangan (misalnya *multipath* dan *noise*).
5. *Channel mismatch* (misalnya penggunaan mikropon yang berbeda *channel* dalam perekaman dan verifikasi).
6. Sakit (misalnya *flu* yang dapat merubah *vocal tract*).
7. Aging (model *vocal tract* dapat berubah berdasarkan usia).

**Klasifikasi Berdasarkan Sinyal Eksitasi.**

Berdasarkan sinyal eksitasi yang dihasilkan pada proses produksi suara, sinyal suara ucapan dapat dibagi menjadi tiga bagian yaitu *silence*, *unvoiced*, dan *voiced* :

1. Sinyal *silence* : sinyal pada saat tidak terjadi proses produksi suara ucapan, dan sinyal yang diterima oleh pendengar dianggap sebagai bising latar belakang.
2. Sinyal *unvoiced* : terjadi pada saat pita suara tidak bergetar, dimana sinyal eksitasi berupa sinyal random.
3. Sinyal *voiced* : terjadi jika pita suara bergetar, yaitu pada saat sinyal eksitasi berupa sinyal pulsa kuasi-periodik.

**Mel Frequency Cepstrum Coefficients (MFCC).**

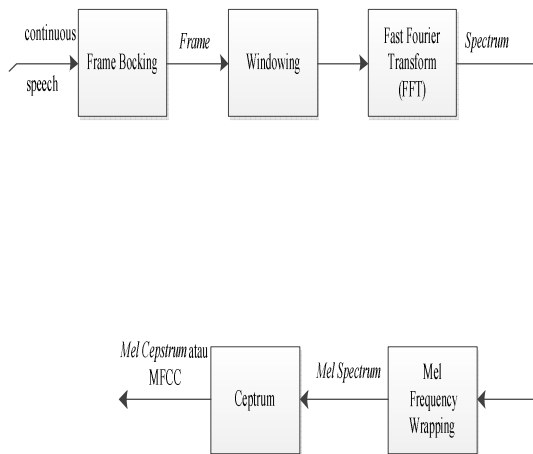
*Feature extraction* (ekstraksi ciri) merupakan proses untuk menentukan satu nilai atau vektor yang dapat dipergunakan sebagai penciri obyek atau individu. Dalam pemrosesan suara, ciri yang biasa dipergunakan adalah *mel frequency cepstrum coefficients* (MFCC) yang menghitung koefisien *cepstrum* dengan mempertimbangkan persepsi sistem pendengaran manusia terhadap frekuensi suara. MFCC didasarkan pada variasi yang telah diketahui dari jangkauan kritis telinga manusia dengan frekuensi. MFCC memiliki 2 jenis filter dimana bersifat linear pada frekuensi dibawah 1000 Hz dan bersifat logaritmik pada frekuensi diatas 1000 Hz. Beberapa keunggulan dari metode ini adalah :

- a. Mampu untuk menangkap karakteristik suara yang sangat penting bagi pengenalan suara, atau dengan kata lain dapat menangkap informasi-informasi

penting yang terkandung dalam *signal* suara.

- b. Menghasilkan data seminimal mungkin, tanpa menghilangkan informasi-informasi penting yang dikandungnya.
- c. Mereplikasi organ pendengaran manusia dalam melakukan persepsi terhadap signal suara. (Aditya,2012)

Blok diagram pemrosesan *mel frequency cepstrum coefficients* (MFCC) dapat dilihat pada gambar 2 dibawah ini :



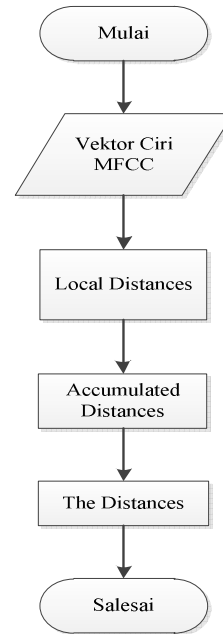
Gambar 2 Blok Diagram *Mel Frequency Cepstrum Coefficients*

