

IDENTIFIKASI SUARA PRIA DAN WANITA BERDASARKAN FREKUENSI SUARA

Binyamin Widi Prasetya⁽¹⁾, Budi Susanto⁽²⁾, Joko Purwadi⁽³⁾

Abstrak:

Suara manusia antara individu yang satu dengan yang lain memiliki perbedaan. Namun demikian, dalam perbedaan-perbedaan tersebut masih terdapat adanya kesamaan antara suara pria yang satu dengan pria yang lain dan antara wanita yang satu dengan wanita yang lain. Melalui Aplikasi ini, penulis mencoba melakukan penelitian untuk mengidentifikasi suara manusia yang sudah disimpan dalam file WAV, sehingga diketahui jenis kelamin pemilik suara tersebut pria atau wanita. Proses identifikasi yang dilakukan berdasarkan frekuensi suara ini hanya diberlakukan untuk suara orang dewasa. Proses perbandingan suara akan dilakukan berdasarkan nilai discrete file suara yang akan dicari dengan algoritma Discrete Fourier Transform dalam frekuensi 50 – 1000 Hz. Sedangkan untuk menggolongkan jenis kelamin akan dilakukan dengan menggunakan metode klasifikasi K-Nearest Neighbor.

Kata Kunci : *Suara, Frekuensi Suara, File WAV Algoritma Discrete Fourier Transform, Algoritma K-Nearest Neighbor.*

1. Pendahuluan

Tuhan menciptakan manusia menjadi dua golongan, yaitu pria dan wanita. Secara biologis, walaupun masing-masing manusia memiliki karakter tersendiri, namun tidaklah sulit untuk digolongkan berdasarkan jenis kelaminnya.

Dalam sekian banyak perbedaan yang ada pada masing-masing individu, ada beberapa kemiripan dari tiap-tiap kaum pria dan kaum wanita yang bisa diamati secara langsung. Salah satunya adalah suara mereka. Walaupun suara masing-masing individu manusia berbeda-beda, namun ada kemiripan antara suara wanita yang satu dengan wanita yang lainnya, juga suara pria yang satu dan suara pria yang lainnya. Bila didengarkan dengan seksama, suara wanita cenderung lebih tinggi bila dibandingkan dengan suara pria. Melalui perbedaan tersebut, bisa dilakukan identifikasi terhadap suara manusia melalui media komputer dengan memperhatikan perbandingan frekuensi suaranya, agar dapat digolongkan ke dalam suara pria atau wanita.

Agar hasil identifikasi suara bisa lebih akurat, maka Aplikasi dibuat dengan batasan masalah :

- 1) Sampel suara yang berlaku adalah suara orang dewasa yang dilafalkan secara normal (tidak dalam kondisi sakit / dibuat-buat) agar tingkat akurasi bisa lebih tinggi.
- 2) Input suara berasal dari hasil perekaman dengan menggunakan microphone dengan Format PCM 8.000 Hz, 8 Bit, 1 channel (mono).
- 3) Durasi waktu untuk file yang akan diproses adalah 1 detik.
- 4) Frekuensi yang digunakan untuk penghitungan nilai DFT (Discrete Fourier Transform) adalah 50 Hz – 1.000 Hz.
- 5) Inputan suara dalam proses perekaman sudah ditentukan, yaitu dengan mengucapkan huruf "a" dalam waktu kurang lebih 2 detik secara kontinyu.
- 6) Noise saat perekaman diabaikan.
- 7) Jumlah sampel suara untuk keperluan klasifikasi terdiri atas 15 suara pria dan 15 suara wanita.

