

Identifikasi Suara dengan MATLAB sebagai Aplikasi Jaringan Syaraf Tiruan

Speech Recognition in MATLAB as Artificial Neural Network Application

John Adler, Muhamad Azhar, Sri Supatmi

Teknik Komputer, Universitas Komputer Indonesia (UNIKOM)

Jl. Dipati Ukur 112-116 Bandung

Email : john.adler007@gmail.com

Abstrak

Suara berguna untuk melakukan suatu komunikasi antar sesama manusia. Dimana masing-masing suara memiliki karakteristik dan tingkatan frekuensi yang berbeda-beda. Suara dihasilkan oleh suatu pita suara dengan kerjasama seluruh organ penghasil suara. Saat ini aplikasi teknologi ini berkembang dengan cepat dan sangat berguna bagi kehidupan manusia, contoh : membantu tunarungu, pembuatan google translate dan lain-lain. Ada beberapa permasalahan dalam pengidentifikasian suara yaitu kecepatan dan ketepatan pengolahan data suara.

Sedangkan untuk mentransformasi sinyal suara sebagai inputan menjadi sinyal frekuensi suara dalam format WAV diperlukan suatu metoda Jaringan Syaraf Tiruan yang dapat meminimalkan masalah, yaitu menggunakan Fast Fourier Transform (FFT) dengan bantuan software Matlab dalam mengenali kekuatan suara (desibel) dan beberapa karakteristik suara lainnya. Tujuan perancangan ini agar mampu mengidentifikasi dan menganalisa karakteristik suara serta mengolah data suara yang sudah tersimpan pada sistem yang sudah dibuat sehingga sesuai dengan keinginan.

Dengan mengambil 8x data pada sepuluh orang yang telah dikenali frekuensi suara, maka persentase suara di malam hari lebih besar dibandingkan siang hari berturut-turut adalah 75%, 50%, 25%, 37%, dan 0% sebanyak tiga orang, dua orang, dua orang, satu orang, dan dua orang. Perbedaan selisih persentase adalah kondisi psikis, organ penghasil suara serta kondisi cuaca di tempat pengambilan data.

Kata Kunci : Suara, Fast Fourier Transform, Matlab (Matrix Laboratory), Jaringan Syaraf Tiruan, Frekuensi

Abstract

Sounds useful to make a communication among human beings. Where each sound have characteristics and levels of different frequency. Sounds produced by vocal cords with cooperation of all sound-producing organs. Currently, the application of this technology develops quickly and was very useful for human life, for example: helping hearing impaired, making google translate and the others. There are several problems in sounds identifying sound, like velocity and accuracy of processing voice data.

While for transform sound signals as input signals into sound frequencies in WAV format required a method of Artificial Neural Network to minimize the problem, using Fast Fourier Transform (FFT) with the help of software Matlab to recognize the power of sound (decibels) and a few other voice characteristics. Design goal is to be able to identifying and analyzing characteristics of sound and voice processing data already saved on the system that have been made so that as you wish.

By taking 8 times data on to ten peoples who have recognized sound frequency, the percentage of sound at night more than during day in a row was 75%, 50%, 25%, 37%, and 0% as many as three people, two people, two people, one people, and two people. Differences to this percentage are psychological condition, sound-producing organs and weather conditions on the set of data.

Keywords : Sound, Fast Fourier Transform, Matlab (Matrix Laboratory), Artificial Neural Networks, frequency

I. PENDAHULUAN

Suara merupakan nikmat yang diberikan Tuhan kepada manusia, dipergunakan untuk berkomunikasi dengan sesama bahkan dengan pencipta-Nya. Manusia memiliki karakteristik suara yang berbeda-beda dan juga tingkatan frekuensi yang berbeda. Frekuensi suara yang dapat didengar oleh manusia antara 20 sampai 20.000 Hz. Range ini berbeda-beda secara individu dan umumnya tergantung usia [1].

Dalam pengolahan suara ini akan ada beberapa permasalahan yang muncul seperti kecepatan pengolahan, ketepatan, dan lain sebagainya. Untuk itu diperlukan suatu metoda Jaringan Syaraf Tiruan yang dapat meminimalisir masalah yang mungkin timbul pada sistem jaringan saraf tiruan yang dibuat, metoda yang akan dipakai untuk membuat aplikasi ini yaitu *Fast Fourier Transform* (FFT), karena *Fast Fourier Transform* adalah metoda yang fungsinya mentransformasi sinyal suara menjadi sinyal frekuensi.

