

**ANALISIS DAN SIMULASI IDENTIFIKASI JUDUL LAGU DARI SENANDUNG
MANUSIA MENGGUNAKAN EKSTRAKSI CIRI DCT
(DISCRETE COSINE TRANSFORM)
ANALYSIS AND SIMULATION IDENTIFICATION OF SONG'S TITLE FROM
HUMAN'S HUMMING USING DCT (DISCRETE COSINE TRANSFORM) FEATURE
EXTRACTION**

Arintyo Archamadi^[1], Ir.Rita Magdalena, MT.

.^[2], I Nyoman Apraz Ramatryana, ST.,MT.^[3]

^{1,2,3}Prodi S1 Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom, Bandung
arintyoarchamadi@yahoo.co.id, ritamagdalen@telkomuniversity.ac.id, ramatryana@telkomuniversity.ac.id

Abstrak

Seiring perkembangan teknologi banyak orang menginginkan sebuah pencarian yang cepat dan praktis. Musik merupakan kumpulan kata yang dapat dinyanyikan bersama irama sebuah nada, sehingga banyak orang dapat menikmatinya. Namun sering terkendala dengan judul lagu yang akan dinyanyikan, dikarenakan hanya hafal beberapa liriknya dan tidak mengetahui judul lagunya, sehingga akan membuat penikmat lagu kesudahan apabila akan menikmati lagu yang mereka maksud. Dengan adanya permasalahan tersebut peneliti telah membuat sebuah solusi yaitu pengaplikasian pengenalan suara yang berupa identifikasi judul lagu melalui *humming* atau senandung manusia. Pada penelitian sebelumnya, pernah dirancang sistem pengenalan suara menggunakan senandung manusia berbasis jaringan syaraf tiruan *back-propagation*.

Pada tugas akhir ini, dilakukan penelitian dengan cara memasukkan suara *humming* untuk mengidentifikasi judul lagu yang sesuai. Maka penulis menerapkan metode ekstraksi ciri dengan metode *Discrete Cosine Transform* (DCT) serta menggunakan klasifikasi DTW (*Dynamic Time Warping*) yang berfungsi untuk mengukur kemiripan pola nada dan membandingkan nilai hasil pengujian ekstraksi ciri yang ditangkap *microphone* dengan suara latihan yang digunakan sebagai *database*. Pada proses klasifikasi data lagu pada database yang memiliki vokal yang lebih menonjol sangat menentukan kecocokan senandung terhadap nada lagu.

Dengan adanya tugas akhir ini dihasilkan sebuah simulasi pengolahan sinyal suara yang dapat mengklasifikasi judul lagu melalui nada masukan berupa *humming* dan menghasilkan keluaran berupa nama judul lagu yang sesuai dengan *database* setelah dilakukan ekstraksi ciri *Discrete Cosine Transform* (DCT). Untuk pengujian data uji dan latihan menggunakan lagu asli memiliki akurasi tertinggi sebesar 100 %. Dan akurasi 80,06% dengan data uji berupa *humming* sejumlah 30 judul lagu dan data latihan berupa lagu asli sejumlah 50 judul lagu. Sedangkan dengan data uji *humming* sejumlah 30 judul lagu dan data latihan berupa lagu dengan file suara orang bernyanyi tanpa menggunakan *instrument* (vokal) sejumlah 50 judul lagu menghasilkan akurasi 92,3%.

Kata Kunci : Senandung Manusia, DCT (*Discrete Cosine Transform*), *Dynamic Time Warping*

Abstract

Along with the development of technology and more people want a quick and practical search. Music is a collection of words that can be sung along with the rhythm of a tone, so that many people can enjoy it. But often constrained by the title of the song that will be sung, because only memorized some of the lyrics and did not know the title of the song, so it would make trouble if the song lovers will enjoy the songs they mean. Given these issues, researchers will create a solution that is in the form of voice recognition application identification through humming the title song or human's humming. In the previous research, has been designed voice recognition system using human's humming based on neural network back-propagation. So in this research the authors will make an identification of song's title from human's humming and compare that with the previous research.

In this thesis, the research done by entering a humming sound to identify the appropriate title track. The authors apply the method of feature extraction methods *Discrete Cosine Transform* (DCT) and using the classification of *Dynamic Time Warping* which serves to measure the similarity of tone patterns and comparing

the value of feature extraction test results are captured microphone with a voice trainer that is used as a database. In the classification process length of the frame and track data on a database that has a more prominent vocal determine the suitability of chanting the tune.

Given this thesis produced by a simulated voice signal processing to classify the title track through humming tones as input and produce output in the form of the name of song titles that match the database after feature extraction Discrete Cosine Transform (DCT). Test data to test and train using the original song has the highest accuracy of 100%. And the accuracy of 80.06% with test data in the form of humming some song titles and 30 training data in the form of the original song number 50 song titles. While the 30 test data humming some song titles and 50 song' titles training data in the form of a song with the singing sound files without using the instrument (vocal) resulted accuracy of 92.3%.

