

Pengenalan Pembicara untuk Menentukan Gender Menggunakan Metode MFCC dan VQ

YOULLIA INDRAWATY N , ANDRIANA , DITA PERMATASARI

Jurusan Teknik Informatika, Fakultas Teknologi Industri
Institut Teknologi Nasional Bandung

Email: youllia@itenas.ac.id

ABSTRAK

Klasifikasi suara berdasarkan gender dibuat dengan tujuan agar komputer mampu mengenali suara laki-laki dan perempuan. Dengan kemampuan komputer yang mampu membedakan suara laki-laki dan perempuan pada pengembangan selanjutnya akan memperkuat tingkat suatu sistem keamanan yang menggunakan password dengan suara. Penelitian ini mengenai pengenalan gender dari pengucap/ pembicara dengan ucapan bergantung teks dan bergantung pembicara, dalam proses pengenalan tersebut digunakan algoritma ekstraksi yang disebut Mel Frequency Cepstral Coefficients (MFCC) digunakan untuk ekstraksi ciri dari sinyal wicara sedangkan proses pengelompokan menggunakan metode Vector Quantization (VQ). Dalam tahap pengenalan, ukuran distorsi berdasarkan minimisasi jarak Euclidean digunakan untuk mencocokkan penutur uji dengan penutur dalam database. Database wicara menggunakan 20 penutur, terdiri dari 10 penutur pria dan 10 penutur wanita dengan tingkat akurasi pria mencapai 90% dan wanita 80%.

Kata kunci : Pengenal pembicara, MFCC, VQ, Klasifikasi Gender, Bergantung Teks.

ABSTRACT

Voice classification of gender is created to recognize the voice of male or female. The ability of computer to recognize the voice of male or female in development will strengthen the level of a security system that uses a password with voice. This study is the introduction of a reciter / speaker with text dependent speech and speaker dependent speech, the recognition process is using extraction algorithm called Mel Frequency cepstral coefficient (MFCC) is used for feature extraction from speech signals and the process of grouping method is using VQ. In the introduction phase, the size of the distortion based on minimization of the Euclidean distance is used to match test speakers with database speakers. Speech database uses 20 speakers, consist of 10 male speakers and 10 female speakers with an accuracy rate of male is 90% and female is 80%.

Keywords: Speaker recognition, MFCC, VQ, gender clasification, text-dependent.

1. LATAR BELAKANG

Klasifikasi suara berdasarkan gender dibuat dengan tujuan agar komputer mampu mengenali suara laki-laki dan perempuan. Dengan kemampuan komputer yang mampu membedakan suara laki-laki dan perempuan pada pengembangan selanjutnya akan memperkuat tingkat suatu sistem keamanan yang menggunakan password dengan suara. Pencocokan password tidak hanya berdasarkan kata saja, namun ditambah dengan pencocokan karakteristik suara sehingga akan lebih aman. Penelitian ini mengenai pengenalan gender dari pengucap/pembicara dengan ucapan bergantung teks dan bergantung pembicara, Terdapat banyak metode yang dapat digunakan untuk pengekstraksian sinyal ucapan. Salah satu teknik pengkodean yang dapat digunakan dalam pengekstraksian sinyal ucapan adalah MFCC (*Mel Frequency Cepstral Coefficient*). MFCC ini menghitung koefisien cepstral. MFCC merupakan salah satu metode yang banyak digunakan dalam bidang *speech technology*, baik *speaker recognition* maupun *speech recognition*.

Pada proses klasifikasi digunakan metode VQ (*Vector Quantization*), yang mengubah hasil ekstraksi analisis MFCC dari masing-masing pengucap menjadi sekumpulan *codebook*. *Codebook* tersebut kemudian dibandingkan dengan hasil ekstraksi koefisien MFCC dari sinyal masukan yang akan dikenali. Pengklasifikasian ditentukan dengan menghitung jarak penyimpangan. Jarak penyimpangan terkecil merupakan identitas dari sinyal tersebut.

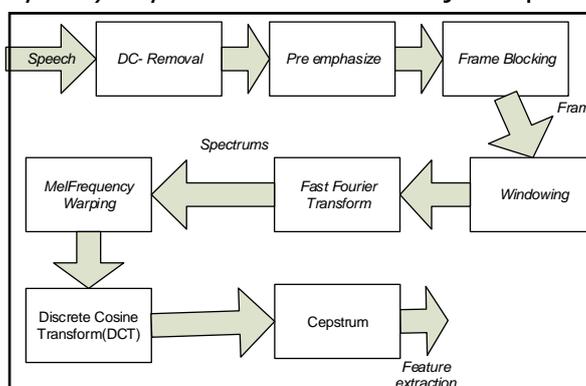
Oleh karena itu penelitian ini mengaplikasikan metode tersebut, dengan mengangkat judul, yaitu "Pengenalan Pembicara untuk Menentukan *Gender* menggunakan metode MFCC dan VQ". Melalui penelitian ini penulis membangun sebuah aplikasi yang berfungsi untuk mengetahui *gender*/jenis kelamin dari suara manusia yang sebelumnya pernah disimpan sebagai data latihan.