**Dynamic Time Warping (DTW).** *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.

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*). *Local distance* antara sebuah vektor ciri dari suatu *template* referensi  $x = (x_1, x_2, x_3, \dots, x_n)$  dengan sebuah vektor ciri dari suara uji  $y = (y_1, y_2, y_3, \dots, y_n)$  yang masuk dirumuskan dengan menggunakan *manhattan distance*, dimana jarak dihitung sebagai selisih nilai  $x_i$  dan  $y_j$  yang diabsolutkan yang dirumuskan dengan persamaan berikut :

$$c(i,j) = |x_i - y_j| \quad ; \quad i = 1 : n \quad \text{dan} \quad j = 1 : m$$

Berikut ini merupakan metode *dynamic time warping* dapat digambarkan dengan *flowchart* dibawah ini :



Gambar 3 *Flowchart Dynamic Time Warping (DTW)*

Keterangan *flowchart* *dynamic time wrapping* (DTW) diatas adalah :

- 1.Data mel cepstrum atau vektor ciri MFCC merupakan input-an
- 2.Local distances adalah jumlah jarak manhattan distance. Setelah didapat local distance (jarak lokal) maka bisa ditentukan cost matrix yang memiliki dimensi cost (nxm).
- 3.Accumulated distances adalah jarak perhitungan antara jarak lokal satu dengan jarak lokal sampingnya. Setelah accumulated cost matrix terbentuk dicarilah warping path. Proses warping path adalah jalur atau path yang melalui matriks yang berisi jarak minimal dari elemen D(i,j) hingga elemen paling akhir D(n,m) yang terdiri dari elemen-elemen D(i,j) itu sendiri.
- 4.The distances adalah jarak akhir yang menunjukkan warping path jarak antara suara uji yang di-masukan dengan suara referensi yang telah tersimpan. Jarak warping path didapat dari accumulated cost matrix minimum pada kolom terakhir. (Rianto, 2011).

## METODOLOGI PENELITIAN

**Metode Penelitian.** Metode yang digunakan dalam penelitian ini adalah pertama pengambilan sampel penutur yang terdiri dari suara laki – laki dan perempuan yang berumur rata-rata 22 - 24 tahun, masing-masing penutur mengucapkan satu pola kata yang sama yaitu kata “telkom”. Kedua adalah memroses MFCC tujuannya untuk mendapatkan ekstraksi ciri dari masing-masing penutur. Ketiga adalah proses pelatihan, pada proses ini pola kata yang sudah ditentukan dimasukkan secara urut mulai penutur 1, penutur 2 seterusnya sampai penutur terakhir. Keempat adalah menguji penutur dengan MFCC dan DTW untuk mengenali penutur. Kelima yaitu menganalisa dan pengambilan kesimpulan.

**Alat Penelitian.** Alat penelitian yang digunakan yaitu :

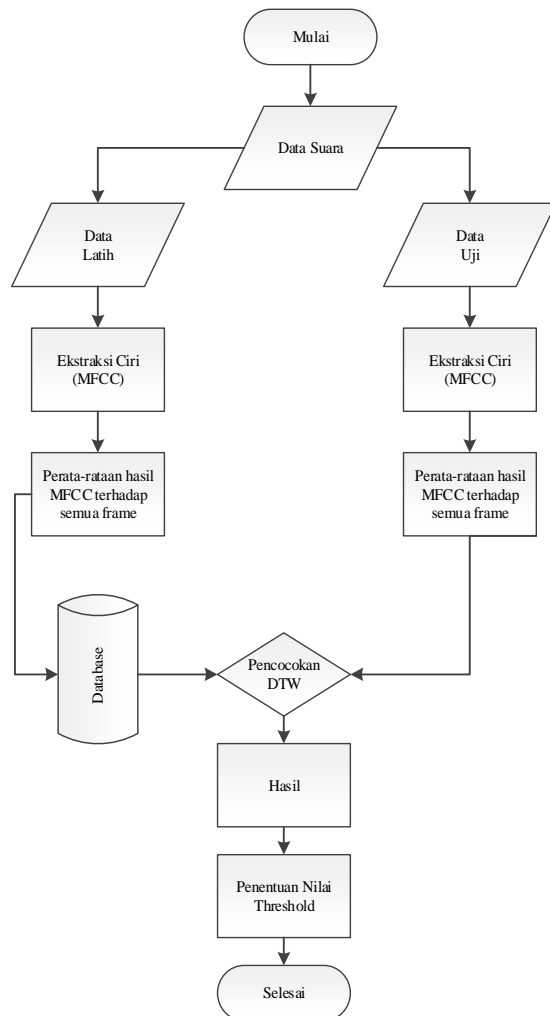
1. Perangkat lunak
  - Sistem operasi *Microsoft Windows 7 Ultimate*
  - Aplikasi *Matlab*
  - Aplikasi *GoldWave*
2. Perangkat keras
  - Laptop dengan spesifikasi *prosesor AMD, RAM 2 GB, hard disk 500 GB*
  - *Microphone*

**Perancangan Sistem.** Dalam proses perancangan suatu sistem dilakukan penelitian dan penganalisaan tentang sistem yang akan dibangun, berikut ini adalah gambar diagram alur penelitian pengenalan penutur (*speaker identification*).