2. Landasan Teori

Beberapa dasar teori yang menjadi landasan penulisan yaitu :

⁽¹⁾ Binyamin Widi Prasetya, Mahasiswa Teknik Informatika, Fakultas Teknik, Universitas Kristen Duta Wacana

⁽²⁾ Budi Susanto, S.Kom., M.T., Dosen Teknik Informatika, Fakultas Teknik, Universitas Kristen Duta Wacana

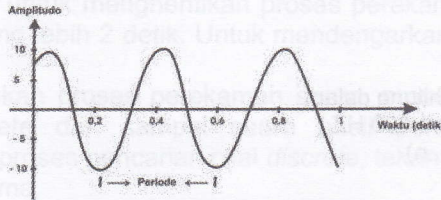
⁽³⁾ Joko Purwadi, S.Kom., M.Kom., Dosen Teknik Informatika, Fakultas Teknik, Universitas Kristen Duta Wacana

1) File WAV

Wav atau Wave adalah singkatan dari *waveform*, dan merupakan format file audio standar Windows. Format ini merupakan hasil turunan dari spesifikasi format RIFF (Resource Interchange File Format). Format RIFF sendiri diadopsi dari format IFF (Interchange File Format) milik Electronic Arts.

2) Suara

Suara atau bunyi merupakan akibat dari suatu getaran yang menghasilkan gelombang, karena perubahan tekanan udara secara cepat. Getaran yang terjadi mempunyai pola gelombang tertentu yang sering disebut dengan istilah *waveform*.



Gambar 1, Waveform

3) Frekuensi

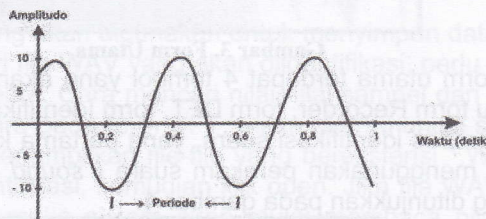
Frekuensi adalah jumlah getaran yang terjadi dalam satu detik. Frekuensi dilambangkan dengan f dan memiliki satuan Hertz (Hz) atau bisa juga dengan satuan *Cycle Per Second* (cps). Berikut ini adalah penggolongan jenis frekuensi berdasarkan besarnya :

Tabel 1. Penggolongan Frekuensi

Golongan	Frekuensi
Infra Sound	0 - 20 Hz
Human Hearing Frequency	20 Hz – 20 kHz
Ultra Sound	20 kHz – 1 GHz
Hyper Sound	1 GHz – 10 THz

4) Sinyal Analog

Sinyal analog adalah sinyal yang kontinu menurut waktu (time-domain). Sinyal Analog merupakan campuran dari sinyal Sinusoida yang jumlahnya bisa tak terbatas. Amplitudo pada sinyal analog untuk sebuah file suara bisa sama dalam interval waktu tertentu, dan bisa juga berubah-ubah sesuai dengan keras atau lemahnya suara pada interval waktu tersebut.



Gambar 2. Sinyal Analog

5) Sampling

Sampling merupakan suatu proses konversi sinyal analog menjadi sekumpulan nilai agar dapat dikenali oleh komputer. Pada proses sampling akan diambil beberapa sampel data dari sinyal analog. Tiap sampel akan mewakili satu nilai dimana posisi pergerakan grafik pada waktu tertentu. Nilai hasil sampling dimasukkan ke dalam fungsi $g(t) = n$, dimana t adalah waktu, dan n adalah nilai sampel pada waktu t .

6) Discrete Fourier Transform

Discrete Fourier Transform merupakan sebuah metode untuk mentransformasikan nilai-nilai sampel yang berbasis waktu (time-domain) menjadi berbasis frekuensi (frequency-domain).

Berikut ini adalah perumusan untuk melakukan perhitungan dengan menggunakan Discrete Fourier Transform :

$$G(f) = \sum_{t=0}^{n-1} g(t) \cos \frac{2\pi \cdot f \cdot t}{n} + i \sin \frac{2\pi \cdot f \cdot t}{n}$$

Di mana :

- G (f) : Transformasi fourier dari g(t)
- n : Jumlah sample
- i : $\sqrt{-1}$
- g(t) : Nilai sampel ke-t
- f : Nilai frekuensi (yang dihitung dalam kasus ini antara 50HZ – 1.000 HZ)
- t : Nomor sampel (0,1,2,3...n)

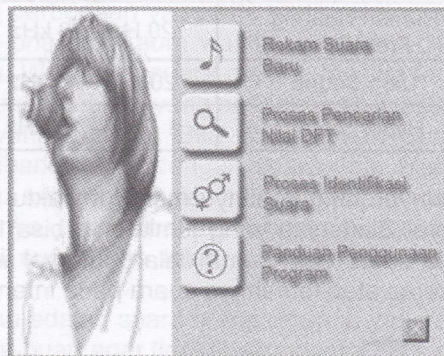
7) K-Nearest Neighbor

K-Nearest Neighbor adalah metode klasifikasi yang sistem kerjanya adalah dengan cara menghitung jarak sebuah data baru dengan data-data sampel yang sudah ada. Semakin dekat jarak antar keduanya, maka kedua data tersebut dikatakan cenderung mirip.

