

# **PENGENALAN DAN REPRESENTASI SIMBOL AKOR MUSIK MENGGUNAKAN HIDDEN MARKOV MODEL DENGAN PENDEKATAN DOUBLY NESTED CIRCLE OF FIFTH**

## **MUSIC CHORD RECOGNITION AND SYMBOL REPRESENTATION USING HIDDEN MARKOV MODEL WITH DOUBLY NESTED CIRCLE OF FIFTH APPROACH**

**Muhammad Miftah Muslim Zulfikar**

Prodi S1 Teknik Informatika, Fakultas Teknik, Universitas Telkom

[muh.miftah.m.z@gmail.com](mailto:muh.miftah.m.z@gmail.com)

---

### **Abstrak**

Tugas akhir ini menjelaskan sistem pengenalan akor berbasis Hidden Markov Model (HMM) dengan pendekatan Doubly Nested Circle of Fifth (DNCOF). Terdapat dua versi HMM dalam sistem ini, trained-HMM dan untrained-HMM. Pada trained-HMM sistem dilatih menggunakan 180 lagu dari 13 album The Beatles untuk membentuk model akor. Sedangkan untrained-HMM menggunakan Chord Template untuk membentuk model akor. Model akor yang digunakan dalam sistem berjumlah 24 akor yang terdiri dari 12 nada dengan kombinasi mayor dan minor. DNCOF digunakan sebagai probabilitas transisi pada sistem HMM. Dalam pengenalan ini, lagu diekstrak ke dalam bentuk chromagram dan tugas dari sistem pengenalan ini adalah untuk mengisi label akor disetiap chromagram frame. Tugas akhir ini mengevaluasi bagaimana pengaruh DNCOF terhadap akurasi dua sistem pengenalan akor berbasis HMM dimana dalam skenario terbaik DNCOF yang menggunakan untrain-HMM menghasilkan akurasi terbaik yaitu 96.94%. Dan juga mengevaluasi perbandingan antara trained-HMM dengan untrained-HMM dimana keduanya memiliki akurasi yang cukup dekat disetiap skenario.

**Kata Kunci:** Akor, PCP, HMM, DNCOF, Chromagram.

---

### **Abstract**

This final project describes chord recognition system based HMM (Hidden Markov Model) with DNCOF (Doubly Nested Circle of Fifth) approach. There are two versions of HMM in this system, trained-HMM and untrained-HMM. On trained-HMM, chord model are trained using 180 songs of The Beatles. On untrained-HMM use chord templates as chord models. Chord models used in this system amounted to 24 chords which consists of 12 tones with two combination of major and minor. DNCOF used as the transition probabilities in HMM systems. In this recognition, the song is extracted into chromagram and task of this recognition is to fill chord labels in every chromagram frames. This final project evaluates how DNCOF influence the accuracy of the two HMM-based chord recognition system where the best DNCOF's scenario that used untrained-HMM produces 96.94% of accuracy. And also evaluates the comparison between trained-HMM and untrained-HMM.

**Keyword:** Chord, PCP, HMM, DNCOF, Chromagram.

---

## **1 Pendahuluan**

Musik merupakan salah satu bagian terpenting dalam kehidupan manusia. Mayoritas orang suka dengan musik, mulai dari musik tradisional hingga musik modern yang banyak berkiblat pada musik barat. Musik itu tersusun dari tempo, ritme, nada dan instrumen yang saling berharmonisasi. Salah satu bagian terpenting dalam musik modern saat ini adalah akor. Akor merupakan sekumpulan dua nada atau lebih yang dibentuk dari tangga nada mayor harmonis yang dimainkan secara bersamaan atau berurutan. Pengenalan akor secara otomatis sangat berguna bagi penikmat musik yang hendak mengetahui komposisi akor pada suatu lagu. Banyak aplikasi yang mampu

melakukan pengenalan akor dalam musik dengan metode yang beragam, beberapa diantaranya merupakan hasil pengembangan dari Pitch Class Profile (PCP) atau Chroma. Chroma merupakan vektor ciri dengan membagi nada ke dalam 12 dimensi tingkatan pitch. Kumpulan Chroma dalam selang waktu tertentu disebut Chromagram, sistem pengenalan akor bertujuan untuk mengisi label akor disetiap chromagram frame.

Hidden Markov Model (HMM) merupakan teknik pattern recognition yang paling umum digunakan untuk mendeteksi pattern yang berbentuk urutan. Metode ini banyak digunakan pada pengenalan bicara (Speech Recognition) lalu berkembang ke pengenalan

musik salah satunya adalah pengenalan akor. Syauiq Riezan [4] menggunakan EM-Trained HMM dengan state berjumlah 24 (12 mayor dan 12 minor), dan semua parameter HMM dilatih menggunakan Expectation Maximation (EM). Adi Ferdian [6] merupakan pengembangan dari [4] dengan menambah jumlah state menjadi 36 dengan penambahan 1 jenis akor Dominan<sup>7</sup>. Algoritma fuzzy digunakan untuk melatih parameter HMM. Lain halnya dengan Jati Perdana [3] mengimplementasikan HMM tidak dilatih (untrained-HMM), sebagai gantinya menggunakan Doubly Nested Circle of Fifth (DNCOF) dan template akor sebagai dasar penentuan parameter-parameter HMM.

Terinspirasi dari beberapa paper yang telah disebutkan diatas, tugas akhir ini berupaya untuk mengembangkan sistem pengenalan akor sehingga dapat digunakan secara luas menggunakan DNCOF sebagai dasar pengetahuan musik berupa probabilitas transisi antar akor yang akan diterapkan kedalam dua sistem berbasis HMM (trained dan untrained). Sistem ini terdiri dari 4 proses utama yaitu ekstraksi fitur chroma, beat tracking (pencarian tempo), measure of fit, dan algoritma most likely sequence.

