

PENGOLAHAN SUARA MENGGUNAKAN TRANSFORMASI WAVELET DAN JARINGAN SYARAF TIRUAN UNTUK PENGENALAN PEMBICARA

Rustati Rahmi

ABSTRAK

Pengenalan pembicara berbeda dari pengenalan ucapan (*speech recognition*), dimana pengenalan ucapan lebih dikonsentrasikan pada ekstraksi dari beberapa bagian informasi pesan yang didalamnya berisikan teks yang diucapkan. Teks yang diucapkan ini mengandung unit-unit linguistic yang lebih dikenal sebagai *Phonem*, kata-kata kunci suatu pesan ataupun pesan yang lengkap. Sedangkan pengenalan pembicara (*speaker recognition*) menekankan pada analisa suara yang diucapkan oleh setiap pembicara sehingga dapat dibedakan karakteristik unik suara dari tiap pembicara tersebut.

Dalam proses pengenalan pembicara akan dilakukan beberapa proses yaitu: akuisisi data suara digital, kemudian dilakukan proses *frame blocking* dan *windowing* (proses segmentasi dan frame) untuk meminimalkan diskontinuitas sinyal pada bagian awal dan akhir sinyal suara. Pada proses selanjutnya dilakukan Ekstraksi ciri sebagai proses pemampatan data agar berdimensi kecil tanpa mengubah karakteristik sinyal suara. Proses ini dilakukan dengan menggunakan Transformasi *Wavelet*, yaitu menghasilkan koefisien prediksi. Kemudian dilakukan proses referensi pembicara dengan melakukan pelatihan terhadap data sinyal suara yang telah di-ekstrak, untuk mendapatkan klasifikasi pola suara pembicara yang berbeda, proses ini dilakukan dengan metode JST Som Kohonen.

Dari penelitian ini telah didapati akurasi pengenalan pembicara dengan tingkat keberhasilan sebesar 96% dalam proses pengenalan pola suara.

Kata kunci: Pengenalan Pembicara (recognition speaker), Transformasi Wavelet, SOM Kohonen

I. PENDAHULUAN

1.1 Latar Belakang

Komunikasi suara (*speech communication*) merupakan salah satu media komunikasi utama manusia yang paling cepat dan tepat dalam menyampaikan suatu informasi. Analisa suara merupakan hal yang penting dalam beberapa proses yang berkaitan dengan pemrosesan suara, misalnya identifikasi ucapan (*speaker identification*) atau pengenalan pembicara (*speaker recognition*).[1]

Pada pengembangannya sistem keamanan, pengolahan sinyal suara mulai ditingkatkan sebagai sistem deteksi sinyal suara pembicara untuk pengenalan pelaku kriminal. Didalam penelitian tersebut masih banyak kendala/hambatan yang dihadapi yaitu tentang metode yang tepat didalam proses sinyal suara untuk pengenalan pembicara secara tepat.[2]

Pengenalan pembicara (*speaker recognition*) menekankan pada analisa suara yang diucapkan oleh setiap pembicara sehingga dapat dibedakan karakteristik unik suara dari tiap pembicara tersebut.[1] Pada penelitian-penelitian sebelumnya tentang pengenalan pembicara khususnya dalam proses pengolahan sinyal suara banyak digunakan metode *Fourier Transformation*, seperti: *Linier Prediction Coding* (LPC), *Mel Frequency Cepstrum Coefficients* (MFCC), *Neural Predictive Coding* (NPC) dan sebagainya, dimana semua metode tersebut berbasiskan Transformasi Fourier dan tingkat pengenalannya cukup baik [2].

Pada penelitian ini akan dicoba metode lain dengan suatu konsep dan pendekatan tanpa berbasis Transformasi *Fourier* yaitu Transformasi *Wavelet*. Transformasi *Wavelet* ini belum banyak digunakan untuk analisa suara.

Transformasi *Wavelet* sinyal suara menghasilkan resolusi waktu yang baik pada frekwensi tinggi dalam menentukan batas awal suara dan parametrisasi ciri suara durasi rendah serta mampu menganalisa sinyal diskontinu secara akurat.[3]

Pada proses pengenalan pembicara, dimana seperti yang diketahui bahwa sinyal suara manusia mempunyai tingkat *variabilitas* yang sangat tinggi, sehingga sinyal suara tersebut mempunyai pola ucapan yang berbeda-beda. Salah satu metode yang dapat mengenali perbedaan pola suara, yaitu menggunakan metode Jaringan Syaraf Tiruan. Salah satu jenis JST yang mampu memberikan unjuk kerja yang bagus adalah JS dengan arsitektur *Multi-Layer Perceptron* (MLP) dan pembelajaran *Backpropagation*. [4] Kedua metode JST ini merupakan jenis *Supervised*. Namun ingin dicoba juga untuk JST yang mempunyai jenis *Unsupervised*, yaitu *SOM Kohonen*. Dimana prinsip pembelajarannya tidak menggunakan data *output* sebagai target pembelajaran.

1.2 Rumusan Masalah

Akurasi pengenalan suara pembicara yang dihasilkan pada penelitian sebelumnya, masih memungkinkan untuk ditingkatkan dengan menggunakan alternative metode lain, sehingga apa yang menjadi pokok masalah dari penelitian ini adalah :

“ Pengenalan suara pembicara pada penelitian sebelumnya sebesar 86% masih perlu ditingkatkan akurasi, sehingga computer dapat lebih mengenali sample suara yang diinputkan dan teregistrasi dalam proses pelatihan”

1.3 Tujuan Penelitian

Tujuan yang akan dicapai oleh penulis dari penulisan tesis ini adalah :

1. Meningkatkan sistem pengenalan suara pembicara dengan pendekatan metode Transformasi *Wavelet* dan JST *Som-Kohonen* untuk meningkatkan keakurasian pengenalan pola sinyal suara.

2. Pengukuran tingkat akurasi sistem dan perbandingan terhadap hasil penelitian sebelumnya.

1.4 Manfaat Penelitian

Dapat digunakan dibidang keamanan dalam pengidentifikasian pembicara yang diaplikasikan pada sistem Kata kunci (*password*) dan juga dapat dimanfaatkan untuk mengenali pola suara pembicara dalam rangka pemberian kepastian bila digunakan dalam kegiatan mata-mata (*spy*) dalam melakukan proses penyadapan dan sebagainya.