Berikut ini adalah rumus untuk menghitung jarak antara dua data dengan menggunakan K-Nearest Neighbor :

$$d(x_1, x_2) = \sqrt{(x_{1,1} - x_{2,1})^2 + (x_{1,2} - x_{2,2})^2 + \dots + (x_{1,n} - x_{2,n})^2}$$

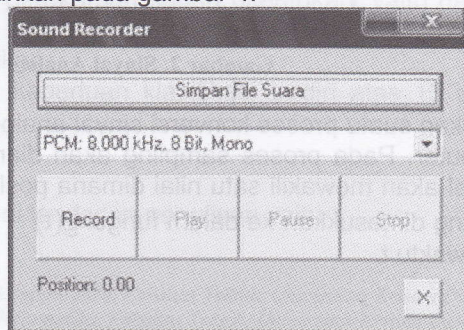
3. Implementasi Sistem



Gambar 3. Form Utama

Pada tampilan form utama terdapat 4 tombol yang akan menghubungkan user dengan form-form yang lain, yaitu form Recorder, form DFT, form Identifikasi dan form Panduan.

Untuk memulai proses identifikasi suara, yang pertama kali harus dilakukan user adalah merekam suara dengan menggunakan perekam suara (*sound recorder*) yang sudah tersedia pada sistem, seperti yang ditunjukkan pada gambar 4.



Gambar 4. Sound Recorder

Form *recording* yang digunakan dirancang dengan menggunakan *WaveAudioPackage* v1.6 yang di download dari situs <http://www.delphiarea.com> (author : Kambiz R. Khojasteh) secara cuma-cuma.

Untuk melakukan proses perekaman suara, user harus menekan tombol Simpan File Suara, kemudian inputkan nama file WAV yang akan direkam. Kemudian pilih format file WAV yang akan digunakan pada saat perekaman, seperti yang tertera pada *combobox*. Dalam kasus ini, user harus menggunakan file WAV dengan format PCM, 8.000 kHz, 8 Bit, Mono, agar proses identifikasi bisa dilakukan. Tekan tombol *record* untuk memulai proses perekaman suara, kemudian tekan tombol *stop* untuk menghentikan proses perekaman suara. Proses perekaman suara dilakukan selama kurang lebih 2 detik. Untuk mendengarkan hasil perekaman suara tekan tombol *play*.

Setelah user melakukan proses perekaman suara, proses selanjutnya yang dilakukan adalah mencari nilai discrete dari sampel suara hasil perekaman, agar nantinya bisa diidentifikasi. Untuk memulai proses pencarian nilai *discrete*, tekan tombol Proses Pencarian Nilai DFT pada tampilan Form Utama.

Gambar 5. Form DFT awal.

Dalam kasus ini, file yang akan digunakan untuk menyimpan data suara adalah file WAV. Untuk mencari nilai *discrete* dari file WAV yang akan diidentifikasi, perlu diketahui terlebih dahulu nilai-nilai sampel file WAV tersebut. Cara mencari nilai-nilai sampel dari file WAV adalah dengan menekan tombol Buka File Suara pada form DFT, maka akan muncul sebuah open dialog yang secara otomatis hanya akan menampilkan file-file yang berekstensi *.WAV saja. Pilih file WAV hasil rekaman yang akan diidentifikasi, kemudian klik open. Jika file WAV yang dipilih memenuhi kriteria yang telah ditentukan, maka sistem akan mulai membaca nilai-nilai sampel pada file WAV, yang kemudian akan ditampilkan pada *ListBox1* (bagian atas) dalam form Proses Cari DFT.