## 2 Tinjauan Pustaka

### 2.1 Teori Musik

Nada merupakan bagian terkecil dari musik. nada dapat diklasifikasikan menjadi 12 macam dengan tiap nadanya memiliki interval  $\frac{1}{2}$  (*semitone*) membentuk sebuah tangga nada kromatik yang diantaranya C, C#, D, D#, E, F, F#, G, G#, A, A#, dan B. tangga nada yang umum digunakan adalah tangga nada diatonik, tangga nada ini terdiri dari 7 nada, sebagai contoh tangga nada C mayor yang terdiri dari C, D, E, F, G, A, B, C dimana C merupakan key note dan interval untuk tangga nada mayor adalah 1-1- $\frac{1}{2}$ -1-1- $\frac{1}{2}$ . Tangga nada diatonik seluruhnya berjumlah 7 dengan nada dan interval yang sama dengan tangga nada mayor tetapi key note yang berbeda (*modes*). Mayoritas musik yang beredar di dunia saat ini yang dapat disebut musik modern diciptakan menggunakan tangga nada diatonik tersebut.

Bagian terpenting dari sebuah lagu pada musik modern saat ini adalah akor. akor pada lagu berfungsi untuk mengiringi lagu tersebut. Akor merupakan sekumpulan nada yang terdiri dari tiga atau lebih nada yang dibentuk dari tangga nada harmonis yang dimainkan secara bersamaan atau satu persatu. Penggunaan akor banyak ditemukan pada musik barat modern, musik Afrika, dan musik Oseania. Bentuk akor yang paling umum adalah triad. Akor triad disusun oleh 3 buah nada (*root*, *third* & *fifth*) yang memiliki aturan interval tertentu. *Root* merupakan


nada dasar akor, *third* merupakan nada ke-3 dan *fifth* merupakan nada ke-5. Akor mayor memiliki aturan dari *root* ke *third* intervalnya 2 sedangkan *third* ke *fifth* intervalnya  $1\frac{1}{2}$ . Untuk akor minor dari *root* ke *third* intervalnya  $1\frac{1}{2}$  sedangkan dari *third* ke *fifth* intervalnya 2. Contoh akor C mayor yaitu C-E-G.

Tempo merupakan ukuran kecepatan sebuah lagu yang biasanya menggunakan satuan *Beat Per Minute* (BPM). BPM menyatakan banyaknya ketukan dalam satu menit, misal terdapat sebuah lagu dengan tempo 60 BPM berarti dalam satu menit terdapat 60 ketukan atau tepat satu ketuk per detik. Begitupun lagu dengan tempo 120 BPM memiliki 120 ketukan dalam satu menit atau satu ketuk sama dengan  $\frac{1}{2}$  detik.

### 2.2 Anotasi Akor Musik

Anotasi akor musik adalah suatu catatan berupa nama atau simbol untuk merepresentasikan komposisi nada dari suatu akor. Berikut ini merupakan sebagian dari ragam anotasi akor yang sering digunakan:

Tabel 2-1 : Ragam anotasi akor musik.

No	Jenis Anotasi
a	
b	C Major : I <sup>7</sup> - ii <sup>7</sup> - IVc - Ivb - VII <sup>7</sup> c - V <sup>7</sup> - I
c	C Major : C <sup>7</sup> - d <sup>7</sup> - F/C - F/A - B <sup>7</sup> /F - G <sup>7</sup> - C
d	CM7 - Dm7 - F/C - F/A - Bdim7/F - G7 - Csus4 - C
e	C:maj7 - D:min7 - F:maj/C - F:maj7/A - B:dim7/F - G:7 - C:sus4 - C

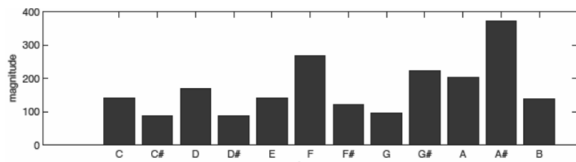
Setiap anotasi akor memiliki kekurangan dan kelebihan tergantung dari kebutuhannya. Notasi yang baik untuk digunakan pada sistem pengenalan akor haruslah sederhana, intuitif untuk dipelajari, terstruktur dan tidak ambigu untuk dapat diuraikan program komputer [6]. pada poin (a) merupakan anotasi yang umum digunakan musik klasik barat. Setiap simbolnya mewakili setiap not yang dibunyikan. Poin (b) merupakan notasi angka romawi, setiap angka romawi mewakili urutan akor berdasarkan basis. Terdapat ambiguitas pada simbol I<sup>7</sup> dan V<sup>7</sup> dimana I<sup>7</sup> merupakan C:maj7 sedangkan V<sup>7</sup> merupakan G:7. Begitu pula pada poin (c) yang juga memiliki ambiguitas yang sama. Pada poin (d) tidak ada ambiguitas untuk membedakan antara akor mayor dengan dominan, tetapi akan menjadi sulit apabila akan diuraikan oleh sistem. Pada poin (e) merupakan hasil penelitian C. Harte. Tidak ada ambiguitas yang

terjadi dan terdapat simbol ‘:’ dan ‘/’ yang memudahkan sistem untuk dapat mengurainya. Untuk itu penulis merujuk kepada anotasi C. Harte untuk diimplementasikan ke dalam sistem.

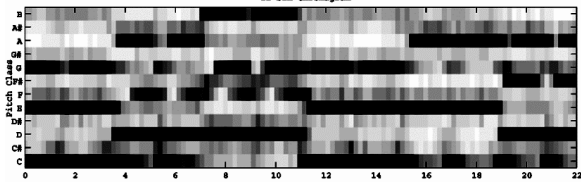
### 2.3 Pitch Class Profile (PCP)

*Pitch* adalah properti perceptual yang memungkinkan nada diurut berdasarkan frekuensi. Setiap nada memiliki frekuensi tertentu yang unik. Setiap frekuensi tersebut memiliki label semitone yang diikuti dengan angka yang menunjukkan oktaf. Contoh: A4 merupakan nada A dengan posisi berada di oktaf ke 4 memiliki frekuensi 440Hz.

PCP atau *chroma* merupakan representasi kekuatan 12 nada semitone tangga nada kromatik. Sebagai contoh, elemen pertama pada PCP menjelaskan bahwa seberapa kuat nada C yang dihasilkan [10]. Kumpulan dari PCP dalam jangka waktu tertentu disebut *chromagram*. Pada saat implementasinya, musik akan dipetakan kedalam bentuk chromagram. Dari situ kita dapat melihat pola kekuatan 12 nada semitone yang dihasilkan dari pemetaan musik tersebut tiap *chromagram frame*.