II. TINJAUAN PUSTAKA

A. Jaringan Syaraf Tiruan

Jaringan Syaraf Tiruan adalah suatu teknik pemrosesan informasi yang terinspirasi oleh sistem sel syaraf biologi. Jaringan Syaraf Tiruan menyerupai otak manusia dalam dua hal yaitu pengetahuan diperoleh jaringan melalui proses belajar dan kekuatan hubungan antar sel syaraf (*neuron*) digunakan untuk menyimpan pengetahuan. Jaringan Syaraf Tiruan mengadopsi dasar dari sistem syaraf biologi, menerima masukan dari suatu data atau dari output sel syaraf pada jaringan syaraf. Setiap masukan datang melalui suatu hubungan yang sudah ada, dan setiap pola-pola informasi masukan dan output yang diberikan kedalam Jaringan Syaraf Tiruan diproses dalam *neuron*. *Neuron-neuron* tersebut terkumpul didalam lapisan-lapisan yang disebut *neuron layers*.

B. Suara

Suara atau gelombang akustik sesungguhnya merupakan kasus khusus dari suatu gelombang elastik pada medium udara atau fluida. Manusia mulai memperhatikan suara sejak lama, bahkan alat musik yang menghasilkan suara sudah ada di negara Mesir, yang kemudian dikembangkan

secara terstruktur oleh Al-Farabi, Al-Kindi dan masyarakat China. Sebuah kenyataan yang cukup unik bahwa pada awalnya, musik adalah sebuah disiplin ilmu yang mempelajari suara dan bunyi-bunyian, kemudian oleh Al-Farabi digolongkan ke dalam ilmu hitung dan bukan ilmu seni [1].

Suara merupakan suatu hal yang unik dan memiliki range frekuensi tertentu dan intensitas suara yang bisa dan tidak bisa didengar oleh manusia. Satuan untuk mengukur intensitas suara tersebut adalah *desibel (dB)* diambil dari nama penemunya yaitu Alexander Graham Bell yang dikenal sebagai penemu telepon, sedangkan satuan dari frekuensi suara adalah Hertz, diambil dari nama seorang Fisikawan, Heinrich Rudolf Hertz untuk menghargai jasa atas kontribusinya dalam bidang elektromagnetik.

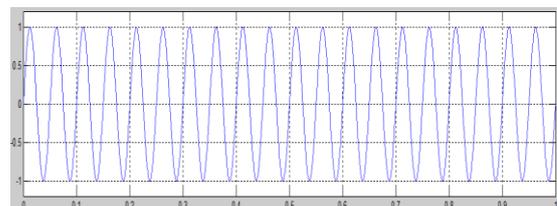
Suara yang sudah tersimpan dapat diubah ke dalam berbagai format audio seperti mp3, wav, flac, real audio, midi, dan sebagainya. Format penyimpanan suara yang dipakai dalam penelitian ini yaitu format wav karena hanya Matlab yang bisa menyimpan format suara dalam bentuk wav.

C. Fast Fourier Transform

FFT (*Fast fourier transform*) adalah suatu metoda untuk mentransformasikan sinyal suara menjadi sinyal frekuensi, artinya proses perekaman suara disimpan dalam bentuk digital berupa gelombang *spectrum* suara berbasis frekuensi.

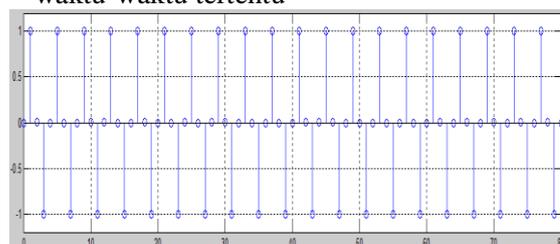
Ada beberapa klasifikasi sinyal yang ada :

1. Sinyal waktu kontinu, yaitu terdefinisi pada setiap waktu



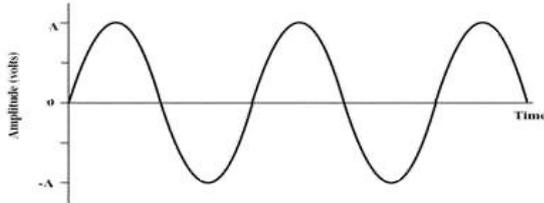
Gambar 1. Sinyal kontinu

2. Sinyal waktu diskrit, yaitu terdefinisi pada waktu-waktu tertentu



Gambar 2. Sinyal diskrit

3. Sinyal analog, yaitu sinyal waktu kontinu dengan amplitudo yang kontinu



Gambar 3. Sinyal analog

4. Sinyal digital, yaitu sinyal waktu diskrit dengan amplitudo bernilai diskrit



Gambar 4. Sinyal digital

Metoda FFT ini akan memecah (*desimasi*) N buah titik pada Transformasi Fourier diskrit menjadi dua (N/2) titik transformasi, kemudian memecah lagi (N/2) titik menjadi (N/4) titik, kemudian satu kumpulan dari nilai-nilai berindeks genap dan nilai-nilai berindeks ganjil dan seterusnya hingga diperoleh titik minimum. Contoh untuk desimasi 16 titik diperlihatkan pada