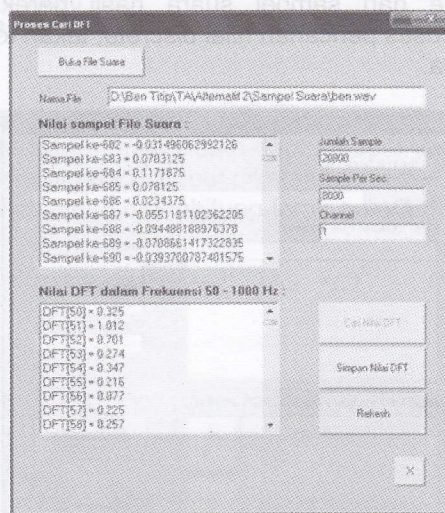
Tabel 2. Kriteria file WAV yang dipakai

Audio Format	PCM
Audio sample rate	8.000 Hz
Audio channel	1 (mono)
Audio sample size	8 bit
Jumlah sampel minimal	9.000 sampel

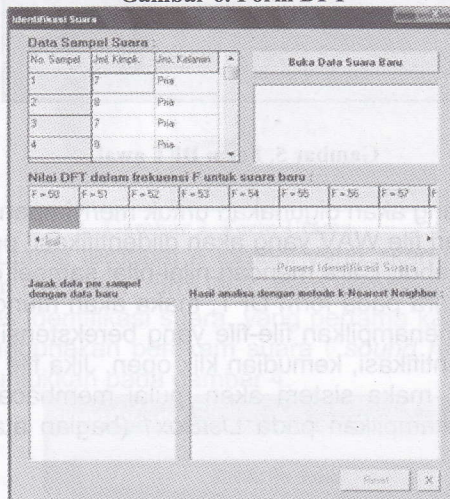
Proses selanjutnya adalah mencari nilai *discrete* dari file WAV yang sudah diketahui nilai-nilai sampelnya, dengan menekan tombol Cari Nilai DFT pada form DFT. Pada kasus ini, proses pencarian nilai *discrete* dilakukan dengan menggunakan algoritma *Discrete Fourier Transform* (DFT). Nilai *discrete* yang dicari adalah nilai *discrete* pada frekuensi 50 Hz hingga 1000 Hz. Penulis menggunakan range frekuensi tersebut karena masih termasuk dalam kategori *Human Hearing Frequency*.

Setelah proses pencarian nilai *discrete* dengan menggunakan algoritma DFT selesai, maka hasilnya ditampilkan pada *ListBox2* dalam form DFT seperti yang terlihat pada gambar 6. Simpan nilai-nilai *discrete* tersebut ke dalam file TEXT dengan cara menekan tombol Simpan Nilai DFT, beri nama file, lalu klik tombol save.

Proses berikutnya adalah proses identifikasi suara. Untuk memulai proses identifikasi suara, tekan tombol Proses Identifikasi Suara pada form Utama. Berikut ini adalah tampilan awal dari form DFT dan form Identifikasi Suara :



Gambar 6. Form DFT



Gambar 7. Form Identifikasi Suara

Pada saat form Identifikasi Suara mulai dijalankan, program juga secara otomatis membuka data sampel suara yang tersimpan dalam database dalam bentuk file TEXT. Dalam data sampel suara terdapat 3 buah *field* yaitu nomor sampel, jumlah kelompok dan jenis kelamin. Data tersebut merupakan data dari sampel suara yang terdiri dari 15 sampel suara pria dan 15 sampel suara wanita.

Yang dimaksud dengan jumlah kelompok dalam data sampel suara adalah jumlah kelompok yang terbentuk dari nilai-nilai DFT dengan nilai yang relatif tinggi yang bergerombol dalam frekuensi tertentu.

Selanjutnya, buka data suara baru yang berupa nilai-nilai *discrete* yang disimpan dalam file TEXT. Untuk melakukan proses tersebut tekan tombol Buka Data Suara Baru, kemudian akan muncul *Open Dialog* yang hanya akan menampilkan file dalam format *.txt saja. Pilih file TEXT yang datanya akan diidentifikasi, kemudian klik tombol Open.