Gambar 2-1 : Bentuk Pitch Class Profile (PCP) [4]



Gambar 2-2 : Chromagram

### 2.4 Measure of Fit

*Measure of Fit* merupakan himpunan probabilitas observasi yang dihitung dari PCP untuk mengetahui seberapa dekat probabilitas suatu PCP dalam setiap *chromagram frame* terhadap setiap state akor yang ada. Probabilitas tersebut dapat dirumuskan dengan:

$$C(t) = \begin{bmatrix} C_C(t) \\ C_{C\#}(t) \\ \vdots \\ C_{B:min}(t) \end{bmatrix}$$

$$C_{P_n}(t) = p(P_n | chroma(t))$$

$$P_1 = C, P_2 = C\#, \dots, P_{24} = B:min$$

Dimana vektor akor dalam suatu frame adalah  $C(t)$  yang memiliki set probabilitas dari  $C_{P_1}$  sampai  $C_{P_{24}}$ . Dan  $chroma(t)$  merupakan PCP pada frame tertentu. Setiap probabilitas dihitung menggunakan

algoritma *mahalanobis distance* untuk menghitung seberapa dekat jarak PCP terhadap *state* akor. Rumus *mahalanobis distance* adalah:

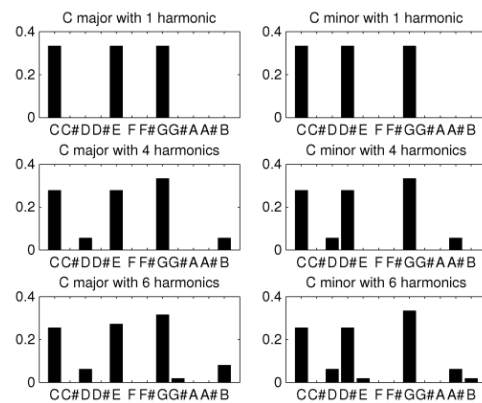
$$d(Chroma, \mu)$$

$$= \sqrt{(chroma - \mu)^T S^{-1} (chroma - \mu)}$$

Variabel *mean* ( $\mu$ ) dan *covariance* ( $S$ ) dihitung saat pelatihan akor.

### 2.5 Template Akor

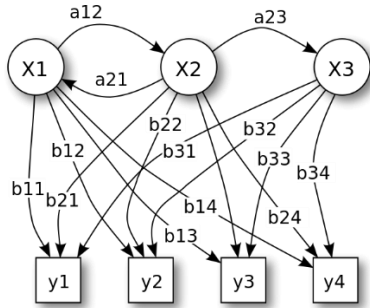
*Template* akor merupakan 12-dimensi PCP yang merepresentasikan kekuatan energi tiap komponen PCP suatu akor berdasarkan teori musik. Misal, akor C terdiri dari nada C, E, dan G, maka bentuk template akornya adalah {1,3,3,3,1,3, 3,1,3,3,3,3}. Dimana 1 adalah komponen yang dimiliki akor C dan 3 adalah komponen lain yang tidak dimiliki oleh akor C. Namun, informasi yang didapat dari chromagram suatu akor tidak hanya intensitas nada akor itu saja tetapi merupakan perpaduan intensitas harmonik untuk setiap nada. Intensitas harmonik dihitung menggunakan rumus  $0.6^{i-1}$  untuk setiap nada akor dengan  $i$  adalah jumlah nada harmonik. Pada tugas akhir ini penulis menggunakan template akor dengan 6 harmonik dikarenakan pada [8] menghasilkan akurasi terbaik diantara harmonik lainnya. Berikut merupakan template akor dengan 1,4, dan 6 harmonik:



Gambar 2-3 : Template Akor pada akor C mayor dengan 1,4, dan 6 harmonik

### 2.6 Hidden Markov Model

HMM merupakan teknik pattern recognition yang paling umum digunakan untuk mendeteksi pattern yang berbentuk urutan. Urutan chord dalam music adalah contohnya. HMM bekerja dengan cara membuat model dari sistem yang terdiri dari beberapa state. State tersebut bersifat hidden, tetapi dapat membentuk rangkaian yang dapat diobservasi. HMM digunakan untuk mengklasifikasi dengan mengurutkan observed sequence dengan observed sequences yang mungkin digenerate dari model tersebut.



Gambar 2-4 : Hidden Markov Model [3]

Pada gambar diatas, X merupakan *State*, setiap *state* mewakili jumlah akor yang akan dikenali setiap *state* memiliki busur kearah *state* lainnya dengan nilai a yang merupakan probabilitas *state* yang satu ke *state* lainnya. y merupakan observasi yang mewakili kekuatan-kekuatan tiap komponen PCP yang muncul. Setiap *state* X memiliki busur kearah observasi dengan nilai b yang merupakan probabilitas kemunculan *State* terhadap observasi yang dilakukan.

### 2.7 Algoritma Most Likely Sequence

Algoritma *most likely sequence* adalah algoritma yang digunakan untuk menentukan urutan *state* dari sebuah peristiwa pada *Hidden Markov Model* berdasarkan pada observasi yang dilakukan [3]. Pada kasus ini algoritma *most likely sequence* yang digunakan adalah algoritma Viterbi. Output dari Algoritma Viterbi ini adalah urutan *state* yang memiliki probabilitas tertinggi untuk terjadi berdasarkan pada observasi yang dilakukan.

Algoritma Viterbi bekerja dengan cara sebagai berikut. Diasumsikan terdapat sebuah *Hidden Markov Model* dengan ruang *state* S, probabilitas awal  $\pi_i$  untuk mulai di *state* i, dan probabilitas transisi  $a_{i,j}$  untuk transisi dari *state* i menuju *state* j. Katakanlah melakukan observasi terhadap  $y_1 \dots y_t$  maka urutan yang paling mungkin untuk menghasilkan observasi tersebut dapat dicari dengan

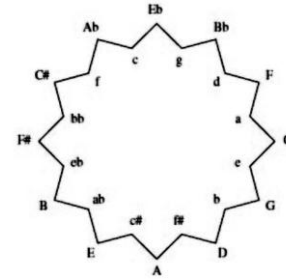
$$V_{1,k} = P(y_1|k). \pi_i$$

$$V_{t,k} = P(y_t|k). \max_{x \in S} (a_{x,k} \cdot V_{t-1,x})$$

Di mana  $V_{t,k}$  merupakan probabilitas dari urutan yang paling mungkin terjadi pada sejumlah t observasi pertama yang memiliki k sebagai *final state*.

