

# unmer 1

*by* Mohammad Zoqi Sarwani

---

**Submission date:** 04-Nov-2019 11:11AM (UTC+0700)

**Submission ID:** 1206397870

**File name:** Vol.1\_No.1-Unmer.doc (378.5K)

**Word count:** 2048

**Character count:** 13003

Article history

Received : xxxx

Revised : xxxx

Accepted : xxxx

# SPEECH TO TEXT PROCESSING FOR INTERACTIVE AGENT OF VIRTUAL TOUR NAVIGATION

Dian Ahkam Sani<sup>1\*</sup>, Muchammad Saifulloh<sup>2</sup>

<sup>1,2</sup>Universitas Merdeka Pasuruan, Pasuruan, 67129, Indonesia

<sup>1</sup>dianahkam@unmerpas.ac.id, <sup>2</sup>ipungyellow1@gmail.com

## ABSTRACT

Berkembangnya ilmu pengetahuan dan teknologi, maka salah satu cara untuk mencapainya adalah bagaimana menggantikan cara interaksi manusia dengan komputer dengan memberikan inputan suara. Konversi suara menjadi bentuk teks dengan metode *Backpropagation* dapat dipahami dan dimengerti melalui ekstraksi [1], diantaranya dengan *Linear Predictive Coding* (LPC). *Linear Predictive Coding* adalah salah satu cara untuk merepresentasikan sinyal dalam memperoleh fitur atau ciri unik dari tiap pola suara. Secara garis besar, cara kerja sistem pengenalan suara ini [4] adalah mula-mula sinyal suara manusia yang diterima dengan menggunakan *microphone* (sinyal analog) dicuplik dengan kecepatan cuplik 8000 Hz sehingga menjadi sinyal digital dengan bantuan *sound card* pada komputer. Sinyal digital hasil cuplikan ini kemudian diproses awal menggunakan *Linear Predictive Coding* (LPC) sehingga didapat beberapa koefisien LPC yang merupakan feature (ciri) dari suara tersebut. Hasil keluaran LPC ini akan di latih dengan metode pembelajaran *Backpropagation*. Dari hasil pembelajaran akan di klasifikasikan dengan sebuah [1] data lalu disimpan di database. Hasil pengujian berupa program pengenalan yang dapat menampilkan plot suara dan teks hasil pengenalan suara dengan prosentase pengenalan suara dari responden di dalam database sebesar 80% dalam pengujian secara *Real Time* dari 100 data.

Keywords : Suara, Linear Predictive Coding (LPC), Backpropagation .

## I. INTRODUCTION

Perkembangan teknologi pada saat ini bisa dikatakan sudah sangat berkembang. Namun dalam perkembangan itu sendiri tidak lepas dari tiga hal pokok seperti suara, penglihatan, dan sentuhan. Sehingga dalam mengembangkan teknologi khusus nya kecerdasan buatan mengacu pada tiga hal tersebut[1].

Banyak kemudahan yang ditawarkan untuk kepentingan interaksi manusia dan komputer. Misalnya dalam pengenalan suara (*Speech Recognition*) dan merubahnya menjadi teks. *Speech Recognition* meru[2]kan konversi sebuah sinyal akustik, yang ditangkap oleh microphone atau telepon, untuk merangkai kata[2], bagi sebagian manusia dalam mengenal sebuah kata mudah saja karena manusia mempunyai pengenalan pola yang sangat baik. Tetapi bagaimana dengan komputer? Untuk itu diperlukan teknologi untuk mengubah suara ke dalam bentuk teks. Sehingga komputer mampu mengenali inputan suara dan menerjemahkannya kedalam bentuk teks.