Keywords : Human humming, DCT (Discrete Cosine Transform), Dynamic Time Warping

1. Pendahuluan

Lagu merupakan suatu kesatuan musik yang terdiri atas susunan nada yang berurutan. Setiap lagu ditentukan oleh panjang-pendek dan tinggi-rendahnya nada, lagu juga merupakan seni yang melukiskan pemikiran dan perasaan manusia lewat keindahan suara. Terkadang kita ingin menyanyikan sebuah lagu tetapi tidak mengetahui judul lagu, sehingga hanya dapat bersenandung dan lirik yang kurang tepat. Maka perlu adanya simulasi untuk mengidentifikasi judul lagu melalui *humming* / bersenandung. Dengan perkembangan teknologi saat ini maka perlu adanya pemanfaatan untuk menerapkan suatu penelitian dengan menciptakan simulasi yang dapat mengidentifikasi judul lagu. Pada penelitian sebelumnya (Iqbal Tawakal, 2014)^[6] pernah dirancang sistem pengenalan suara menggunakan senandung manusia berbasis jaringan syaraf tiruan *back-propagation*.

Sehingga dari permasalahan tersebut peneliti merancang simulasi identifikasi judul lagu yang memanfaatkan masukan berupa *humming*. Simulasi ini memerlukan ekstraksi ciri DCT (*Discrete Cosine Transform*), yang merupakan teknik untuk merubah suatu sinyal menjadi komponen frekuensi dasar, dengan memperhitungkan nilai riil dari hasil transformasi, sedangkan transformasi ideal digunakan untuk proses kompresi data, dengan kemampuan memampatkan data hingga 99% lebih kecil dari sinyal masukannya. Kemudian untuk klasifikasi akan digunakan metode *Dynamic Time Warping* , cara kerja metode ini untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran dengan mengukur kemiripan pola paling dekat dengan objek tersebut. Metode ini akan mengklasifikasikan setiap suara dengan masukan nada *humming* yang akan direkam menggunakan *microphone* dan kemudian akan dicocokkan dengan kumpulan lagu yang disimpan di *database*, setelah itu program akan menampilkan keluaran hasil berupa judul lagu yang sesuai.

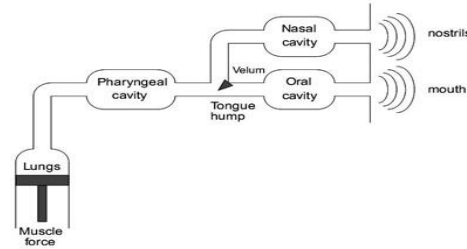
Dengan adanya simulasi ini dapat membantu pengguna untuk mendapatkan informasi judul lagu yang diinginkan sesuai dengan masukan nada senandung yang sudah teridentifikasi.

2. Dasar Teori dan Metodologi Perancangan

2.1 Dasar teori

a. Produksi Suara [7]

Produksi Suara Gelombang suara bervariasi sebagaimana variasi tekanan media perantara seperti udara. Suara diciptakan oleh getaran dari suatu obyek, yang menyebabkan udara disekitarnya bergetar. Yang pertama adalah Produksi suara karena suara adalah mekanisme terjadinya suara. suara manusia dapat di golongan kedalam kelompok alat musik tiup. sebelum menyanyi, kita harus memompa atau menghirup udara melalui hidung (*inhalasi*) masuk ke dalam paru – paru di bantu oleh otot perut, otot dada, otot sisi tubuh dan otot diafragma. kemudian paru – paru mengalirkan kembali udara keluar atau di hembuskan (ekshalasi) sedemikian rupa sehingga membentur pita suara yang terdapat di dalam *larynx* (tenggorokan) bentuk pita suara ini seperti selaput yang berbelah di bagian tengahnya. pita suara terbuka pada saat menghirup udara dan akan menutup dan bergetar pada saat kita bersuara (bernyanyi/berbicara) menjadi suara yang jelas dan indah di dalam rongga mulut. sebenarnya pita suara ini tidak menutup secara total tetapi masih ada celah kecil sehingga akibat tekanan udara dari bawah membuat pita suara ini bergetar. getaran ini di perkuat dan di perbesar oleh rongga resonansi yang ada pada tubuh kita .



Gambar 1. Alur Kerja Produksi Suara Manusia

b. Struktur Lagu [8]

Struktur lagu merupakan susunan dan hubungan antara unsur-unsur musik dalam suatu lagu sehingga menghasilkan suatu lagu yang bermakna. Bentuk atau struktur lagu itu merupakan susunan dan hubungan antara unsur musik dalam suatu lagu yang menghasilkan suatu komposisi. Lagu memiliki tujuh komponen berikut struktur yang terdapat pada lagu.

b1. Intro

Intro adalah awal dari sebuah lagu, yang merupakan pengantar lagu. Intro juga berfungsi memberikan waktu untuk penyanyi dan pendengar mempersiapkan diri sebelum lagu benar-benar dimainkan. Musik pengisi intro berupa musik instrumental atau suara vokal yang nadanya diambil dari reff lagu. Intro pada lagu pun terdiri dari 3, yaitu intro awal, intro tengah dan intro akhir.

b2. Verse/Bait

Verse adalah pengantar sebuah lagu sebelum lagu masuk ke bagian chorus. Verse sering disebut adalah bagian “basa-basi” dari sebuah lagu. sebuah lagu yang baik bahkan memiliki verse yang kuat secara melodik dan harmonik yang tidak kalah dengan bagian reff-nya. Dan rata-rata detail lagu terletak di bagian Verse yang menceritakan lebih banyak daripada bagian Reff/Chorus.

