

BIOMETRIK SUARA DENGAN TRANSFORMASI *WAVELET* BERBASIS *ORTHOGONAL DAUBENCHIES*

Ketut Agustini⁽¹⁾

Abstract: Biometric as one of identification or recognition person techniques that based on uniquely part of human body. Voice one of uniquely human has. Voice signal that come out from different speakers give different voice pattern. Because of high variations used neural network (NNW) for matching the patterns. Before voice data is processed using NNW it's processed using digital processed through feature extraction phase using discrete wavelet orthogonal base 4 orders with 10 and 15 decomposition rate. The result of NNW is processed by model decision maker that determine speaker identification. The result of experiment shows system biometric built can identify as high as 86%

Keywords: Voice Biometric, Wavelet Transformation, Orthogonal Daubenchies, NNW, Decomposition Rate

Biometrik merupakan studi tentang metode otomatis untuk mengenali atau mengidentifikasi manusia berdasarkan satu atau lebih bagian tubuh manusia atau kelakuan dari manusia itu sendiri. Dalam dunia teknologi informasi, biometrik relevan dengan teknologi yang digunakan untuk menganalisa fisik dan kelakuan manusia untuk autentifikasi. Contohnya dalam pengenalan fisik manusia yaitu dengan pengenalan sidik jari, retina, iris, pola dari wajah (*facial patterns*), tanda tangan dan cara mengetik (*typing patterns*) serta suara. Beberapa hal yang mendorong penggunaan identifikasi secara biometrik adalah biometrik bersifat universal (terdapat pada setiap orang), unik (tiap orang mempunyai ciri khas tersendiri), dan tidak mudah dipalsukan (Xafopoulos, 2001). Dengan teknik biometrik seseorang tidak harus membawa

suatu alat identifikasi seperti pada teknik konvensional.

Proses biometrik (selanjutnya menggunakan kata “identifikasi”) dengan suara memiliki keunggulan secara ekonomis dibandingkan dengan karakteristik yang lain. Identifikasi dengan suara hanya membutuhkan alat tambahan berupa mikrofon dan kartu suara, sedangkan karakteristik-karakteristik yang lain misalnya sidik jari atau wajah membutuhkan alat tambahan seperti *scanner*. Hal ini sedikit banyak dapat menekan biaya pengembangan sistem.

Identifikasi melalui suara termasuk dalam masalah *nonalgorithmic* (Fu, 1994). Walaupun sirkuit digital (komputer) mempunyai kecepatan yang jauh lebih tinggi daripada otak manusia tetapi dalam memproses masalah-masalah *nonalgorithmic* otak

⁽¹⁾ Ketut Agustini, Manajemen Informatika Fakultas PTK Undiksha Singaraja, Kampus Tengah Jl. Udayana Singaraja, Email: adisti_cecilia@yahoo.com

manusia lebih unggul. Suatu teknik yang dibuat dengan memodelkan otak manusia adalah Jaringan Syaraf Tiruan (JST) atau *artificial neural network*. Seperti pada otak manusia, JST 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 suatu pola melalui pembelajaran dan diharapkan dapat memecahkan masalah-masalah yang bersifat *nonalgoritmik*.

Penelitian mengenai identifikasi pembicara telah banyak dilakukan dengan berbagai metode pemrosesan sinyal, seperti *Linier Prediction Coding (LPC)*, *Mel Frequency Cepstrum Coefficients (MFCC)*, *Neural Predictive Coding (NPC)*, dan sebagainya, yang mana keseluruhan metode diatas berbasiskan Transformasi Fourier, dan tingkat identifikasinya telah mencapai 100%, berikut adalah metode-metode yang telah diterapkan dengan tingkat identifikasi yang telah dicapainya, dituangkan dalam bentuk Tabel 1 (Chetouani, 2004).