Dalam penelitian terdahulu pernah dilakukan oleh [3] yang berjudul "Pengenalan Suara Manusia Menggunakan Metode Linier Predictive Coding (LPC)". Menyatakan bahwa kesulitan user atau pengguna komputer untuk menggunakan sejumlah tools yang terdapat di [4] satu aplikasi komputer, keinginan manusia untuk mempermudah penggunaan suatu alat dan (dalam dunia kerja). Sehingga pengenalan karakteristik alam [4] manusia dapat digunakan. Suara merupakan salah satu bentuk biometric yang dapat digunakan sebagai person identification, pengenalan suara pembicara (*speech recognition*) tidak membutuhkan peralatan khusus dan karena karakteristik suara manusia yang berbeda.

Selanjutnya pada penelitian yang dilakukan oleh [4] yang berjudul "Pengenalan Pengucap Tak Bergantung [1] teks dengan Metode Vector Quantization (VQ) Melalui Ekstraksi Linear Predictive Coding (LPC)". Menyatakan bahwa Dalam ilmu pengetahuan dan teknologi khususnya bidang komunikasi, pengolahan sinyal memegang peranan yang penting. Penelitian yang intensif dalam bidang pengolahan sinyal menyebabkan teknologi komunikasi berkembang dengan pesat. Salah satunya adalah pengenalan pengucap. Pengenalan pengucap merupakan cara yang digunakan untuk mengetahui identitas seseorang yang mengucapkan sinyal informasi. Hal ini bisa dilakukan karena masing-masing individu memiliki karakteristik-karakteristik sinyal ucapan yang spesifik. Karakteristik ucapan dapat dibedakan melalui ekstraksi dengan suatu teknik pengkodean. Teknik pengkodean yang umum digunakan dalam pegekstraksian sinyal ucapan adalah LPC (Linear Predictive Coding). Analisa LPC menghasilkan suatu estimasi parameter ucapan dasar, antara lain pitch, formant, persamaan area jalur vokal, dan untuk mendapatkan (kompresi) sinyal ucapan agar didapat bit-rate rendah untuk keperluan transmisi atau penyimpanan.

Selanjutnya pada penelitian yang dilakukan oleh [5] yang berjudul "Pengkonversian Data Analog Menjadi Data Digital Dan Data Digital Menja[5] Data Analog Menggunakan Interface Ppi 8255 Dengan Bahasa Pemrograman Borland Delphi 5.0". Menyatakan bahwa Perkembangan teknologi secara modern sangat diperlukan pada kondisi saat ini untuk menyempurnakan teknologi F. Sistem kontrol dengan berbasis PC (Personal Computer) merupakan salah satu penerapan dari teknologi modern dimana banyak dari aplikasi komputer ini dapat membantu manusia dalam menyelesaikan pekerjaannya. Salah satu aplikasi

penggunaan sistem kontrol berbasis komputer PC adalah mengontrol proses konversi analog ke digital dan konversi digital ke analog yang ditujukan untuk menggantikan pengontrolan secara manual sehingga dapat meningkatkan efisiensi waktu dan tenaga. Oleh karenanya maka di buatlah teknologi untuk mengubah suara ke dalam bentuk text. Sehingga komputer mampu mengenali inputan suara dan menerjemahkannya kedalam bentuk teks..

## II. METHOD

Pengenalan ucapan dengan metode backpropagation dapat dimengerti melalui ekstraksi ciri dengan tujuan untuk memperjelas ciri tiap pola suara, diantaranya dengan Linier Predictive Coding (LPC) yang bersumber pada buku dan literature internet.

### A. Diagram Alur Sistem

Alur penelitian berisi proses penginputan secara realtime sehingga dapat menghasilkan output berupa teks seperti pada Fig. 1. Dalam sub bab ini berisi penjelasan diagram alur berupa data training dan data testing dari penelitian yang dilakukan. Berikut diagram alur yang digunakan adalah sebagai berikut:

