

## PERANCANGAN SISTEM PENGENALAN PEMBICARA PADA TEXT DEPENDENT MENGUNAKAN OVERLAPPING DEKOMPOSISI WAVELET PACKET TRANSFORM DAN JARINGAN SYARAF TIRUAN PROPAGASI BALIK

Nanang Saputra<sup>1</sup>, Suyanto<sup>2</sup>, Retno Novi Dayawati<sup>3</sup>

<sup>1</sup>Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

---

### Abstrak

Pada suatu kebutuhan tertentu, suatu sistem dituntut untuk dapat mengenali suara seseorang. Sistem yang dapat mengenali suara manusia itu dinamakan speaker recognition. Masing-masing manusia memiliki pola/fitur tertentu yang khas. Pola inilah yang dikenali suatu sistem untuk membedakan siapa yang mengucapkannya. Masalah yang timbul adalah bagaimana sistem memperoleh fitur yang dapat menggambarkan seseorang dengan lebih tepat dan detail secara perorangan. Dimana suara manusia menempati frekuensi 300 - 3400 Hz yang tidak menutup kemungkinan suara antara dua orang yang berlainan akan diartikan sistem sebagai satu orang. Karena itu pendeteksian orang yang diinginkan bisa salah. Pada penelitian ini dibahas tentang pengenalan suara seseorang (speaker recognition) menggunakan Overlap Dekomposisi Paket Wavelet (ODPW) dan Jaringan Syaraf Tiruan Backpropagation (JST BP). Suara seseorang dibagi-bagi spektralnya menggunakan overlap dekomposisi paket wavelet, selanjutnya discrete cosine transform digunakan pada energi tiap subband yang didapat dan kemudian dikenali dengan Jaringan Syaraf Tiruan Backpropagation (JST BP).

**Kata Kunci :** Speaker Recognition, ODPW, JST BP, discrete cosine transform.

---

### Abstract

In some case, system must able to differentiate human voice. Syatem which can differentiate that human voice was called speaker recognition. Each of person has some unique feature. Those feature that were known by system to differentiate who was speaking. The problem appear is how system get the feature which can descriptive individual more actual and detail every one person. Human voice has spectral frequency between 300 - 3400 Hz which may be voice from personal one person defined as person else and voice from two person defined as one person. So that, speaker recognition is false. In that research explain about speaker recognition use Overlapping Decomposition Packet Wavelet (ODWP) and neural network Backpropagation (JST BP). Voice from some one would be decomposed by ODWP with scenario which would be planed, further discrete cosine transform would be used at subband energy and then feature would be differentiated/classified by Neural Network Backpropagation (JST BP)

**Keywords :** Speaker Recognition, ODPW, JST BP, discrete cosine transform.

---

# 1. Pendahuluan

## 1.1 Latar belakang

Meskipun banyak pengolahan suara seperti *speech and speaker recognition* yang dicapai dengan level performansi tertentu dan meskipun sudah banyak produk yang dikeluarkan pada saat ini, ternyata masih belum didapatkan solusi yang sempurna, sehingga banyak riset dilakukan. Misalnya masalah didalam mendapatkan parameter yang paling tepat dan memuaskan dari sebuah sinyal suara pada *speaker recognition* [9].

*Mel-Frequency Cepstral Coefisien* (MFCC) adalah fitur suara yang sering dipakai untuk proses *speaker recognition* [10]. Tujuan *speaker recognition* adalah untuk membedakan suara manusia dengan mengolah informasi pada spektral suara manusia, semakin detil informasi yang didapat maka perbedaan suara akan semakin baik. Pada dasarnya transformasi wavelet merupakan sebuah teknik pemrosesan sinyal beresolusi ragam. Dengan sifat penskalaanya, wavelet dapat memilah-milah suatu sinyal data berdasarkan komponen frekuensi yang berbeda-beda. Dengan demikian masing-masing bagian dapat dipelajari berdasarkan skala resolusi yang sesuai. Sehingga diperoleh gambaran data secara keseluruhan dan lebih detil. *Dekomposisi Wavelet Packet Transform* (DWPT) merupakan kondisi khusus dari *Overlapping Dekomposisi Wavelet Packet Transform* (ODWPT). Penggunaan ODWPT sebagai alternatif untuk ekstraksi fitur suara akan digunakan pada *speaker recognition*.