b3. Chorus/Reff

Chorus/reff adalah bagian pada lagu yang sering diulang-ulang dan merupakan bagian utama dari sebuah lagu. *chorus* merupakan bagian yang paling ditunggu-tunggu dalam sebuah lagu dan klimaks dari sebuah lagu. *Chorus* memiliki nilai excitement yang lebih tinggi dari Verse. Sementara Reff lebih sederhana dari pada *Chorus*. Reff yang berisi pengulangan kata yang biasanya menggunakan bagian lain dari lagu (biasanya Verse) untuk diulang di bagian ini. Notasi pengulangannya dan syairnya sama.

b4. Bridge

Bridge merupakan sebuah bagian lagu yang bukan merupakan verse atau chorus. Bridge ini biasanya dipakai untuk menghubungkan antara bagian-bagian lagu. misalnya menghubungkan antara chorus dengan verse, atau menghubungkan antara reff dengan *chorus*.

b5. Ending

Ending adalah merupakan bagian penutup dari sebuah lagu. *ending* berfungsi agar lagu berakhir lancar, *smooth* (mulus), dan tidak berhenti secara mendadak. Dapat merupakan bagian intro yang diulang, dapat juga berupa bagian akhir lagu yang diulang-ulang, atau dapat juga berupa instrumen yang berbeda yang sengaja dibuat untuk akhiran dari lagu.

c. Discrete Cosine Transform [1]

Discrete Cosine Transform (DCT) merupakan suatu teknik untuk merubah suatu sinyal menjadi komponen frekuensi dasar, dengan memperhitungkan nilai riil dari hasil transformasi. Discrete Cosine Transform merupakan transformasi yang berhubungan dengan transformasi Fourier yang memberikan fungsi diskrit dengan hanya mengambil nilai cosinus dari eksponensial kompleks. *Discrete Cosine Transform* adalah transformasi ideal untuk proses kompresi data, dengan kemampuan memampatkan data hingga 99% lebih kecil dari sinyal masukannya. DCT dapat digunakan untuk mendapatkan nilai koefisien. Hal ini dikarenakan DCT mampu menghasilkan sinyal berfrekuensi rendah lebih sedikit dan frekuensi tinggi yang banyak. DCT merekonstruksi urutan data dengan DCT koefisien, parameter yang berguna yang digunakan untuk reduksi data. Persamaan DCT dapat dilihat pada persamaan di bawah.

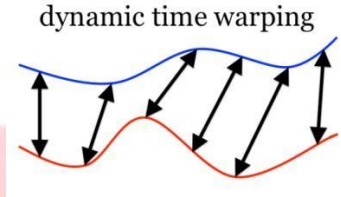
$$y_k = \sum_{n=0}^{N-1} x_n \cdot \frac{\cos\left(\frac{(n+\frac{1}{2})k\pi}{N}\right)}{2}$$

Dimana $\left(\begin{matrix} \\ \end{matrix} \right)$ $\left\{ \begin{matrix} \\ \end{matrix} \right.$ k=1

$$- -2k N \dots\dots\dots (1)$$

d. Dynamic Time Warping [2] [9]

Dynamic Time Warping algorithm (DTW) adalah algoritma yang menghitung *optimal warping path* antara dua waktu. Metode ini untuk menentukan kecocokan optimal antara dua deret waktu dengan beberapa pembatasan tertentu. Berikut ini contoh gambaran pencocokan 2 deret:



Gambar 2. Penyelarasan Sistem DTW

Algoritma ini menghitung baik antara nilai *warping path* dari dua waktu dan jaraknya. Misalnya, kita memiliki dua sekuens numerik (a₁, a₂, ..., a_m) dan (b₁, b₂, ..., b_m). Dengan pemisalan ini, maka dapat dikatakan bahwa panjang dua sekuens ini bisa saja berbeda. Algoritma ini memulai dengan penghitungan jarak lokal antara elemen dari sekuens menggunakan tipe jarak yang berbeda. Frekuensi yang paling banyak menggunakan metode untuk penghitungan jarak adalah jarak absolut antar nilai dua elemen. Jika dalam matriks maka dapat ditulis dengan memiliki n garis dan m kolom, secara umum:

$$() \sqrt{\sum ()} \dots\dots\dots (2)$$

Langkah selanjutnya merupakan point-point yang harus kita lewati untuk menemukan *warping path*.

- Baris pertama

$$() \sum () , - \dots\dots\dots (3)$$

- Kolom pertama

$$() \sum () , - \dots\dots\dots (4)$$

- Elemen lainnya

$$() () * () () + , - , - \dots (5)$$

Warping path merupakan jalur/lintasan atau *path* yang melalui matriks yang berisi jarak minimal (*minimal distance*) dari elemen () hingga elemen () yang terdiri dari elemen-elemen () itu sendiri, ongkos *warping path* secara global dari dua sekuen.

$$GC = \sum$$

Dimana *W_i* adalah elemen yang dimiliki *warping path* dan p adalah jumlahnya. Penghitungannya dibuat untuk dua sekuens diperlihatkan pada gambar dibawah dan *warping path* diberi *highlight*.