Penjelasan dari gambar 4 diagram alur penelitian:

### 1. Data Suara

Data yang berupa sinyal suara penutur yang direkam dengan menggunakan mikropon yang terhubung dengan laptop. Perekaman suara dilakukan dengan bantuan program *GoldWave*, dengan frekuensi *sampling 8000Hz, 16 bit, mono channel*. Kemudian suara tersebut disimpan dalam bentuk file yang bereksitensi “.wav”. Suara yang akan diproses pada penelitian ini dan dijadikan referensi dalam pembentukan sistem identifikasi penutur diambil dari 12 orang sumber penutur yang berbeda terdiri dari 6 orang laki-laki dan 6 orang perempuan.



Gambar 4 Diagram alur penelitian identifikasi penutur

### 2. Data Latih dan Data Uji

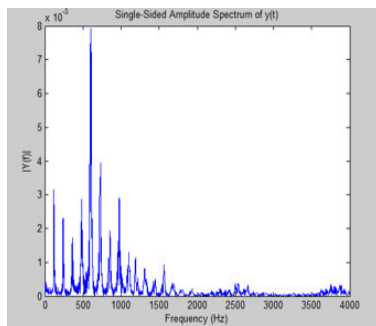
Tahap pelatihan digunakan untuk membentuk pola yang mewakili karakter suara untuk setiap penutur yang akan diidentifikasi. Tahap pengujian digunakan untuk melihat kinerja dari sistem yang telah dirancang. Pada penelitian ini akan ada 3 (tiga) jenis pengujian yang akan dilakukan yaitu:

- Pengujian data latih.
- Pengujian setiap penutur dengan volume suara berbeda.
- Pengujian penutur diluar penutur dalam *database*.

3. Ekstraksi Ciri (MFCC)

Ekstraksi ciri sinyal suara pada penelitian ini menggunakan MFCC. Parameter dari MFCC yaitu :

- a. *Input*, yaitu masukan suara yang berasal dari tiap pembicara.
- b. *Sampling rate*, yaitu banyaknya nilai yang diambil dalam satu detik. Dalam penelitian ini digunakan *sampling rate* sebesar 8000 Hz.
- c. *Time frame*, yaitu waktu yang diinginkan untuk satu *frame* (dalam milidetik). *Time frame* yang digunakan adalah 30 ms.
- d. *Lap*, yaitu *overlapping* yang terdapat dari N/2.
- e. *Cepstrum coefficient*, yaitu jumlah *cepstrum* yang diinginkan sebagai *output*. *Cepstrum coefficient* yang digunakan sebanyak 13. Nilai koefisien 13 didapat dari nilai *spectrum* pada gambar 5 yang mencakup nilai frekuensi dari data suara. (Abriyono, 2012).



Gambar 5 Hasil rekaman suara di FFT

- 4. Hasil ekstraksi ciri menggunakan MFCC memiliki hasil berupa matriks ciri  $n \times k$ ,  $n$  adalah jumlah frame dan  $k$  adalah koefisien. Untuk menghasilkan matriks yang berukuran sama di setiap suara, yaitu matriks berukuran  $l \times k$ , dilakukan perataan koefisien pada setiap baris.
- 5. Database digunakan untuk menyimpan ciri dari data pelatihan.
- 6. Pencocokan DTW. Tahapan berikutnya yang dilakukan setelah didapatkan ciri melalui ekstraksi ciri adalah identifikasi. Metode yang digunakan pada tahapan ini adalah dynamic time warping (DTW). Pada pengenalan ini sistem akan

membandingkan antara suara uji (data uji) dengan suara referensi (data latih) dengan cara menghitung jarak yang paling minimum.

