
PERANCANGAN SISTEM AKSES PERSONEL MENGGUNAKAN SPEECH RECOGNITION

Ariyawan Sunardi¹⁾, Djoko Hari Nugroho²⁾

¹⁾Prodi Teknik Elektro UNPAM

Jln. Puspiptek Raya No 46 Buaran, Setu - Tangerang Selatan 15310

²⁾Prodi Teknik Elektro - ISTN

¹⁾Email : dosen00332@unpam.ac.id

ABSTRAK

Pengenalan personel dalam meningkatkan keamanan sebuah objek menjadi hal yang penting. Berbagai metode digunakan untuk pengenalan personel melalui alat bantu kartu maupun biometrik. Suara merupakan biometrik populer yang dapat diterapkan dalam hardware. Telah dilakukan perancangan pengenalan personel menggunakan fasilitas Graphic User Interface (GUI) Matlab R2010a. Tujuan penelitian ini untuk mendesain system pengenalan personel dimulai dari input speech sampai output identitas personel. Sampel suara dengan frekuensi mono 8000 Hz dan 8 bit per sampel ditangkap melalui mikrofon kemudian disimpan sebagai file .wav. Sampel suara kemudian dilakukan proses pre prosesing melalui dekomposisi. Selanjutnya identifikasi personel menggunakan fasilitas fuzzy dari software Matlab. Identitas personel disimpan sebagai database sebagai data uji terhadap file suara input. Didapatkan database 10 personel yang direkam masing-masing sebanyak 5 kali. Untuk pengujian dilakukan 10 kali dengan file suara yang berbeda diperoleh keberhasilan 70,00% dan rata-rata running time terlama 6,832 detik.

Kata kunci : pengenalan personel, GUI, Matlab, speech recognition, safety dan security.

ABSTRACT

Personnel Access System Design For Safety And Security Using Speech Recognition Identification. *The introduction of personnel in increasing the security of an object is important. Various methods are used for the introduction of personnel through card and biometric tools. Sound is a popular biometric that can be applied in hardware. The design of the introduction of personnel using the Graphic User Interface (GUI) facility Matlab R2010a has been carried out. The purpose of this study is to design a personnel recognition system starting from input speech to the output of personnel identity. Sound samples with mono 8000 Hz frequencies and 8 bits per sample are captured through microphones and then saved as .wav files. Sound samples are then subjected to a pre-processing process through decomposition. Furthermore, identification of personnel uses fuzzy facilities from the Matlab software. Personality standards are stored as databases as test data against input sound files. A database of 10 personnel was recorded each of 5 times. For testing carried out 10 times with different sound files 70.00% success was achieved and the longest running time was 6.832 seconds.*

Keywords : personnel identification, GUI, Matlab, speech recognition, safety and security

PENDAHULUAN

Speech recognition adalah proses pengenalan suara manusia yang dilakukan oleh komputer. Suara merupakan salah satu sistem biometrik yang paling populer dan mudah dalam penerapan *hardware*. Suara manusia tidak dapat dikenali begitu saja ketika dimasukkan ke dalam komputer. Sinyal suara manusia yang diterima melalui mikropon berupa sinyal analog sehingga harus diubah menjadi sinyal digital.

Dari sinyal digital yang didapat, diperlukan beberapa tahap untuk mengenalinya, diantaranya adalah pre prosesing, ekstraksi fitur dan pengenalan pola. Pre prosesing dilakukan untuk mendapatkan ciri dasar dari sinyal suara. Dalam ekstraksi fitur digunakan transformasi terhadap suara untuk mengubah domain waktu pada suara menjadi domain frekuensi. Pencarian pola dilakukan dengan membandingkan data kata personal yang ada di basis data dan menentukan pola yang paling cocok dengan suara tersebut.