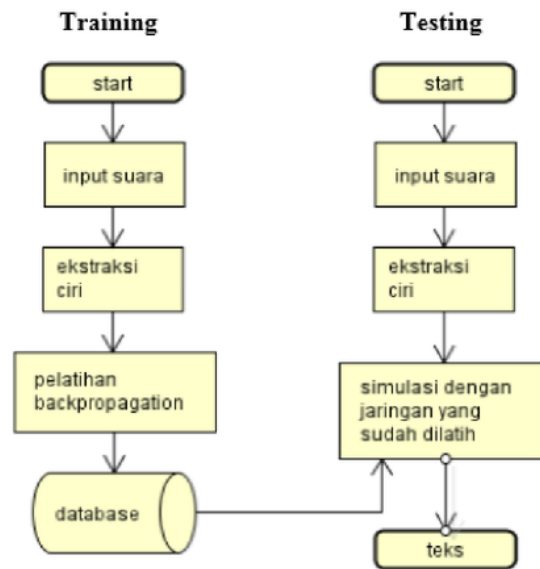


Fig. 1. Diagram Alir Pengenalan Suara

Sistem ini bermula dari pencuplikan sinyal analog menjadi sinyal digital dengan proses perubahan, penyaringan, dan estimasi sinyal audio menjadi data binary dengan memperoleh nilai dari setiap karakteristik sinyal audio diperlukan pemrosesan sinyal digital. Dari sinyal digital inilah maka proses - proses pemberian parameter suara ke dalam deretan feature vector yang berupa ringkasan dan informasi yang relevan dari suara tersebut dapat dilakukan dengan proses ekstraksi ciri Linier Prediktive Coding (LPC). Keluaran hasil LPC yang berupa nilai koefisien LPC inilah yang merupakan feature (ciri) dari suara yang diucapkan.

Untuk proses pengenalan dan pengambilan keputusan dapat digunakan metode Backpropagation. Metode Backpropagation akan mengklasifikasikan setiap suara dengan sebuah kata dan akan disimpan di database. Pada metode pencarian suara yang masuk akan dibandingkan dengan yang sudah ada di database, kemudian keluaran yang didapatkan akan ditampilkan dengan bentuk teks.

### B. Linier Prediktive Coding (LPC)

*Feature Extraction* adalah penggambaran suara secara singkat supaya informasi yang relevan dan penting untuk subsekuen pengenalan tetap terjaga. Tujuan akhir dari proses *feature extraction* adalah pemberian parameter suara ke dalam deretan feature vector yang berupa ringkasan dan informasi yang relevan dari suara tersebut. Feature yang diekstrak diharapkan mampu membedakan suara yang serupa, sehingga model dapat dibuat tanpa memerlukan data training yang besar[2].

Suara yang telah dirubah dari sinyal analog menjadi sinyal digital melalui microphone dan masuk ke sound card kemudian diekstraksi ciri sehingga didapatkan ciri yang merepresentasikan suara tersebut sehingga suara 1 dengan yang lain terlihat perbedaan lebih jelas[6].

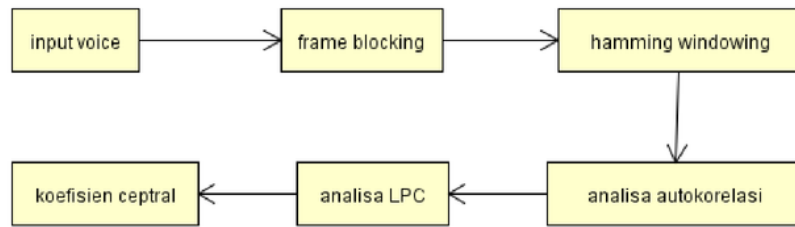


Fig.2. Block Diagram Ekstraksi Ciri LPC

Fig. 2 menjelaskan *Block Diagram Ekstraksi Ciri LPC*, dengan langkah-langkah sebagai berikut:

1. Input Suara

Suara manusia yang direkam dalam format WAV dengan menggunakan mikrofon akan diproses ke tahap selanjutnya.