Data Sampel Suara

No. Sampel	Jml. Klompok	Jns. Kelamin
1		Pria
2	8	Pria
3	7	Pria
4	8	Pria

Buka Data Suara Baru

Keterangan Data Suara Baru

Jumlah Kelompok : 8
Jenis Suara : ?

Nilai DFT dalam frekuensi F untuk suara baru :

F = 50	F = 51	F = 52	F = 53	F = 54	F = 55	F = 56	F = 57
1.012	0.701	0.274	0.347	0.216	0.077	0.225	

Proses Identifikasi Suara

Jarak data per sampel dengan data baru

Hasil analisa dengan metode k-Nearest Neighbor

Reset

Gambar 8. Form Identifikasi Suara Setelah Proses Buka Data Suara Baru

Proses selanjutnya adalah melakukan identifikasi suara. Untuk melakukan proses identifikasi suara, tekan tombol Proses Identifikasi Suara pada form Identifikasi Suara. Dalam proses identifikasi suara akan dilakukan klasifikasi data dengan menggunakan metode *K-Nearest Neighbor*. Atribut yang akan digunakan sebagai acuan dalam proses klasifikasi hanya ada satu, yaitu atribut Jumlah Kelompok.

Data Sampel Suara

No. Sampel	Jml. Klompok	Jns. Kelamin
1		Pria
2	8	Pria
3	7	Pria
4	8	Pria

Buka Data Suara Baru

Keterangan Data Suara Baru

Jumlah Kelompok : 8
Jenis Suara : "PRIA"

Nilai DFT dalam frekuensi F untuk suara baru :

F = 50	F = 51	F = 52	F = 53	F = 54	F = 55	F = 56	F = 57
1.012	0.701	0.274	0.347	0.216	0.077	0.225	

Proses Identifikasi Suara

Jarak data per sampel dengan data baru

Hasil analisa dengan metode k-Nearest Neighbor

Sampel ke - 1 = 1
Sampel ke - 2 = 0
Sampel ke - 3 = 1
Sampel ke - 4 = 0
Sampel ke - 5 = 1
Sampel ke - 6 = 2
Sampel ke - 7 = 2
Sampel ke - 8 = 0
Sampel ke - 9 = 1
Sampel ke - 10 = 1
Sampel ke - 11 = 0
Sampel ke - 12 = 0
Sampel ke - 13 = 0
Sampel ke - 14 = 1

Sampel terdekat ke - 4 :
Sampel ke - 11
Jarak dengan data baru : 0
Jenis suara : Pria

Sampel terdekat ke - 5 :
Sampel ke - 12
Jarak dengan data baru : 0
Jenis suara : Pria

Kesimpulan :
Jenis suara anda tergolong suara : PRIA

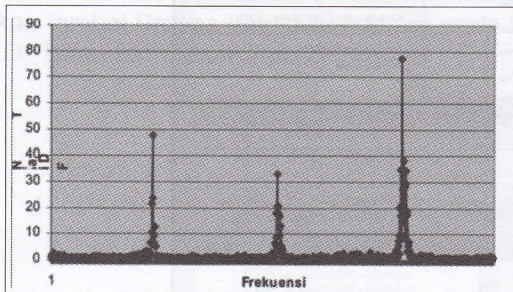
Reset

Gambar 9. Tampilan Form Identifikasi Suara Sesudah Proses Identifikasi

Setelah proses identifikasi suara selesai, tampilan form Identifikasi Suara akan berubah seperti pada gambar 9. Pada *ListBox1* akan ditampilkan jarak data per sampel dengan data suara baru, sedangkan pada *ListBox2* ditampilkan 5 buah data yang jaraknya paling dekat dengan data suara baru. Untuk jenis suara baru, hasilnya akan ditampilkan di *ListBox2* dan *ListBox3*.

