

Pengenalan Suara Menggunakan Linear Predictive Coding dan Self Organizing Maps

Putri Sahayu
Magister Informatika Fakultas Ilmu Komputer
Universitas Sriwijaya
Palembang, Sumatera Selatan
sahayuputri@gmail.com

Gita Fadila Fitriana
Magister Informatika Fakultas Ilmu Komputer
Universitas Sriwijaya
Palembang, Sumatera Selatan
gitafadila@gmail.com

Abstract—Pada penelitian ini dibangun sebuah perangkat lunak untuk mengenali suara manusia dengan mengimplementasikan metode *Linear Predictive Coding* (LPC) dan *Self Organizing Maps* (SOM). Sebagai data masukan, mula-mula suara direkam dengan frekuensi 8000 Hz dan disimpan menjadi data digital dengan format .wav (*Wave Audio Format*). LPC digunakan untuk ekstraksi ciri parameter suara sehingga mendapatkan koefisien cepstral ini digunakan sebagai inoutan jaringan syaraf tiruan SOM untuk melakukan pelatihan dan pengenalan suara. Hasil pengujian pengenalan suara menunjukkan keberhasilan 100 % untuk pengujian data latih dan 78% data uji.

Keywords—Speech Recognition, Linear Predictive Coding, Self Organizing Maps, Cepstral coefficient

I. PENDAHULUAN

Suara merupakan sebuah pola, manusia dapat mengenali ucapan karena adanya pengetahuan. Komputer juga dapat mengenali ucapan manusia dengan menggunakan jaringan syaraf tiruan. JST merupakan salah satu representasi buatan dari otak manusia yang selalu mencoba untuk mensimulasikan proses pembelajaran pada otak manusia. Sinyal suara merupakan sinyal yang tidak terbatas dalam domain waktu (*infinite time interval*). Suara manusia akan menghasilkan sinyal analog yang terus kontinyu. Sinyal suara analog tidak dapat langsung direpresentasikan pada komputer [6].

Pengenalan suara merupakan perluasan dari teknik dan sistem yang memungkinkan komputer untuk menerima input suara. Pengenalan suara ini dapat diimplementasikan pada pengendalian peralatan, *software*, mengubah ucapan ke teks, otomatisasi layanan operator, dan alat bantu pengenalan suara bagi penyandang cacat [1]. Hal ini akan mudah bagi siapa saja dalam mengoperasikan komputer jika mengetahui cara komunikasi manusia dengan suara. Namun hal ini menjadi latar belakang masalah penelitian ini, hal ini karena suara memiliki sinyal suara yang kompleks dari variasi tekanan udara yang melewati pita suara dan vocal tract, yaitu mulut, lidah, gigi, bibir dan langit-langit [2]

Dalam penelitian ini akan mengembangkan sistem untuk pengenalan suara yang berfokus pada angka dengan menerapkan LPC dan SOM. Sistem yang dirancang pada penelitian ini

adalah aplikasi pengenalan suara menggunakan metode metode *Linear Predictive Coding* untuk ekstraksi fitur serta SOM (*Self Organizing Maps*) untuk membangun perangkat yang dapat mengenali suara manusia. Aplikasi ini dapat mengenali suara yang berfokus pada angka. Hasil penelitian ini akan mengkonversi sinyal suara manusia ke dalam bentuk teks.

II. METODE PENELITIAN

A. Linear Predictive Coding (LPC)

Tahapan dalam pengenalan ucapan terdiri dari *preprocessing*, ekstraksi ciri, dan klasifikasi [10]. Salah satu pendekatan yang digunakan pada tahap klasifikasi adalah jaringan syaraf tiruan. Menurut penelitian yang dilakukan [5]. LPC merupakan cara yang digunakan untuk mendapatkan sebuah pendekatan sinyal suara. LPC secara khusus merupakan metode yang cocok dalam pengolahan sinyal suara. Menurut [3], *Linear Predictive Coding* (LPC) merupakan salah satu teknik analisis sinyal percakapan yang paling *powerful* dan menyediakan ekstraksi fitur yang berkualitas baik dan efisien untuk digunakan dalam perhitungan. Terdapat 6 proses pada tahap pengekstrasian ciri LPC [3], antara lain yaitu *preemphasis* yang berguna untuk meratakan spectral sinyal kata yang telah di cuplik dengan menggunakan persamaan rumus persamaan II.1.

$$\tilde{s}(n) = \hat{s}(n) - \hat{a}s(n-1) \quad (II.1)$$

Kemudian dilanjutkan dengan *frame blocking* yang bertujuan untuk memotong sinyal suara menjadi *frame-frame* yang memuat N cuplikan kata dan *frame-frame* yang berdekatan dipisahkan sejauh M cuplikan.

Kemudian *frame-frame* yang telah terpotong dilakukan proses *windowing* pada masing-masing *frame* dengan menggunakan rumus persamaan II-2. Hal ini bertujuan agar meminimalkan diskontinuitas sinyal awal dan akhir pada masing-masing *frame*. Terdapat beberapa tipe window yang sering digunakan pada pemrosesan sinyal, yaitu *Hanning window* dan *Hamming window*. Adapun model window yang digunakan adalah *Hamming Window*.

$$w(n) = 0.54 - 0.46 \left(\frac{2\pi n}{N-1} \right), 0 \leq n \leq N-1$$

(II.2)

Setiap *frame* hasil *windowing* akan dicari nilai autokorelasinya berdasarkan nilai Orde dari LPC yang bersesuaian dengan frekuensi ucapan. Nilai autokorelasi yang diperoleh akan digunakan kembali pada proses penentuan koefisien LPC dan Error. Kemudian dilanjutkan dengan pengkonversian koefisien LPC menjadi koefisien *cepstral*. Koefisien ini adalah sebagai inputan pada *Self Organizing Maps*.