gambar 5. Jika mempunyai N titik maka akan menghasilkan $^2\log N$ tingkat sampai mendapat satu titik. Untuk N=16, berarti memerlukan $^2\log 16=4$ tingkat, untuk N=512 memerlukan tujuh tingkat, dan seterusnya. Ilustrasi dari *desimasi Fast fourier Transform* untuk N=16 ditunjukkan melalui Gambar 5[2].

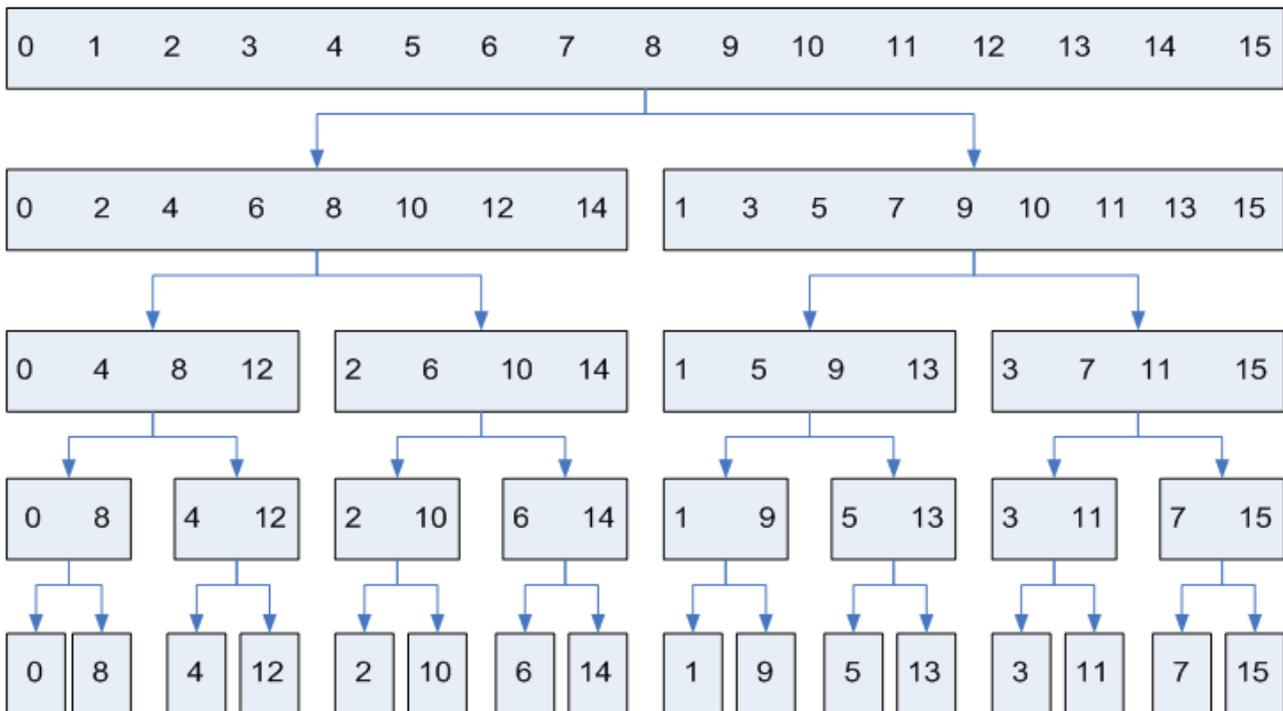
Jika melakukan *desimasi* seperti Gambar 5 akan terasa sulit, maka ada cara lain untuk mendapatkan desimasi sampai satu titik adalah dengan melakukan pembalikan bit, satu desimal direpresentasikan dengan empat bit[2].

Dengan menggunakan rumusan di bawah ini, maka diperoleh :

$$X[k] = \sum_{\substack{n=1 \\ n \text{ genap}}}^{N-2} x[n] \cdot W_N^{kn} + \sum_{\substack{n=2 \\ n \text{ ganjil}}}^{N-1} x[n] \cdot W_N^{kn} \tag{1}$$

dengan:

x[n] menyatakan berapa titik sinyal yang dihitung; x[k] adalah jumlah sinyal; W_N^{kn} merupakan faktor sinyal.



