

PENERAPAN METODE ALGORITMA GENETIK DALAM PENGENALAN POLA SINYAL SUARA MANUSIA

Muhammad Yassir
Pemda Kabupaten Pangkajene dan Kepulauan
yassirkoe@yahoo.co.id

Abstrak

Sinyal suara yang masuk diekstraksi dengan metode LPC lalu dikodekan menjadi 24 string biner (1 dan 0) yang kemudian menjadi data masukan pada Algoritma Genetik (AG). Metode AG bekerja berdasarkan proses pencocokan pola sinyal suara masukan dengan pola sinyal suara yang telah tersimpan dalam data base. Pengambilan keputusan AG bergantung pada pencarian nilai fitness (kecocokan) terbaik dengan cara memanfaatkan teknik randomisasi dengan melibatkan penyalinan string dan pertukaran string. Pada penelitian ini digunakan suara dari 5 orang target yang dimasukkan dalam 3 data base yang berbeda. Sistem pertama menggunakan masing-masing 10 data suara tiap orang, sistem kedua menggunakan masing-masing 7 data suara tiap orang, dan sistem ketiga menggunakan masing-masing 5 data suara tiap orang. Hasil penelitian menunjukkan pengujian sinyal suara yang telah tersimpan dalam data base adalah 100% dikenali, dan pengujian untuk sinyal suara baru dari kelima target adalah 72%.

Kata kunci : *sinyal suara, pengenalan pola, Algoritma Genetik (AG)*

I. PENDAHULUAN

Suara merupakan modal besar yang dimiliki oleh manusia untuk dapat mengkomunikasikan maksud dan tujuan dalam kehidupannya. Berbagai hal dapat dilakukan dengan hanya mengeluarkan suara, misalnya memberikan perintah kepada sesamanya. Dengan mudahnya manusia dapat mengerti apa yang diucapkan oleh manusia yang lain tanpa mengalami kesulitan sedikitpun. Beberapa tahun terakhir perkembangan dunia teknologi informatika dan komunikasi telah banyak mengalami kemajuan yang sangat pesat, salah satunya adalah penelitian di bidang suara dimana sinyal

informasi yang dikirimkan tidak hanya berupa data teks tetapi juga suara manusia.

Dalam aplikasinya suara pengucap digunakan untuk mengidentifikasi suatu kondisi serta kontrol akses, seperti *voice dial, phone banking, phone shopping, data base access service, voice mail, security control, dan remote access* terhadap komputer. Namun berbagai persoalan akan muncul saat perintah suara diberikan kepada mesin. Hal ini dikarenakan perbedaan karakter suara dari tiap individu[1].

Suara manusia merupakan sinyal yang kompleks, sehingga diperlukan suatu metode analisis yang dapat mengekstraksi parameter atau ciri khas suara individu

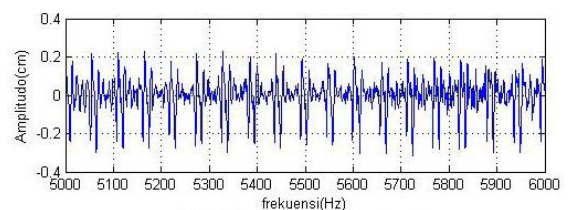
menjadi lebih sederhana tetapi tidak menghilangkan karakter khasnya. Metode yang digunakan dalam penelitian ini adalah Linear Predictive Coding (LPC) yang sangat efisien untuk pengolahan sinyal suara serta untuk pengenalan polanya digunakan metode Algoritma Genetik yang dapat dipakai untuk mendapatkan solusi yang tepat untuk masalah optimal dari satu variabel atau multi variable. Algoritma ini biasa dipakai pada aplikasi bisnis, teknik maupun pada bidang keilmuan. Algoritma Genetik pada dasarnya adalah program komputer yang mensimulasikan proses evolusi. Dalam hal ini kumpulan string yang diistilahkan sebagai populasi dari kromosom dihasilkan secara random dan memungkinkan untuk berkembang biak sesuai dengan hukum-hukum evolusi dengan harapan akan menghasilkan individu kromosom yang prima. Kromosom ini pada kenyataannya adalah kandidat penyelesaian dari masalah, sehingga bila kromosom yang baik berkembang, solusi yang baik terhadap masalah diharapkan akan dihasilkan[2]. Penelitian ini diharapkan dapat digunakan untuk melakukan pengenalan terhadap pola suara seseorang melalui kata-kata yang diucapkan orang tersebut. Sistem ini antara lain bermanfaat untuk melakukan identifikasi semacam aplikasi absensi, kontrol akses untuk fasilitas tertentu, *remote* akses untuk jaringan komputer, *forensik*, dan lain-lain, serta untuk pengembangan ilmu pengetahuan khususnya dalam pengolahan sinyal suara.