2. Frame Blocking

Sinyal suara hasil perekaman dibagi ke dalam beberapa frame yang terdiri dari N sampel suara, dengan jarak antara frame yang berdekatan dipisahkan oleh M-sampel. Jika  $M \leq N$ , beberapa frame yang berdekatan akan saling overlap dan hasil estimasi spektral LPC akan berkorelasi dari frame ke frame. Sebaliknya, jika  $M > N$ , tidak akan ada overlap antara frame yang berdekatan[7].

Dimana :

jumlah sample tiap frame (N) =  $(8000 \times 0.04) 320$  sample

jumlah overlapping tiap frame (M) =  $N/2 (320/2) = 160$  sample

jumlah framenya  $(i-N/M)+1 = (8000-320/160)+1 = 49$  Frame

3. Windowing

Windowing memiliki fungsi untuk menghilangkan efek diskontinuitas yang diakibatkan oleh proses sebelumnya. Sehingga sampel yang telah dibagi menjadi beberapa frame perlu dijadikan suara kontiniu[6]. Jenis window yang biasa digunakan adalah Hamming window yang mempunyai bentuk umum seperti pada formula (1) :

$$w(n) = 0,54 - 0,46 \cos \left( \frac{2\pi n}{N-1} \right), 0 \leq n \leq N-1 \quad (1)$$

Dimana :

$w(n)$  : Hamming window

N : jumlah sample suara

4. Analisa Autokorelasi

Setiap frame dari sinyal setelah melalui proses windowing, kemudian dilakukan analisis autokorelasi dengan nilai autokorelasi tertinggi p adalah orde LPC yaitu 12 karena Nilai p biasanya antara 8 sampai 16 dengan menggunakan formula (2).

$$r_l(m) = \sum_{n=0}^{N-1-m} \tilde{x}_l(n) \tilde{x}_l(n+m), m = 0,1,\dots,p. \quad (2)$$

Dimana :

$r_l(m)$  : Koefisien autokorelasi

$\tilde{x}_l$  : Sinyal masukan

P : Orde analisa LPC

5. Analisa LPC

Proses selanjutnya adalah analisis LPC, yang mengubah setiap frame autokorelasi p+1 ke dalam bentuk parameter parameter LPC atau yang biasa disebut dengan koefisien LPC dengan menggunakan formula (3).

$$a_m = \alpha_m^{(p)} \quad 1 \leq m \leq p \quad (3)$$

Dari data ini kemudian dijadikan masukan bagi jaringan syaraf tiruan Backpropagation.



### C. <sup>1</sup>ackPropagation.

Setelah melalui tahap-tahap diatas selanjutnya parameter-parameter yang dihasilkan dimasukkan ke dalam jaringan saraf tiruan dengan menggunakan metode pembelajaran BackPropagation. Diagram alir proses pelatihan jaringan dapat dilihat pada Fig.3.