### 2.8 Doubly Nested Circle of Fifth

DNCOF merupakan dasar pengetahuan musik yang merepresentasikan hubungan ketetangaan antara tangga nada mayor dan minor. DNCOF digunakan sebagai dasar perhitungan probabilitas transisi akor.



Gambar 2-5 : DNCOF [2]

Gambar diatas menggambarkan hubungan ketetangaan antar tangga nada mayor dan minor dimana terdapat 12 akor mayor dengan simbol huruf kapital dan 12 akor minor dengan simbol huruf kecil. Tangga nada mayor yang bersebelahan memiliki tepat satu nada yang berbeda, begitupun tangga nada minor. Sebagai contoh, tangga nada C dan G memiliki tepat satu nada yang berbeda yaitu F untuk tangga nada C major dan F# untuk tangga nada G major. Terdapat 24 akor yang dapat terbentuk mewakili 24 tangga nada yang saling terhubung dengan jarak terjauh adalah 12 [2,7].

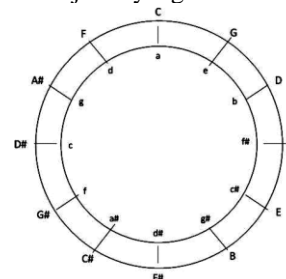
Adapun probabilitas transisi dari masing-masing akor ditentukan dengan formula:

$$a_{i,j} = \frac{12 - Cd(i,j) + \epsilon}{144 + 24\epsilon}$$

Di mana  $a_{i,j}$  merupakan probabilitas transisi dari *state* i ke *state* j,  $\epsilon$  merupakan konstanta non-negatif, dan  $Cd(i,j)$  merupakan sebuah fungsi yang menggunakan nomor *state* di mana terjadi transisi, i dan j, kemudian nilai antara 0 sampai 12 sesuai dengan nilai jarak antara akor I dan akor j pada DNCOF dikurangi dengan 12. Sebagai contoh, jarak antara C# dan F adalah 0, maka perhitungannya menjadi:

$$\frac{12 - Cd(C\#,F) + \epsilon}{144 + 24\epsilon} = \frac{12 - 12 + \epsilon}{144 + 24\epsilon} = \frac{0 + \epsilon}{144 + 24\epsilon}$$

Adapun DNCOF hasil rancangan Jati[3] merupakan peningkatan dari metode sebelumnya dengan ide dasar setiap tangga nada yang bersebelahan memiliki tepat satu nada yang berbeda sehingga harusnya memiliki jarak yang lebih dekat.



Gambar 2-6 : Modifikasi DNCOF [3]

Dengan adanya modifikasi tersebut jarak terjauh berubah menjadi 7 dengan rumus perhitungan menjadi:

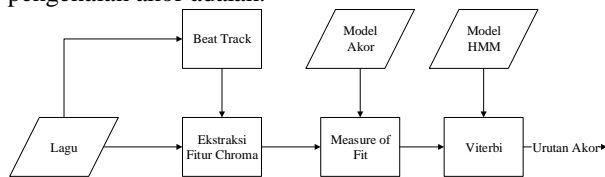
$$a_{i,j} = \frac{7 - Cd(i,j) + \epsilon}{84 + 24\epsilon}$$

Menurut Jati [3] hasil dari kedua metode tersebut tidak terlalu signifikan karena modifikasi tersebut dilakukan hanya untuk menyesuaikan dengan dasar theory yang ada.

### 3 Perancangan Sistem

#### 3.1 Gambaran Umum Sistem

Terdapat dua pendekatan yang digunakan dalam sistem pengenalan akor ini yaitu trained-HMM dan untrained-HMM. Secara umum diagram blok sistem pengenalan akor adalah:



Gambar 3-1 : Diagram Blok Sistem Pengenalan Akor

Pada sistem pengenalan akor ini lagu yang akan dikenali dirubah ke dalam bentuk *chromagram* melalui proses ekstraksi fitur *chroma* dengan lebar *frame* ditentukan melalui proses *beat track*. Selanjutnya, setiap *frame* akan dihitung probabilitas observasinya berupa vektor akor melalui proses *measure of fit*. selanjutnya menentukan urutan akor terbaik berdasarkan vektor akor dan model HMM melalui proses viterbi. hasil dari proses viterbi berupa urutan akor.

##### 3.1.1 Lagu

Lagu merupakan inputan pada sistem ini. Terdapat tiga buah dataset yang digunakan pada sistem pengenalan akor ini. Dataset 1 berisi lagu-lagu yang dibuat menggunakan musik berformat MIDI yang kemudian di-*ekspor* keformat mp3. Dataset 2 berisi lagu-lagu yang digunakan adalah dataset mirex tahun 2009 yang terdiri dari 180 lagu *The Beatles* berformat mp3 dengan spesifikasi: *mono*, *bit rate* 32kbps, *sampling rate* 16KHz. Dan dataset 3 berisi lagu-lagu band *zweieck* yang terdiri dari 18 lagu dengan *bit rate* 24kbps, *sampling rate* 11025Hz.

##### 3.1.2 Beat Track

Proses *beat track* merupakan proses untuk mencari tempo pada sebuah lagu. Proses *beat track* digunakan untuk menentukan lebar *frame* pada *chromagram*. Pertama-tama, lagu akan dicari besaran temponya. Kemudian dihitung lebar *frame* berdasarkan tempo yang didapat.

Dimisalkan lagu yang digunakan memiliki durasi 60 detik (1 menit) dengan tempo yang didapat dari proses pencarian tempo adalah 120 bpm. Maka jumlah *frame* yang dibentuk sebanyak 120 *frame* dengan masing-masing *frame* memiliki durasi 0.5 detik.