Pada penelitian ini suara analog diubah menjadi digital dengan format *.wav dengan data masukan (pada basis data) berupa sinyal suara yang diambil dari 20 orang responden (10 pria dan 10 wanita) dengan rentang usia 18- 40 tahun dalam keadaan sehat (tidak *flu* atau batuk). *Sampling rate* yang digunakan ialah 44100Hz, *channel* mono, 16 bit. Bahasa yang digunakan dalam proses latihan dan proses uji adalah Bahasa Indonesia dan diucapkan oleh orang Indonesia.

2. METODOLOGI PENELITIAN

2.1 *Mel Frequency Cepstral Coefficient (MFCC)*

Tahapan proses *Mel Frequency Cepstral Coefficient* ditunjukkan pada Gambar 1



Gambar 1 Tahapan proses MFCC
(Sumber: Putra, Darma dan Adi Resmawan, 2012)

Tahapan proses *Mel Frequency Cepstral Coefficient* dijelaskan sebagai berikut ini:

DC-Removal

Tujuan dari *DC-Removal* adalah mendapat normalisasi dari data suara input.

Proses *DC-Removal* ini dapat dituliskan dalam persamaan (1) berikut:

$$y[n] = x[n] - \bar{x}, 0 \leq n \leq N-1 \quad \dots\dots(1)$$

Keterangan :

$y[n]$ = Sample Signal hasil DC Removal

$x[n]$ = Sample signal asli

\bar{x} = Nilai rata-rata sample signal asli

N = Panjang signal

Pre-emphasize Filtering

Pre-emphasize Filtering merupakan salah satu jenis *filter*. *Filter* ini mempertahankan frekuensi- frekuensi tinggi pada sebuah *spectrum*, yang umumnya tereliminasi pada saat proses produksi suara.

Proses *Pre-emphasize* ini dapat dituliskan dalam persamaan (2) berikut:

$$y[n] = s[n](1-\alpha), 0.9 \leq \alpha \leq 1.0 \quad \dots(2)$$

Dimana:

$y[n]$ = *signal* hasil *pre-emphasize filter*

$s[n]$ = *signal* sebelum *pre-emphasize filter*

Frame Blocking

Pada langkah ini, sinyal ucapan yang terdiri dari S sampel ($X(S)$) dibagi menjadi beberapa *frame* yang berisi N sampel, masing-masing *frame* dipisahkan oleh M ($M < N$).

Windowing

Proses *windowing* yang dipakai adalah proses *Hamming Window*, proses tersebut dapat dituliskan dalam Persamaan (3) berikut :

$$w(n) = 0.54 - 0.46 \cos \frac{2\pi n}{M-1} \quad (3)$$

Dimana:

N = Jumlah sample pada masing- masing Frame

$n = 0,1,2,3, \dots .N-1$

FFT

Fast Fourier Transform (FFT), yang berfungsi untuk mengubah domain waktu menjadi domain frekuensi.

Proses FFT dapat dituliskan dalam persamaan berikut:

$$f(n) = \sum_{K=0}^{N-1} (Y_k) e^{-2\pi jkn/N}, n = 0,1,2,, \dots, N - 1 \quad (4)$$

Dengan:

$f(n)$ = frekuensi

$k = 0,1,2,\dots,(N-1)$

N = jumlah sample pada masing-masing frame

j = bilangan imajiner ($\sqrt{-1}$)
 n = 1,2,3, . . . (N-1)

Mel- frequency wrapping

Persamaan yang digunakan untuk menghitung skala mel untuk frekuensi dalam Hz ditunjukkan dalam Persamaan (5) berikut :

$$Mel f = 2595 * \log_{10}\left(1 + \frac{f}{700}\right) \quad (5)$$

Cepstrum

Untuk mengubah ke dalam domain waktu, maka diperlukan DCT (Discrete Cosine Transform) dengan persamaan dituliskan dalam Persamaan (6) berikut

$$C_n = \sum_{k=1}^K (\log \tilde{S}_k) \cos \left[n \left(k - \frac{1}{2} \right) \frac{\pi}{k} \right], n = 1, 2, \dots, k \quad (6)$$

Dimana:

C_n = koefisien cepstrum *mel-frequency*

\tilde{S}_k = mel frekuensi

k = jumlah coefficient

2.2 VQ (Vector Quantization)

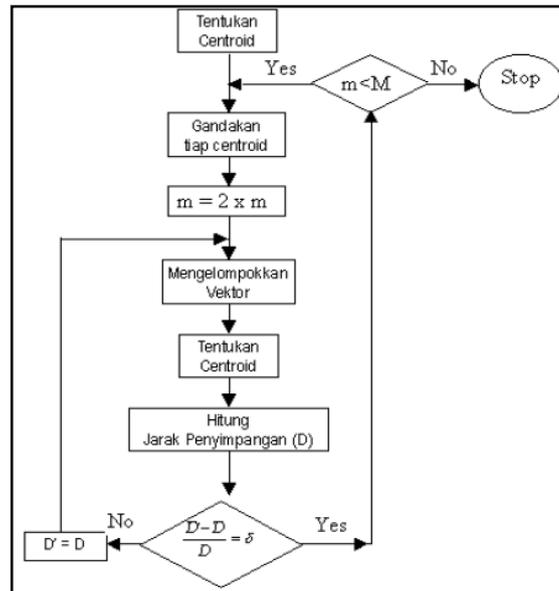
VQ adalah proses dari pemetaan vektor dari ruang vektor yang besar menjadi sebuah wilayah yang terbatas.

2.3 Algoritma K-Means

Salah satu metode dasar yang digunakan untuk disain bukode adalah algoritma iterasi yang dikenal sebagai algoritma Lloyd (K-means algorithm atau Lloyd algorithm) [2].

Algoritma tersebut diimplementasikan dalam prosedur berulang seperti pada Gambar 2 dan langkah- langkah berikut:

1. Menentukan vektor *codebook* pertama, yang merupakan *centroid* dari keseluruhan vektor ciri.(tidak ada pengulangan pada tahap ini)
2. Gandakan ukuran *codebook*
3. Pencarian *Nearest-Neighbour* : untuk tiap vektor ciri, temukan *codeword* di dalam *codebook* tersebut (*codebook* saat itu) yang paling dekat(jarak penyimpangannya paling kecil), dan tempatkan vector tersebut dalam kelompok *codeword* tersebut.
4. Pembaharuan *centroid* : memperbaharui *codeword* pada tiap kelompok dengan menggunakan *centroid* dari vektor ciri terletak di sel tersebut.
5. Iterasi I : mengulang langkah 3 dan 4 sampai diperoleh jarak penyimpangan rata rata (D) yang besarnya dibawah batasan yang telah ditentukan (δ). D' merupakan nilai distorsi awal yang nilainya ditentukan pada saat inialisasi pada awal program.



Gambar 2. Algoritma K-means

2.4 Perhitungan Jarak Penyimpangan

Untuk menghitung jarak penyimpangan antara dua vektor maka digunakan *Euclidean distance* (jarak Euclidean) [4]. Persamaan untuk menghitung jarak Euclidean Distance ditunjukkan oleh Persamaan (7) berikut :

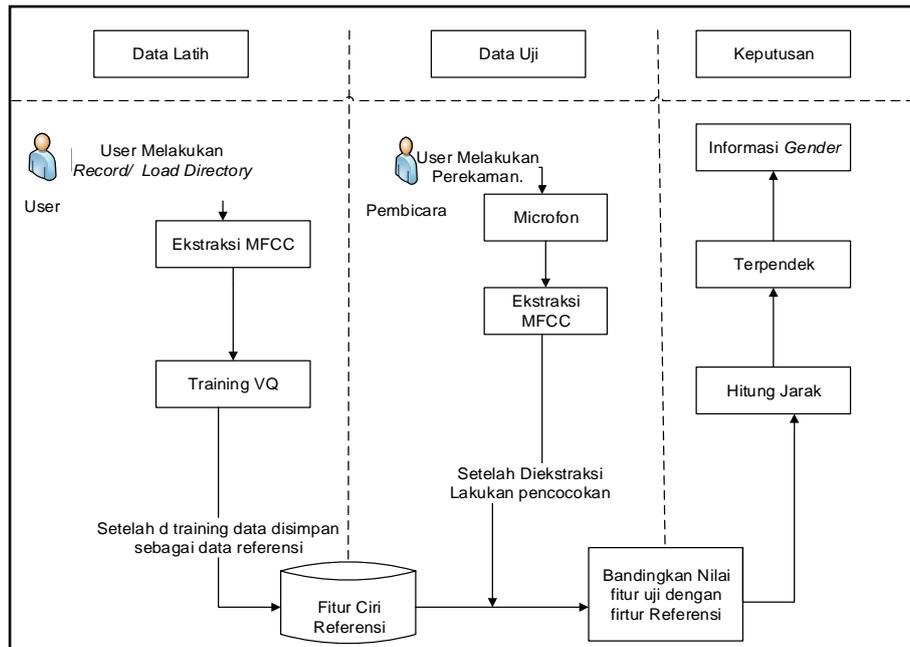
$$d_E(x, c) = \sqrt{\sum_{i=1}^{dim} (X_i - C_i)^2} \quad (7)$$