- 7. Hasil. Pada tahap ini, dikatakan cocok apabila antara data uji dengan data latih (database) dikenali dengan melihat nilai jaraknya terkecil (minimum) dari semua database. Dan apabila data uji dengan data latih tidak cocok (tidak dikenali) menunjukkan nilai jarak yang terbesar (maksimum) dari semua database.
- 8. Penentuan nilai threshold. Pada pencocokan DTW proses yang digunakan untuk mencari jarak adalah dengan cara membandingkan satu penutur uji dengan semua penutur yang berada di dalam database, sehingga apabila terdapat 12 penutur di dalam database maka proses ini akan diulang sebanyak 12 kali. Untuk mengatasi kemungkinan penutur yang tidak terdapat di dalam database tersebut dikenali, maka jarak terpendek itu harus memenuhi nilai threshold yang diberikan dan apabila nilai threshold tersebut tidak terpenuhi maka penutur tersebut akan ditolak.

Untuk menentukan nilai threshold terlebih dahulu harus dilakukan pengujian untuk semua data uji. Setelah didapatkan semua jarak dari data uji maka jarak tersebut akan disorting dari jarak yang bernilai terkecil hingga jarak yang bernilai terbesar.

**Perhitungan Tingkat Pengenalan.** Tingkat pengenalan sistem akan dihitung untuk mengevaluasi hasil penelitian. Untuk setiap data yang diuji, akan dilihat apakah data tersebut teridentifikasi dengan benar atau tidak. Persentase tingkat pengenalan dihitung menggunakan rumus sebagai berikut:

$$\text{Tingkat pengenalan} = \frac{\sum \text{suara benar}}{\sum \text{suara uji}} \times 100\%$$

**HASIL DAN PEMBAHASAN**

**Pengolahan Data Suara.** Perekaman dilakukan terhadap penutur dengan durasi waktu rata-rata 1 (satu) detik pada *sampling rate* ( $F_s$ ) sebesar 8000 Hz, mono *channels*, dan resolution 16-bit. Dalam tugas akhir ini digunakan 20 penutur yang akan dibagi menjadi 12 orang sebagai penutur didalam *database* dan 8 orang sebagai penutur diluar *database*. Penutur didalam *database* terdiri dari 6 orang laki-laki dan 6 orang perempuan yang mengucapkan kata “telkom” dengan perulangan sebanyak 13 kali dimana 3 data

sebagai data yang dilatih ke dalam program dan 10 data sebagai data yang akan diujikan. Penutur diluar *database* terdiri dari 4 orang laki-laki dan 4 orang perempuan mengucapkan kata “telkom” dengan perulangan sebanyak 10 kali.

Untuk membuktikan *performance* dari sistem identifikasi penutur ini, ada 3 (tiga) jenis pengujian yaitu :

1. Pengujian data latih.
2. Pengujian setiap penutur dengan volume suara berbeda.
3. Pengujian penutur diluar penutur dalam *database*.

**Persentase Tingkat Pengenalan.**

Berdasarkan hasil semua pengujian yang telah diperoleh, maka dapat dihitung persentase tingkat pengenalan dari pengujian data latih, pengujian untuk volume suara normal, volume suara rendah dan volume suara tinggi pada setiap penutur dengan *threshold* dan tingkat penolakan terhadap penutur yang diluar penutur dalam *database* menggunakan persamaan :

$$\text{tingkat pengenalan} = \frac{\sum \text{suara benar}}{\sum \text{suara uji}} \times 100\%$$

Hasil persentase tingkat pengenalan dan tingkat penolakan penutur dapat dilihat pada Tabel di bawah ini :

- a. Tingkat pengenalan data latih

Tabel 1 Hasil tingkat pengenalan data latih

Nama penutur	Tingkat engenalan%)
Dika	100%
Dwi	100%
Hari	100%
Man	100%
Tunas	100%
Umam	100%
Baiq	100%
Atiq	100%
Dian	100%
Muna	100%
Rina	100%
Ully	100%
<b>Rata-rata</b>	<b>100%</b>

Berdasarkan hasil Tabel 1 dapat disimpulkan bahwa pengujian yang dilakukan dengan data yang sama