Gambar 5. Desimasi sinyal

D. Mikrofon

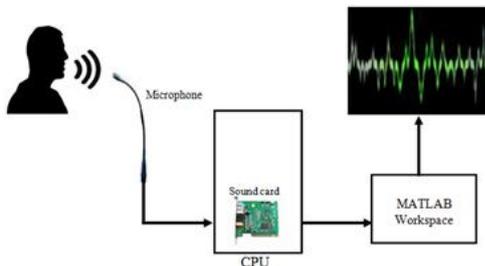
Mikrofon yang digunakan dalam riset ini terbuat dari bahan kristal aktif, dan bahan ini dapat menimbulkan tegangan sendiri saat menangkap adanya getaran dari luar sehingga tidak membutuhkan pencatu daya.



Gambar 6. Mikrofon Piezoelektris

III. Metode dan Teknik Pengukuran

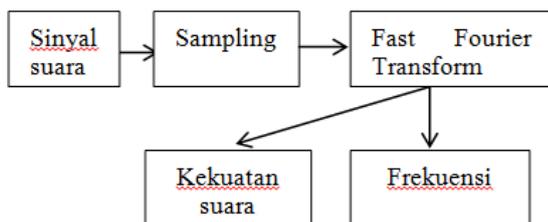
Sistem pengambilan suara pada riset ini diawali pengambilan sinyal suara langsung memakai *microphone* ke PC dan filenya disimpan dalam format wav. Setelah itu dapat dilihat bentuk frekuensi dan sebagainya dari visualisasi yang terlihat dalam sistem aplikasi yang telah dirancang. Pemrograman proses perekaman suara memakai program berbasis Matlab.



Gambar 7. Ilustrasi cara kerja aplikasi

Setelah mengambil sinyal suara melalui *microphone*, didalam CPU komputer suara diubah menjadi suara digital dan diolah oleh sistem aplikasi Matlab sehingga ditampilkan sinyal frekuensinya pada GUI aplikasi yang sudah dibuat.

Untuk memudahkan memahami gambar 7, maka berikut ditampilkan diagram blok cara kerja dari proses perekaman sampai ditampilkan bentuk frekuensi dari data suara yang sudah disimpan adalah sebagai berikut:



Gambar 8. Diagram blok cara kerja

E. Flowchart Perancangan Program

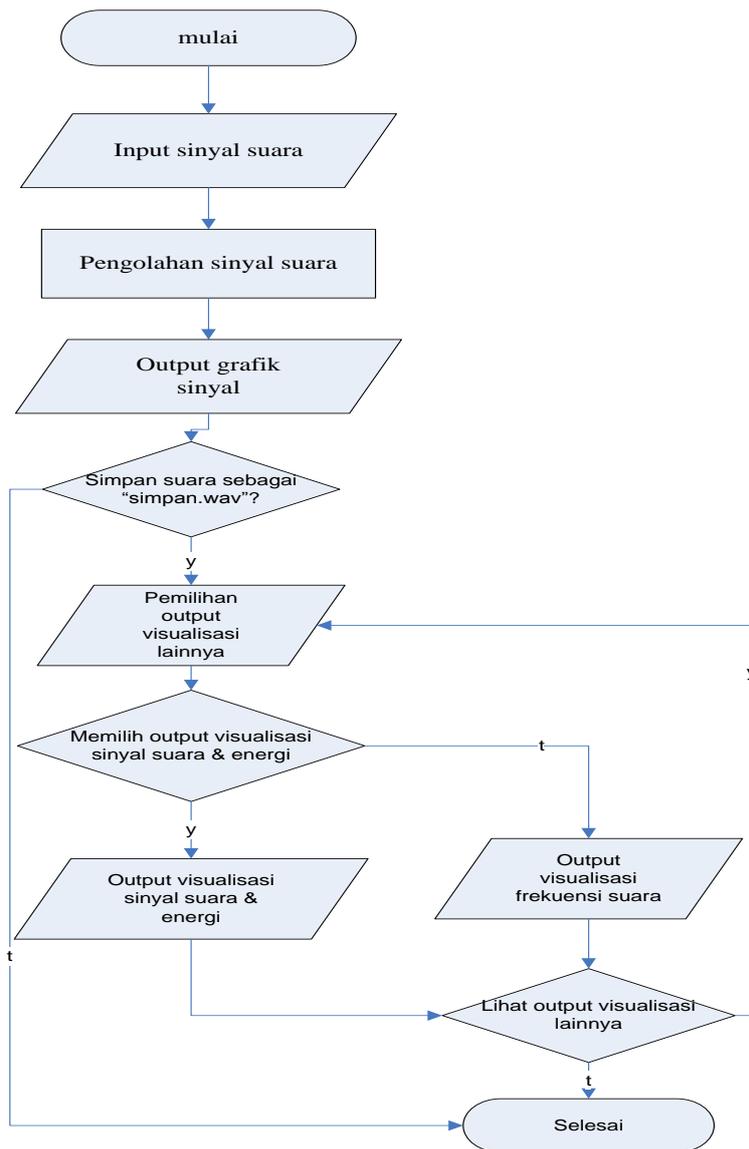
Model ini dipergunakan untuk memodelkan perangkat lunak yang dibuat. Model *Flowchart* adalah suatu metodapenggambaran secara grafik dari langkah-langkah dan urutan prosedur dari suatu sistem program yang dibuat. Dengan *flow chart*, proses dari alur program yang dibuat digambarkan dengan sederhana menggunakan simbol yang mudah untuk dimengerti. *Flowchart* program ditunjukkan pada Gambar 9.