Dimana x merupakan vektor ciri dan c merupakan vector dari suatu *codebook*. Pada pengenalan pengucap dengan menggunakan metode VQ, jarak penyimpangan Euclidean ini digunakan untuk menghitung jarak

Penyimpangan antara masing-masing vektor ciri dengan *codeword* pada tiap-tiap *codebook*, sehingga dapat diketahui *codeword* mana yang memiliki jarak penyimpangan terdekat dengan vektor ciri.

2.5 Cara Kerja Sistem

Untuk menggambarkan langkah-langkah dan urutan prosedur dari aplikasi Pengenalan pembicara untuk menentukan *gender* secara umum maka dibuat sebuah model kerja sistem aplikasi secara umum seperti pada Gambar 3 .



Gambar 3. Model kerja system aplikasi secara umum

Pertama *user* melakukan pengambilan sample suara, pengambilan tersebut dapat dilakukan dengan cara *record*:

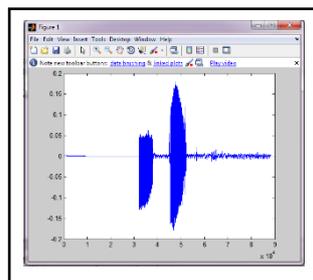
```
recObj = audiorecorder(Fs, nbits, 1, dev_id);
```

atau

load sound

```
[nama_file nama_folder] = uigetfile({'*.wav'}, 'Ambil file suara');  
[y Fs] = wavread ([nama_folder nama_file]);
```

Aplikasi membaca sinyal suara yang dipilih oleh *user*, kemudian aplikasi memilih *axes* yang digunakan sebagai grafik. Sehingga terbentuk sebuah grafik sinyal suara seperti Gambar 4



Gambar 4. Grafik sinyal suara

Kemudian suara yang masuk dilakukan proses ekstraksi ciri menggunakan metode MFCC, dengan tahapan sebagai berikut:

- *DC Removal*
Remove DC Components bertujuan untuk menghitung rata-rata (nilai mean) dari data sampel suara (*y_filter*), dan mengurangi setiap nilai sampel suara dengan nilai rata-rata tersebut kemudian disimpan dalam variable *y_norm*.

```
nilaimean = mean(y_filter);  
y_norm = y_filter-nilaimean;
```

- *Pre-emphasize*

Filter ini mempertahankan frekuensi- frekuensi tinggi pada sebuah *spectrum* dengan cara memfilter nilai koefisien (1) dari pre-emphasize dengan nilai alpha (0.9375) kemudian disimpan kedalam *variable* sinyal suara (*xx*).

```
xx=filter([1 -0.9375],1,xx);
```

- *Frame Blocking*

Frame blocking dilakukan dengan menggunakan fungsi yang telah disediakan oleh Matlab, kemudian hasilnya disimpan ke dalam *variable* *fr*.

```
fr=framing(xx);
```

- *Windowing*

Proses ini dibutuhkan untuk meminimalisir diskontinuitas antar frame yang menyebabkan kehilangan informasi yang terdapat dalam suatu sinyal.

Windowing dilakukan dengan menggunakan fungsi yang telah disediakan oleh matlab, kemudian hasilnya disimpan kedalam *variable* *s*.

```
s = y .* hamming(256);
```

- FFT

Fast Fourier Transform (FFT), yang berfungsi untuk mengubah domain waktu menjadi domain frekuensi. FFT dilakukan dengan menggunakan fungsi yang telah disediakan oleh matlab, kemudian hasilnya disimpan kedalam *variable* *t*

```
t = abs(fft(s));
```

- *Mel- frequency wrapping*

Mel- frequency wrapping dilakukan dengan menggunakan fungsi yang telah disediakan oleh matlab, dengan ketentuan yang digunakan pada penelitian ini yaitu jumlah *mel bank*=64, panjang *frame*= 256 dan jumlah sampling 44100 kemudian hasilnya disimpan kedalam *variable* *bank*.