II. LANDASAN TEORI

1.5 Sinyal Suara

Sinyal dapat dijelaskan sebagai deskriptif fisik yang bervariasi seiring waktu atau variabel bebas lainnya yang menyimpan suatu informasi. Pengklasifikasian sinyal dapat dibedakan menjadi beberapa jenis yaitu: sinyal waktu kontinyu, sinyal waktu diskrit, sinyal nilai kontinyu, sinyal nilai diskrit, sinyal *random*, dan sinyal *nonrandom*. Sinyal waktu kontinyu dengan nama lain sinyal analog adalah sinyal yang belum melalui proses apapun.

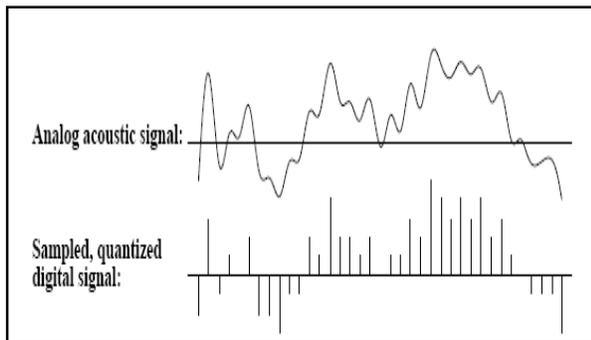
Sedangkan sinyal nilai diskrit atau sinyal digital adalah sinyal analog yang telah melalui proses *sampling*, *quantization*, dan *encoding*.

Sampling adalah proses mengambil nilai-nilai sinyal pada titik-titik diskrit sepanjang variabel waktu dari sinyal waktu

kontinu, sehingga didapatkan sinyal waktu diskrit. Jumlah titik-titik yang diambil setiap detik dinamakan sebagai *sampling rate*. Dalam melakukan *sampling*, perlu diperhatikan kriteria *Nyquist* yang menyatakan bahwa sebuah sinyal harus memiliki *sampling rate* yang lebih besar dari $2f_m$, dengan f_m adalah frekuensi paling tinggi yang muncul disebuah sinyal.

Quantization adalah proses memetakan nilai-nilai dari sinyal nilai kontinu menjadi nilai-nilai yang diskrit, sehingga didapatkan sinyal nilai diskrit.

Encoding adalah proses mengubah nilai-nilai sinyal ke menjadi bilangan biner.



Gambar 1. Diagram Sinyal Analog dan Sinyal Digital[5]

1.6 Transformasi Wavelet

Transformasi adalah proses merepresentasikan suatu sinyal ke dalam *domain* / kawasan lain. Tujuan dari transformasi adalah untuk lebih menonjolkan sifat atau karakteristik sinyal tersebut. Definisi wavelet (secara harfiah berarti “gelombang kecil”) adalah himpunan fungsi dalam ruang vektor L^2I , yang mempunyai sifat-sifat sebagai berikut [1]:

1. Berenergi terbatas

Academy of Finland. SOM menggunakan metode pembelajaran *unsupervised*, artinya di dalam melakukan pembelajaran tidak menggunakan data *output* sebagai target pembelajaran.

2. Merupakan fungsi *band-pass* pada domain frekuensi
3. Merupakan hasil penggeseran (*translasi*) dan penskala (*dilatasi*) dari sebuah fungsi tunggal (induk), yaitu:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \dots (2.1)$$

dengan $a, b \in R$ (bilangan nyata), dan $a \neq 0$. Dalam hal ini a adalah parameter penskala dan b adalah parameter penggeser posisi terhadap sumbu t . Faktor normalisasi $|a|^{-1/2}$ digunakan untuk memastikan bahwa $\|\psi_{a,b}(t)\| = \|\psi(t)\|$.

Pada dasarnya, Transformasi *Wavelet* merupakan sebuah teknik pemrosesan sinyal *multiresolusi*. Dengan sifat penskalaannya, wavelet dapat memilah-milah suatu sinyal data berdasarkan komponen frekuensi yang berbeda-beda. Dengan demikian tiap-tiap bagian dapat dipelajari berdasarkan skala resolusi yang sesuai, sehingga diperoleh gambaran data secara keseluruhan dan *detail*.

Keuntungan Transformasi *Wavelet* adalah bahwa jendelanya bervariasi. Untuk mengisolasi ketidakkontinuan sinyal, dapat digunakan fungsi basis yang sangat pendek. Pada saat yang sama, untuk analisis frekuensi secara terperinci, dapat digunakan fungsi basis yang sangat panjang.