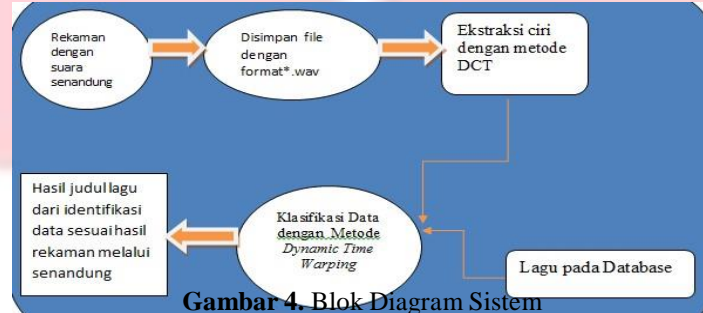
	-2	10	-10	15	-13	20	-5	14	2
3	5	12	25	37	53	70	78	89	90
-13	16	28	15	43	37	70	78	105	104
14	32	20	39	16	43	43	62	62	74
-7	37	37	23	38	22	49	45	66	71
9	48	38	42	29	44	33	47	50	57
-2	48	50	46	46	40	55	36	52	54

Gambar 3. *Warping Path*

2.2 Metodologi Perancangan

a. Blok Diagram Sistem

Perancangan sistem ini yaitu proses identifikasi lagu berdasarkan *humming* yang memiliki beberapa tahap, tahap yang pertama yaitu merekam sebuah lagu dengan bersenandung yang menggunakan *recorder* dengan format rekaman *.wav dan frekuensi samplingsnya 44,1 KHz. Tahap selanjutnya yaitu melakukan ekstraksi ciri pada hasil rekaman dengan menggunakan metode *Discrete Cosine Transform (DCT)* yang memiliki beberapa tahapan yaitu : *pre-emphasis*, *sampling*, *frame blocking*, *windowing*, *DC Removal*, *similarity measure*. Tahap berikutnya yaitu melakukan pengkalsifikasian hasil dari rekaman tersebut menggunakan metode *Dynamic Time Warping* , metode ini akan mendeteksi hasil yang hampir mirip / mendekati dengan menghitung pola pada data suara yang telah direkam dengan data yang sudah disimpan di *database*. Kemudian sistem akan mengidentifikasi lagu beserta judulnya sesuai hasil rekaman suara bersenandung yang dilakukan sebelumnya



Gambar 4. Blok Diagram Sistem

b. Ekstraksi Ciri

Discrete Cosine Transform (DCT) digunakan untuk merubah suatu sinyal menjadi komponen frekuensi dasar, dengan memperhitungkan nilai riil dari hasil transformasi. *Discrete Cosine Transform* merupakan transformasi yang berhubungan dengan transformasi *Fourier* yang memberikan fungsi diskrit. Sehingga *Discrete Cosine Transform* merupakan transformasi ideal untuk proses kompresi data. *DCT* dapat digunakan untuk mendapatkan nilai koefisien. Hal ini dikarenakan *DCT* mampu menghasilkan sinyal berfrekuensi rendah lebih sedikit dan frekuensi tinggi yang banyak. *DCT* merekonstruksi urutan data dengan *DCT* koefisien, pada tahap ekstraksi ciri ini melalui beberapa tahap yaitu :

b1. Pre-emphasize

Merupakan proses pengolahan sinyal suara, prosesnya yaitu dilakukan untuk mendapatkan bentuk *spectral* frekuensi sinyal suara yang lebih halus dan bertujuan untuk menyamarkan *noise* yang ada pada suatu sinyal. Dengan memperkecil nilai amplitudo yang rendah yang berkemungkinan sebagai *noise* dan memperbesar nilai amplitudo yang konstan tinggi yang bisa diambil sebagai informasi pada suatu sinyal. *Pre-emphasis* didasari oleh hubungan input/output dalam domain waktu yang dinyatakan dalam persamaan :

$$y(n) = \alpha \cdot x(n) - (\alpha - 1) \cdot x(n-1) \dots\dots\dots (6)$$

dimana α merupakan konstanta filter *pre-emphasis* yang biasanya bernilai $0,9 < \alpha < 1,0$
 keterangan :

$y(n)$ = Sinyal Hasil Penjumlahan

$x(n)$ = Sinyal Waktu Diskrit

α = Konstanta Filter

b2. Frame Blocking

Merupakan proses yang digunakan untuk pembagian sinyal audio menjadi beberapa *frame*, pada perancangan sistem ini menggunakan beberapa nilai *frame* diantaranya yaitu dengan nilai *2 second*, *1 second*, *0.75 second*, *0.5 second* dan *0.25 second*. Semakin besar nilai *frame*, maka semakin cepat waktu komputasinya di karenakan ciri yang dicocokkan menjadi sedikit oleh sistem. Satu *frame* terdiri dari beberapa sampel tergantung tiap berapa detik suara yang akan di-sampling dan berapa besar frekuensi sampling-nya. Sinyal hasil dari framing adalah sinyal terpotong yang diskontinu. Sinyal terpotong tersebut akan dilanjutkan dalam proses *windowing*. Berikut persamaan *frame blocking* :