Salah satu penerapan *speech recognition* adalah identifikasi personal berdasarkan suara personal tersebut yang telah ada di database. Penulis melihat adanya peluang pengembangan sistem akses personal untuk menjamin keselamatan dan keamanan suatu instalasi. Pengembangan sistem ini melalui penerapan *speech recognition*. Dalam penelitian ini dibatasi pada sistem identifikasi personal dan adanya false alarm untuk menjamin keselamatan dan keamanan. Ide untuk merancang sistem *speech recognition* yang menggunakan transformasi *wavelet* sebagai ekstraksi fitur dan *neuro fuzzy* sebagai metode klasifikasi.

TEORI

Speech Recognition

Speech Recognition atau pengenalan suara menjadi salah satu hal potensial yang menjadi perhatian interaksi antara manusia dan komputer. Desain dari sistem *Speech Recognition* membutuhkan perhatian dan penanganan khusus karena terdapat beberapa kendala : realisasi akustik fonem, keragaman akustik, keragaman pengucapan/dialek bahasa, kondisi emosional pembicara, kecepatan berbicara dan lingkungan.

Pemrosesan Sinyal Digital

Sinyal digital berbeda dengan sinyal analog karena sudah *disampled* dan *diquantized*. Keduanya membatasi berapa banyak informasi yang ada di sinyal digital. Pengolahan sinyal digital adalah pemrosesan sinyal yang mempunyai kaitan dengan penyajian perubahan bentuk dan manipulasi dari sisinya dan informasi dalam bentuk digital.

Proses pengubahan gelombang suara menjadi data digital ini dinamakan *Analog-to-Digital Conversion* (ADC), dan kebalikannya, pengubahan data *digital* menjadi gelombang suara dinamakan *Digital-to-Analog Conversion* (DAC).

Preprosesing Sinyal

Dalam tahapan preprosesing, semua sinyal suara yang pertama dikonversi ke tingkat sampling 8 kHz dengan resolusi 16-bit. Dalam prakteknya, suara pembicara akan terkontaminasi dengan komponen noise/gangguan. Kondisi rekaman suara, perangkat maupun lingkungan bertindak sebagai tambahan pengganggu sumber yang tidak diketahui. Oleh karena itu, perlu disaring oleh sinyal digital penyaringan sebagai pra-pengolahan untuk analisis dan klasifikasi suara pembicara.

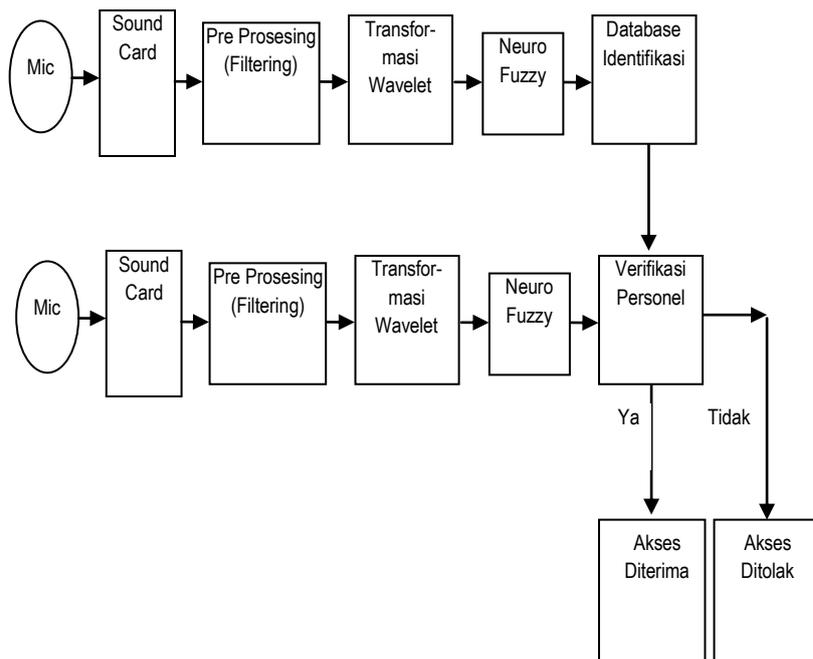
$$X = FS \cdot dt \text{ (detik)} \cdot \left(\frac{\text{bit}}{8}\right) \cdot j \quad (2.1)$$