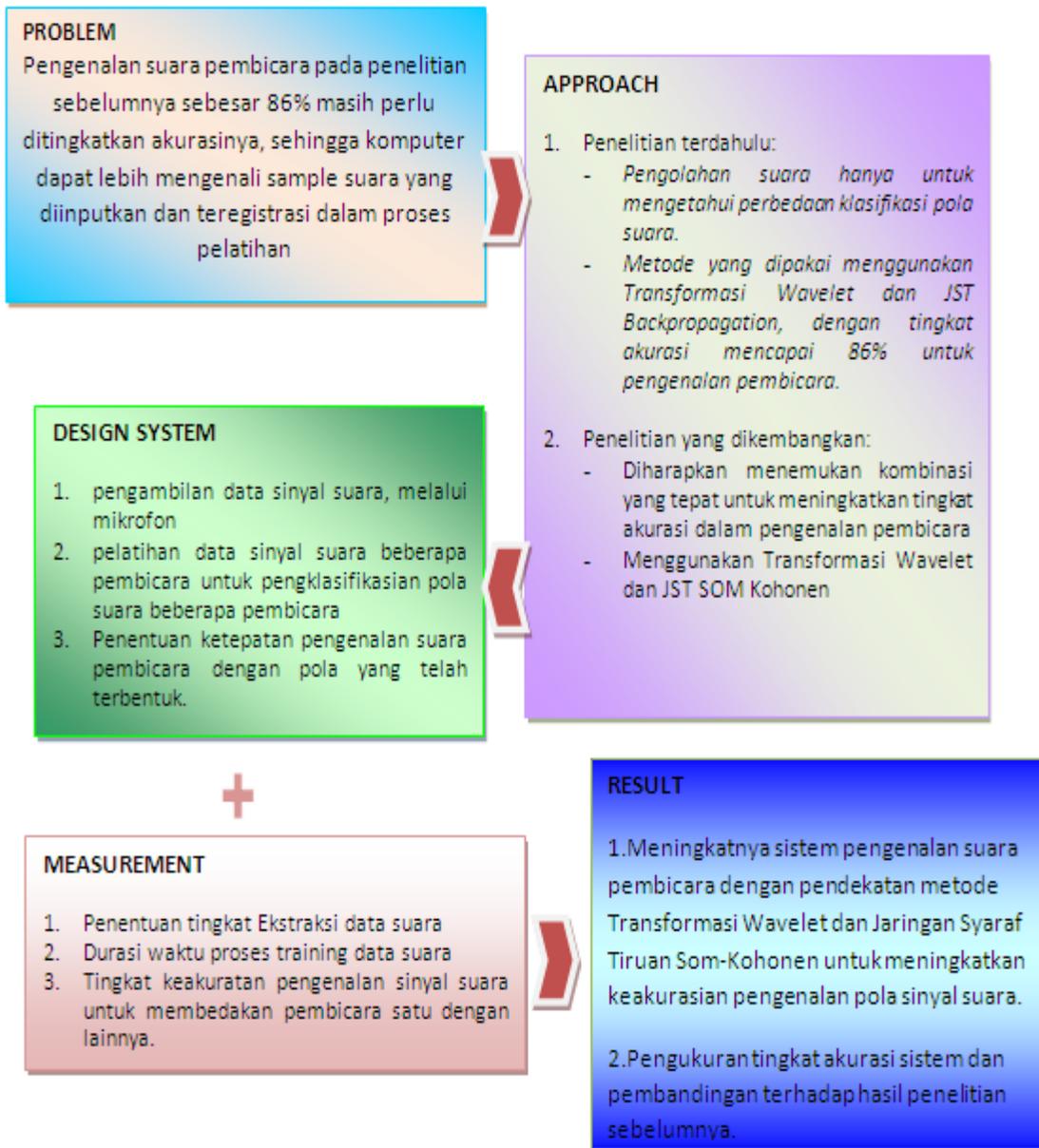
1.7 Jaringan Syaraf Tiruan SOM Kohonen

Jaringan *Self Organizing Maps* (SOM) dikembangkan pada tahun 1982 oleh Teuvo Kohonen, seorang profesor dari *The*

Sewaktu melakukan pembelajaran, unit kelompok yang vektor bobotnya memiliki jarak yang paling dekat dengan vektor masukan akan dipilih sebagai unit pemenang. Jarak tersebut biasanya ditentukan dengan menggunakan *Euclidian Distance*. Unit pemenang dan unit tetangganya lalu diperbaharui bobotnya. Unit tetangga ditentukan berdasarkan topologi yang digunakan untuk unit pengelompokan, misalnya: topologi linear, *rectangular*, *hexagonal*, *diamond*, dan lain-lainnya.

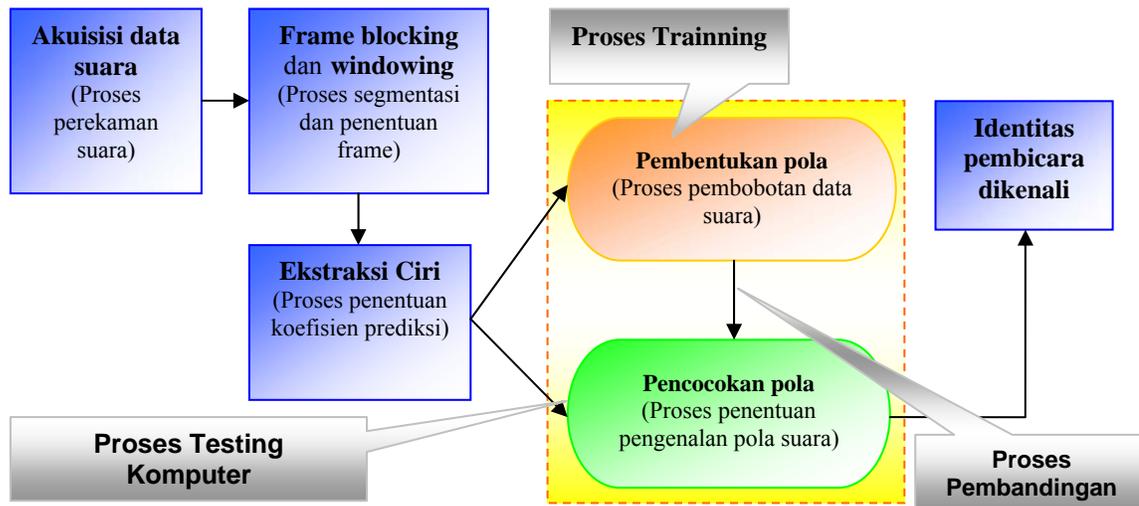
III. METODE PENELITIAN

3.1 Kerangka Pemikiran



Gambar 2. Kerangka Pemikiran

3.2 Prosedur Perancangan Sistem

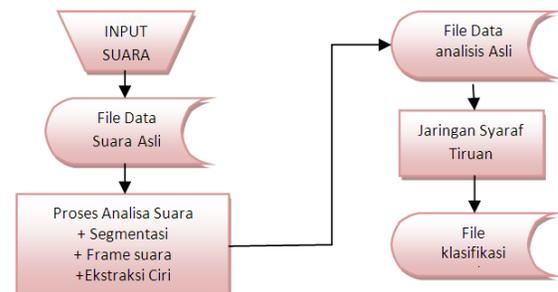


Gambar 3. Tahapan Proses Sistem Pengenalan Pembicara

3.3 Pengolahan Data Suara

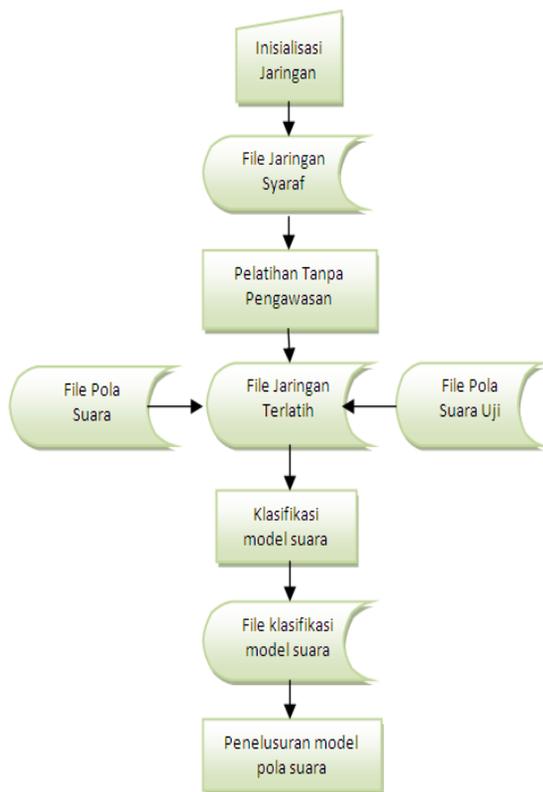
Pada sistem pengenalan pembicara ini, proses yang dirancang meliputi 4(empat) modul program yang dibentuk, yaitu terdiri dari:

1. Modul Akuisisi : berfungsi untuk pengambilan suara dan perekaman suara serta proses sampling suara
2. Modul Verifikasi : berfungsi untuk penentuan awal dan akhir kata dari penentuan daerah suara, pengkodean pola suara serta pembuatan file vector pola suara sebagai data masukkan pada jaringan syaraf tiruan. Modul ini berisi proses pembuatan frame, segmentasi dan ekstraksi file.
3. Modul Network : berfungsi untuk menciptakan file jaringan syaraf dan melakukan pelatihan jaringan syaraf dengan menggunakan SOM Kohonen.
4. Modul Klasifikasi : berfungsi untuk membentuk model pola suara untuk pengujian sistem pengenalan suara.