#### 3.1.3 Ekstraksi Fitur Chroma

Proses ekstraksi fitur *Chroma* merupakan proses untuk merubah file audio menjadi PCP. Setelah itu dilakukan proses segmentasi kedalam bentuk *frame* dengan lebar *frame* sesuai dengan hasil proses *beat track*. *Chroma* atau *Pitch Class Profile* adalah vektor berdimensi 12 yang menunjukkan distribusi energi dari kelas-kelas *pitch* dasar yang membentuk suatu akor musik. *Chroma* dipakai sebagai fitur dalam sistem ini karena mempunyai ciri unik yang dapat membedakan nada-nada yang satu dengan yang lain ke dalam satu oktaf yang sama. Akor sendiri dapat dilihat secara langsung dari distribusi energi *pitch* terbesar yang ada pada vektor *chroma*. Ekstraksi fitur *chroma* dilakukan menggunakan *ISP (Intelligent Sound Processing) Toolbox*. *ISP Toolbox* adalah sebuah *toolbox* untuk Matlab yang biasa digunakan para periset *Music Information Retrieval* untuk mengekstrak *chroma* dari potongan lagu.

#### 3.1.4 Measure of Fit

Pada proses ini Setiap PCP pada *Chromagram frame* akan dihitung tingkat kedekatannya terhadap 24 macam model akor. *Mahalanobis distance* digunakan pada proses ini untuk menghitung tingkat kedekatan PCP terhadap setiap model akor. Sebagai ilustrasi, sebuah *chromagram* berupa matriks berukuran 5x12 akan dihitung nilai kedekatannya terhadap model akor dengan matriks berukuran 24x12 yang akan menghasilkan matriks 5x24.

#### 3.1.5 Model HMM

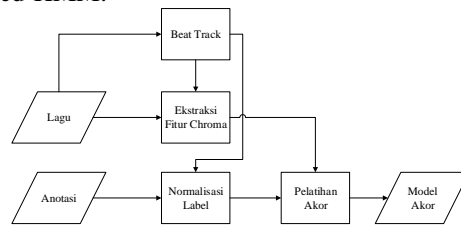
Berisi parameter-parameter HMM yaitu: probabilitas awal dan probabilitas transisi. Probabilitas awal merupakan matriks berukuran 1x24 sedangkan probabilitas transisi merupakan matriks berukuran 24x24. Pada sistem trained-HMM semua parameternya dilatih menggunakan dataset anotasi sedangkan pada untrained-HMM semua parameternya didapat dari DNCOF.

#### 3.1.6 Viterbi

*Chromagram*, probabilitas awal, probabilitas transisi, dan probabilitas observasi menjadi inputan pada proses ini. Algoritma *Viterbi* digunakan untuk menentukan urutan akor terbaik berdasarkan parameter-parameter yang diinputkan. Pelatihan Akor

#### 3.2 Pelatihan Akor

Berikut merupakan diagram blok pelatihan akor untuk trained-HMM:



Gambar 3-2 : Diagram Blok Pelatihan Akor

Pada tahap ini lagu data latih akan dirubah ke dalam bentuk chromagram melalui ekstraksi fitur chroma dan Anotasi akan dinormalisasi sesuai dengan jumlah akor yang akan dikenali. Proses beat track digunakan pada kedua proses tersebut untuk menentukan jumlah frame dalam satu lagu. Selanjutnya chromagram dan label akan dilatih membentuk model akor melalui proses pelatihan akor. Untuk metode untrained-HMM model akor berupa template akor yang telah dijelaskan pada bab 2.

### 3.2.1 Normalisasi Label

Proses Normalisasi Label merupakan proses untuk menyederhanakan anotasi akor hasil pelatihan dengan 24 macam state akor yang terdiri dari 2 bentuk akor mayor dan minor ditambah dengan 1 macam state 'no chord' dalam bentuk angka 0-24. Beberapa dataset anotasi diidentifikasi memiliki variasi akor lebih dari 2 macam bentuk akor yang akan dikenali sistem ini sehingga proses penyederhanaan kedalam 2 macam bentuk akor perlu dilakukan. Berikut beberapa bentuk akor yang disederhanakan:

**Mayor** : mayor, mayor7, dominan7.

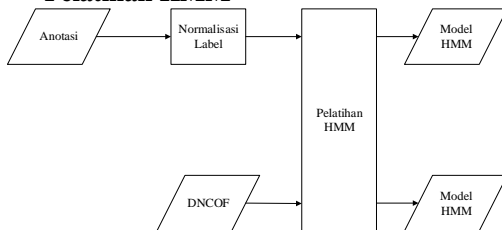
**Minor** : minor, minor7, minor9, minor sus4, diminished.

Setelah dilakukan proses penyederhanaan akor, dilakukan proses pelabelan chromagram dengan menyesuaikan anotasi dengan jumlah frame yang ada.

### 3.2.2 Proses Pelatihan Akor

Proses pelatihan akor merupakan proses pembentukan model akor yang terdiri dari *mean vector* (akor *mean*) dan *covariance* yang didapat dari *chromagram* yang telah diberi label setiap framenya.

### 3.3 Pelatihan HMM



Gambar 3-3 : Diagram blok pelatihan HMM

Terdapat dua macam pendekatan dalam tahap ini. pertama, model HMM dilatih menggunakan anotasi. kedua, model HMM dilatih menggunakan DNCOF matriks. Keluaran dari proses ini merupakan 24 macam model akor yang terdiri dari 2 parameter HMM yaitu probabilitas awal dan probabilitas transisi. Pada trained-HMM, pelatihan dilakukan menggunakan dataset anotasi yang telah dilakukan normalisasi terlebih dahulu. Probabilitas awal dihitung dengan mencari frekuensi kemunculan suatu akor pada dataset anotasi tersebut. Sedangkan probabilitas transisi dihitung dengan mencari frekuensi transisi dari state-i menuju ke state-j. Sedangkan pada untrained-HMM, diasumsikan setiap state memiliki peluang yang sama untuk berada di urutan awal sebuah sequence,

sehingga probabilitas awalnya  $\frac{1}{24}$ . Untuk probabilitas transisinya merupakan hasil perhitungan rumus DNCOF yang telah dipaparkan pada bab 2.