II. TINJAUAN PUSTAKA

1. Sinyal Suara Manusia

Suara merupakan salah satu media komunikasi yang paling sering dan umum digunakan oleh manusia. Manusia dapat

memproduksi suaranya dengan mudah tanpa memerlukan energi yang besar. Selain itu, setiap manusia mempunyai bentuk suara yang berbeda-beda sehingga dapat dibedakan dengan mudah sesuai dengan karakter suaranya masing-masing. Hal ini semakin memicu kebutuhan akan adanya identifikasi orang (*person identification*). Suara manusia merupakan salah satu bentuk biometrik yang dapat digunakan untuk identifikasi orang (*person identification*) selain itu, pengenalan suara (*voice recognition*) cukup mudah untuk digunakan atau dilakukan oleh manusia. Perangkat lunak pengenalan suara (*voice recognition*) adalah suatu aplikasi yang memungkinkan manusia untuk menggunakan teknologi, khususnya komputer dan tanpa perlu berhubungan secara langsung. Dengan adanya perangkat lunak pengenalan suara (*voice recognition*) manusia dimudahkan untuk berinteraksi dengan komputer, manusia cukup memberikan perintah-perintah secara lisan/suara. Bentuk gelombang sinyal suara mempunyai bentuk yang unik seperti pada gambar berikut.

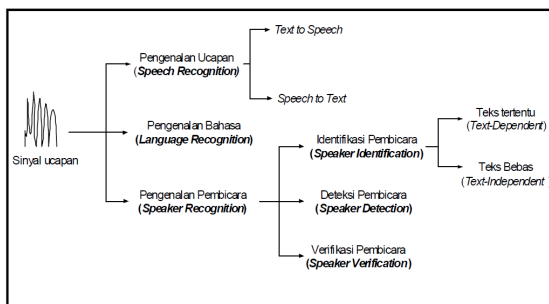


Gambar I.1 *Contoh Sinyal suara manusia untuk bunyi vokal "a"*

2 Proses Pengenalan Pembicara

Pengenalan pembicara dapat diklasifikasikan ke dalam tiga tahap yaitu identifikasi, deteksi dan verifikasi. Identifikasi pembicara merupakan proses untuk menentukan identitas pembicara melalui suara yang telah diucapkan,

sedangkan deteksi pembicara merupakan proses penemuan suara pembicara dari sekumpulan suara, dan verifikasi pembicara merupakan proses untuk memverifikasi kesesuaian suara pembicara dengan identitas yang diklaim oleh pembicara. Pengenalan pembicara lebih menitik beratkan pada pengenalan suara pembicara dan tidak pada pengenalan ucapan pembicara[4].



Gambar I.2 Klasifikasi Sistem Pengolahan Sinyal Suara.

Metode identifikasi pembicara yang merupakan bagian dari pengenalan pembicara (Gambar I.2), dapat dibagi ke dalam metode *text-independent* dan *text-dependent*. Pada sistem *text-independent*, model pembicara mengambil karakteristik ucapan seseorang melalui sinyal ucapan dengan mengabaikan apa yang diucapkannya, dalam artian kata-kata yang diucapkan sembarang (bebas). Sebaliknya pada sistem *text-dependent*, pengenalan identitas pembicaranya didasarkan pada ucapan seseorang dengan kata-kata yang spesifik atau telah disepakati, seperti *password*, *card numbers*, kode PIN dan sebagainya[4].

3 Prinsip Dasar Proses Pengenalan Suara

Semua sistem identifikasi pembicara melalui dua proses penting yaitu *feature extraction* dan *feature matching*. *Feature*

extraction yang merupakan proses mengekstraksi data hasil akuisisi sehingga dihasilkan data yang berdimensi lebih kecil, yang nantinya digunakan untuk merepresentasikan tiap-tiap pembicara. *Feature matching* menyangkut prosedur yang mengidentifikasi pembicara yang tidak dikenal dan membandingkan fitur ekstraksi suara yang dimasukkan dengan salah satu dari himpunan pembicara yang telah dikenal[4].

4 Autokorelasi Linear Predictive Coding LPC

Linear Predictive Coding (LPC) merupakan salah satu teknik analisis sinyal percakapan yang menyediakan ekstraksi fitur yang berkualitas baik dan efisien untuk digunakan dalam perhitungan. LPC pertama kali digunakan pada tahun 1978 untuk membuat alat sintesis sinyal percakapan. LPC melakukan analisis dengan cara memperkirakan formant, memisahkan formant dari sinyal, yang dinamakan proses inverse filtering, lalu mengestimasi intensitas dan frekuensi dari sinyal percakapan yang tersisa, yang disebut *residue*. Karena sinyal percakapan bervariasi seiring waktu, estimasi tersebut dilakukan untuk setiap potongan kecil dari sinyal, yang dinamakan *frame*[4].