Tabel 1 Metode-metode yang Digunakan Penelitian Sebelumnya

Metode pemrosesan sinyal suara	Tingkat identifikasi (%)
LPC (Linier Predictive Coding)	90.61
LPCC	96.73
MFCC (Mel Frequency Cepstrum Coefficient)	97.55
PLP	86.12
NPC (random initialization)	61.63
NPC (linier initialization)	100

Sumber: Universite Pierre&MarieCurie, LA Science A Paris, 2004

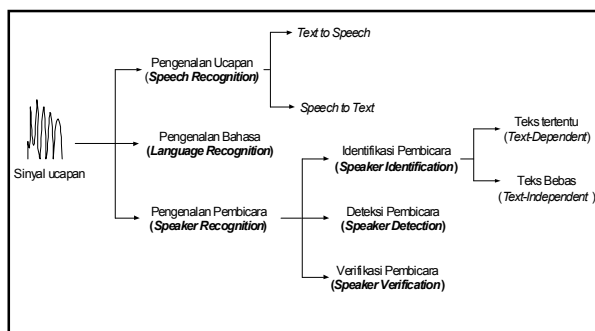
Namun, masih banyak kelemahan yang dimiliki transformasi *fourier* diantaranya, kurang mampu

memberikan informasi sinyal dalam *domain* waktu dan frekuensi secara bersamaan dan menganalisis sinyal yang tidak *stationer*, untuk itu ingin dikembangkan suatu konsep atau pendekatan lain dalam pemrosesan sinyal tanpa berbasiskan *transformasi fourier* yaitu dengan *transformasi wavelet*.

Transformasi Wavelet merupakan sarana yang mulai populer untuk pemrosesan sinyal, seperti citra dan suara, dan transformasi ini belum banyak diaplikasikan untuk analisis suara, khususnya untuk identifikasi pembicara menggunakan teks berbahasa Indonesia. Dalam praktek, *Transformasi Wavelet* digunakan untuk ekstraksi ciri dalam sistem pengenalan suara karena mempunyai karakter khusus yang sesuai untuk analisis sinyal, termasuk sinyal suara. *Transformasi wavelet* sinyal suara menghasilkan resolusi waktu yang baik pada frekuensi tinggi dalam menentukan lokasi awal suara dan parameterisasi ciri suara durasi pendek serta mampu menganalisis sinyal diskontinu (*non stationary*) secara akurat (Krisnan, 1994).

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 menitikberatkan pada pengenalan suara pembicara dan tidak pada pengenalan ucapan pembicara (Ho, 1998).

Metode identifikasi pembicara yang merupakan bagian dari pengenalan pembicara (Gambar 1), dapat dibagi ke dalam metode *text-independent* dan *text-dependent*. Pada sistem *text-independent*, model pembicara meng-*capture* 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 (Mudry,1997) .



Gambar 1 Klasifikasi Sistem Pengolahan Sinyal Suara

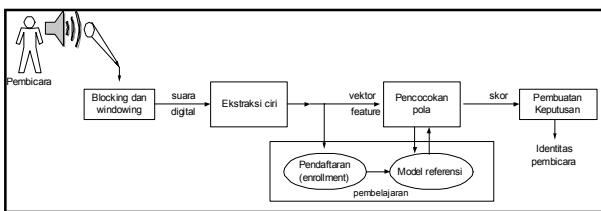
Semua sistem identifikasi pembicara melalui dua proses penting yaitu *feature extraction* dan *feature matching*. *Feature extraction* merupakan proses mengekstraksi data hasil akuisisi sehingga dihasilkan data yang berdimensi lebih kecil, yang nantinya digunakan untuk merepresentasikan tiap-tiap pembicara. *Feature matching* menyangkut prosedur aktual yang mengidentifikasi pembicara yang tidak dikenal dan membandingkan fitur ekstraksi suara yang dimasukan dengan salah satu dari himpunan pembicara yang telah dikenal.

Sistem Identifikasi pembicara juga menyajikan dua sesi yang berbeda, yang pertama menunjukkan

sesi pendaftaran (*enrollment sessions*) atau fase training, sedangkan yang kedua menunjukkan sesi operasi atau fase *testing*. Di dalam fase *training*, tiap pembicara yang telah terdaftar memasukkan contoh (*sampel*) suaranya sehingga sistem dapat mulai dibangun atau dilatih berdasarkan *reference* model pembicara tadi.