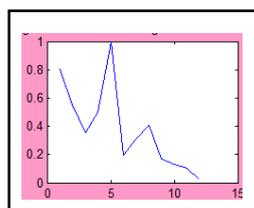
```
bank=melbankm(64,256,44100,0,0.5,'t');
```

- *Cepstrum*

Pada tahap ini, akan menghasilkan nilai MFCC yang merupakan hasil konversi log mel cepstrum ke dalam domain waktu menggunakan DCT.

```
dctcoef(k,:)=cos((2*n+1)*k*pi/(2*24));
```

aplikasi memproses suara yang telah dipilih oleh user menggunakan metode MFCC. Sehingga akan terbentuk grafik seperti Gambar 5



Gambar 5. Grafik proses MFCC

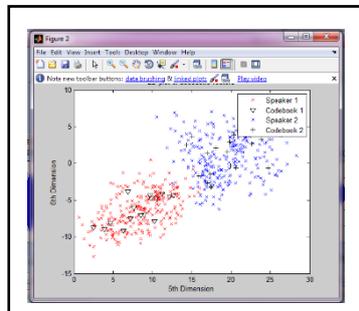
setelah dilakukan ekstraksi ciri, kemudian dilakukan training *vector quantization* dengan menggunakan algoritma *k-menas* untuk melakukan pengelompokan *gender* berdasarkan kemiripan/ kedekatan dari vektor ciri yang telah didapat dari proses ekstraksi, vektor ciri hasil *training VQ* disimpan ke dalam database sebagai data referensi. *Training VQ* digunakan dengan menggunakan fungsi yang telah terdapat pada matlab dimana nilai MFCC yang dihasilkan dari ekstraksi sebelumnya yang disimpan pada *variable d* dan jumlah *centroid* yang disimpan pada *variable k*. Nilai hasil pengelompokan disimpan ke dalam file *kmeans* disimpan kedalam

`kmeans.mat`

`R= vqlbg(d, k)`

`Save kmeans.mat`

Dengan menggunakan perintah tersebut maka aplikasi akan menampilkan grafik ciri yang telah ter *cluster* seperti pada Gambar 6



Gambar 6. Hasil Cluster ciri pembicara

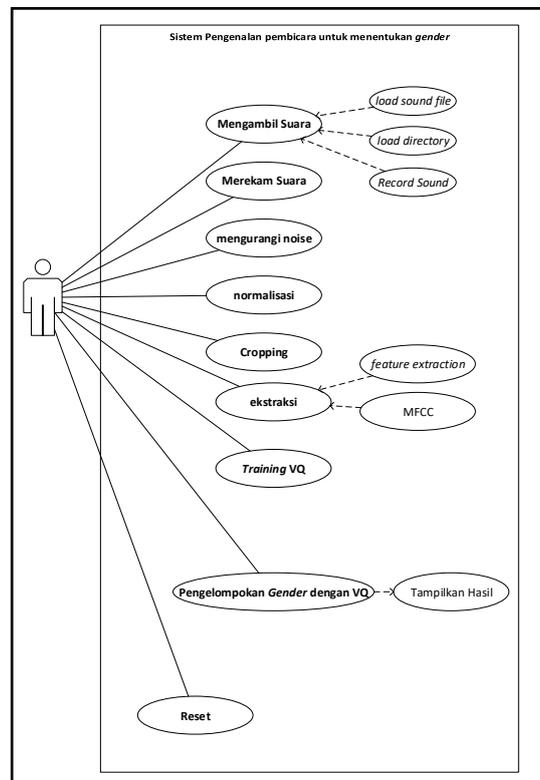
Fase tersebut dinamakan dengan fase *training* (proses latihan). Setelah dibuat data referensi pada fase latihan, selanjutnya dilakukan fase testing (pengujian). Pada fase ini user kembali melakukan perekaman suara melalui mikrofon kemudian suara yang masuk diekstraksi ciri nya. Hasil dari ekstraksi ciri pada fase pengujian, nilai tersebut dibandingkan dengan data referensi yang terdapat pada *database*, proses pencocokan ini menggunakan metode *VQ(vector Quantization)* dengan menghitung jarak terdekat dari setiap data latihan (x_1) dengan data uji (x_2) menggunakan perhitungan jarak *Euclidean Distance*. Jarak yang terpendek adalah jarak yang diidentifikasi sebagai *gender* tersebut. Berikut ini adalah perintah perhitungan menggunakan *Euclidean Distance* pada matlab:

```
Dist= Euclidean(x1,x2)
Dist=sqrt(sum((x1(:)-x2(:)).^2));
```

3. ANALISIS DAN PEMBAHASAN

3.1 Fungsionalitas Sistem

Seluruh fungsionalitas yang terdapat pada aplikasi akan dinyatakan pada *use case* diagram. *Use case* diagram dapat dilihat pada Gambar 7.