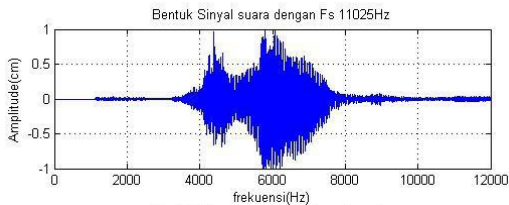
5 Algoritma Genetik (AG)

Algoritma genetik merupakan evolusi/perkembangan dunia komputer dalam bidang kecerdasan buatan (*artificial intelligence*). AG adalah algoritma yang diinspirasi teori evolusi yang kemudian diadopsi menjadi algoritma komputasi untuk mencari solusi suatu permasalahan. Salah satu aplikasi AG adalah pada permasalahan optimasi kombinasi, yaitu mendapatkan suatu nilai solusi optimal

terhadap suatu permasalahan yang mempunyai banyak kemungkinan solusi[3].

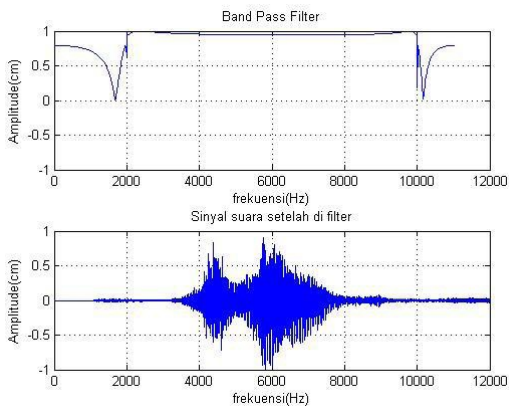
III. HASIL DAN PEMBAHASAN

Jumlah sinyal suara yang ditetapkan sebagai target adalah 5 orang responden yang masing-masing dibagi atas 10 masukan, 7 masukan, dan 5 masukan sinyal suara tiap orang yang disimpan dalam format WAV kemudian diekstraksi ciri dengan menggunakan LPC, untuk selanjutnya diidentifikasi dengan menggunakan metode Algoritma Genetik pada Matlab. Suara yang direkam adalah hasil dari pengucapan kata “HALO” dengan pola masukan yang sama. Gambar II.1 di bawah ini memperlihatkan gambar *sample* sinyal suara masukan yang telah direkam.



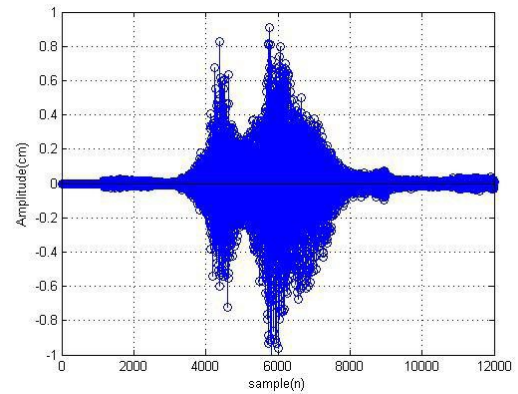
Gambar II.1 Sinyal Suara Yas1

Hasil *Filtering*:



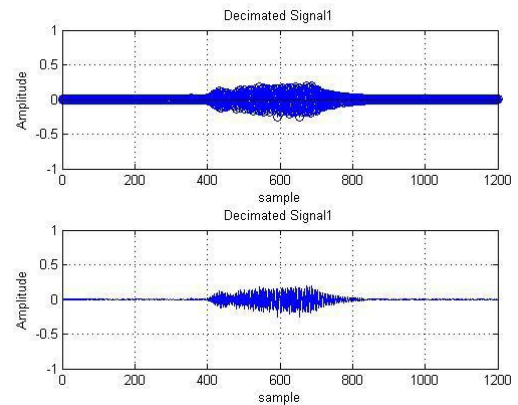
Gambar II.2 Proses Filter Suara Yas1

Hasil *Sampling*:



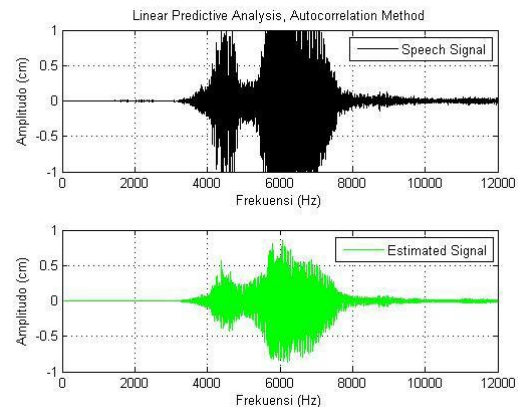
Gambar II.3 Sampling Suara Yas1

Hasil Desimasi:



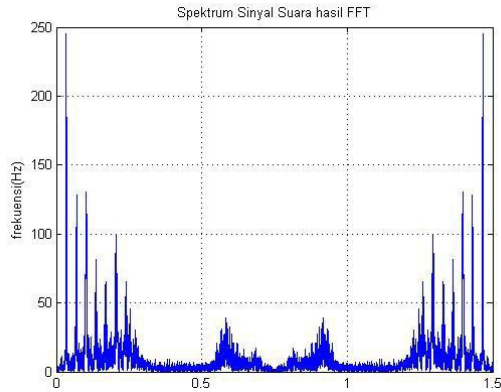
Gambar II.4 Desimasi Suara Yas1

Setelah selesai dinormalisasi, maka untuk mendapatkan vektor cirinya digunakan Linear Predictive Coding (LPC) dengan cara memperkirakan formant, memisahkan formant dari sinyal, lalu mengestimasi intensitas dan frekuensi dari sinyal percakapan yang tersisa.



Gambar II.5 Hasil LPC Suara Yas1

Setelah melewati tahap LPC maka diteruskan ke tahap FFT yaitu metode untuk mentransformasikan sinyal domain waktu menjadi sinyal domain frekuensi.



Gambar II.6 Spektrum Suara Yas1

Setelah itu, vektor suara hasil ekstraksi ciri LPC kemudian dinormalisasi menggunakan nilai maksimum dari deretan vektor ciri sebagai penormalisasi, sehingga diperoleh hasil vektor ciri ternormalisasi dengan nilai maksimum adalah 1 (satu). Karena nilai vektor ciri yang telah dinormalisasi cukup beragam, maka nilai *threshold* di set ke beberapa nilai tertentu agar vektor ciri yang dihasilkan berbeda-beda, proses ini disebut pengkodean.

Tabel II.1 Contoh Hasil Proses Normalisasi dan Pengkodean Vektor Ciri Suara “Yas1”

No	Vektor ciri	Vektor ciri ternormalisasi	Hasil Pengkodean
1.	62.3038	0.4395	1
2.	62.3676	0.4400	1
3.	62.3710	0.4400	1
4.	62.4341	0.4405	1
5.	62.4798	0.4408	1
6.	62.4296	0.4404	1
7.	62.3109	0.4396	1
8.	62.0964	0.4381	1
9.	61.5777	0.4344	1
10.	60.8255	0.4291	0
11.	60.0242	0.4235	0

12.	59.2892	0.4183	0
13.	58.7870	0.4147	0
14.	59.5177	0.4199	0
15.	62.2811	0.4394	1
16.	67.0646	0.4731	1
17.	73.4474	0.5182	1
18.	82.3304	0.5808	1
19.	93.0611	0.6565	1
20.	103.4091	0.7295	1
21.	113.4723	0.8005	1
22.	123.7419	0.8730	1
23.	133.1082	0.9391	1
24.	141.7469	1.0000	1
Total		13.0689	
Rata-rata		0.5445	

Hasil pengkodean vektor ciri inilah yang dijadikan sebagai input pada Algoritma Genetik.

$$\text{Nilai fitness :} \\ = \frac{\sum \text{gen cocok}}{24}$$

$$\times 100$$

Dari hasil evaluasi nilai *fitness* pada masukan suara “Jan1” dapat diketahui bahwa nilai *fitness* terbaik adalah 100%.

Tabel II.3 Pengenalan suara untuk AG dengan 10 suara data base

No	Input uji	Dikenali sebagai				
		achy	ida	ichal	linda	yas
1	Yas					✓
	Yas					✓
	Yas					✓
	Yas					✓
	Yas					✓
	Yas					✓
	Yas					✓
	Yas					✓
	Yas					✓