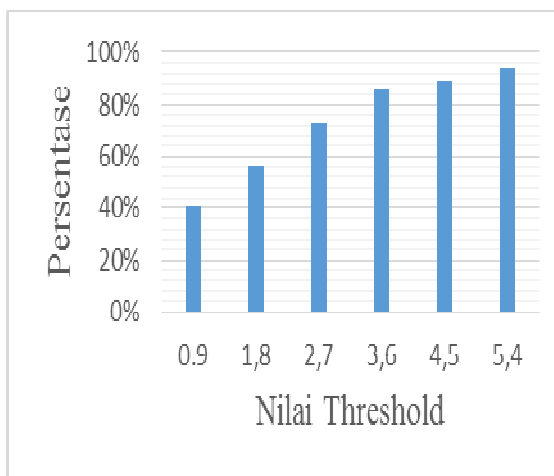
dengan data dalam *database* akan menghasilkan persentase pengenalan sebesar 100%.

- b. Tingkat pengenalan setiap penutur dengan volume suara berbeda
  - Tingkat pengenalan untuk volume suara normal

Tabel 2 Hasil persentase pengenalan untuk volume suara normal

Nama Penutur	Tingkat pengenalan tiap <i>threshold</i> (%)					
	0,9	1,8	2,7	3,6	4,5	5,4
Dika	30%	40%	90%	100%	100%	100%
Dwi	30%	80%	90%	100%	100%	100%
Hari	30%	30%	40%	80%	80%	80%
Man	60%	70%	70%	70%	80%	80%
Tunas	80%	80%	80%	80%	80%	100%
Umam	70%	70%	70%	90%	90%	90%
Baiq	10%	30%	90%	90%	90%	90%
Atiq	30%	40%	70%	90%	90%	90%
Dian	0%	40%	60%	100%	100%	100%
Muna	30%	40%	70%	80%	100%	100%
Rina	100%	100%	100%	100%	100%	100%
Ully	20%	50%	50%	50%	60%	100%
<b>Rata-rata</b>	<b>41%</b>	<b>56%</b>	<b>73%</b>	<b>86%</b>	<b>89%</b>	<b>94%</b>

Berdasarkan hasil Tabel 2 dapat disimpulkan bahwa pengujian yang dilakukan untuk volume suara normal dengan nilai *threshold* tertinggi yaitu *threshold* 5,4 rata-rata persentase pengenalannya sebesar 94% dan dari hasil Tabel diatas kemudian dibuat grafik perbandingannya sebagai berikut:



Gambar 6 Grafik persentase pengenalan untuk volume suara normal

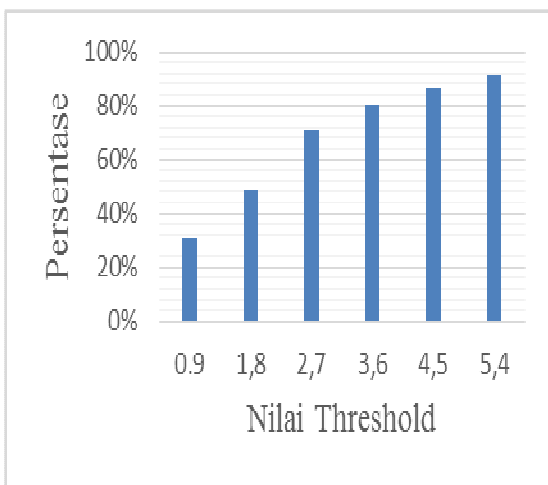
Dari grafik diatas dapat dilihat bahwa untuk volume suara normal semakin tinggi nilai *threshold* maka persentase pengenalannya akan semakin tinggi.

- Tingkat pengenalan untuk volume suara rendah

Tabel 3 Hasil persentase pengenalan untuk volume suara rendah

Nama Penutur	Tingkat pengenalan tiap <i>threshold</i> (%)					
	0,9	1,8	2,7	3,6	4,5	5,4
Dika	30%	60%	100%	100%	100%	100%
Dwi	30%	80%	100%	100%	100%	100%
Hari	30%	30%	40%	80%	80%	80%
Man	80%	100%	100%	100%	100%	100%
Tunas	50%	70%	100%	100%	100%	100%
Umam	20%	20%	30%	40%	60%	70%
Baiq	10%	40%	60%	70%	70%	70%
Atiq	30%	40%	60%	80%	100%	100%
Dian	0%	30%	90%	100%	100%	100%
Muna	30%	30%	50%	80%	90%	90%
Rina	60%	60%	60%	60%	80%	90%
Ulyy	0%	30%	60%	60%	60%	100%
<b>Rata-rata</b>	<b>31%</b>	<b>49%</b>	<b>71%</b>	<b>81%</b>	<b>87%</b>	<b>92%</b>