Tabel 1. Perhitungan Sampling Data Sinyal

Frekuensi (Hz)	Bits/sampel	Channel	Perhitungan	Data sampel
8000	8	Mono	$8000 \times 0.8 \times (8/8) \times 1$	6400
11025	8	Mono	$11025 \times 0.8 \times (8/8) \times 1$	8820
22050	16	Stereo	$22050 \times 0.8 \times (16/8) \times 2$	70560
44100	16	Stereo	$44100 \times 0.8 \times (16/8) \times 2$	141120

METODE

Sistem identifikasi personel yang dirancang dalam penelitian ini dapat dilihat pada gambar 1.



Gambar 1. Diagram Blok Sistem Identifikasi Personel

Pada gambar 1 dapat dilihat proses kerja sebagai berikut :

- a. *Mikrofon* menangkap suara sampel yang kemudian diubah menjadi tegangan elektrik yang dimengerti oleh perangkat elektronik.

- b. *Sound card* mengubah tegangan elektrik dari mikrofon menjadi sinyal digital.
- c. Preprosesing atau filtering melakukan filter pada sinyal input untuk menghilangkan sinyal DC (nol).
- d. Transformasi *wavelet* akan melakukan dekomposisi untuk mengekstraksi informasi yang ada didalam sinyal suara, dengan membagi atau memecah sinyal tersebut ke dalam *band-band* frekuensi.
- e. Hasil fitur dari dekomposisi *wavelet* akan menjadi input untuk pembuatan database personel untuk pengujian menggunakan *neuro fuzzy*.
- f. Verifikasi personel dilakukan dengan mencocokkan dengan database personel.

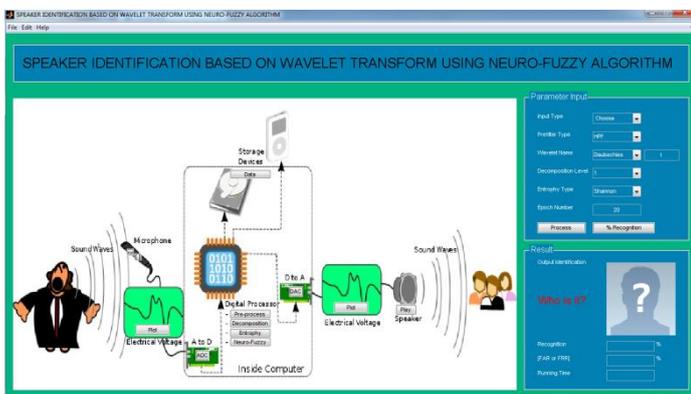
Untuk meningkatkan jaminan keamanan dan keselamatan personel dan instalasi, maka diperlukan adanya sistem *false alarm*. *False alarm* sebagai indikator untuk akses personel yang ditolak.

HASIL DAN PEMBAHASAN

Hasil perancangan system akses personel ditunjukkan pada Gambar 2. Ini menggunakan fasilitas GUI dari software Matlab R2010a. Tampilan sistem pengenalan suara dengan menggunakan antarmuka pengguna secara grafis, agar lebih interaktif, lebih mudah dan lebih menarik.

Suara dari personel yang akan melakukan akses ditangkap oleh mikrofon. Mikrofon akan mengubah menjadi energi listrik. Energi listrik yang masih berupa sinyal analog akan diubah ke sinyal digital oleh sound card. Sound card akan memberikan inputan sinyal digital kepada software Matlab R2010a untuk dilakukan preprosesing, transformasi wavelet dan

neuro fuzzy. Hasil dari software Matlab R2010a ini yang akan kembali memberikan identifikasi personel. Hasil dari identifikasi dengan masukan file wav dari masing-masing personel sebanyak 5 kali akan digunakan menjadi database.