Secara umum sistem identifikasi pembicara mempunyai tahapan sebagai berikut dengan diagram bloknya diilustrasikan pada Gambar 2 (Campbell,1997), (a) akuisisi data suara digital, yaitu proses untuk mengakuisisi ucapan pembicara (dalam sinyal analog) dan mengubahnya menjadi sinyal digital. Sinyal digital yang terbentuk berupa suatu vektor yang merepresentasikan suara pembicara, (b) *frame blocking dan windowing*, yaitu *frame blocking* merupakan proses segmentasi sinyal suara digital yang telah diakuisisi ke dalam durasi tertentu, sedangkan *frame windowing* adalah proses yang bertujuan untuk meminimalkan diskontinuitas (*non-stationary*) sinyal pada bagian awal dan akhir sinyal suara, (c) ekstraksi ciri (*feature extraction*), yaitu mengekstrak data hasil akuisisi sehingga dihasilkan data yang berdimensi lebih kecil tanpa merubah karakteristik sinyal suara, (d) pembentukan model referensi pembicara, merupakan tahapan pembelajaran dan akan membentuk suatu model referensi agar sistem dapat mengenali pembicara. Tahap ini memerlukan data berupa vektor-vektor ciri hasil dari ekstraksi ciri yang mencakup seluruh pembicara, model referensi yang terbentuk akan digunakan dalam pencocokan pola, pembentukan model referensi pembicara merupakan tahapan khusus yang dilakukan pada waktu awal sebelum sistem siap digunakan, tahap ini hanya dilakukan sekali dan setelah dilakukan maka sistem siap untuk digunakan,

(e) pencocokan pola (*pattern matching*), yaitu proses pencocokan pola dengan menerima data yang telah diolah oleh ekstraksi ciri sebagai data *input*, proses tersebut akan mencocokkan pola data masukan (*input*) dengan model referensi dan akan memberikan hasil berupa besarnya skor kesesuaian data *input* dengan pola-pola referensi yang ada, (f) Pembuatan keputusan yaitu yang akan menerima skor hasil pencocokan pola. Pada sistem identifikasi, pembuatan keputusan akan menentukan identitas pembicara. Seperti terlihat pada Gambar 2.



Gambar 2 Tahapan Identifikasi Pembicara

Berdasarkan latar belakang diatas, maka rumusan pertanyaan pada penelitian ini adalah “Bagaimanakah rancangan model prototipe sistem biometrik suara yang dibangun menggunakan transformasi wavelet berbasis *orthogonal daubenchies*?”, “bagaimanakah implementasi model *prototype* sistem biometrik suara yang dibangun menggunakan transformasi wavelet berbasis *orthogonal daubenchies*?”, dan “Berapakah tingkat identifikasi (generalisasi) tertinggi yang dicapai dari sistem biometrik suara yang dibangun menggunakan transformasi wavelet berbasis *orthogonal daubenchies*?”.

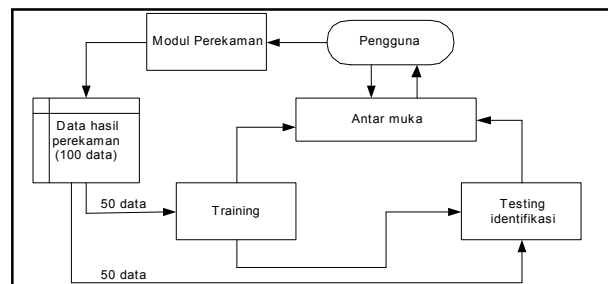
Pada dasarnya tujuan dari penelitian ini adalah untuk mendapatkan jawaban dari pertanyaan penelitian yang telah dirumuskan diatas. Sehingga tujuan penelitian ini adalah (1) diperolehnya rancangan model prototipe sistem biometrik suara yang

dibangun menggunakan transformasi wavelet berbasis *orthogonal daubenchies*, (2) diimplementasikannya model prototipe sistem biometrik suara yang dibangun menggunakan transformasi wavelet berbasis *orthogonal daubenchies* serta 3). Diperolehnya tingkat identifikasi (generalisasi) tertinggi yang dicapai dari sistem biometrik suara yang dibangun menggunakan transformasi wavelet berbasis *orthogonal daubenchies*.