Berdasarkan hasil Tabel 3 dapat disimpulkan bahwa pengujian yang dilakukan untuk volume suara rendah dengan nilai *threshold* tertinggi yaitu *threshold* 2,3 rata-rata persentase pengenalannya sebesar 45% dan dari hasil Tabel diatas kemudian dibuat grafik perbandingannya sebagai berikut:



Gambar 7 Grafik persentase pengenalan untuk volume suara rendah

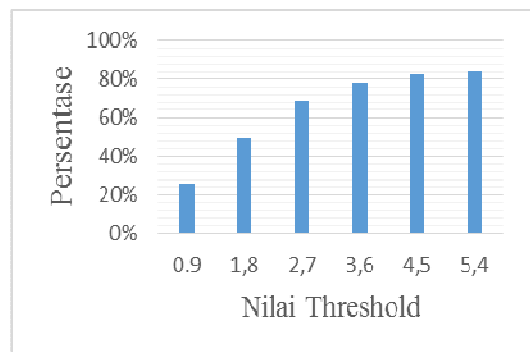
Dari grafik diatas dapat dilihat bahwa untuk volume suara rendah semakin tinggi nilai *threshold* maka persentase pengenalan akan semakin tinggi.

- Tingkat pengenalan untuk volume suara tinggi

Tabel 4 Hasil persentase pengenalan untuk volume suara tinggi

Nama Penutur	Tingkat pengenalan tiap <i>threshold</i> (%)					
	0,9	1,8	2,7	3,6	4,5	5,4
Dika	30%	40%	60%	90%	100%	100%
Dwi	10%	40%	80%	90%	90%	90%
Hari	30%	40%	80%	90%	90%	90%
Man	10%	80%	80%	80%	80%	80%
Tunas	30%	30%	40%	50%	50%	50%
Umam	30%	60%	100%	100%	100%	100%
Baiq	30%	40%	50%	70%	100%	100%
Atiq	0%	70%	90%	100%	100%	100%
Dian	30%	50%	80%	90%	90%	90%
Muna	30%	40%	60%	70%	70%	70%
Rina	70%	70%	70%	70%	80%	100%
Ulyy	10%	30%	40%	40%	40%	40%
<b>Rata-rata</b>	<b>26%</b>	<b>49%</b>	<b>69%</b>	<b>78%</b>	<b>83%</b>	<b>84%</b>

Berdasarkan hasil Tabel 4 dapat disimpulkan bahwa pengujian yang dilakukan untuk volume suara tinggi dengan nilai *threshold* tertinggi yaitu *threshold* 2,3 rata-rata persentase pengenalannya sebesar 48% dan dari hasil Tabel diatas kemudian dibuat grafik perbandingannya sebagai berikut :



Gambar 8 Grafik persentase pengenalan untuk volume suara tinggi

Dari grafik diatas dapat dilihat bahwa untuk volume suara tinggi semakin tinggi nilai *threshold* maka persentase pengenalan akan semakin tinggi.

c. Tingkat penolakan penutur diluar database dengan nilai threshold

Tabel 5 Hasil persentase penolakan penutur diluar database

Nama Penutur	0,9	1,8	2,7	3,6	4,5	5,4
Agustin	100%	100%	100%	100%	90%	0%
Kadek	100%	100%	100%	100%	100%	100%
Arifin	100%	100%	100%	100%	80%	50%
Dila	100%	100%	100%	100%	100%	100%
Ozi	100%	100%	100%	100%	100%	100%
Ita	100%	100%	100%	100%	90%	80%
Henri	100%	100%	100%	100%	100%	100%
Aji	100%	100%	100%	100%	100%	100%
<b>Rata-rata</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>	<b>95%</b>	<b>79%</b>

Berdasarkan Tabel 5 dapat disimpulkan bahwa sistem cukup mampu menolak atau tidak mengenali penutur yang tidak terdapat dalam database.