F. Perancangan Antar Muka Sistem

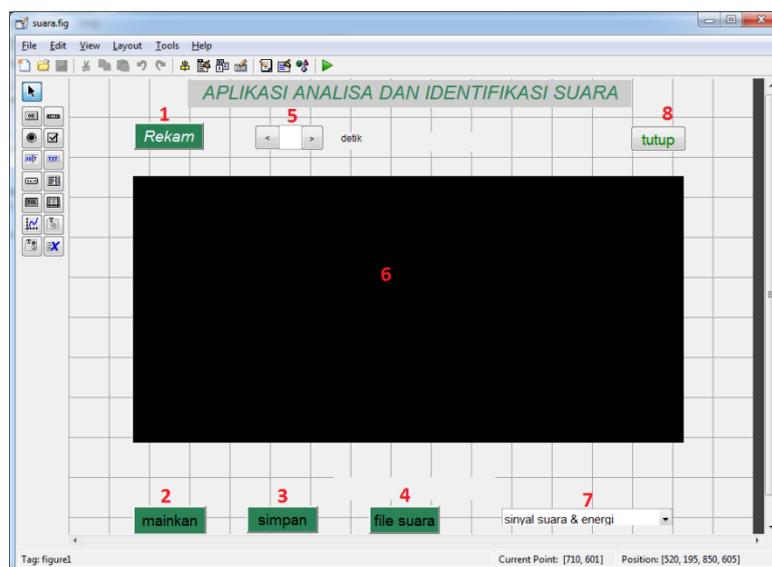
Perancangan dilakukan dengan menggunakan tools yang sudah tersedia di Matlab, tools tersebut tidak dapat diakses langsung tetapi harus diketikkan '*guide*' sehingga muncul window untuk membuat dan memprogram antarmuka aplikasi. Menu aplikasi yang dirancang pada tahap ini yaitu menu rekam, mainkan, simpan, ambil data, tutup aplikasi, menu pengaturan lama perekaman, tutup aplikasi dan menu untuk plot frekuensi suara. Perancangan Menu Aplikasi ditunjukkan pada Gambar 10.

Perancangan aplikasi *Graphycal User Interface (GUI)* dari gambar 10 terdiri dari delapan menu dengan keterangan sebagai berikut:

1. Menu rekam berfungsi untuk merekam sinyal suara dan menyimpannya dalam bentuk format wav.
2. Menu untuk menjalankan file data suara wav yang sudah disimpan dan menampilkan karakteristik fekuensinya pada layar visualisasi frekuensi (nomor 6).
3. Menu untuk menyimpan file hasil rekaman pada harddisk laptop.
4. Menu untuk mengambil data suara yang sudah disimpan untuk ditampilkan kembali dalam bentuk frekuensinya.
5. Menu untuk mengatur lama waktu perekaman.
6. Menu untuk visualisasi frekuensi suara.
7. Menu untuk memilih visualisasi frekuensi.
8. Menu yang berfungsi untuk keluar dari aplikasi yang dijalankan.