Gambar 7. Usecase fungsionalitas sistem

Tampilan Halaman Menu Utama

Form utama merupakan antarmuka yang ditampilkan pada awal dari aplikasi pengenalan Gender melalui suara. Form ini berisi 3 *button* untuk mengakses *form* Latih, *form* uji, *form real time* ditampilkan pada Gambar 8

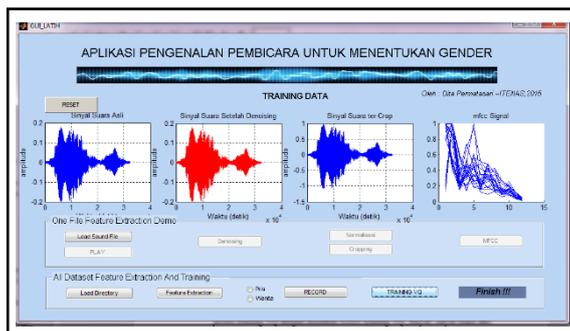


Gambar 8. Halaman Menu

Tampilan Halaman Latih Suara

Form Latih suara seperti pada Gambar 9 merupakan antarmuka dari program pengenalan pembicara untuk menentukan Gender. Dalam implementasinya antar muka ini pengguna melakukan proses ekstraksi dan training VQ. Suara yg digunakan untuk melakukan poses

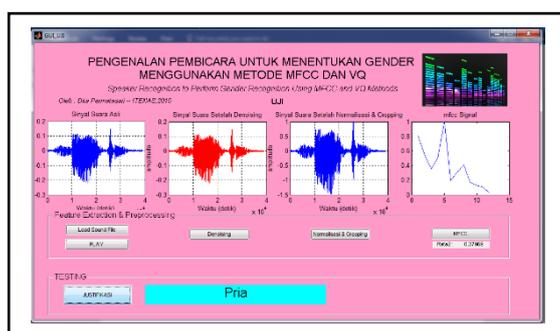
pelatihan ini dalam format *.wav*. proses pengambilan sample suara ini dapat dilakukan dengan melakukan *load file* dan untuk mendengarkan *file* suara yang sudah di *load* maka dibuat tombol *play*. Kemudian untuk melakukan proses ekstraksi dilakukan dengan menekan *button* MFCC yang sebelumnya telah dilakukan proses normalisasi, *denoising* dan *cropping*. Atau dapat dilakukan dengan menekan tombol *button load sound directory* untuk melakukan ekstraksi satu *folder* sekaligus dengan menekan tombol *feature extraction*. Dan untuk melakukan pengelompokan sinyal suara yang telah dilakukan ekstraksi dilakukan proses training VQ dengan menekan *button* training VQ. Sinyal- sinyal suara dari proses- proses tersebut juga ditampilkan pada form ini. Untuk menghilangkan sinyal- sinyal suara yang telah diproses disediakan *button reset*.



Gambar 9. Halaman Training Suara

Tampilan Halaman Uji Suara

Form pengujian ucapan suara seperti pada Gambar 10 diperuntukan untuk melakukan pengujian pembicara yang sebelumnya pernah terdaftar di dalam database. *Form* pengujian ini di tampilkan pada Gambar. Dalam implementasinya pengguna melakukan *load file* untuk menentukan suara mana yang akan diujikan sebagai pengenalan *gender*. Kemudian pengguna melakukan proses *denoising* dengan menekan *button denoising* dan melakukan proses *normalisasi & cropping* dengan menekan tombol *normalisasi & cropping*. Suara yang telah di *load* dan melewati proses tersebut kemudian diekstraksi menggunakan MFCC, untuk melakukan ekstraksi tersebut dapat dilakukan dengan menekan *button* MFCC. Pengguna yang telah melakukan tahapan tersebut kemudian dilakukan perbandingan dengan suara yang ada di referensi.

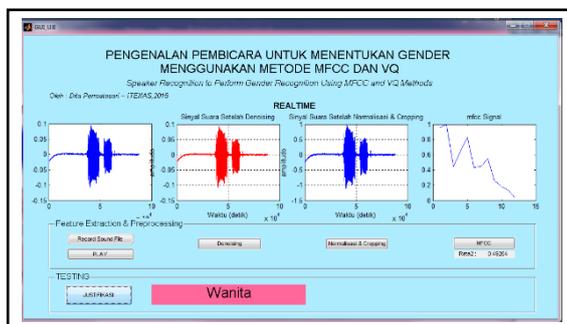


Gambar 10. Halaman Uji Suara

Tampilan Halaman Real Time

Form real time seperti pada Gambar 11 ini sama dengan *form* pengujian ucapan yang fungsinya juga sama diperuntukan untuk melakukan pengujian pembicara yang sebelumnya pernah terdaftar di dalam database. Dalam implementasinya pengguna melakukan *recod sound file* untuk dijadikan suara yang akan diujikan untuk pengenalan *gender*. Kemudian pengguna melakukan proses *denoising* dengan menekan *button denoising* dan melakukan proses *normalisasi & cropping* dengan menekan tombol *normalisasi & cropping*. Suara yang

telah di *load* dan melewati proses tersebut kemudian diekstraksi menggunakan MFCC, untuk melakukan ekstraksi tersebut dapat dilakukan dengan menekan *button* MFCC. Pengguna yang telah melakukan tahapan tersebut kemudian dilakukan perbandingan dengan suara yang ada di referensi.