Tabel 6 Hasil kompromi antara tingkat diterima dan ditolak volume suara normal

Nama Penutur	Hasil kompromi tiap threshold (%)					
	0,9	1,8	2,7	3,6	4,5	5,4
Diterima	41%	56%	73%	86%	89%	94%
Ditolak	100%	100%	100%	100%	95%	79%

Dari Tabel ini bisa dilihat bahwa nilai threshold antara 3,6 sampai 4,5 memiliki tingkat terima dan ditolak paling baik.

## KESIMPULAN

Berdasarkan hasil uji coba dan analisis yang telah dilakukan, maka dapat diambil beberapa kesimpulan antara lain sebagai berikut:

1. Persentase pengenalan untuk penutur sebagai data latih sebesar 100%
2. Tingkat pengenalan volume suara normal dengan semua nilai threshold 0,9 - 5,4 digabung maka rata-rata persentase pengenalannya sebesar 73%, tingkat pengenalan volume suara rendah dengan semua nilai threshold dari 0,9 - 5,4 digabung maka rata-rata persentase pengenalannya sebesar 68%, tingkat pengenalan volume suara tinggi dengan semua nilai threshold 0,9 - 5,4 digabung maka rata-rata persentase pengenalannya sebesar 65%,
3. Pengujian sistem dalam mengidentifikasi penutur dengan volume suara normal dengan threshold 5,4 rata-rata persentase

pengenalannya sebesar 94%, volume suara rendah dengan threshold 5,4 rata-rata persentase pengenalannya sebesar 92%, volume suara tinggi dengan threshold 5,4 rata-rata persentase pengenalannya sebesar 84% dan untuk penutur diluar database rata-rata persentase penolakannya sebesar 79%.

4. Kompromi antara tingkat diterima dan ditolak yang memiliki tingkat akurasi optimum yaitu pada nilai threshold diantara 3,6 sampai 4,5 dengan tingkat terima diantara 86% sampai 89% dan tingkat ditolak diantara 100% sampai 95%.
5. Nilai threshold mempengaruhi berhasil atau tidaknya sistem pengenalan penutur, nilai threshold diantara 3,6 sampai 4,5 adalah nilai threshold yang bisa digunakan untuk pengenalan penutur.

## SARAN

Pada penelitian ini masih terdapat beberapa kekurangan oleh karena itu saran yang diberikan untuk penelitian lebih lanjut adalah sebagai berikut :

1. Melakukan penambahan variasi noise pada data suara.
2. Melakukan pengujian secara langsung atau real time.
3. Melakukan pengujian sistem dengan menggunakan kalimat

## DAFTAR PUSTAKA

- Abriyono, 2012, Pengenalan Ucapan Suku Kata Bahasa Lisan Menggunakan Ciri LPC, MFCC, dan JST, Tugas Akhir Teknik Informatika, STMIK
- Aditya, R., 2012, Prototipe Pengenalan Suara sebagai Penggerak Dinamo Starter Pada Mobil, Tugas Akhir Teknik Mesin Fakultas Teknologi Industri Universitas Gunadarma.
- Munawar, B., 2010, *Pengidentifikasian kata dengan menggunakan metode Hidden Markov Model (HMM) melalui ekstraksi ciri Linear Predictive Coding (LPC)*, Tugas Akhir Teknik Komputer Fakultas Teknik Universitas Komputer Indonesia Bandung.
- Prasetyo, M.E.B., 2010, *Teori Dasar Hidden Markov Model*, Program Studi Sistem dan Teknologi Informasi, Sekolah Teknik Elektro dan Informatika, Institut Teknologi Bandung.



Rianto, J., 2011, Perangkat Lunak Pengenalan Suara (Voice Recognition) Untuk Absensi Karyawan Dengan Menggunakan Metode Dynamic Time Warping (DTW), Tugas Akhir Teknik Informatika Fakultas Teknik Dan Ilmu Komputer Universitas Komputer Indonesia Bandung



**Sri Agustini Yumiati**, lahir di Mataram pada tanggal 07 Agustus 1991, Menempuh Pendidikan Program Strata 1 (S1) di Fakultas Teknik Universitas Mataram sejak tahun 2010

Widya Dharma.Munawar, B., 2010, *Pengidentifikasian Kata Dengan Menggunakan Metode Hidden Markov Model (HMM) Melalui Ekstraksi Ciri Linear Predictive Coding (LPC)*, Jurusan Teknik Informatika, Fakultas Teknik Dan Ilmu Komputer, Universitas Komputer Indonesia.