Gambar 4. Flowchart Proses Pemrosesan Sinyal Suara

Pada proses pelatihan dan pengujian, data-data inputan (berupa koefisien prediksi yang dihasilkan dari proses pengolahan data suara menggunakan transformasi wavelet) diolah pada jaringan syaraf dengan melakukan pelatihan terhadap bobot data pelatihan dengan inialisasi yang telah ditentukan berupa ukuran layar kompetitif dengan ukuran 10 x 10, sebagai tempat pemetaan kluster dan beberapa parameter yang diperlukan sebagai penentu pelatihan seperti *learning rate*, *threshold*, *jarak tetangga (DNeighbor)*, *nilai alpha* dan *banyak iterasi*.

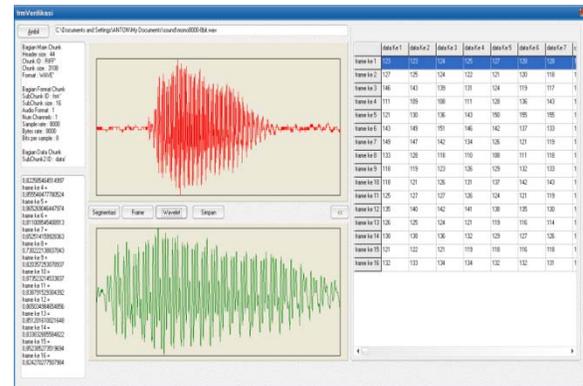


Gambar 5. Flowchart Pelatihan dan Pengujian Pengenalan Suara

File data suara yang dihasilkan tersebut merupakan himpunan nilai-nilai amplitudo dengan panjang data sesuai dengan lama perekaman. Nilai-nilai amplitudo tersebut berkisar antara 0 – 255, data ini dapat dilihat dari plot data grafik waktu diskrit. Dari panjang data suara tersebut hanya sebagian yang merupakan gelombang suara bergetar (*voice*) dan yang lainnya merupakan gelombang acak (*random*) berupa desis. Salah satu hasil segmentasi terhadap sebuah file data suara sebagai berikut:

Nama File:

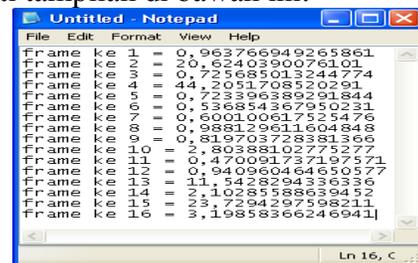
C:\WINDOWS\Media\chimes.wav dengan awal kata pada titik = *awal kata 8* dan akhir kata pada titik = *akhir kata 6958* dengan panjang data = *panjang data 6950*



Gambar 6. Gambar Proses Frame Sinyal Suara Yang Tersegmentasi

Pada bagian selanjutnya adalah proses penentuan frame dari data yang telah tersegmentasi. Pada bagian ini bertujuan untuk penentuan penyederhanaan data spectrum yang panjang untuk diringkas menjadi 16 data nilai yang akan digunakan sebagai data pelatihan.

Berikut ini adalah hasil perhitungan koefisien prediksi dari contoh data suara perekaman yang ada. Dari 16 frame yang ada, maka hasil perhitungan koefisien prediksi terdiri dari 16 angka prediksi yang dihitung berdasarkan algoritma *Wavelet*, seperti tampilan di bawah ini:

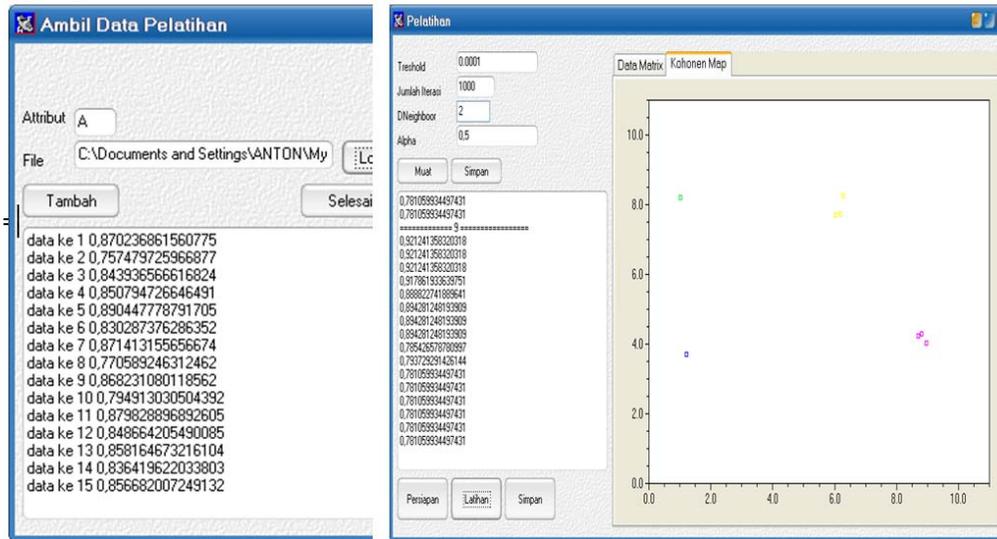


Gambar 7. Hasil perhitungan koefisien prediksi dari data frame

3.4 Proses Pelatihan dan Pengujian Pengenalan

Vektor masukan (X) untuk jaringan ditentukan dari hasil perhitungan *Wavelet*, bobot (w) untuk tiap koneksi titik ditetapkan

dengan nilai acak antara 0 sampai dengan 1. Layar keluaran jaringan ditentukan dengan ukuran 10x10 dengan nilai *learning rate* (α) ditetapkan antara 0.3 sampai dengan 0,5. Proses pelatihan dari jaringan akan berhenti apabila bobot jaringan konvergen dengan ditunjukkan oleh nilai *Training Error*. Penjelasan ini dapat dilihat pada gambar di bawah ini. Hasil keluaran dari sistem jaringan syaraf berupa peta (*mapping*) hasil pengelompokkan (*cluster*) pola suara-pola suara yang dilatihkan.