Gambar 11. Halaman Real Time

3.2 Pengujian Offline

Pengujian dilakukan oleh 20 orang yang terdiri dari 10 pria dan 10 wanita dengan melakukan *load* data. Masing- masing orang diambil *sample* suaranya dengan mengucapkan kata yang sama yaitu kata “**baju**”.

Hasil pengujian suara pria ditunjukkan oleh Tabel 1.

Table 1. hasil uji pria mengucapkan kata sama dengan data latih

No	Nama	Usia	pekerjaan	Gender yang terdeteksi
1	Adhe	22th	Mahasiswa	P
No	Nama	Usia	pekerjaan	Gender yang terdeteksi
2	Andri	22th	Mahasiswa	P
3	Asep	20th	Mahasiswa	P
4	Eka	24th	Mahasiswa	W
5	Fahmi	20th	Mahasiswa	P
6	Hafiz	20th	Mahasiswa	P
7	Rifky	22th	Mahasiswa	P
8	Toni	23th	Mahasiswa	P
9	Yoga	20th	Mahasiswa	P
10	M.eka	22th	Mahasiswa	P

Keterangan Tabel 1.

P : Pria W : Wanita

Hasil pengujian suara wanita ditunjukkan oleh Tabel 2.

Table 2. hasil uji wanita mengucapkan kata sama dengan data latih

No	Nama	Usia	pekerjaan	Gender yang terdeteksi
1	Dinda	18th	Mahasiswa	W
2	Dita	22th	Mahasiswa	W
3	Fuzna	23th	Pegawai	P
4	Intan	18th	Mahasiswa	W
5	Isty	23th	Megawai	P
6	Nadia	18th	Mahasiswa	W
7	Novi	23th	Pegawai	W
8	Nur	22th	Pegawai	W
9	Dinda2	25th	Mahasiswa	W
10	Eva	22th	Pegawai	W

Keterangan Tabel 2.
P : Pria W : Wanita

Total Hasil Pengujian

Tabel 1 dan tabel 2 menunjukkan hasil pengujian dari *sample* suara. Tingkat kesuksesan dalam persen(%) dapat dihitung dengan cara :

Presentase matching=

$$\frac{\text{Hasil Yang Sesuai/Benar}}{\text{Total Seluruh Uji Coba}} \times 100 \%$$

Tingkat kesuksesan dalam persen(%) untuk pengujian yang terdiri dari 10 orang wanita yang pembicara dan kata nya sama dengan data latih adalah sebagai berikut:

$$\frac{8}{10} \times 100 = 80 \%$$

Sedangkan pada 10 orang pria adalah sebagai berikut:

$$\frac{9}{10} \times 100 = 90 \%$$

Pengujian Online

Pengujian dilakukan oleh 20 orang yang terdiri dari 10 pria dan 10 wanita dengan melakukan *record* secara *real time*. Masing- masing orang diambil *sample* suaranya sebanyak 10 kali dengan mengucapkan kata yang sama yaitu kata "**baju**".

Hasil pengujian suara pria ditunjukkan oleh Tabel 7.

Table 3. hasil uji online pria yang mengucapkan kata yang sama dengan data latih

No	Nama	Usia	pekerjaan	Gender yang terdeteksi										Jumlah Benar	
				1	2	3	4	5	6	7	8	9	10		
1	Adhe	22th	Mahasiswa	P	P	P	P	P	P	P	P	P	P	P	10
2	Andri	22th	Mahasiswa	W	W	P	W	P	W	W	P	P	P	P	5
3	Asep	20th	Mahasiswa	P	P	W	P	P	P	W	W	P	P	P	7
4	Eka	24th	Mahasiswa	W	P	W	P	P	W	P	P	P	P	P	7
5	Fahmi	20th	Mahasiswa	P	W	P	P	W	P	W	W	W	W	W	4
6	Hafiz	20th	Mahasiswa	P	W	P	P	P	P	P	P	P	P	P	9
7	Rifky	22th	Mahasiswa	P	P	P	P	P	W	P	P	P	P	P	9
8	Toni	23th	Mahasiswa	P	P	P	P	P	P	W	P	P	P	P	8
9	Yoga	20th	Mahasiswa	P	W	p	P	W	W	P	W	W	W	W	4
10	M.eka	22th	Mahasiswa	P	P	P	W	P	P	W	W	W	W	W	4

Hasil pengujian suara wanita ditunjukkan oleh Tabel 8.