### 3.4 Perhitungan Akurasi

Perhitungan akurasi yang dilakukan adalah dengan membandingkan hasil pengenalan dengan nilai sebenarnya. Akurasi dapat dihitung dengan persamaan berikut :

$$Akurasi = \frac{\sum data\_uji\_benar}{\sum data\_keseluruhan} \times 100\%$$

## 4 Pengujian dan Analisa

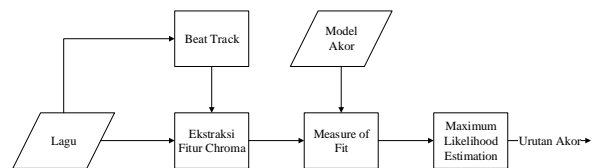
### 4.1 Skenario Pengujian

Pengujian pada sistem pengenalan akor ini didasarkan pada tujuan yang telah tercantup pada bagian pendahuluan diantaranya untuk mengimplementasikan HMM dalam sistem dengan menggunakan DNCOF sebagai parameter probabilitas transisi. Selanjutnya adalah untuk mengetahui dan menganalisa hasil perbandingan performansi sistem antara trained-HMM dengan untrained-HMM. Terdapat 3 macam dataset yang digunakan dalam pengujian ini yaitu:

- Dataset 1** : 36 buah lagu pendek yang dibuat menggunakan aplikasi *Band-in-a-Box* untuk menguji performansi sistem menggunakan data uji sederhana. Terdiri dari 3 variasi progresi akor I-V-VI-IV, VI-I-V-II, dan I-IV-V-I yang diterapkan ke dalam 12 key note yang berbeda.
- Dataset 2** : 17 lagu *zweieck* digunakan untuk mengetahui apakah kedua metode tersebut dapat diterapkan pada lagu lain.

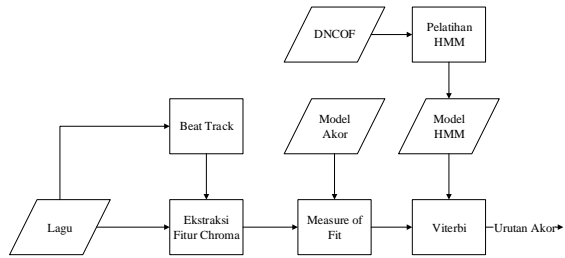
Dari 2 macam dataset tersebut, Pengujian dilakukan dengan menggunakan 3 skenario diantaranya:

- Skenario 1** : Pengujian dengan langsung mencari nilai maximum likelihood bertujuan untuk mengetahui performansi model akor yang dibangun. Skenario ini menggunakan 2 macam skema. Skema 1-1 menggunakan *trained-model* dan skema 2-1 menggunakan *untrained-model*.



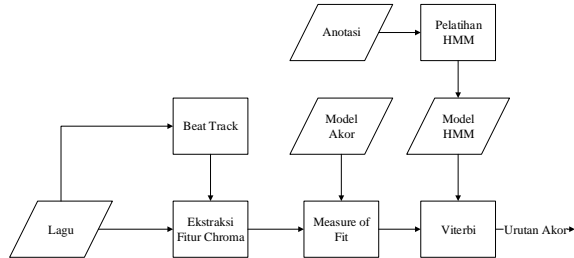
Gambar 4-1 : Skenario Pengujian 1

- Skenario 2** : Pengujian dengan menggunakan DNCOF sebagai probabilitas transisi untuk mengetahui apakah pengetahuan musik dapat diterapkan dan menghasilkan akurasi yang baik. Skenario ini menggunakan 2 macam skema. Skema 1-2 menggunakan *trained-model* dan skema 2-2 menggunakan *untrained-model*.



Gambar 4-2 : Skenario Pengujian 2

**3. Skenario 3** : Skenario ini bertujuan untuk membandingkan hasil akurasi sistem antara sistem *untrained*-HMM berbasis DNCOF dengan sistem *trained*-HMM berbasis anotasi. Skenario ini menggunakan 2 macam skema yaitu skema 2-1 menggunakan trained-model dengan *untrained*-HMM berbasis DNCOF dan skema 3-1 menggunakan *trained*-model dengan *trained*-HMM berbasis anotasi.



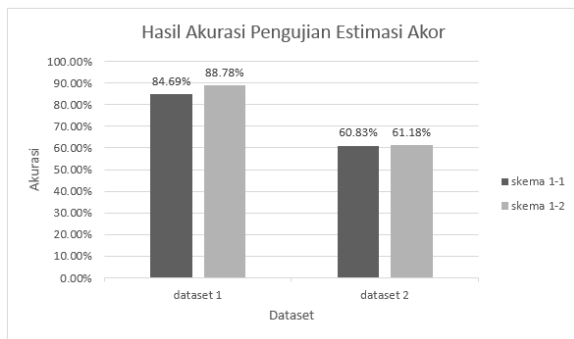
Gambar 4-3 : Skenario Pengujian 3

## 4.2 Pengujian dan Analisa

### 4.2.1 Pengujian dan Analisa Estimasi Akor

Tabel 4-1 : Hasil akurasi Pengujian estimasi akor

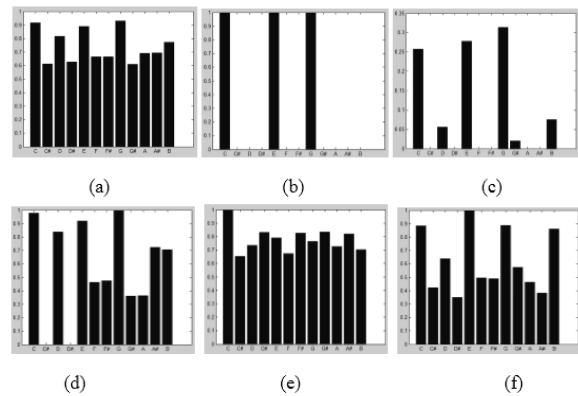
	Skema 1-1	Skema 2-1
Dataset 1	84.69%	88.78%
Dataset 2	60.83%	61.18%
rata-rata	72.76%	74.98%



Gambar 4-4 : Grafik hasil pengujian estimasi akor Tabel 4-1 merupakan hasil akurasi antara skema 1-1 yang menggunakan *trained*-model dengan skema 2-1 yang menggunakan *template* akor dan Gambar 4-1 merupakan grafik akurasi yang dapat menginformasikan perbandingan performansi antara skema 1-1 dengan skema 2-1. Pada dataset 1 terlihat bahwa keduanya menghasilkan akurasi yang sangat

baik dan tidak berbeda jauh. Begitu pula pada dataset 2, kedua skema tersebut memiliki performansi yang tidak berbeda jauh, hanya terputus sekitar 1% walau akurasi rendah pada pengujian dengan dataset yang kompleks.