	Yas					✓
2	Achy	✓				
	achy	✓				
	achy	✓				
	achy	✓				
	achy	✓				
	achy	✓				
	achy	✓				
	achy	✓				
	achy	✓				
	achy	✓				
3	ichal			✓		
	ichal			✓		
	ichal			✓		
	ichal			✓		
	ichal			✓		
	ichal			✓		
	ichal			✓		
	ichal			✓		
4	ida		✓			
	ida		✓			
	ida		✓			
	ida		✓			
	ida		✓			
	ida		✓			

	ida		✓			
	ida		✓			
	ida		✓			
	ida		✓			
	ida		✓			
5	linda				✓	
	linda				✓	
	linda				✓	
	linda				✓	
	linda				✓	
	linda				✓	
	linda				✓	
	linda				✓	
	linda				✓	
	linda				✓	

Jumlah sample = 50

Sukses = 50

Gagal = 0

Tabel II.4 Pengenalan 25 suara input baru untuk AG dengan 10 suara data base

No	Input uji	Dikenali sebagai				
		ica	ida	ita	aji	yas
1	Yas					✓
	Yas					✓
	Yas					✓
	Yas					✓
	Yas					✓
2	achy	✓				
	achy	✓				

	achy	✓				
	achy	✓				
	achy				✓	
	ichal			✓		
	ichal			✓		
3	ichal			✓		
	ichal				✓	
	ichal			✓		
	Ida				✓	
	Ida				✓	
4	Ida		✓			
	Ida		✓			
	Ida					✓
	linda				✓	
	linda				✓	
5	linda				✓	
	linda			✓		
	linda					✓

Jumlah sample = 25

Sukses = 18

Gagal = 7

Dari tabel II.4 dapat diketahui bahwa sistem dapat mengenali suara uji baru dari kelima target dengan benar sebanyak 18 dari 25 *sample* suara yang diujikan untuk 10 data base.

Kegagalan sistem dalam mengenali suara bisa terjadi karena:

1. Penggunaan parameter desimasi (n) yang di set ke nilai 10 terlalu kecil untuk mengubah jumlah sample suara

yang beragam, yakni dari 12000 menjadi 1200 sehingga menghasilkan vektor ciri yang sangat beragam.

2. Adanya kesamaan antara kromosom suara uji dan kromosom suara pada data base. Hal ini dikarenakan saat pengkodean, rata-rata jumlah vektor ciri yang hampir sama menyebabkan penentuan nilai threshold menghasilkan bit biner yang sama.
3. Diperoleh nilai *fitness* yang memenuhi kriteria terministik pada pada populasi ke-n setelah beberapa generasi.

Analisis Performansi Sistem:

Proses pengujian RSD dilakukan untuk fungsi aktivasi sigmoid bipolar :

$$RSD1 = \frac{\text{Data sukses}}{\text{Jumlah Data}} \times 100\%$$

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

Uji dengan 10 data base

$$\frac{\text{Data sukses}}{\text{Jumlah Data}} \times 100\%$$

$$\frac{18}{25} \times 100\% = 72\%$$

IV. KESIMPULAN

Tingkat keberhasilan sistem berdasarkan pengujian yang dilakukan pada pengujian suara data base dengan 10, 7, dan 5 data input adalah 100%. Tingkat keberhasilan sistem untuk pengujian 5 suara input baru dari kelima target untuk 50 suara data base adalah 72%, untuk 35 suara data base yaitu 60% , dan untuk 25 suara data base yaitu 53%.

SARAN

1. Penggunaan metode LPC (*Linier Predictive Coding*), sebagai ekstraksi

ciri dapat diganti dengan metode yang lain misalnya dengan gabor Wavelet, atau keluarga wavelet yang lain, atau metode-metode transformasi dan filter yang lain.

2. Penggunaan Algoritma Genetik dalam mengidentifikasi suara ini belum mencapai hasil yang maksimal, maka untuk pengembangan sistem selanjutnya dapat digunakan kecerdasan buatan lainnya seperti Logika Fuzzy, Sistem Pakar, Jaringan Syaraf Tiruan (JST), dan lain sebagainya.

V. DAFTAR PUSTAKA

- [1]. Identifikasi Suara.
<http://www.digilib.itb.ac.id/gdl.php?mod=browse&op=read&id=jbptitbpp-gdl-fenikhairu-2417&q=identifikasi%20suara>. Akses : 20 Desember 2011
- [2]. Razmayana R, dan Safri T, 2009. *Speech Recognition Menggunakan Gabor Wavelet dan Genetic Algorithms untuk Sistem Keamanan*. Fakultas Teknik Unhas. Makassar.
- [3]. Sistem Produksi Suara Manusia.
<http://yoanesbandung.wordpress.com/2008/05/26/sistem-produksi-suara-manusia/>. Akses : 23 Desember 2011
- [4]. Rudy Adipranata, Resmana. 1999. *Pengenalan Suara Manusia Dengan Metode LPC dan Jaringan Saraf Tiruan Propagasi Balik*. Surabaya.