Gambar 8. Data Uji Suara dan Peta Pola Pelatihan dari masing-masing data suara pembicara

4. PENGUJIAN PERANGKAT LUNAK

Pengujian perangkat lunak merupakan salah satu proses rangkaian dari pengembangan perangkat lunak. Pengujian dilakukan dengan dua metode yaitu: dengan pengujian dengan metode *White box* dan metode *Black box*.

4.1 Pengujian *White Box*

Metode *white box* ini adalah suatu metode desain test case yang menggunakan struktur kontrol desain prosedural untuk memperoleh test case. Dengan menggunakan metode pengujian *white box*, perancangan sistem dapat melakukan test case yang dapat:

- a. *Semua modul telah digunakan minimal satu kali*
- b. *Menggunakan semua keputusan logis pada sisi true dan false*

c. *Mengeksekusi semua loop (perulangan) pada batasan mereka dan pada batas operasional mereka*

Sistem pengenalan suara pembicara yang dibangun ini, terdiri dari beberapa modul/prosedur/fungsi yang mempunyai fungsi proses masing-masing, sebagai sampel pengujian white box pada penelitian ini adalah modul Proses Segmentasi, karena modul ini dianggap paling rumit. Modul/prosedur proses segmentasi dibangun dengan menggunakan algoritma seperti dibawah ini, maka pengujian pada modul/prosedur ini dapat ditunjukkan dengan menelusuri logika pada algoritma ini untuk melihat efisien dan kompleksitas logika yang dibangun. Pengujian tersebut dapat dilihat pada gambar penelusuran jalur *independent* di bawah ini.

```

procedure Segmentasi(filename : string);
const treshl = 5;

var i,j,itel,ite2:integer;
    tawal,takhir:boolean;
    tmpstr :string;
    stream2:TFilestream;
    min,max,temp:integer;
tawal:=false;
takhir:=false;
min := wavedata[0].Data[0];
max := wavedata[0].Data[0];
for i := 0 to high(wavedata[0].Data) do >> 2
    if ((max - min) > treshl) then >> 3
        if (tawal = false) then >> 4a
            awal := i; >> 4b
            tawal := true;
        else >> 5
            akhir := i;
        if (wavedata[0].Data[i] > max) and (wavedata[0].Data[i] < min)
            then >> 6
                max := wavedata[0].Data[i];
                min := wavedata[0].Data[i];
            else >> 7
                if (wavedata[0].Data[i] > max) and (wavedata[0].Data[i] > min)
                    then >> 8
                        max := wavedata[0].Data[i]
                    else >> 9
                        min := wavedata[0].Data[i];
                end if; >> 10
            end if; >> 11
        end if; >> 12
    end For; >> 13
temp:=awal;

min := wavedata[0].Data[high(wavedata[0].Data)];
max := wavedata[0].Data[high(wavedata[0].Data)];

for i := high(wavedata[0].Data) downto 0 do >> 15
    if ((max - min) > treshl) then >> 16
        if (takhir = false) then >> 17
            akhir := i; >> 18
            takhir := true;
        else >> 19
            awal := i;
        if (wavedata[0].Data[i] > max) and (wavedata[0].Data[i] < min)
            then >> 20
                max := wavedata[0].Data[i];
                min := wavedata[0].Data[i];
            else >> 21
                if (wavedata[0].Data[i] > max) and (wavedata[0].Data[i] > min)
                    then >> 22
                        max := wavedata[0].Data[i]
                    else >> 23
                        min := wavedata[0].Data[i];
                end if; >> 24
            end if; >> 25
        end if; >> 26
    end for; >> 27

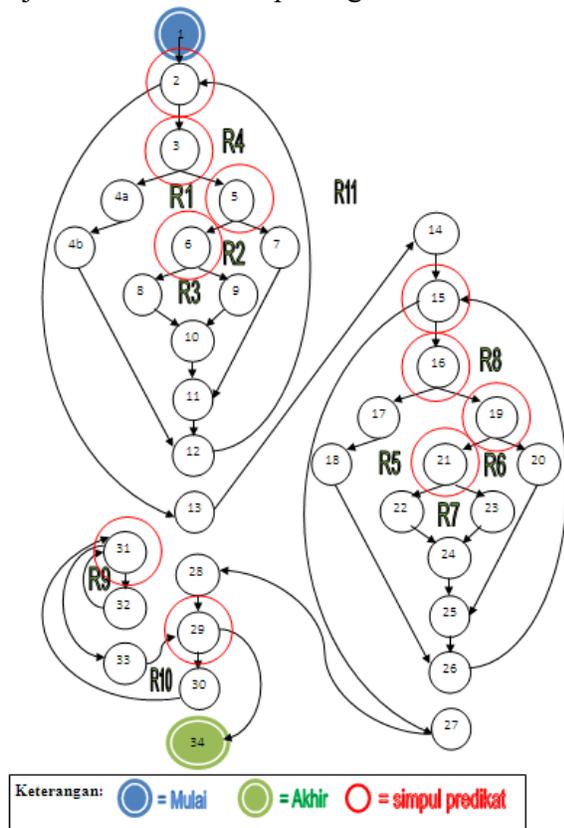
awal := temp; >> 28

for j := awal to akhir do >> 29
    buf[j] := wavedata[0].data[j]; >> 30
    frmVerifikasi.series2.Clear;
    for i := awal to akhir do >> 31
        frmVerifikasi.Series2.AddXY(i/36000, round(buf[i]), '', >> 32
            clGreen);
    end for; >> 33
end for; >> 34

```



Dari listing prosedur di atas, dapat dijabarkan dalam simpul logika berikut:



Gambar 9. Simpul Logika Aplikasi

Dari jalur logika di atas maka dapat ditentukan beberapa jalur logica/ jalur independent yang membangun prosedur segmentasi di atas. Oleh karena itu maka untuk menguji apakah benar jalur logika yang dibentuk, diperlukan pencarian nilai kompleksitas siklomatisnya, seperti di bawah ini:

Berdasarkan flow graph diatas, maka kompleksitas cyclomatic-nya dapat di hitung sebagai berikut :

1. Flow graph diatas mempunyai 11 region
2. $V(G) = 44 \text{ edges} - 35 \text{ nodes} + 2 = 11$
3. $V(G) = 10 \text{ predicates nodes} + 1 = 11$

Hasil kompleksitas siklomatis mendapatkan besaran nilai 11, hal ini menunjukkan prosedur segmentasi dibangun dengan 11 logika yang merupakan jalur proses sebagai unit proses untuk aktifitas segmentasi

spectrum suara dari data suara yang ada hasil perekaman. Nilai kompleksitas siklomatis lebih dari (>10) menunjukkan bahwa pada prosedur tersebut lebih banyak proses yang diperlukan untuk aktifitas yang diperlukan sesuai fungsi yang ditetapkan.