Penelitian ini diharapkan dapat digunakan untuk melakukan identifikasi seseorang melalui kata-kata yang diucapkan orang tersebut. Hasil yang diberikan pada identifikasi berupa identitas pengguna sistem. 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.

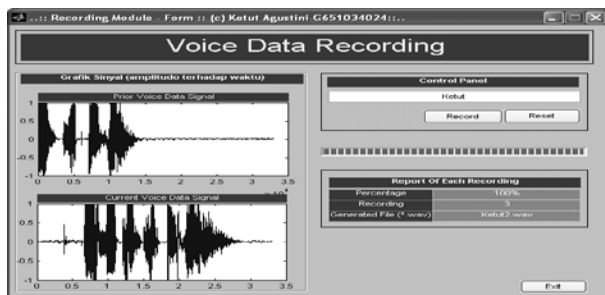
Metode Penelitian

Perancangan model sistem dibangun untuk memudahkan pengguna di dalam pengolahan data dan melihat hasil yang diperoleh dari model sistem tersebut. Sistem yang akan dikembangkan disajikan pada Gambar 3. Sistem tersebut terbagi ke dalam tiga modul yaitu modul perekaman, modul *training* (pelatihan), dan modul *testing* (pengujian)



Gambar 3 Perancangan Model Sistem

identifikasi. Modul training dan modul testing identifikasi dalam satu *interface* atau satu submenu, sedangkan modul perekaman terdapat dalam submenu yang berbeda (Gambar 4). Pada modul perekaman didalamnya terdapat suatu tahapan praproses (*preprocessing*) dan data hasil perekaman yang dihasilkan seluruhnya adalah 100 data suara.



Gambar 4 *Interface* dari Modul Perekaman Suara

Pengguna sistem akan mengucapkan kata yang telah ditentukan sebelumnya yaitu “Ilmu Komputer”. Data audio yang diperoleh akan diubah menjadi bentuk digital (vektor) menggunakan proses *sampling* dengan perangkat lunak MATLAB 7.0.1. Perekaman dilakukan selama 3 detik (1 detik sama dengan 1000 ms) dengan frekuensi *sampling* 20kHz (dalam 1 detik diperoleh data sebanyak 22.050 data). Akuisisi data dilakukan pada beberapa tahap. Pada tahap pertama dilakukan akuisisi data untuk pembelajaran sistem. Pada tahap kedua akuisisi data dilakukan untuk menguji sistem identifikasi.

Dalam penelitian ini menggunakan *frame* (n) dengan lebar waktu 30 ms di mana tiap *frame* menyimpan data sebanyak 661 (hasil pembulatan dari 661,5) sampel dengan *overlap* (m) 50%, sehingga diperoleh jumlah *frame* dengan waktu perekaman selama 1 detik sebesar 65 *frame* (dengan tiap *frame* mengandung data sebanyak 22050 data). Dengan diperolehnya dalam 1 detik 65 *frame* maka

perekaman yang dilakukan selama 3 detik menghasilkan 195 *frame*. Proses ekstraksi ciri pada penelitian ini, adalah data yang telah terbagi ke dalam *frame-frame* dan telah dikalikan dengan *Hamming window*. Masing-masing dari proses ekstraksi ciri diatas akan menghasilkan koefisien-koefisien (koefisien *detail* dan perkiraan) yang diperoleh dari hasil dekomposisi pada *level* 10 dan 15. Pada penelitian ini koefisien yang diambil sebagai masukan ke proses selanjutnya adalah koefisien yang dihasilkan dari frekuensi rendah yaitu koefisien perkiraan (*approximation*) karena bagian penting dari suatu sinyal terletak pada frekuensi tersebut, yang mampu memberikan identitas dari suatu sinyal. Koefisien yang dihasilkan akan membentuk suatu vektor. Algoritma berikut adalah untuk mencari koefisien *detail* dan perkiraan pada proses *multiple* dekomposisi:

- Sinyal yang masuk difilter ke dalam sinyal frekuensi rendah (*low-pass filter*) dan sinyal frekuensi tinggi (*high-pass filter*)
- Lakukan *downsampling* pada ke dua sinyal tersebut
- *Low-pass* frekuensi hasil *downsampling* selanjutnya melalui proses seperti pada tahap pertama
- Lakukan ulang sampai pada *level* yang diinginkan

Pembentukan model referensi pembicara dan pencocokan pola dilakukan menggunakan JST Propagasi Balik. Arsitektur yang digunakan untuk JST Propagasi Balik adalah *Multi Layer Perceptron*, dengan satu lapisan tersembunyi. JST terlebih dahulu dilatih untuk membentuk model referensi pembicara. Setelah tahap pembelajaran selesai dilakukan, JST dapat digunakan untuk melakukan pencocokan pola.