Gambar 9. Flowchart cara kerja program



Gambar 10. Perancangan menu aplikasi

IV. HASIL DAN PEMBAHASAN

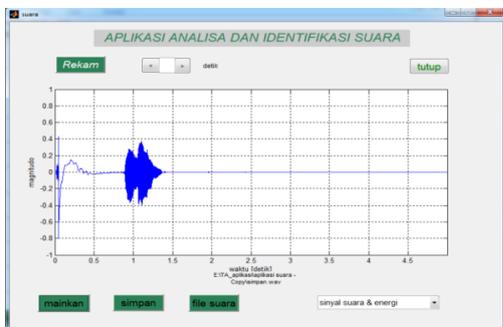
A. Pengujian

Langkah-langkah pengujian yang dilakukan yaitu setelah program Matlab dijalankan, adalah dengan mengetikkan kata “guide” pada menu *command prompt* Matlab. Kemudian tekan enter, sehingga muncullah tampilan *GUI window* yaitu *Guide Quick Start* yang terdiri dari menu tab *Create New GUI* dan *Open Existing GUI*. Pilih *open existing GUI* yang berisi aplikasi Matlab yang sudah kita buka sebelumnya atau memilih *button browse* untuk membuka aplikasi yang tersimpan di folder komputer. Untuk menjalankan aplikasi GUI yang sudah dibuat tekan tombol segitigawarna hijau di sudut kanan atas (gambar 10). Setelah itu akan tampil aplikasi Matlab yang siap untuk digunakan.

Langkah-langkah pengujian pada Gambar 10 sesuai dengan nomor yang sudah ada pada gambar tersebut yaitu:

1. Tentukan waktu lamanya perekaman (dalam satuan detik)
2. Tekan *button* rekam sehingga perekaman akan berjalan
3. Setelah itu mainkan suatu file suara yang sudah direkam
4. Tekan *button* simpan untuk menyimpan file suara
5. Untuk mengambil file suara yang sudah kita simpan
6. Berfungsi untuk menampilkan bentuk sinyal

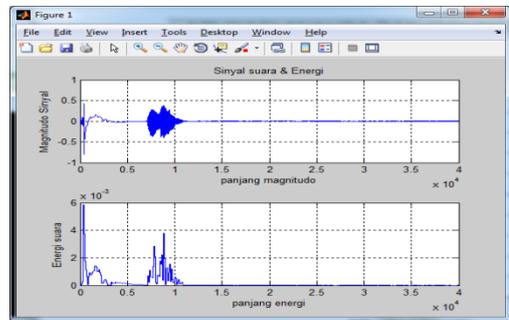
Berikut ditampilkan hasil pengujian yang sudah dilakukan. Pengujian dilakukan dengan mengambil *sampel* suara manusia dengan suatu kata “halo”, proses perekaman dilakukan selama 5 detik dengan frekuensi sampling 8 KHz.



Gambar 11. Tampilan interface

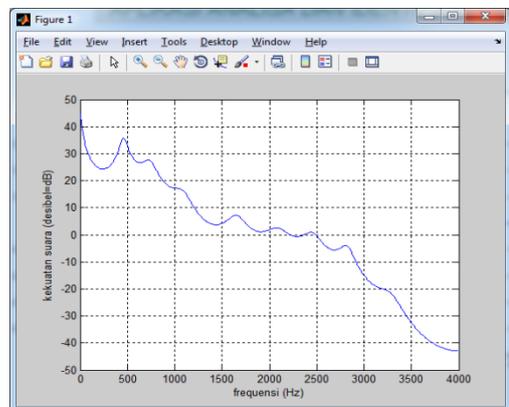
Pada Gambar 11, terlihat bentuk frekuensi dari file suara yang sudah disimpan dan lamanya waktu perekaman yang dilakukan adalah 5

detik. Dari gambar diatas dapat diambil data dan dilihat bentuk energi serta *magnitudo* dari sinyal suara yang sudah disimpan dengan menekan menu *dropdown* yang ada pada aplikasi, apabila memilih tombol *button* sinyal suara & energi maka akan ditampilkan data seperti pada Gambar 12.

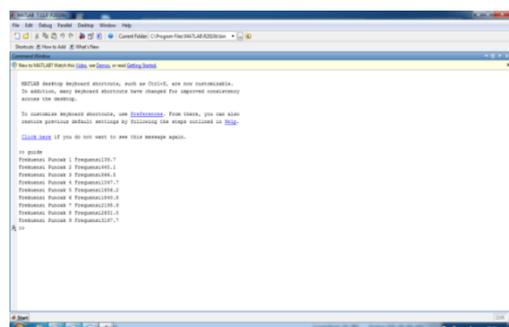


Gambar 12. Grafik sinyal suara dan energy

Pada gambar 13, dibawah terlihat tingkatan data kekuatan suara (satuan desibel) terhadap frekuensi (satuan hertz) yang secara otomatis frekuensi puncak akan muncul dan diperhalus. Grafik frekuensi tersebut muncul ketika memilih tombol *button* frekuensi suara pada *dropdown* menu, menu frekuensi suara ini juga menampilkan secara otomatis data frekuensi puncak pada *command prompt* Matlab.