**Analisa** : Penggunaan *template* akor dengan 6 harmonik pada skema 2-1 terbukti dapat menghasilkan akurasi yang lebih baik dibandingkan dengan yang dibentuk menggunakan pelatihan dataset 2 dengan selisih rata-rata 1,57%. Ini dikarenakan penggunaan *template* akor 6 harmonik pada skema 2-1 mendekati bentuk akor *mean* standar tanpa adanya gangguan berupa *noise* atau lainnya.



Gambar 4-5 : model akor C mayor (atas) pada trained model (a) template akor (b) dan template akor dengan 6 harmonik (c), beberapa contoh kandidat vektor chroma C mayor (bawah) yang terdapat pada dataset 1 (d) dataset 2 (e) dan dataset 3 (f).

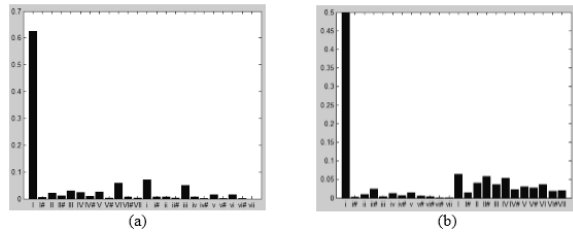
Pada gambar 4-7 sumbu x merupakan variasi 12 nada pada PCP dan sumbu y merupakan *amplitude* yang dihasilkan tiap nada. Kebanyakan kesalahan pada pengenalan akor ini terjadi karena akor-akor yang bertetangga memiliki struktur yang mirip satu sama lain. beberapa relasi akor yang memiliki satu atau dua nada yang sama diantaranya: *parallel* minor/major, *relative* minor/major, *subdominan*, *dominan* dan *median*.

Tabel 4-2 : Daftar relasi akor yang memiliki struktur akor yang mirip

Akor referensi	C:maj (I)	C:min (i)
Pararel	C:min (i)	C:maj (I)
Relative	A:min (vi)	G#:maj (VI)
Median	E:min (iii)	D#:maj (III)
subdominan	F:maj (IV)	F:min (iv)
dominan	G:maj (V)	G:min (v)

Contoh kasus pengenalan akor skema 1-1 dengan dataset 2 terlihat bahwa kebanyakan kesalahan terjadi pada akor yang memiliki struktur mirip dengan akor yang seharusnya. Paling banyak kesalahan tersebut terjadi pada relasi *parallel* major/minor dikarenakan relasi *parallel* memiliki dua nada yang sama, contoh C:maj memiliki nada C-E-G sedangkan

C:min memiliki nada C-Eb-G yang sama-sama memiliki nada C dan G di kedua akor tersebut.



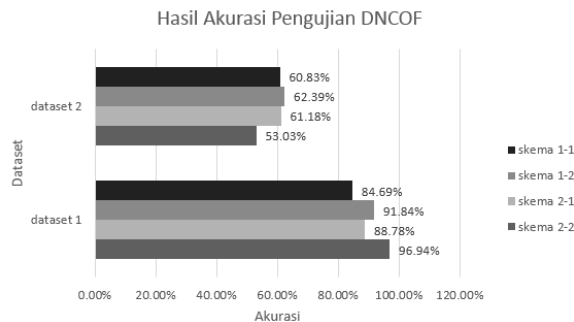
Gambar 4-6 : Grafik Confussion bar akor pada skema 1-1 dengan dataset 3 (a) akor mayor (b) akor minor

Pada gambar 4-8 sumbu x merupakan tingkatan akor dilihat dari sudut pandang akor sebenarnya (referensi) dan sumbu y merupakan prosentase hasil pengenalan dimana urutan pertama merupakan hasil yang benar sedangkan yang lainnya merupakan error.

#### 4.2.2 Pengujian dan Analisa Penggunaan Matriks Transisi DNCOF

Tabel 4-3 : Hasil Pengujian pada penggunaan matriks transisi DNCOF

	skema 1-1	skema 2-1	skema 1-2	skema 2-2
dataset 1	84.69%	88.78%	91.84%	96.94%
dataset 2	60.83%	61.18%	62.39%	53.03%
rata-rata	72.76%	74.98%	77.12%	74.99%



Gambar 4-7 : Grafik hasil pengujian pada penggunaan matriks transisi DNCOF

Pada pengujian ini terlihat bahwa penggunaan DNCOF dapat meningkatkan akurasi sistem. Pada kasus *trained-model-HMM*, penggunaan DNCOF dapat meningkatkan akurasi rata-rata sebesar 4.36%. Sedangkan pada kasus *untrained-model-HMM*, penggunaan DNCOF dapat meningkatkan akurasi rata-rata sebesar 0.01%.

**Analisa :** Pada hasil pengujian penggunaan matriks transisi DNCOF menunjukkan bahwa matriks transisi DNCOF dapat meningkatkan akurasi baik pada skema observasi yang menggunakan pelatihan maupun skema observasi yang menggunakan template akor. DNCOF dapat memperkecil kesalahan yang banyak terjadi pada beberapa akor yang memiliki struktur yang mirip. Pada gambar 4-7 dan 4-8 merupakan grafik *confussion*

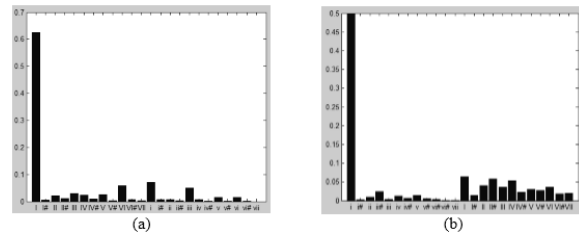
*bar* yang menginformasikan hasil prosentase pengenalan akor terhadap akor sebenarnya (referensi), terlihat bahwa terjadinya peningkatan akurasi pada penggunaan DNCOF dan juga terjadinya penurunan kesalahan pada relasi akor paralel, ini dikarenakan walaupun akor paralel memiliki struktur yang mirip tetapi menurut teori DNCOF memiliki jarak 4 dari jarak terjauh adalah 7. Sedangkan untuk relasi akor lainnya tidak adanya penurunan kesalahan dikarenakan menurut teori DNCOF jarak transisi akor pada relasi akor tersebut juga dekat sehingga masih sulit dibedakan.