Pada proses identifikasi, pembuatan keputusan dilakukan dengan metode nilai maksimum. Jika *neuron output* ke-*n* merupakan *neuron* dengan nilai maksimum, maka data yang masuk dikenali sebagai pembicara ke-*n*. Sebagai contoh jika *neuron* pertama pada lapisan *output* bernilai 1 dan yang lainnya 0, maka *input* diidentifikasi sebagai pembicara pertama. Fungsi yang digunakan untuk metode nilai maksimum didalam *syntax* matlab adalah *Competitive Transfer Function (COMPET)*. *Syntax* matlab dari penggunaan fungsi tersebut, di mana *testing* merupakan hasil simulasi dengan jaringan, *numkenal_test* adalah jumlah pola yang dikenal, dan *jum_pola* adalah jumlah pola keseluruhan, adalah sebagai berikut,

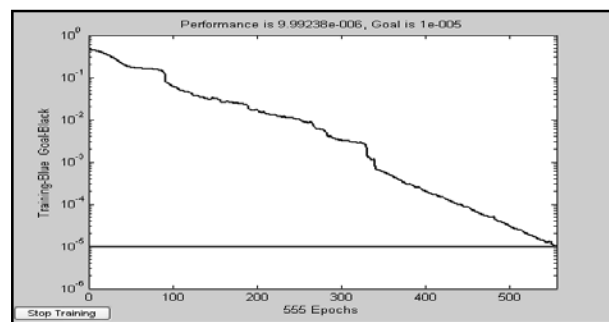
```
hasil_test=compet(ytesting);
numkenal_test=recognize(hasil_test);
generalisasi_test=numkenal_test/
    jum_pola*100
```

HASIL DAN PEMBAHASAN

Pengujian yang telah dilakukan merupakan (1) tahap praproses meliputi proses akuisisi data suara digital, proses *frame blocking* dan *windowing*, serta proses ekstraksi ciri dengan *wavelet* tipe *orthogonal* yaitu *Daubenchies* dengan orde 4 pada *level* 10 dan 15. Pada proses akuisisi data suara digital dan proses *frame blocking* serta *windowing*, kombinasi suara yang digunakan dengan lamanya durasi perekaman, memberikan hasil untuk tiap data suara berupa vektor, dengan dimensi 66.150 (22.050 x 3). Jumlah pembicara yang diambil sebagai data training adalah 50 (lima puluh) data suara yang dipilih dengan *index* ganjil dari 100 (seratus) data hasil perekaman, kemudian sisanya dengan *index* genap digunakan untuk data *test* identifikasi. Hal ini dilakukan agar

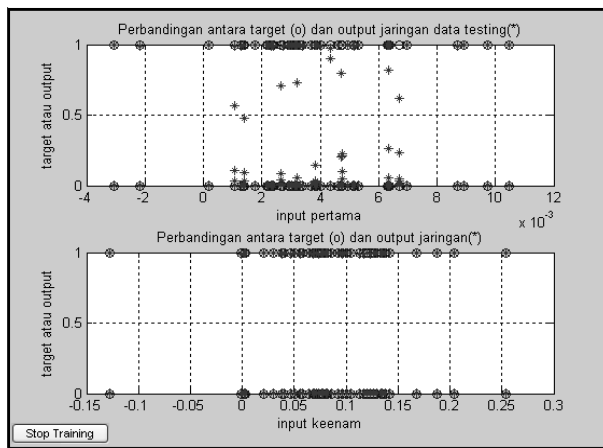
menghasilkan pengidentifikasian yang lebih akurat, mengingat sinyal suara sangat terpengaruh (*sensitive*) terhadap waktu.