Independent path adalah alur manapun dalam program yang memperkenalkan sedikitnya satu kumpulan perintah pemrosesan atau kondisi baru. Adapun independent path dari gambar flow graph diatas adalah sebagai berikut:

- Path 1 : 1-2-13-14-15-27-28-29-34
- Path 2 : 1-2-3-4a-4b-12-2-13-14-15-16-17-18-26-15-27-28-29-30-31-33-29-34
- Path 3 : 1-2-3-5-6-8-10-11-12-2-13-14-15-16-19-21-22-24-25-26-15-27-28-29-30-31-33-29-34
- Path 4 : 1-2-3-5-6-9-10-11-12-2-13-14-15-16-19-21-23-24-25-26-15-27-28-29-30-31-33-29-34
- Path 5 : 1-2-3-5-7-11-12-2-13-14-15-16-19-20-25-26-15-27-28-29-30-31-33-29-34
- Path 6 : 1-2-3-4a-4b-12-2-3-5-7-11-12-2-13
- Path 7 : 1-2-3-5-6-8-10-11-12-2-3-5-6-9-10-11-12-2-13
- Path 8 : 1-2-13-14-15-16-17-18-26-15-16-19-20-25-26-15-27
- Path 9 : 1-2-13-14-15-16-19-21-22-24-25-26-15-16-19-21-23-24-25-26-15-27
- Path 10 : 1-2-13-14-15-27-28-29-30-31-33-29-34
- Path 11 : 1-2-13-14-15-27-28-29-30-31-32-31-33-29-34

Tabel 1. Pengujian

Path	Input	Output	Hasil
Path 1	Data suara kosong	Tidak terjadi proses segment	Sesuai
Path 2	Data suara ke-i lebih besar dari threshold	Terdeteksi spectrum bergetar	Sesuai
Path 3	Data suara ke-i lebih besar dari max dan lebih kecil dari min	Max=data ke-i dan min data ke-i	Sesuai
Path 4	Data suara ke-i	Max=data ke-i dan min	Sesuai

	lebih besar dari max dan lebih besar min	data sebelumnya	
Path 5	Data suara ke-i	Proses penentuan awal dan akhir suara bergetar	Sesuai
Path 6	Data suara ke-i	Proses penentuan awal suara bergetar	Sesuai
Path 7	Data suara ke-i	Proses penentuan akhir suara bergetar	Sesuai
Path 8	Data suara ke-i	Pembentukan spectrum awal suara bergetar	Sesuai
Path 9	Data suara ke-i	Pembentukan spectrum akhir suara bergetar	Sesuai
Path 10	Data suara ke-i	Proses pembentukan spectrum suara bergetar	sesuai
Path 11	Data suara ke-i	Proses penentuan nilai spectrum suara bergetar	Sesuai

4.3.2 Pengujian Black Box

Metode *black box* adalah pengujian *user interface* dimana setelah aplikasi sistem identifikasi pembicara diberikan ke pengguna untuk dapat dioperasikan dengan baik. Metode pengujian ini akan diterapkan dengan menggunakan tabel referensi masukan keluaran untuk menguji perilaku sistem saat diberikan masukan tertentu,

apabila keluaran yang dihasilkan sesuai yang diharapkan maka dapat dikatakan bahwa sistem lolos dari pengujian *black box*.

Dalam sistem ini pengujian dilakukan dengan memberikan data-data contoh dan dibandingkan dengan informasi yang dihasilkan. Modul yang akan diuji secara *Black Box* adalah: Pelatihan data suara dan Hasil pengenalan pembicara.

Tabel 2. Hasil Uji Black Box Sistem Identifikasi Pembicara

Input (Event)	Output	Fungsi	Hasil Uji (Sesuai atau tidak sesuai)
Input pengguna dan password yang telah terdaftar	Form Login Identif tampil	Menampilkan form login identifikasi pembicara	Sesuai
Input pengguna dan password yang belum terdaftar	Form pesan kesalahan pemakai sistem tampil	Menampilkan pesan kesalahan pemakai sistem	Sesuai
Klik menu Akuisisi	Form Aplikasi Perekaman Suara Manusia dengan Pengaturan Parameter cuplik suara tampil	Menampilkan aplikasi perekaman suara manusia dengan pengaturan parameter cuplik suara	Sesuai
Klik menu verifikasi segmentasi	Form Aplikasi untuk melakukan pilihan segmentasi, dan simpan tampil	Menampilkan aplikasi untuk melakukan segmentasi dan hasilnya	Sesuai
Klik menu verifikasi frame	Form aplikasi untuk melakukan proses frame dan hasilnya	Menampilkan aplikasi untuk melakukan proses frame dan hasilnya	Sesuai
Klik menu verifikasi perhitungan wavelet	Form aplikasi untuk melakukan perhitungan wavelet dan hasilnya	Menampilkan aplikasi untuk melakukan perhitungan wavelet dan hasilnya	Sesuai
Klik menu simpan	Form aplikasi untuk melakukan penyimpanan data perhitungan wavelet	Menampilkan aplikasi untuk melakukan penyimpanan data perhitungan wavelet	Sesuai

Klik menu jaringan pelatihan	Form aplikasi untuk melakukan pelatihan data suara dari nilai koefisien prediksi dan tampilan pemetaan klasifikasi pola suara	Menampilkan aplikasi untuk melakukan pelatihan data suara dari nilai koefisien prediksi dan tampilan pemetaan klasifikasi pola suara	Sesuai
Klik menu jaringan pengujian	Form Aplikasi untuk memasukkan data suara yang mau di uji dan menghasilkan identifikasi pembicara bila suara uji pernah dilatih	Memasukkan data suara yang mau di uji dan menghasilkan identifikasi pembicara bila suara uji pernah dilatih	Sesuai
Data bobot hasil pelatihan	Pelatihan klasifikasi pola suara	Hasil perhitungan nilai bobot pemenang dari proses pelatihan	Sesuai
Data suara uji	Pembandingan beberapa pola suara dengan nilai bobot pemenang proses pelatihan	Hasil pengenalan pembicara yang mempunyai data suara yang telah mengalami pelatihan	Sesuai

yang telah ditentukan dan menghasilkan 30 (tiga puluh) data prediksi yang digunakan untuk input jaringan. Untuk selanjutnya diidentifikasi menjadi 30 data suara pembicara. Contoh salah satu hasil perhitungan koefisien prediksi menggunakan transformasi wavelet dari 30 (tiga puluh) orang pembicara dapat dilihat pada data di bawah ini.

```
[general]
namafile=agung1.wav
[Detail]
frame1=14,4528128629258
frame2=13,1898163926032
frame3=34,6185596865345
frame4=0,970489474221717
frame5=1,09861581373799
frame6=0,961142055473891
frame7=1,00548502823213
frame8=0,944394099234442
frame9=14,5642174827821
frame10=0,939924609174206
frame11=0,981128607101161
frame12=0,951677017022268
frame13=0,92707939435886
frame14=0,857421980621375
frame15=0,733837779599622
frame16=0,808417203913459
```