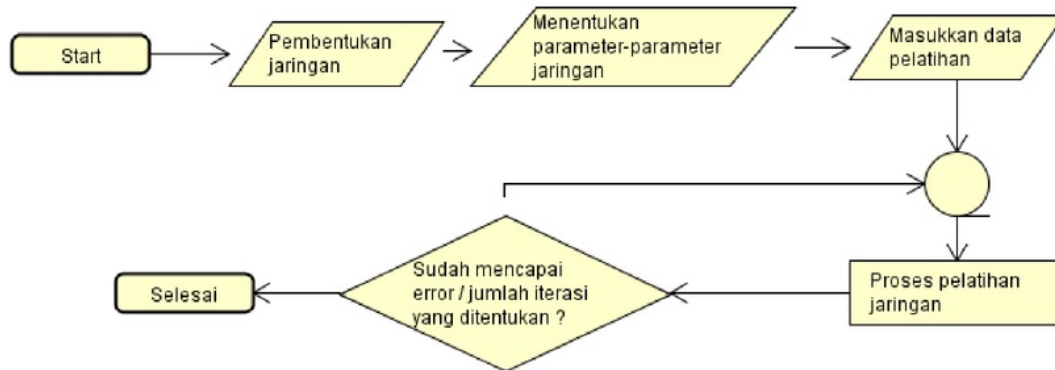


Fig. 3. Diagram alir BackPropagation

Dimana pembentukan jaringan dengan menetapkan kesalahan yang ditargetkan dan iterasi maximal, nilai bias=1 untuk semua lapis input <sup>2</sup> tersembunyi (hidden) pada data akhir. Tetapkan target pola output dengan memasukkan 100 output yang telah ditentukan. Inisialisasi bobot awal secara acak kecil, nilai acak di antara -0.5 dan 0.5 untuk efisiensi dapat digunakan aturan pemilihan bobot awal *Nguyen widrow*. Yaitu bobot disesuaikan dengan banyaknya *neuron* tersembunyi. Setting parameter iterasi awal (epoch) sama dengan 1. Selama epoch kurang dari maksimal iterasi akan berjalan. Lakukan perhitungan kesalahan antara pola output *Neural Network* dan pola target. Apabila sudah mencapai error / jumlah maksimal iterasi maka proses pembelajaran selesai.

Data latihan yang akan diujikan disini adalah 700 data suara baru yang terdiri dari 7 suara untuk satu kata, disampling dalam bentuk matrik kemudian matrik tersebut disimulasikan ke jaringan yang sudah dilatih sebelumnya untuk mendapatkan hasil pengenalan masing-masing kata.

### III. RESULT AND DISCUSSION

Metode *Backpropagation* yang digunakan pada ujicoba untuk konversi suara menjadi bentuk teks seperti pada Fig.4, terdapat satu button saja yang dapat digunakan untuk *record* untuk memulai program pengenalan suara dan diikuti dengan mengatakan salah satu suara.



Fig. 4. Uji Metode *Backpropagation*

Hasil uji coba terhadap pengenalan suara pada Fig. 5 pada kotak Static Text 1 dan plot perekaman pada kotak Axes1.User dapat mengulangi kembali proses pengenalan suara dengan menekan tombol "RECORD". Program pengenalan dapat diakhiri dengan menekan tombol "X" untuk keluar dari program pengenalan suara.



Fig.5. Pengenalan Suara

Pengujian sistem Konversi Suara Menjadi Bentuk Teks Menggunakan Metode *Backpropagation* ini dilakukan dengan menggunakan *microphone*. Pengujian menggunakan 100 data. Hasil pengujian sistem identifikasi pengenalan karakter plat nomor kendaraan disajikan dalam bentuk tabel 1.

TABLE I  
PENGUJIAN KONVERSI SUARA

Input	Output
Berjalan	Benar
Berlari	Benar
Melihat	Benar
Membaca	Benar
Membayar	Salah
Membersihkan	Salah
Memikir	Benar
Mencatat	Benar
Mendaftar	Benar
Mendengarkan	Salah

Dari hasil pengujian 100 data terdapat jumlah keberhasilan identifikasi sebanyak 80 data, sementara yang salah berjumlah 20 data.

$$A = \frac{\text{Data benar}}{\text{jumlah data}} \times 100\%$$

$$A = \frac{80}{100} \times 100\%$$

$$A = 80\%$$

Tingkat akurasi dari pengenalan suara adalah 80%. Berdasarkan data sampel yang telah dilakukan pengujian, kegagalan pengenalan disebabkan oleh beberapa hal yaitu:

1. Kondisi ramai, bising dsb.
2. Terjadi banya noise.
3. Suara tidak ada di dalam databse.

#### IV. CONCLUSION

Sistem pengenalan suara menjadi teks telah bekerja sesuai dengan perancangan, dapat mengkonversi suara menjadi teks dan menampilkannya pada GUI MATLAB. Sistem ini dapat mengenali suara dengan benar sesuai dengan suara yang di inputkan. Dimana tingkat pengenalan menjadi lebih baik jika yang menginputkan suara adalah orang yang suaranya ada dalam database, dan pada penelitian ini menghasilkan tingkat pengenalan terbaik yaitu 80% dari 100 data. Namun tingkat pengenalan suara menjadi turun ketika kata-kata yang diucapkan oleh orang lain yang tidak ada di dalam database, di ruangan yang ramai dan bising.