$$\text{Jumlah Frame} = ((I-N)/M)+1 \dots\dots\dots (7)$$

Keterangan : I = *Sample rate*

N = *Sample point (sample rate*waktu framing (s))*

M = $N/2$

b3. Windowing

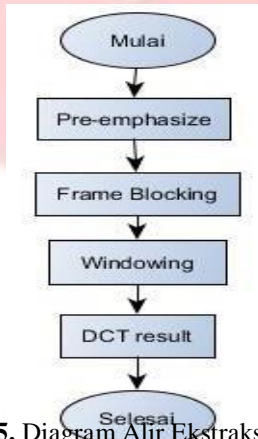
Merupakan proses yang digunakan untuk mereduksi puncak tiap segmen mulai dari awal hingga akhir pada suatu frame). Fungsi windowing yang digunakan adalah *hamming window* karena fungsi ini dapat membuat data pada awal *frame* dan akhir *frame* mendekati nilai 0 dengan baik. Jika didefinisikan *window* yang digunakan sebagai $w(n)$ maka output dari proses *window* adalah sebagai berikut :

dan model *window* yang digunakan adalah *Hamming window* yaitu:

$$w(n) = \frac{1}{2} \left(1 + \cos\left(\frac{2\pi n}{N-1}\right) \right) \quad (8)$$

Keterangan : $x(n)$ = Sinyal Input
 $w(n)$ = Tipe *windowing* (*Hamming*)

Berikut diagram alir ekstraksi ciri menggunakan DCT :



Gambar 5. Diagram Alir Ekstraksi Ciri DCT

c. Pre-Processing

Pre-processing merupakan proses untuk memperbaiki kualitas data sebelum dilakukan ekstraksi ciri. Pre-processing terdiri dari beberapa tahap sebagai berikut:

c1. Filtering

Tahap pertama pada pre-processing adalah *Filtering*. Data akan di filter menggunakan BPF (*Band Pass Filter*) untuk meredam sinyal selain sinyal suara senandung yang akan diteliti. Pada penelitian ini BPF yang digunakan memiliki frekuensi rentang dari 300-3400 Hz dikarenakan frekuensi suara manusia berada pada rentang frekuensi tersebut.

Rumus BPF :

$$|H(f)| = \frac{1}{\sqrt{1 + \left(\frac{f - F_T}{B}\right)^2}} \quad (9)$$

Keterangan : F_T = Frekuensi Tengah
 F = Frekuensi Sampling
 F_L = Frekuensi Rendah
 F_H = Frekuensi Tinggi

c2. Resample

Pada tahap ini data akan disampling ulang untuk mempersingkat kerja sistem dengan frekuensi sampling yang lebih sedikit dari frekuensi sampling aslinya tanpa menghilangkan data informasi, agar sistem komputasi tidak berat. Frekuensi sampling pada *resample* di dapat dari *teorema Nyquist*.
Teorema Nyquist : $f_s \geq 2 \cdot f_{max}$, dimana $f_s = 8000$ Hz , $f_{max} = 3400$ Hz

c3. Convert to Mono

Pada tahap ini data akan diubah dari *stereo* ke *mono* dengan mencari nilai rata-rata kedua kanal pada data *stereo* untuk diubah menjadi satu kanal saja (*mono*).

Rumus : $Output = \frac{k_1 + k_2}{2}$ (10)

Keterangan : k_1 = kanal pertama
 k_2 = kanal kedua

c4. DC Removal

Pada tahap ini data yang telah dikonversikan ke mono akan dilakukan DC Removal. DC Removal adalah proses pengurangan sinyal dengan nilai rata-rata sinyal atau bisa juga penghilangan sinyal DC yang mengganggu.

$$D[i] = S[i] - \frac{\sum_{i=1}^N S[i]}{N} \quad ()$$

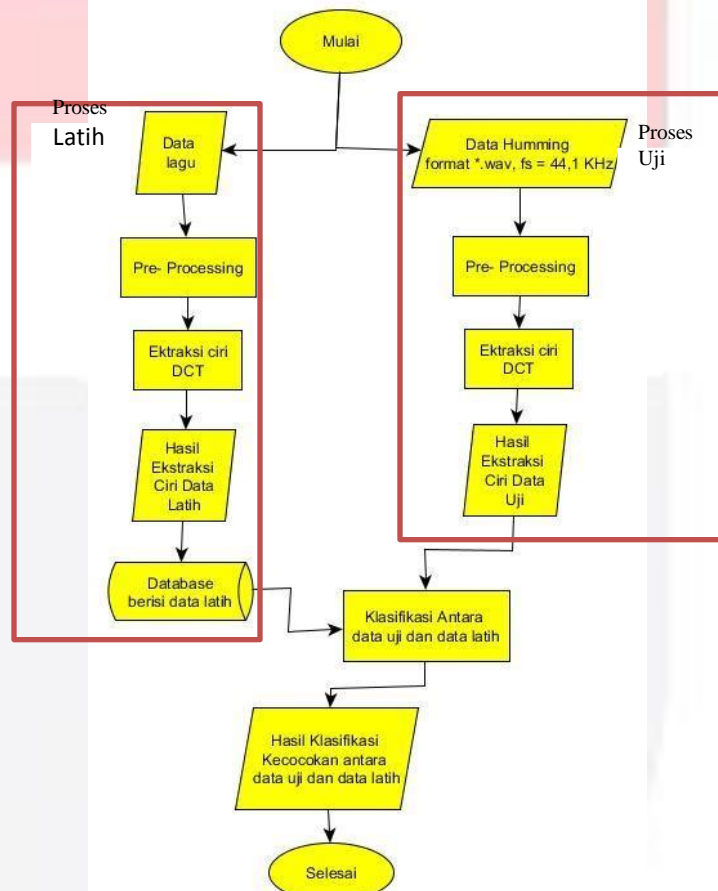
Keterangan : D[i] = Hasil sinyal ke-i setelah di lakukan DC Removal.