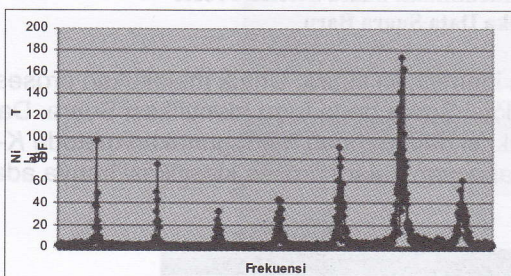
4. Analisis Sistem

Untuk melakukan analisa, penulis memasukkan data nilai DFT yang sudah disimpan dalam file TEXT ke dalam Microsoft Office Excel, sehingga diperoleh sebuah grafik seperti pada gambar 10.

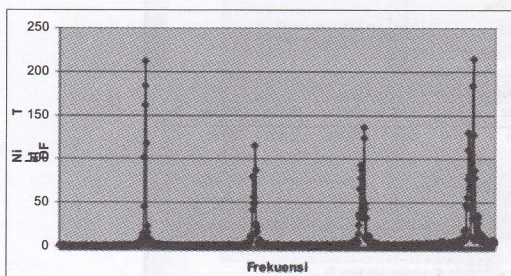


Gambar 10. Grafik Nilai DFT

Dari grafik tersebut dapat dilihat adanya perbedaan yang sangat signifikan antara sampel suara pria dan sampel suara wanita. Pada sampel suara pria jumlah peak (dalam kasus ini disebut sebagai jumlah kelompok) akan lebih banyak jika dibandingkan dengan jumlah peak yang ada pada sampel suara wanita.

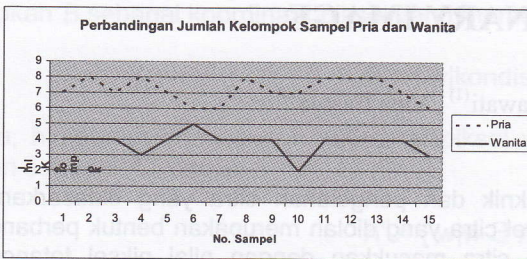


Gambar 11. Grafik Nilai DFT suara pria



Gambar 12. Grafik Nilai DFT suara wanita

Dari penelitian sejumlah 30 sampel suara yang terdiri atas 15 sampel suara pria dan 15 sampel suara wanita diperoleh perbandingan sebagai berikut :



Gambar 13. Gambar perbandingan jumlah Kelompok pria dan wanita

5. Kesimpulan

- Untuk mempermudah pengolahan file suara dengan komputer, format suara yang berbasis waktu (*time domain*) sebaiknya diubah menjadi berbasis frekuensi (*frequency domain*) salah satu caranya adalah dengan menggunakan algoritma *Discrete Fourier Transform* (DFT).
- Jumlah pengelompokan nilai DFT (*Discrete Fourier Transform*) pada sampel suara pria lebih banyak jika dibandingkan dengan pengelompokan nilai DFT pada sampel wanita. Oleh karena itu, proses identifikasi dapat dilakukan berdasarkan perbedaan tersebut.
- Pada kasus tertentu, tidak menutup kemungkinan suara seorang pria akan terdeteksi sebagai suara wanita, dan juga sebaliknya. Hal tersebut dapat terjadi karena secara Biologis terdapat beberapa kemiripan, misalnya warna suara, frekuensi suara, susunan pita suara, dan lain-lain.

6. Daftar Pustaka

- Benford, Tom, *Welcome to... PC Sound, Music and MIDI*. (New York : Management Information Source, Inc, First Edition - 1993)
- Juhara, Zamronny P., *Pemrograman DirectSound*, PC Media terbitan bulan Agustus 2007.
- Penerbit Andi, Wahana Komputer, *Panduan Praktis Pemrograman Borland Delphi 7.0*, Yogyakarta : Andi Offset.
- Steinmetz, Ralf, Nahrstedt, klara, *Multimedia: Computing, Communications and Applications*. (New Jersey : Prentice-Hall, Inc)
- _____, "*discrete-fourier-transform*", dalam <http://www.answers.com/topic/discrete-fourier-transform.htm>; 11 November 2007.
- _____, "*K-nearest neighbor algorithm*", dalam <http://kiew.cs.uni-dortmund.de:8001/mlnet/instances/81d91e93-df4da3c279>;

13 Maret 2008