Gambar 10. Contoh data suara yang telah di verifikasi

5. HASIL DAN PEMBAHASAN

Penelitian tesis ini diperlukan sebanyak 30 pembicara dengan perlakuan tiap-tiap pembicara merekam sinyal suara sebanyak masing-masing 5 kali perekaman dengan kata yang sama.

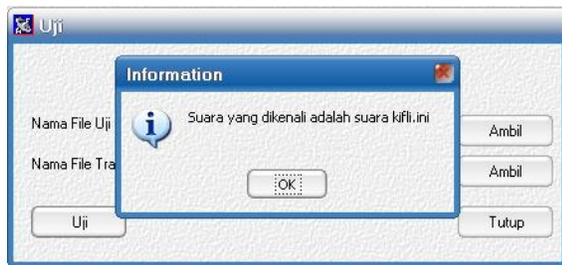
Pengujian yang telah dilakukan pada tahap praproses dengan menetapkan proses ekstraksi ciri dengan wavelet tipe orthogonal yaitu Symlet, dengan ukuran vector (1 x 16) yang merupakan penjabaran dari proses frame. Jumlah pembicara yang diambil sebagai data training adalah 30 (tiga puluh) data suara hasil perekaman (data asli), kemudian dilakukan perhitungan koefisien prediksi menggunakan transformasi wavelet dengan ukuran vector

Proses belajar jaringan syaraf tiruan *SOM-Kohonen* yang ditentukan oleh parameter-parameter jaringan (*vector* masukkan, layar kompetisi, nilai α /*learning rate*, dan banyaknya iterasi). Hasil penelitian jaringan syaraf tersebut menunjukkan prosentase keberhasilan dan kegagalan dari proses pelatihan beberapa file suara sebagai data pelatihan, dalam mengenali dan mengklasifikasikan pola-pola suara dari beberapa suara yang diuji. Dari hasil pengujian dengan nilai parameter yang terdiri dari *Threshold learning rate*, *iterasi*, *Dneighbor*, nilai α . Dimana kombinasi pemakaian data input sangat mempengaruhi hasil pengenalan. Dari penentuan nilai $threshold = 0,0001$, *Iterasi* = 10000, *Dneighbor* = 3, dan nilai $\alpha(\alpha) =$

0,3, menghasilkan akurasi pengenalan yang lebih baik, dibandingkan dengan kombinasi parameter yang lain. Dari hasil uji coba, maka didapat hasil dengan inputan 30 (tiga puluh) orang pembicara menghasilkan tingkat akurasi 90,67% tingkat keberhasilan. Namun dengan data input 5 (lima) orang pembicara dihasilkan tingkat akurasi sebesar 96% tingkat keberhasilan.

	1						
	/						
	5						
	=						
	2						
Jml	0						1/25=
TK/%	%	-	-	-	-	-	4%

Threshold : 0,0001
 Iterasi : 10000
 Dneighbor : 3
 Alpha : 0.3



Gambar 11. Informasi hasil pengujian pengenalan suara

Tabel uji coba pengenalan suara menggunakan 5 orang dan 30 orang pembicara .

Threshold : 0,0001
 Iterasi : 10000
 Dneighbor : 3
 Alpha : 0.3

Nama	su	suara	suara	suara	suara	Dikenali	Tidak Dikenali
	ara						
	1	2	3	4	5		
Akbar	K	K	K	K	K	5	0
Agung	K	K	K	K	K	4	1
Azzam	K	K	K	K	K	5	0
Sita	K	K	K	K	K	5	0
Made	K	K	K	K	K	5	0
Jml	4						
K/%	/						
	5						
	=						
	8						
	0	5/5=	5/5=	5/5=	5/5=		
	%	100%	100%	100%	100%	24/25=96%	-

Oran	suara	suara	suara	suara	suara	Dik ena	Tidak
	1	2	3	4	5		
g						li	Dikenali
1	TK	K	K	K	K	4	1
2	K	K	K	K	K	5	0
3	K	K	K	K	K	5	0
4	K	K	K	K	K	5	0
5	K	K	K	K	K	5	0
6	TK	K	K	K	TK	3	2
7	K	K	K	K	K	5	0
8	K	K	K	TK	K	4	1
9	K	K	K	K	K	5	0
10	K	K	K	K	K	5	0
11	K	K	TK	TK	K	3	2
12	K	K	K	K	K	5	0
13	TK	K	K	K	K	4	1
14	K	TK	K	K	K	4	1
15	K	K	K	K	K	5	0
16	K	K	K	K	K	5	0
17	K	TK	K	K	K	4	1
18	TK	TK	K	K	K	3	2
19	K	TK	TK	K	K	3	2
20	K	K	K	K	K	5	0
21	K	K	K	K	K	5	0
22	K	K	K	K	K	5	0
23	K	K	K	K	K	5	0
24	K	K	K	K	K	5	0
25	K	K	K	K	K	5	0
26	TK	K	K	K	K	4	1
27	K	K	K	K	K	5	0
28	K	K	K	K	K	5	0
29	K	K	K	K	K	5	0

30	K	K	K	K	K	5	0
Jm						136	
I		26=8	28=9	28=9	29=9	=90	
K/	25=83,33	6,67	3,33	3,33	6,66	,67	
%	%	%	%	%	%	%	-
Jm							
I							
TK	5=16,67	4=19,	2=6,6	2=6,6	1=3,3		14=9,33
/%	%	33%	7%	7%	3%	-	%

5. KESIMPULAN DAN SARAN

5.1. Kesimpulan

Dari hasil penelitian ini, dapat ditarik kesimpulan bahwa untuk membentuk suatu sistem pengenalan identitas pembicara berdasarkan suara dibutuhkan suatu sistem “*pintar*” yang mempunyai kemampuan proses yang tinggi, hal ini diperlukan suatu metode pengolahan data suara dan metode pengenalan pola suara yang tepat.