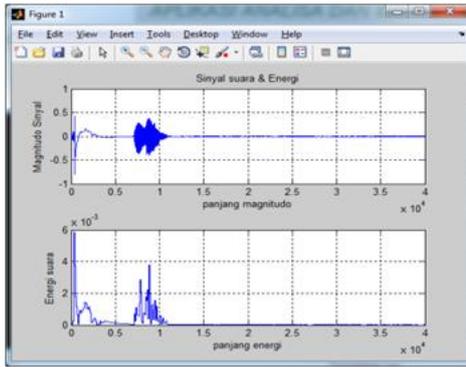


Gambar 13. Frekuensi kekuatan suara

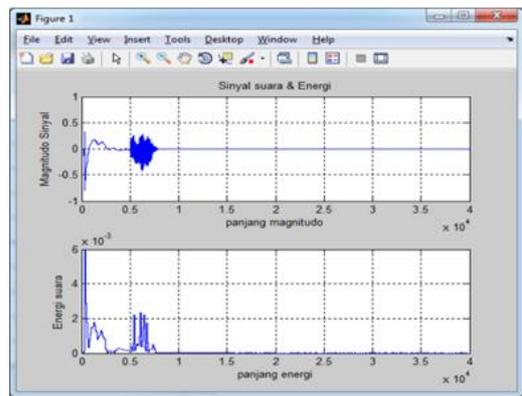
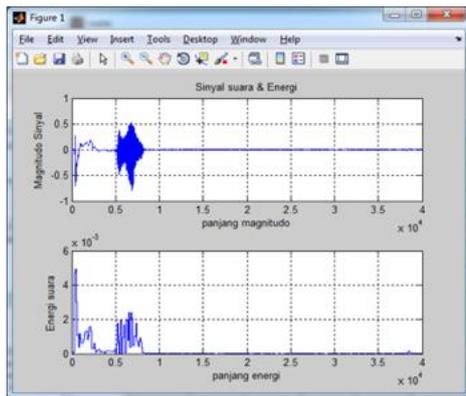
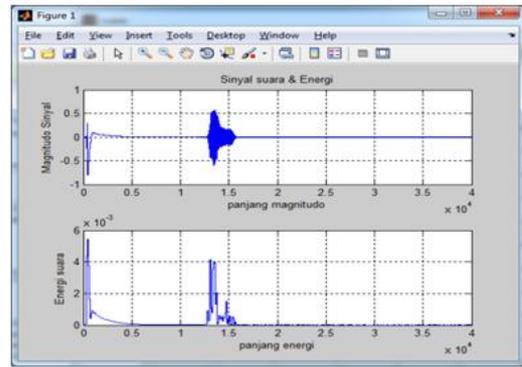


Gambar 14. Delapan buah data frekuensi suara

Grafik data ke-1

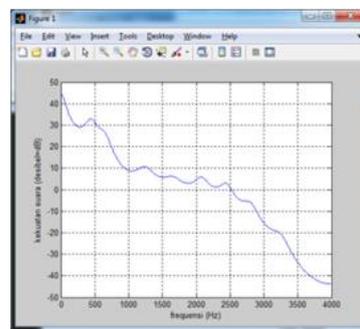
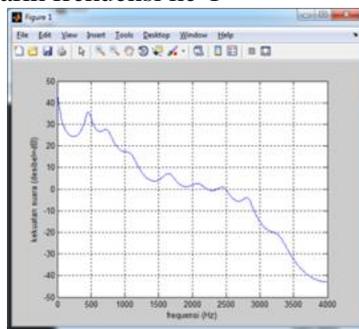


Grafik data ke-2

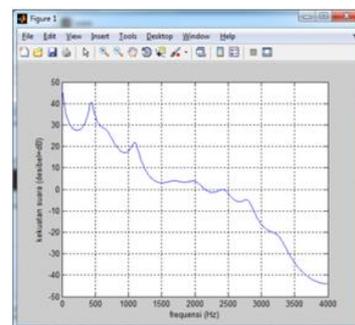
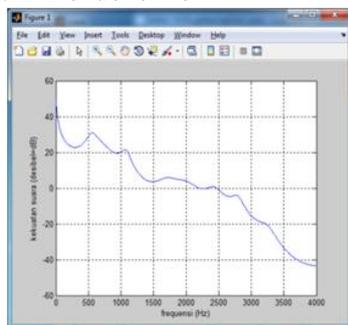


Gambar 15. Grafik sinyal suara dan energy
(a) orang ke-1 (b) orang ke-2

Grafik frekuensi ke-1



Grafik frekuensi ke-2



Gambar 16. Frekuensi suara
(a) orang ke-1 (b) orang ke-2

B. Analisa

Gambar 15 adalah bentuk *sampel* sinyal hasil pengujian yang dilakukan dengan merekam suara 10 orang. Perekaman data ke-1 dilakukan pada siang hari jam 13.00 dan perekaman data ke-2 pada malam hari jam 23.00.