Hasil ekstraksi fitur dari ODWPT kemudian diklasifikasikan dengan metode Jaringan Syaraf Tiruan Backpropagation. Jaringan Syaraf Tiruan adalah Suatu teknik yang dibuat dengan memodelkan otak manusia. Seperti pada otak manusia, Jaringan Syaraf Tiruan terdiri atas neuron - neuron yang saling berhubungan yang dapat bekerja sama satu dengan yang lainnya untuk membentuk suatu sistem. Jaringan Syaraf Tiruan dapat belajar untuk mengenali pola pada sistem pengenalan pembicara melalui proses training.

Keluaran ekstraksi fitur suara dari metode ODWPT yang mendekomposisi sinyal suara berdasarkan frekuensi dapat menjadi inputan untuk neuron - neuron sebagai proses pelatihan pembelajaran untuk metode Jaringan Syaraf Tiruan, sehingga diharapkan dapat memperoleh hasil pengenalan pembicara dengan akurasi yang tinggi.

## 1.2 Perumusan masalah

Masalah yang di teliti berdasarkan latar belakang masalah di atas adalah sebagai berikut :

1. Bagaimana implementasi metode ODWPT pada perancangan sistem pengenalan pembicara.
2. Bagaimana implementasi metode JST pada perancangan sistem pengenalan pembicara.
3. Bagaimana menganalisis hasil pengujian sistem.

Dalam penelitian Tugas akhir ini, lingkup batasan masalah yang di gunakan adalah sebagai berikut :

1. Pengenalan pembicara lebih menitik beratkan pada pengenalan suara pembicara tidak pada pengenalan ucapan pembicara.
2. Sistem Pengenalan Pembicara dengan metode *text dependent* dimana sampel kata yang diucapkan pada proses pelatihan dan kata data uji sama, dan menggunakan ejaan Bahasa Indonesia
3. Data input yang digunakan menggunakan kata “Buka” dan “Tutup”
4. Jumlah suara test (fitur tes) yang digunakan adalah 140 suara wanita dan 140 suara pria dari 8 orang. Jumlah database suara (fitur model) yang dibutuhkan untuk simulasi ini adalah 60 suara pria dan 60 suara wanita dari 6 orang.
5. Lama perekaman suara tergantung lamanya pengucapan satu kata.
6. *Mother wavelet* yang digunakan pada ekstraksi fitur suara ini adalah Daubechies 4.
7. Proses klasifikasi fitur suara menggunakan metode JST Backpropagation.

### 1.3 Tujuan

Tujuan yang di capai dalam tugas akhir ini adalah sebagai berikut :

1. Mengimplementasikan ODWPT dalam perancangan sistem pengenalan pembicara.
2. Mengimplementasikan JST dalam perancangan sistem pengenalan pembicara.
3. Menganalisis dan mengkaji pengaruh parameter frekuensi pada ODWPT.
4. Menganalisis hasil dari penggunaan JST Backpropagation dalam pengenalan suara pembicara, berdasarkan *Learning Rate*, dan *Iterasi*.

### 1.4 Metodologi penyelesaian masalah

Metodologi yang dilakukan untuk menyelesaikan tugas akhir ini mencakup hal – hal sebagai berikut :

1. Mengumpulkan bahan-bahan referensi yang akan menunjang proses penelitian, seperti jurnal-jurnal atau buku-buku tentang Wavelet, ODWPT dan JST.
2. Studi literatur tentang Wavelet, ODWPT dan Jaringan Saraf Tiruan.
3. Identifikasi permasalahan yang akan muncul pada saat melakukan penelitian ini, seperti data inputan, teknik training dan pengujian Sistem.
4. Membuat rancangan sistem pengenalan pembicara dan menyusun algoritma program yang diterapkan.
5. Melakukan training untuk mendapatkan pola inputan dan target, yang nantinya digunakan sebagai data training.
6. Pengimplementasian program berdasarkan algoritma ODWPT dan JST Backpropagation yang telah dibuat.
7. Melakukan analisa pengujian hasil implementasi untuk mengetahui tingkat kesuksesan sistem pengenalan pembicara.
8. Membuat kesimpulan dari hasil penelitian.