B. Self Organizing Maps (SOM)

Self Organizing Map (SOM) adalah jenis jaringan syaraf tiruan yang dilatih menggunakan pembelajaran tanpa pengawasan [4]. Jaringan syaraf tiruan Kohonen memiliki perbedaan yang cukup signifikan dengan algoritma JST propagasi balik, perbedaan terletak di cara bagaimana jaringan ini melakukan pelatihan dan cara memanggil kembali pola. Jaringan syaraf tiruan Kohonen tidak menggunakan fungsi aktivasi. Lebih jauh lagi, JST Kohonen tidak memiliki bobot bias [9] Kaidah dari JST ini adalah generalisasi dari proyeksi topografi langsung antara 2 struktur yang berlapis-lapis yang dikenal sebagai *retinotectal mapping* [8].

C. Hasil Penelitian

Pengujian pertama adalah pengujian dengan data latih (*training data set*) menunjukkan hasil pengujian yaitu pengenalan mencapai 100%. Hasil akurasi yang tinggi ini menunjukkan bahwa ekstraksi fitur dengan menggunakan metode LPC mampu membuat nilai-nilai fitur yang baik sehingga SOM dapat mengenali ucapan saat diuji dengan data yang sama.

Pengujian kedua adalah pengujian dengan data uji (*blind data set*). Pada tabel IV-6 menunjukkan hasil pengujian yang cukup baik yaitu sebesar 82%. Untuk kasus uji angka “enam” dan “delapan” tidak terdapat kesalahan, sedangkan pengujian suara angka “satu”, “dua”, “tiga”, “empat” “lima” ,”tujuh”, “sembilan” dan “sepuluh” terdapat kesalahan. Pengenalan ucapan yang salah pada pengujian kedua ini disebabkan fitur yang didapat dari proses ekstraksi ciri dengan metode LPC tidak begitu baik, dimana fitur-fitur tersebut memiliki kemiripan bunyi sehingga jaringan SOM sulit untuk

membedakan kemiripan bunyi tersebut pada data yang belum di latih.

Bagi kedua sampel suara yang hanya memiliki perbedaan nilai fitur sebesar 14% yang berarti persentasi kemiripan suara mencapai 86%. Pengenalan ucapan yang salah ini disebabkan fitur yang didapat dari proses ekstraksi ciri dengan metode LPC tidak begitu baik, dimana fitur-fitur tersebut memiliki kemiripan bunyi sehingga jaringan SOM sulit untuk membedakan 86% kemiripan bunyi.

KESIMPULAN

Berdasarkan hasil analisis dan implementasi yang telah dilakukan pada penelitian ini, maka dapat diambil kesimpulan sebagai berikut :

1. *Linear Predictive Coding* (LPC) dapat mengembangkan Jaringan Syaraf Tiruan untuk mengenali suara.
2. Tingkat akurasi dalam pengenalan suara ini belum stabil.
3. SOM dapat dengan baik mengenali suara untuk batas perbedaan nilai fitur LPC yang lebih besar dari 14%.

REFERENCES

- [1] Dr. R. L. K. Venkateswarlu, Dr. R. Vasantha kumari, dan G.Vani Jayasri. 2011. *Speech Recognition Using Radial Basis Function Neural Network*. IEEE.
- [2] Arman, Arry Akhmad.(2002). Proses Pembentukan dan Karakteristik Sinyal Ucapan. Bandung: Teknik Elektro Department ITB.
- [3] Resmana, and Adipranata, Rudy. (1999). *Pengenalan Suara Manusia dengan Metode LPC dan Jaringan Syaraf Tiruan Backpropagation*. *Proceeding Seminar Nasional 1 Kecerdasan Komputasional* ; Indonesia University, 20-21 July 1999. 1-5.
- [4] Chandar, S., Suriyanarayanan and Manikandan. (2012). *Speech Recognition Using SOM and Actuation Via Network in Matlab*. *International Technology Research Letter*. 1(1): 88-92.
- [5] Ajub Ajulian Z, Achmad Hidayatno, and Muhammad Widyanto Tri Saksono. (2008). Aplikasi Pengenalan Ucapan Sebagai Pengatur Mobil Dengan Pengendali Jarak Jauh. *Transmisi, Teknik Elektro*, 10: 21-26.
- [6] Rabiner, Lawrence and Juang, Biing-Hwang.1993. *Fundamental Of Speech Recognition*. New Jersey: Prentice-Hall International Inc.
- [7] Kohonen, T. (1990). *Self Organizing Maps*. *Proceedings of The IEEE*. 78(9): 1464-1480.
- [8] Kohonen, T. (1982). Self-Organized Formation of Topologically Correct Feature Maps. *Biological Cybernetics* , 59-69.
- [9] Heaton, Jeff. (2007). *Introducing the Kohonen Neural Network*. Missouri: Heaton Research.
- [10] Aprijani ,Dwi Astuti. and Sufandi, Unggul Utan. (2011). Aplikasi Jaringan Syaraf Tiruan Untuk Mengenali Tulisan Tangan Huruf A, B, C, dan D Pada Jawaban Soal Pilihan Ganda. *Jurnal of Matematika, Saint dan Teknologi*, 12(1): 11-17.