Proses Ekstraksi ciri dari *Daubenchies* orde 4 (*daub4*) pada *level* 10 memberikan hasil tiap satu data *file* (satu file *.wav) pembicara (dari 10 data *file* setiap pembicara) menghasilkan koefisien sebanyak 134 elemen dengan *elapsed time* praprosesnya rata-rata sebesar 1,061 detik, dan pada saat 50 data suara yang digunakan untuk training diproses, membutuhkan waktu 96,479 detik. Dari 50 data suara tersebut akan membentuk suatu matriks dengan dimensi 134 x 50, yang digunakan sebagai *input* untuk proses pembentukan referensi pembicara dan pencocokan pola. Sedangkan pada *level* 15, satu *file* pembicara (dari 10 *file* setiap pembicara) menghasilkan koefisien sebanyak 10 elemen dengan *elapsed time* praprosesnya sebesar 1,092 detik, sedangkan untuk 50 data yang digunakan untuk training menghabiskan waktu sebesar 71,933 detik. Pengujian selanjutnya adalah (2) tahap referensi pembicara dan pencocokan pola dengan menggunakan laju pembelajaran 0.01 dan 0.3 serta toleransi galat sebesar 0.00001, mencapai kekonvergenan pada *epoch* ke-555 dengan waktu 8.07 detik dan *mean square error*-nya sebesar 9.99238×10^{-6} , (Gambar 5).



Gambar 5 Proses Pembelajaran *Daubenchies* Orde 4 *Level* 10

Generalisasi yang dihasilkan terhadap data *testing* sebesar 86% (43 pembicara dari 50 data pembicara dikenali). Hal ini terlihat pada Gambar 6, yang menunjukkan perbandingan antara target JST dengan *output* yang dihasilkan dari data *testing*



Gambar 6 Hasil Simulasi Data Pengujian Pada Epoch ke-555

setelah disimulasikan. Hasil keseluruhan pengujian identifikasi ditunjukkan dalam Tabel 2. Dari keseluruhan hasil, terlihat bahwa laju pembelajaran 0.3 memberikan tingkat generalisasi yang lebih baik dibandingkan dengan laju pembelajaran 0.01. Dengan metode *trial and error* dalam menentukan besarnya *neuron* tersembunyi, maka diperoleh hasil terbaik pada banyaknya *neuron* tersembunyi 10 dan waktu proses tersingkat yaitu 6,98 detik dengan tingkat generalisasi tertinggi 86%.

Pada percobaan yang dilakukan terhadap wavelet tipe Daub4 *level* 15, dengan laju pembelajaran 0.3 dan *neuron* tersembunyi 80 memberikan tingkat pengenalan yang paling baik, namun proses pembelajarannya berjalan sangat lambat dan kinerja tujuannya (10^{-5}) secara keseluruhan belum tercapai. Generalisasi terbaik yang dicapai yaitu

Tabel 2 Hasil Pengujian Identifikasi db4 *Level* 10 dengan $err = 0.00001$

Neuron tersembunyi	Pembelajaran		Mse x 10 ⁻⁶	Pengujian		Wkt proses (dt)
	Lr	epoch		Di-indentifikasi dg benar	Generalisasi (%)	
10	0.3	555	9.992	43/50	86	6.98
	0.01	935	9.922	41/50	82	16.424
20	0.3	598	9.989	42/50	84	8.022
	0.01	633	9.818	42/50	84	8.582
30	0.3	603	9.879	41/50	82	8.722
	0.01	635	9.938	40/50	80	9.063
40	0.3	660	9.980	41/50	82	10.065
	0.01	575	9.860	40/50	80	8.793
50	0.3	688	9.915	42/50	84	11.366
	0.01	833	9.654	41/50	82	13.380
60	0.3	642	9.882	41/50	82	11.336
	0.01	860	9.893	42/50	84	14.851
70	0.3	800	9.954	41/50	82	14.611
	0.01	915	9.953	42/50	84	16.574
80	0.3	777	9.731	41/50	82	15.012
	0.01	810	9.977	41/50	82	16.173
90	0.3	764	9.878	41/50	82	15.612
	0.01	810	9.813	42/50	84	16.694
100	0.3	888	9.900	42/50	84	19.107
	0.01	986	9.714	42/50	84	20.850

Sumber: Hasil Percobaan

84%, menurun dibandingkan dengan Daub4 level 10 yang memiliki koefisien lebih banyak.