Table 4. hasil uji online wanita yang mengucapkan kata yang sama dengan data latih

No	Nama	Usia	pekerjaan	Gender yang terdeteksi										Jumlah Benar
				1	2	3	4	5	6	7	8	9	10	
1	Dinda	18th	Mahasiswa	W	W	W	P	W	P	W	W	P	W	7
2	Dita	22th	Mahasiswa	W	W	P	P	W	W	P	W	W	P	6
3	Fuzna	23th	Pegawai	W	P	W	W	P	P	W	W	W	P	6
4	Intan	18th	Mahasiswa	W	W	P	P	W	P	P	W	P	W	5
5	Isty	23th	Megawai	W	W	W	W	W	P	W	W	P	W	8
6	Nadia	18th	Mahasiswa	W	P	W	W	P	W	W	W	P	W	7
7	Novi	23th	Pegawai	P	W	P	W	P	W	W	P	W	W	6
8	Nur	22th	Pegawai	P	W	W	W	W	W	W	W	W	W	9
9	Dinda2	25th	Mahasiswa	W	P	P	W	W	P	P	P	W	P	4
10	Eva	22th	Pegawai	P	W	P	P	P	W	P	P	W	P	3

Tingkat kesuksesan dalam persen(%) untuk pengujian yang terdiri dari 10 orang wanita yang pembicara dan kata nya sama dengan data latih adalah sebagai berikut:

$$\frac{50}{100} \times 100 = 61 \%$$

Sedangkan pada 10 orang pria adalah sebagai berikut:

$$\frac{48}{100} \times 100 = 67 \%$$

4. KESIMPULAN

Berdasarkan hasil penelitian serta pengujian sistem yang telah dilakukan kepada 10 orang pria dan 10 orang wanita, maka dapat disimpulkan bahwa tingkat kesuksesan pengenalan pembicara yang dapat dikenali pada pengujian *offline* pria mencapai 90% dan wanita mencapai 80%. Sedangkan pada pengujian *online* pria mencapai 67% dan wanita mencapai 61%. Data yang sebelumnya tidak pernah dijadikan data latih akurasi nya tidak lebih baik dari akurasi data yang sebelumnya pernah dijadikan data latih. Sehingga metode MFCC dan VQ terbukti cukup baik untuk memperkuat tingkat suatu sistem keamanan dengan suara. Kegagalan penelitian dapat diakibatkan oleh keadaan lingkungan atau keadaan dari pembicara.

DAFTAR PUSTAKA

- [1] Agustini, Ketut, 'Biometrik Suara Dengan Transformasi Wavelet'. Di unduh 18 Maret 2015.
- [2] Elfitri, Ikhwana, 'Kuantisasi Vektor : Definisi, Design dan Kerja'. No.29, Volume 1, Tahun XV, April 2008, Di unduh 18 April 2015.
- [3] Lawrence Rabiner, dkk. 1993. 'Fundamental of Speech Recognition'. Prentice Hall PTR
- [4] Mustofa, Ali, 'Sistem Pengenalan Penutur dengan metode Mel-Frequency Wrapping'. Jurnal Teknik Elektro vol.7, No.2, September 2007:88-96.
- [5] Putra, Darma dan Adi Resmawan, Juni 2011 , 'Verifikasi Biometrika Suara Menggunakan Metode MFCC dan DTW'.
- [6] " _____", 'Pengenalan Pola'. https://id.wikipedia.org/wiki/Pengenalan_pola. Diunduh pada tanggal 09 Agustus 2015.

- [7] Setiawan, Arif, 'Analisis Klasifikasi Suara Berdasarkan *Gender* Dengan Format Wav Menggunakan Algoritma K-Means'. http://eprints.umk.ac.id/79/1/ANALISIS_KLASIFIKASI_SUARA_BERDASARKAN_GENDER.pdf. Diunduh pada tanggal 20 Agustus 2015.
- [8] Aditya, Eko, dkk. 'Klustering Suara Laki-Laki dan Perempuan Menggunakan Algoritma K-Means Berdasarkan Hasil Ekstraksi FFT (Fast Fourier Transform)'. <http://yudistira.lecture.ub.ac.id/files/2014/04/Klustering-Suara-Laki-Laki-dan-Perempuan-Menggunakan-Algoritma-K-Means-Berdasarkan-Hasil-Ekstraksi-FFT.pdf>. Diunduh pada tanggal 25 November 2015
- [9] Bagas, Susetyo dan Altedzar Riedho W. Seminar Nasional Aplikasi Teknologi Informasi 2012 (SNATI 2012). Yogyakarta, 15-16 Juni 2012. 'Aplikasi Pengenalan *Gender* Menggunakan Suara '. Diunduh pada tanggal 10 November 2015