S[i] = Sinyal awal ke-i

N = Jumlah Sampel , n>0

d. Alir Kerja Sistem

Pada perancangan tugas akhir ini memerlukan alur Sistem kerja dari identifikasi judul lagu menggunakan suara senandung yang akan ditunjukkan pada flowchart di bawah ini.



Gambar 6. Diagram Alir Sistem

Dari *flowchart* diatas alur sistem kerja pada identifikasi judul lagu terdiri dari 2 proses yaitu proses latih dan proses uji, proses latih terdiri dari data lagu, *pre-processing*, ekstraksi ciri DCT, hasil ekstraksi ciri lagu, *database*. Sedangkan pada proses uji terdiri dari data *humming*, *pre-processing*, ekstraksi ciri DCT, hasil ekstraksi ciri *humming*, klasifikasi *Dynamic Time Warping* dan hasil identifikasi lagu. Berikut penjelasan dari proses simulasi yang akan di rancang :

d1. Data Lagu dan Data Humming

Data lagu yang digunakan pada proses latih yaitu file lagu dengan format *.wav Sedangkan pada proses uji data yang digunakan yaitu data hasil rekaman dengan *humming*/bersenandung yang disimpan dengan format *.wav dan menggunakan frekuensi sampling 44,1 KHz.

d2. Pre-processing

Pre-processing digunakan pada proses latih dan proses uji, tahap ini bertujuan untuk memperbaiki kualitas data sebelum dilakukan ekstraksi ciri. Tahapan *pre-processing* terdiri dari *filtering, resample, convert to mono, DC Removal*.

d3. Database

Pada proses latih, data dari *database* dipergunakan untuk acuan pada proses uji, yang terdiri dari judul lagu dengan suara orang bernyanyi tanpa menggunakan *instrument* (vokal), judul lagu dengan suara dari penyanyi asli menggunakan *instrument*, dan judul lagu dengan suara senandung manusia, dengan format *.wav dan *frekuensi sample* 44,1 KHz 16 bit *stereo*. *Database* uji dan latih dalam 1 file judul lagu terdiri dari *reff* dan *verse*. *Database* ini sebelumnya telah dilakukan ekstraksi ciri sehingga mempermudah pengklasifikasian pada proses uji.

d4. Hasil Identifikasi Lagu

Tahapan ini adalah tahap akhir dari proses identifikasi judul lagu dimana akan menampilkan hasil judul yang sesuai dengan yang diinginkan dengan sinyal masukan nada senandung / *humming*.

e. Akurasi dan Error

Pengujian sistem dilakukan dengan menggunakan parameter akurasi dan *error*. Akurasi merupakan ukuran ketelitian sistem dalam mengenali *input* yang diberikan sehingga menghasilkan *output* yang benar. Akurasi sistem secara keseluruhan dapat ditulis sebagai berikut :

$$\frac{\text{Jumlah data yang dikenali dengan benar}}{\text{Jumlah data}} \dots\dots\dots (12)$$

Sedangkan *error* merupakan tingkat kesalahan sistem dalam mengenali *input* yang diberikan terhadap jumlah data secara keseluruhan. *Error* secara matematis dapat ditulis sebagai berikut :

$$\frac{\text{Jumlah data yang salah dikenali}}{\text{Jumlah data}} \dots\dots\dots (13)$$

3. Hasil Pengukuran dan Simulasi

Dari hasil pengukuran dan simulasi yang telah didapat,

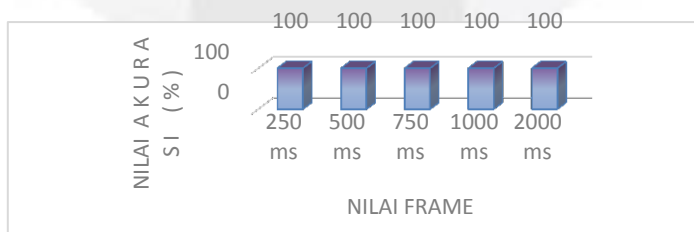
3.1 Membandingkan data latih dan data uji terhadap akurasi dan nilai *framing* pada sistem antara :

- a. Pengujian dan analisis pengaruh nilai *frame* antara data latih dengan data uji yang sama-sama berupa lagu asli.
- b. Pengujian dan analisis pengaruh nilai *frame* antara data latih yang berisi *humming* dengan data uji berupa *humming*.
- c. Pengujian dan analisis pengaruh nilai *frame* antara data latih yang berisi lagu asli dengan data uji berupa *humming*.
- d. Pengujian dan analisis pengaruh nilai *frame* antara data latih yang berisi vokal dengan data uji berupa *humming*.
- e. Pengujian dan analisis pengaruh nilai *frame* antara data latih dengan data uji yang sama-sama berupa vokal.