Transformasi Wavelet Diskret dapat digunakan dalam proses identifikasi pembicara pada bagian pemrosesan awal (praproses) sinyal untuk mendapatkan informasi (ciri) sinyal tersebut. Transformasi Wavelet sebagai pengolah data suara digital dan Jaringan Syaraf Tiruan SOM kohonen dapat memberikan hasil pengenalan yang baik dengan tingkat akurasi mencapai 96%. Namun hal ini juga dipengaruhi oleh banyaknya data input suara pada proses pengenalan. Pada data suara berjumlah 30 (tiga puluh) orang mencapai tingkat akurasi 90,67%, dan dengan inputan sebesar 5 (lima) orang pembicara tingkat akurasi meningkat mencapai 96% tingkat kebenarannya. Hal ini disebabkan pada jaringan SOM Kohonen hanya mempunyai 1 (satu) layer sebagai input dan output untuk pengklasifikasian pola suara, menyebabkan terjadi tumpukan pola klasifikasi dari setiap pembicara bila jumlah inputan lebih banyak, sehingga adanya identifikasi yang bias pada pada proses pengenalan (pembicara satu bias masuk pada identifikasi pembicara yang lain).

5.2. Saran-saran

Saran yang bisa diberikan sebagai bahan pertimbangan setelah dilakukan penelitian dengan algoritma wavelet dan sistem pengenalan menggunakan jaringan syaraf tiruan SOM Kohonen adalah untuk meningkatkan proses pengenalan identitas pembicara diperlukan analisa yang teliti dalam menentukan metode yang akan dipakai. Perlu ada pengkajian lebih lanjut mengenai Transformasi Wavelet, untuk mendapatkan tingkat pengenalan yang lebih tinggi.

Dalam pembuatan sistem pengenalan suara, proses analisa dan pengolahan data diharapkan menjadi focus utama karena pada bagian ini merupakan proses untuk menentukan nilai prediksi pengolahan suara yang akan digunakan dalam proses selanjutnya. Untuk itu perlu ditentukan metode yang tepat untuk proses analisa nilai prediksi, yang dapat menghasilkan nilai yang lebih sederhana tanpa mengurangi nilai yang terdapat pada data suara tersebut. Pemilihan metode pengenalan suara dari jaringan syaraf tiruan, harus dianalisa secara sistematis beberapa keunggulan yang dimiliki oleh setiap metode yang terdapat dalam jaringan syaraf tiruan. Sehingga nantinya dapat menunjukkan performa dari proses pengenalan secara tepat.

DAFTAR PUSTAKA

- [1] Agustini, Ketut. 2006. *Speaker Identification with Discret Wavelet Transformation of Pre-processing and Neural Network*.
- [2] Leon, C.G.K, 2009, Robust Computer Voice Recognition Using Improved MFCC Algorithm, Proc. IEEE
- [3] Raymond S. Wagner, *An Architecture for Distributed Wavelet Analysis and Processing in Sensor Networks*, IPSN’06, April 19–21, 2006, Nashville, Tennessee, USA. Copyright 2006 ACM 1595933344/06/0004

- [4] Pekka-Henrik Niemenlehto, Martti Juhola, And Veikko Surakka, (2006), *Detection of Electromyographic Signals from Facial Muscles with Neural Networks*, ACM Transactions on Applied Perception, Vol. 3, No. 1, January 2006, Pages 48–61
- [5] Agustini, Ketut, 2007, *Biometrik Suara dengan Transformasi Wavelet Berbasis orthogonal Daubenchies*, 50 Gematek Jurnal Teknik Komputer, Volume 9 Nomor 2.
- [6] Khairulvani, Feni, 2007, *Identifikasi Individu Melalui Suara Ucapan dengan Ekstraksi Ciri Mel-Frequency Cepstral Coefficient (MFCC) sebagai Input Jaringan Syaraf Tiruan*, Central Library ITB.
- [7] Campbell, J.P.,JR, 1997, *Speaker Recognition: A Tutorial*. Proc. IEEE, vol. 85, no 9,pp1437-1462, 1997
- [8] Krishnan, M., Neophytou C.P., Prescott G.,(1994),*Wavelet Transform Speech Recognition Using Vector Quantization, Dynamic Time Warping and Artificial Neural Network*, Center of excellence in computer aided sistem engineering and Telecommunication & Information Science Laboratory 2291 Irving Hill Drive,Lawrence, KS 66045.
- [9] Candra, Heru K.,(2000), *Voice Pattern Of Recognition Sistem With The SOM Kohonen,digilib-ITS*, Theses Informatics Engineering RT 006.3 Cans,
- [10] Lawrence B. Holder,*Speech Recognition (Briefly)*,
<http://www.cs.berkeley.edu/~russell/classes/cs188/s05/slides/chapter15b.pdf>, 16 Mei 2010
- [11] Burrus, C.S, Gopinath R.A., Guo, H. (1998), *Introduction to Wavelet and Wavelet Transform A Primer*, International Edition, Prentice-Hall International, Inc.
- [12] Dayhoff, Judith E., (1990),*Neural Network Architectures: An Introduction, Van Nostrand Reinhold*, New York.
- [13] Embree, P.M., Kimble, B., (1991),*C Language Algoritms For Digital Signal Processing*,Prentice Hall International, Inc.
- [14] Fauset, L. (1994),*Fundamentals of Neural Network*,Prentice Hall, Eaglewood Cliffs, NJ.
- [15] Kosko, Bart, (1992), *Neural Network For Signal Processing*,Prentice-Hall International,Inc
- [16] Parson, Thomas, W.,(1986),*Voice and Speech Processing*,McGraw-Hill,USA
- [17] Richard, H. David, M.G., Ivan, L.,(1993),*Sound Blaster: The Official Book.*, McGraw.
- [18] Wael Al-Sawalmeh, Khaled Daqrouq, Omar Daoud, Abdel-Rahman Al-Qawasni,2010, *Speaker Identification Sistem-based Mel Frequency and Wavelet Transform using Neural Network Classifier*, European Journal of Scientific Research.
- [19] Firoz, Shah, A, Raji, Sukumar, A and Babu, Anto.P , 2010 , *Discrete Wavelet Transforms and Artificial Neural Networks for Speech Emotion Recognition*, International Journal of Computer Theory and Engineering, Vol 2, No.3 June
- [20] L.Salhi, M., Talbi, and A.Charif, 2008, *Voice Disorders Identification Using Hybrid Approach : Wavelet Analysis and Multilayer Neural Networks*, World Academy of Science, Engineering and Technology.

Penulis:

Dra. Ec. Hj. Rustati Rahmi, M.Kom.
Dosen Tetap Yayasan Mandiri
pada STMIK BANJARBARU