#### ACKNOWLEDGMENTS

Pada sistem pengenalan suara menggunakan metode backpropagation berbasis Matlab ini tentunya masih jauh dari kesempurnaan. Oleh karena itu diharapkan dapat membantu pengembangan sistem pengenalan suara ini agar semakin baik. Penelitian lanjutan yang dapat dilakukan yaitu menggunakan metode yang berbeda dalam sistem pengenalan sehingga dapat

dibandingkan hasil tingkat pengenalan suaranya dan mengembangkan lebih lanjut menjadi aplikasi untuk pembelajaran pengucapan berbagai macam kata dalam Bahasa.

#### REFERENCES

- [1] F. A. AHDA, "ANALISIS SUARA ALPHABET MENGGUNAKAN JARINGAN SYARAF TIRUAN PROPAGASI BALIK," *Anal. SUARA Alph. MENGGUNAKAN Jar. SYARAF TIRUAN PROPAGASI BALIK*, vol. 60, no. 4, pp. 982–992, 2010.
- [2] M. B. Gunawan, "KONVERSI SUARA KE TEKS MENGGUNAKAN METODE HIDDEN MARKOV MODEL," *KONVERSI SUARA KE TEKS MENGGUNAKAN Metod. HIDDEN MARKOV Model*, p. 45, 2010.
- [3] K. Anam, "Pengenalan suara manusia menggunakan metode," 2013.
- [4] F. AN, "Pengenalan Pengucap Tak Bergantung Teks dengan Metode Vector Quantization ( VQ ) Melalui Ekstraksi Linear Predictive Coding ( LPC )." pp. 1–8, 2004.
- [5] R. A. SRI MELATI SAGITA, SITI KHOTIJAH, "PENGKONVERSIAN DATA ANALOG MENJADI DATA DIGITAL DAN DATA DIGITAL MENJADI DATA ANALOG MENGGUNAKAN INTERFACE PPI 8255 DENGAN BAHASA PEMROGRAMAN BORLAND DELPHI 5 . 0," *ISSN 1979-276X*, vol. 6, no. 2, pp. 168–179, 2013.
- [6] M. Irfandy, "Aplikasi Pengenalan Ucapan Dengan Jaringan Syaraf Tiruan Propagasi Balik Untuk Pengendalian Robot Bergerak," *Apl. Pengenalan Ucapan Dengan Jar. Syaraf Tiruan Propagasi Balik Untuk Pengendali. Robot Berger.*, pp. 1–7, 2004.
- [7] and A. A. Z. Sigit Nur Rohman, Achmad Hidayatno, "APLIKASI PENCIRIAN DENGAN LINEAR PREDICTIVE CODING UNTUK BALIK Landasan Teori," *Apl. PENCIRIAN DENGAN LINEAR Predict. CODING UNTUK PEMBELAJARAN PENGUCAPAN NAMA HEWAN DALAM Bhs. Ingg. MENGGUNAKAN Jar. SARAF TIRUAN PROPAGASI BALIK*, pp. 151–158, 2012.

# unmer 1

---

## ORIGINALITY REPORT

---

**33%**

SIMILARITY INDEX

**34%**

INTERNET SOURCES

**0%**

PUBLICATIONS

**13%**

STUDENT PAPERS

---

## PRIMARY SOURCES

---

**1**

**eprints.undip.ac.id**

Internet Source

**9%**

---

**2**

**www.docstoc.com**

Internet Source

**8%**

---

**3**

**ejournal.undip.ac.id**

Internet Source

**7%**

---

**4**

**theses.uin-malang.ac.id**

Internet Source

**6%**

---

**5**

**journal.lppmunindra.ac.id**

Internet Source

**4%**

---

Exclude quotes  On

Exclude matches  < 2%

Exclude bibliography  On