Gambar 16 adalah bentuk *sampel* grafik frekuensi hasil pengujian yang dilakukan dengan merekam suara 10 orang.

Persentase selisih data ke-1 dengan data ke-2 yang bernilai positif adalah sebagai berikut:

1. Orang pertama : $\frac{6}{8} 100 = 75\%$
2. Orang kedua: $\frac{4}{8} 100 = 50\%$
3. Orang ketiga: $\frac{6}{8} 100 = 75\%$
4. Orang keempat: $\frac{6}{8} 100 = 75\%$
5. Orang kelima: $\frac{2}{8} 100 = 25\%$
6. Orang keenam: $\frac{4}{8} 100 = 50\%$
7. Orang ketujuh: $\frac{2}{8} 100 = 25\%$
8. Orang kedelapan: $\frac{3}{8} 100 = 37.5\%$
9. Orang kesembilan : $\frac{0}{8} 100 = 0.0\%$
10. Orang kesepuluh: $\frac{0}{8} 100 = 0.0\%$

Dari data di atas persentase selisih data orang ke-1, ke-3 dan ke-4 adalah 75%, paling tinggi bila dibandingkan dengan persentase yang lainnya. Sedangkan orang ke-2 dan orang ke-6 adalah 50%, orang ke-5 dan ke-7 adalah 25%, orang ke-8 37.5% dan yang terakhir orang ke-9 dan ke-10 adalah 0.0%. Yang membedakan persentase selisih masing-masing orang pastinya adalah kondisi jiwa, kondisi organ penghasil suara orang tersebut serta kondisi cuaca ditempat pengambilan data suara.

Dari persentase selisih data orang pertama sampai orang kesepuluh yang diambil siang dan malam hari seperti pada data diatas dapat diketahui bahwa total persentase data positif lebih kecil daripada data negatif.

V. KESIMPULAN

1. Telah berhasil dibuat Jaringan Syaraf Tiruan yang mampu memvisualisasikan suatu proses sinyal suara menjadi sinyal frekuensi.
2. Persentase selisih frekuensi bernilai positif adalah 75% sebanyak tiga orang yaitu orang ke-1,3, dan 4, selisih frekuensi 50% dua orang yaitu pada orang ke-2 dan ke-6, selisih frekuensi 25% ada pada data orang ke-5 dan ke-7, selisih frekuensi 37% ada pada orang ke-8, dan selisih yang terkecil yaitu 0% sebanyak dua orang pada data orang ke-9 dan ke-10.
3. Dari nilai-nilai persentase data positif, dapat disimpulkan bahwa frekuensi suara di malam hari lebih besar daripada di siang hari.

REFERENCES

- [1] Ishaq, M. (2007). Fisika Dasar. Bandung: Graha Ilmu.
- [2] Gunawan, D., & Juwono, F. H. (2011). Pengolahan Sinyal Digital Dengan Pemrograman Matlab. Jakarta: Graha Ilmu.
- [3] Siang, J. J. (2005). Jaringan syaraf tiruan dan pemrogramannya menggunakan Matlab, Penerbit Andi, Yogyakarta.
- [4] Away, G. A. (2010). Matlab Programming. Bandung: Informatika.

KETERANGAN PENULIS

Penulis dilahirkan di Bukittinggi. Menyelesaikan sekolah SD sampai SMA di kota yang sama. Kemudian tahun 1992 melanjutkan kuliah di kota Bandung. Tahun 1999 dan 2005 menyelesaikan kuliah Sarjana dan Magister di ITB, Bandung. Sejak tahun 2000, menjabat dosen di jurusan Teknik Komputer, Universitas Komputer Indonesia (Unikom) dengan Kelompok Bidang Keahlian (KBK) Kecerdasan Buatan konsentrasi pada *Image Processing* [Pengolahan Citra atau gambar]. Selain itu juga mengajar mata kuliah Fisika Dasar dan Praktikumnya, Algoritma Pemrograman dan praktikum, Pengolahan Citra, serta Elektronika Dasar dan Lanjut.