Skenario pengujian yang dilakukan dengan mengubah nilai *frame* ekstraksi ciri DCT sebanyak lima kali perubahan, yaitu 250 ms, 500 ms, 750 ms, 1000 ms dan 2000 ms. Dengan jumlah data latih dan data uji berbeda pada ke lima skenario pengujian. Dari hasil pengujian, dilakukan analisa akurasi menggunakan rumus 3.3. berikut adalah hasil dari analisis dari pengujian :

- a. Pengaruh Nilai *Frame* Ekstraksi Ciri terhadap Nilai Akurasi dengan Data Latih Lagu Asli dan Data Uji Lagu Asli

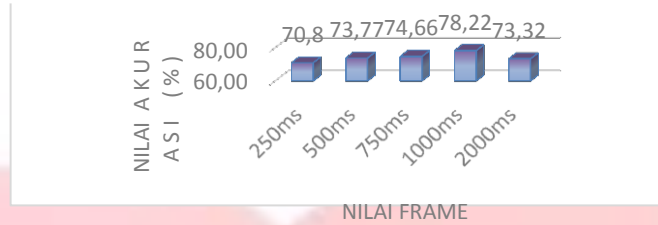
Lagu disini yaitu file yang berformat *.wav dengan frekuensi sampling 44,1 KHz 16 bit stereo. Lagu ini terdiri dari 50 judul dengan 100 file yang terdiri dari *reff* dan *verse*. Lagu dengan 100 file dijadikan data uji dan data latihnya juga lagu yang sama. Dari hasil yang sudah dijalankan akurasi yang didapat 100%. Dalam skenario pengujian ini dimaksud untuk membuktikan sistem telah bekerja dengan benar.



- b. Pengaruh Nilai *Frame* Ekstraksi Ciri terhadap Nilai Akurasi dengan Data Latih *Humming* dan Data Uji *Humming*.

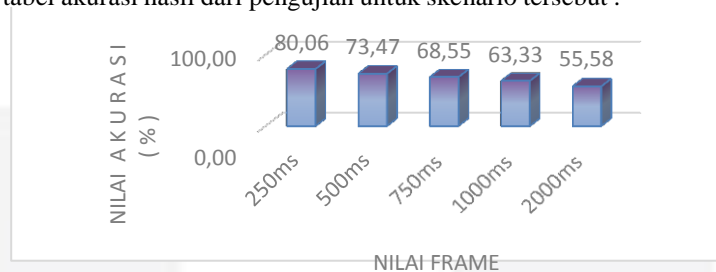
Skenario pengujian ini yaitu data uji terdiri dari data dua orang yang melakukan perekaman dengan judul yang sama dengan data latih. Tetapi pada skenario ini melakukan

penguji antara data latih dengan data uji yang keduanya berupa senandung manusia dengan jumlah data latih sebanyak 30 judul lagu dan data uji sebanyak 15 judul lagu. Dari hasil pengujian, dilakukan analisis akurasi dengan mengubah-ubah nilai *frame* sebanyak 5 kali perubahan pada ekstraksi ciri. Berikut tabel akurasi hasil dari pengujian untuk skenario tersebut :



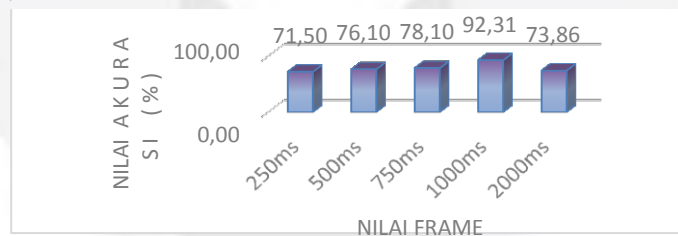
c. Pengaruh Nilai *Frame* Ekstraksi Ciri terhadap Nilai Akurasi dengan Data Latih *Humming* Lagu Asli dan Data Uji *Humming*.

Untuk mengetahui keberhasilan akurasi sistem yang dirancang maka dilakukan pengujian anatara data latih dan data uji yang sebenarnya. Data latih yaitu lagu dengan format *.wav dengan *frekuensi sample* 44,1 KHz 16 bit *stereo* yang Pada skenario pengujian yang ke tiga, pengujian menggunakan data latih berupa lagu asli sebanyak 50 judul lagu, dan data uji berupa senandung manusia sebanyak 30 judul lagu. Dari hasil pengujian, dilakukan analisis akurasi dengan mengubah-ubah nilai *frame* sebanyak 5 kali perubahan pada ekstraksi ciri. Berikut tabel akurasi hasil dari pengujian untuk skenario tersebut :



d. Pengaruh Nilai *Frame* Ekstraksi Ciri terhadap Nilai Akurasi dengan Data Latih Vokal Lagu Asli dan Data Uji *Humming*.

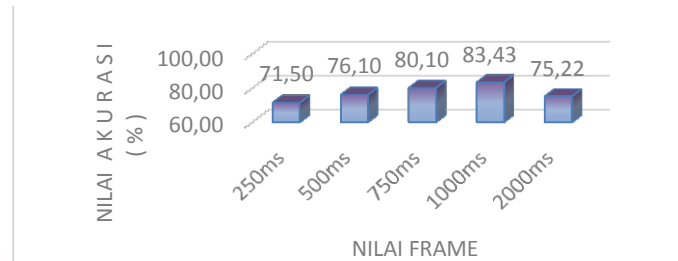
Pada skenario pengujian yang ke empat, pengujian menggunakan data latih berupa vokal sebanyak 50 judul lagu, dan data uji berupa senandung manusia sebanyak 30 judul lagu. Dari hasil pengujian, dilakukan analisis akurasi dengan mengubah-ubah nilai *frame* sebanyak 5 kali perubahan pada ekstraksi ciri. Berikut tabel akurasi hasil dari pengujian untuk skenario tersebut :



e. Pengaruh Nilai *Frame* Ekstraksi Ciri terhadap Nilai Akurasi dengan Data Latih Vokal Lagu Asli dan Data Uji Vokal.