Hal ini menunjukkan bahwa untuk proses ekstraksi ciri menggunakan metode Transformasi *Wavelet*, teori yang menyatakan “semakin besar jumlah *input* yang diterima maka semakin besar pula waktu pembelajaran yang dibutuhkan JST “, kurang berkenaan terhadap kasus ini, karena semakin tinggi tingkat (*level*) dekomposisi, yang berimplikasi terhadap jumlah koefisiennya (koefisien yang dihasilkan semakin tereduksi) maka proses pembelajaran JST berjalan sangat lambat dan menurunkan tingkat pengenalan jaringannya (dapat ditunjukkan pada hasil generalisasi).

SIMPULAN

Transformasi *Wavelet* berbasis orthogonal *daubenchies* dapat digunakan dalam proses identifikasi pembicara pada bagian pemrosesan awal (praproses) sinyal untuk mendapatkan informasi (ciri) sinyal tersebut. Sistem identifikasi pada tahap praproses Daub4 level 10 menghasilkan tingkat generalisasi tertinggi sebesar 86%. Pada transformasi wavelet, tingkat (*level*) dekomposisi mempengaruhi tingkat pengenalan jaringan yaitu makin tinggi tingkat dekomposisi maka tingkat pengenalan jaringan semakin menurun. Jaringan syaraf tiruan propagasi balik baik digunakan untuk pembentukan referensi pembicara dan pencocokan pola.

RUJUKAN

- Burrus, CS, Gopinath, RA, and Guo, H. 1998. *Introduction to Wavelets and Wavelet Transforms A Primer*. International Edition. New Jersey: Prentice-Hall International, Inc.
- Campbell, JPJR. 1997. *Speaker Recognition: A Tutorial*. Proc. IEEE, vol.85, no 9, pp1437-1462.
- Fausett, L. 1994. *Fundamentals of Neural Network*. Englewood Cliffs. New Jersey: Prentice Hall.
- Fu, L. 1994. *Neural Network in Computer Intelligence*, Singapore: McGraw-Hill.
- Gonzalez RC, and Wood, RE. 2002. *Digital Image Processing*, Second Edition, New Jersey: Prentice Hall, Inc.
- Ho, CE. 1998. *Speaker Recognition System*, Project Report. California: California Institut of Technology.
- Jang, JSR, Sun, CT, and Mizutani, E. 1997. *Neuro Fuzzy and Soft Computing, A Computational Approach to Learning and Machine Intelligence*. International Edition. New Jersey: Prentice-Hall International Inc.
- Krishnan, M, Neophytou, CP, and Prescott, G. 1994. *Wavelet Transform Speech Recognition Using Vector Quantization, Dynamic Time Warping and Artificial Neural Networks*, Lawrence, KS 66045:Center of excellence in computer aided systems engineering and Telecommunication & Information Sciences Laboratory 2291 Irving Hill Drive.
- Mallat, S. 1999. *A Wavelet Tour Of Signal Processing*, Second Edition, Academic Press 24-28 Oval Road, London NW1 7DX UK, <http://www.hbuk.co.uk/ap/>
- Mudry, AH. 1997. *Speaker Identification using Wavelet Transform*, Tesis Master of Engineering, Ontario: Ottawa-Carleton Institut of Electrical Engineering.
- Proakis, JG, dan Manolakis, DG. 1997. *Pemrosesan Sinyal Digital*, edisi Bahasa Indonesia jilid 1, Jakarta: Prenhallindo.
- Rabiner, LR, Juang BH. 1993. *Fundamentals of Speech Recognition*. New Jersey: Prentice Hall. ISBN 0-13-015157-2.
- Xafopoulos, A. 2001. *Speaker Verification (an overview)*, TUT-TICSP presentation, TICSP (Tampere International Center for Signal Processing), Tampere: TUT (Tampere Univ).