Gambar 2. Tampilan Antarmuka Pengguna Secara Grafis

Untuk pengujian, file file suara yang telah ada dimasukkan ke dalam PC untuk diolah oleh software Matlab R2010a yang akan menghasilkan identitas personel. Hal ini dilakukan dengan 10 kali percobaan. Didapatkan hasil seperti pada Tabel 2.

Tabel 2. Pengujian Dengan File Suara Personel Yang Sama

No.	File suara ke	Berhasil / tidak
1.	Personel 1-1	Berhasil
2.	Personel 1-2	Tidak
3.	Personel 1-3	Berhasil
4.	Personel 1-4	Berhasil
5.	Personel 1-5	Berhasil
6.	Personel 1-1	Tidak
7.	Personel 1-2	Berhasil
8.	Personel 1-3	Tidak
9.	Personel 1-4	Berhasil
10.	Personel 1-5	Berhasil

Selain itu juga diuji tingkat kecepatan sistem untuk mengenali personel yang bersangkutan. Disini faktor Personel Computer (PC) sangat menentukan. Adapun PC yang digunakan memiliki prosesor Intel Pentium CPU P6300 @ 2,27GHz & 2,26GHz, RAM 2 GB. Hasil pengujian juga ditunjukkan tingkat prosentase pengenalan personel.

Tabel 3. Pengujian Waktu Pengenalan

Level Dekomposisi	Running Time (detik)	Tingkat Pengenalan (%)
1	0,998	14,55
2	2,046	51,82
3	3,232	82,73
4	4,814	98,18
5	6,832	100,00

Didapatkan untuk pengenalan 100 % memerlukan waktu selama 6,832 detik dengan menggunakan level dekomposisi tertinggi yaitu 5.

KESIMPULAN

Berdasarkan hasil penelitian tersebut, maka dapat disimpulkan sebagai berikut : Telah dibuat sistem perancangan sistem akses personel menggunakan *speech recognition*. Tingkat keberhasilan pengujian file suara terhadap database didapatkan 70%. Dan untuk mengenali secara penuh 100% memerlukan waktu 6,832 detik menggunakan level dekomposisi 5.

DAFTAR PUSTAKA

1. Anny Tandyo, Martono, Adi Widyatmoko. 2005. **Speaker Identification Menggunakan Transformasi Wavelet Diskrit Dan Jaringan Saraf Tiruan Back-**

-
- Propagation**. Jakarta : Universitas Bina Nusantara
2. Ariman. 2008. **Pengklasifikasian Modulasi Digital PSK Berbasis Entropi Shannon Menggunakan Wavelet**. Jakarta : ISTN
 3. Ashish Kumar Panda, Amit Kumar Sahoo. 2011. **Study Of Speaker Recognition Systems**. Departement Of Electronics And Communications National Institute Of Technology, Rourkela
 3. A.Vijay kumar, Aruna, M.Vijayapal Reddy. 2011. **A Fuzzy Neural Network for Speech Recognition**. India : ARPN Journal of Systems and Software
 4. Binanto, Iwan. 2010. **Multimedia Digital Dasar Teori dan Pengembangannya**. Yogyakarta: Andi.
 6. Intan Verona. 2006. **Perbandingan Tingkat Akurasi Antara Algoritma Momentum Back Propagation dan Learning vector Quantization Neural Network Untuk Speech Recognition**. Jakarta : Universitas Tarumanegara
 5. Joseph P. Campbell, Jr. **Speaker Recognition**. IEEE : Department of Defense Fort Meade.
 7. Ni Wayan Sumartini Saraswati. **Transformasi Wavelet Dan Thresholding Pada Citra Menggunakan Matlab**. Universitas Udayana : Jurnal GENERIC.