## 5. KESIMPULAN DAN SARAN

### 5.1 Kesimpulan

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

1. Dalam pengenalan pembicara menggunakan metoda *Overlapping Dekomposisi Wavelet Packet Transform* (ODWPT) dengan klasifikasi *JST Backpropagation*, konfigurasi yang optimal adalah *hidden neuron* 95, *learning rate* 0,1 dan *epoch* 100.000 untuk kata “Buka” dengan tingkat akurasi 100%, dan *hidden neuron* 95, *learning rate* 0,15 dan *epoch* 100.000 dengan tingkat akurasi 96,67% untuk kata “Tutup”.
2. Dari hasil Analisa gambar 4.6 FFT frame terpilih yang ternormalisasi terlihat bahwa untuk frekuensi sinyal suara pada kelas yang sama memiliki kemiripan, sedangkan untuk hasil FFT antar kelas menghasilkan frekuensi sinyal suara yang berbeda
3. Dari hasil Analisa gambar 4.7 koefisien wavelet pada subband terpilih terlihat filter wavelet untuk frekuensi sinyal suara pada kelas yang sama memiliki kemiripan, sedangkan untuk hasil filter wavelet antar kelas menghasilkan frekuensi sinyal suara yang berbeda.
4. Ekstraksi fitur dengan menggunakan metode ini mempunyai beberapa kelebihan, diantaranya adalah sistem bersifat *time invariant*. Metode ini tidak tergantung pada perbedaan lama waktu perekaman atau pengambilan data suara. Data yang panjang maupun pendek akan dapat langsung masuk sistem. Kelebihan lain juga terletak pada sifat yang dimiliki oleh metode ODWP yang tidak mengharuskan data yang diproses memiliki jumlah vektor kelipatan  $2^{j_0}$ . Hal ini jelas berbeda apabila metode yang digunakan adalah LPC yang mengharuskan data inputan memiliki panjang vektor yang sama, atau metode DPW yang mengharuskan data inputan memiliki jumlah vektor kelipatan  $2^{j_0}$ .

### 5.2 Saran

Saran-saran untuk pengembangan tahap selanjutnya antara lain:

1. Untuk mendapatkan tingkat akurasi yang lebih aktual dapat dilakukan dengan menambah data suara latih (data untuk mendapatkan Knowledge sistem) dan juga menambah data testing (data untuk pegetesan).
2. Untuk mengurangi waktu komputasi dapat dilakukan dengan lebih menspesifik-kan tujuan sistem sehingga hanya *subband-subband* tertentu yang dikomputasi.
3. Dengan inputan kata yang lebih panjang memungkinkan untuk mendapatkan tingkat akurasi yang lebih tinggi karena lebih terlihat sebaran frekuensi dari sinyal suaranya.

## Daftar Pustaka

- [1]. Siafarikas, M., Ganchev, T., and Fakotakis, N., 2004, "*Wavelet Packet Based Speaker Verification, in Proc of the ISCA Speaker and Language Recognition Workshop – Odyssey*", Toledo, Spain pp. 257-264.
- [2]. Gopinath, R. A., and C. S. Burrus, 2002, "*A Tutorial Overview of Filter banks, Wavelet and Interrelation, Proceeding ISCAS-93, IEEE, 1993LC, CRC Press, The RF Transmission System Handbook*", California.
- [3]. Percival, D., 2000, "*An Introduction to The Wavelet Analysis of Time Series*", University of Washington, Seattle.
- [4]. Freeman, James A. and Skapura, David M, 1992, "*Neural Networks; Algorithms, Applications, and Programming Techniques*". Addison-Wesley Publishing Company, Inc.
- [5]. Carpenter G.A. y Grossberg S, 1987, "*ART2: Self-organizing of stable category recognition codes for analog input patterns*". Applied Optics, Vol. 16.
- [6]. Hudson, Donna L, and Cohen, Maurice E, 1999, "*Neural Networks and Artificial Intelligence for Biomedical Engineering* ", IEEE Press.
- [7]. Fletcher, H., 1940, "*Auditory Pattern, Reviews of Modern Physics*", no. 12 47-65.
- [8]. Zwicker, E., 1961, "*Subdivision of the Audible Frekuensi Range into Critical Bands (Frequenzgruppen)*", The J. of Acoustical society of America, vol.33 248-249.
- [9]. Siang Jong, M.Sc., 2005, "*Jaringan Syaraf Tiruan dan Pemrogramannya Menggunakan Matlab*", Yogyakarta, ANDI.
- [10]. Chetouani, 2004. "*Universite Pierre&MarieCurie*", LA Science A Paris.
- [11]. Xafopoulos, A. 2001, "*Speaker Verification (an overview), TUT-TICSP presentation, TICSP (Tampere International Center for Signal Processing)*", Tampere: TUT (Tampere Univ).
- [12]. Suyanto, 2005, "*Artificial Intelligence*", Informatika, Bandung, Indonesia.
- [13]. <http://www.mathworks.com>.