Skenario pada penguji tahap ini, pengujian antara data latih dan data uji yang keduanya berupa vokal, tetapi pada data uji terdiri dari data dua orang yang melakukan perekaman dengan judul yang sama dengan data latih. Jumlah data pada data latih sebanyak 30 judul lagu dan data uji sebanyak 15 judul lagu.

Dari hasil pengujian, dilakukan analisis akurasi dengan mengubah-ubah nilai *frame* sebanyak 5 kali perubahan pada ekstraksi ciri. Berikut tabel akurasi hasil dari pengujian untuk skenario tersebut :



4. Kesimpulan dan Saran

4.1 Kesimpulan

Dari hasil analisis pengujian sistem Tugas Akhir ini, dapat diambil kesimpulan sebagai berikut:

1. Pada pengujian pertama menggunakan Data Lagu Asli sebagai Data Latih dan Data Lagu Asli sebagai Data Uji, dapat disimpulkan bahwa tiap *frame window* memiliki akurasi yang sama yaitu 100% dengan data latih sebanyak 100 dan data uji sebanyak 100 dikarenakan data latih dan data ujinya sama.
2. Pada pengujian kedua menggunakan Data *Humming* sebagai Data Latih dan Data Lagu *Humming* sebagai Data Uji, dapat disimpulkan bahwa pengujian ini memiliki akurasi tertinggi 78,22 pada *frame window* 1000ms dengan data latih sebanyak 60 dan data uji sebanyak 60.
3. Pada pengujian ketiga menggunakan Data Lagu Asli sebagai Data Latih dan Data Lagu Senandung sebagai Data Uji, dapat disimpulkan bahwa akurasi tertinggi terdapat pada *frame window* 250ms dengan akurasi 80,06 % dengan data latih sebanyak 100 dan data uji sebanyak 60.
4. Pada pengujian menggunakan Data Lagu Vokal sebagai Data Latih dan Data Lagu Senandung sebagai Data Uji, dapat disimpulkan bahwa akurasi tertinggi terdapat pada *frame window* 1000ms dengan akurasi 92,31 % dengan data latih sebanyak 100 dan data uji sebanyak 60.
5. Pada pengujian menggunakan Data Lagu Vokal sebagai Data Latih dan Data Lagu Vokal sebagai Data Uji dapat disimpulkan bahwa akurasi tertinggi terdapat pada *frame window* 1000ms dengan akurasi 83,43 %, dengan data latih sebanyak 60 dan data uji sebanyak 60.

4.2 Saran

Adapun rekomendasi atau saran untuk Tugas Akhir selanjutnya adalah:

1. Pengembangan cara kerja sistem untuk mengidentifikasi judul lagu mampu lebih real time tanpa harus melakukan perekaman terlebih dahulu.
2. Dapat menggunakan ekstraksi ciri dan klasifikasi lain yang mampu meningkatkan akurasi yang lebih baik.
3. Saat perekaman data uji dapat dilakukan di tempat yang kedap suara, agar tidak memiliki *noise*.

DAFTAR PUSTAKA

- [1] Andhini, Rayani Budi, Dwi. 2013. "Aplikasi Pengenalan Judul Lagu Suara Menjadi Teks Menggunakan Metode Jaringan Syaraf Tiruan Backpropagation". Bandung: Institut Teknologi Telkom.
- [2] Pampalk, E., Dixon, S., Widmer, G.: Exploring Music Collections by Browsing Different Views. Proceedings ISMIR 2003, 201–208
- [3] Riska Yessivirna, Marji dan Dian Eka Ratnawati. 2014. *Klasifikasi Suara Berdasarkan Gender (Jenis Kelamin) Dengan Metode K-Nearest Neighbor*. Malang : Universitas Brawijaya Malang
- [4] Sukma, Alvian. 2014. "Dynamic Time Warping Sistem Temu Kembali Informasi". Surabaya. Universitas Airlangga.
- [5] Tan, Li . Digital Signal Processing Fundamentals and Application
- [6] Tawakal, Iqbal. 2014. "Analisis Penggunaan Algoritma Genetika untuk Meningkatkan Performansi dari Aplikasi Ketepatan Lagu dari Senandung Manusia Berbasis Jaringan Syaraf Tiruan *Back-Propagation*". Bandung : Universitas Telkom.
- [7] Wardana, Nyoman Kusuma. 2010. *Mekanisme Pembentukan Suara*. [Online] Available at: <http://waterfilling.blogspot.co.id/2010/12/mekanisme-pembentukan-suara.html> [diakses pada tanggal 08 November 2015]
- [8] <https://id.scribd.com/doc/100801150/Dynamic-Time-Warping-Distance> (diakses pada tanggal 17 Juni 2016)