Kebanyakan kesalahan pada pengenalan akor ini terjadi karena akor-akor yang bertetangga memiliki struktur yang mirip satu sama lain. beberapa relasi akor yang memiliki satu atau dua nada yang sama diantaranya: *parallel* minor/major, *relative* minor/major, *subdominan*, *dominan* dan *median*.

Tabel 4-4 : Daftar relasi akor yang memiliki struktur akor yang mirip

Akor referensi	C:maj (I)	C:min (i)
Pararel	C:min (i)	C:maj (I)
Relative	A:min (vi)	G#:maj (VI)
Median	E:min (iii)	D#:maj (III)
subdominan	F:maj (IV)	F:min (iv)
dominan	G:maj (V)	G:min (v)

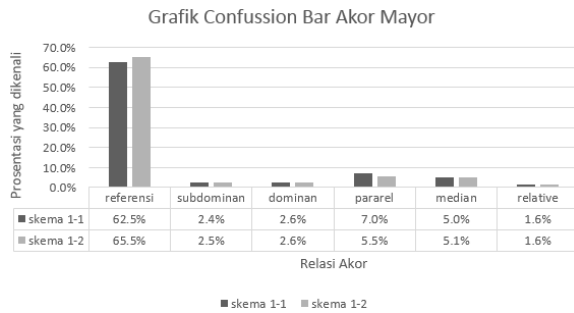
Contoh kasus pengenalan akor skema 1-1 dengan dataset 2 terlihat bahwa kebanyakan kesalahan terjadi pada akor yang memiliki struktur mirip dengan akor yang seharusnya. Paling banyak kesalahan tersebut terjadi pada relasi paralel major/minor dikarenakan relasi paralel memiliki dua nada yang sama, contoh C:maj memiliki nada C-E-G sedangkan C:min memiliki nada C-Eb-G yang sama-sama memiliki nada C dan G di kedua akor tersebut.



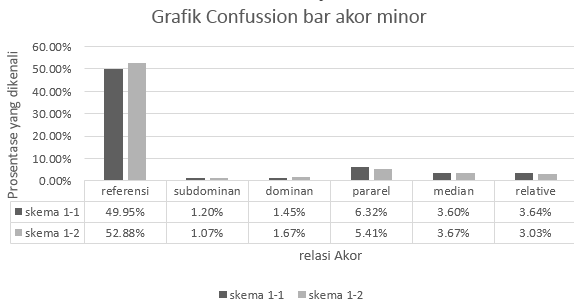
Gambar 4-8 : Grafik Confussion bar akor pada skema 1 dengan dataset 3 (a) akor mayor (b) akor minor

Pada gambar 4-9 sumbu x merupakan tingkatan akor dilihat dari sudut pandang akor sebenarnya (referensi) dan sumbu y merupakan prosentase hasil pengenalan dimana urutan pertama merupakan hasil yang benar sedangkan yang lainnya merupakan *error*.





Gambar 4-9 : Grafik perbandingan *Confussion bar* akor mayor



Gambar 4-10 : Grafik perbandingan *confussion bar* akor minor

Pada gambar 4-10 dan 4-11 dapat dilihat bahwa adanya peningkatan pengenalan akor referensi (akor yang benar) dalam penggunaan DNCOF pada skema 3 dengan peningkatan rata-rata sebesar 2.97% terhadap skema 1-1. Dan memperkecil kesalahan pengenalan terhadap akor parallel dengan penurunan error relasi akor *parallel* rata-rata sebesar 1.2%. Ini dikarenakan pada akor *parallel* walaupun memiliki kemiripan struktur tetapi pada kenyataannya memiliki hubungan ketetangaan yang jauh berdasarkan teori musik *Circle of Fifth*.

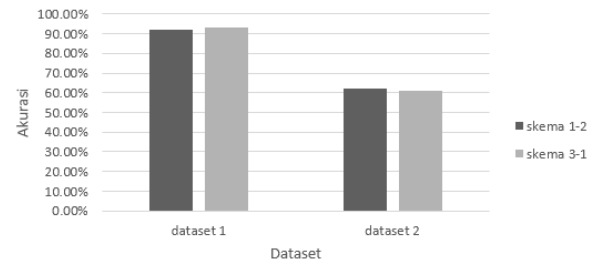
Pada pengujian terhadap dataset 2 dimana DNCOF dapat meningkatkan akurasi untuk skema yang menggunakan *trained*-model dengan perbandingan 60.83% menjadi 62.39%. Tetapi dengan pengujian yang sama yaitu terhadap dataset 2 dengan menggunakan *untrained*-model, penggunaan DNCOF malah menurunkan akurasi dari 61.18% menjadi 53.03%. ini dikarenakan DNCOF dibangun berdasarkan teori *Circle of Fifth* (COF), dimana COF merupakan hubungan ketetangaan tangga nada diatonik, dari 7 buah akor yang dapat dibentuk dari tangga nada tersebut hanya 2 akor yang digunakan yaitu akor pertama untuk mayor dan akor ke-enam untuk minor, sehingga tidak semua hubungan ketetangaan COF dapat dijelaskan pada DNCOF. Sedangkan pada dataset 2 yang merupakan kumpulan 17 lagu zweieck memiliki karakteristik kaya akan akor, walaupun beberapa akor mengalami normalisasi ke bentuk standar, tetapi transisi yang terdapat pada akor tersebut masih ada.

### 4.2.3 Pengujian dan Analisa Perbandingan Matriks Transisi

Tabel 4-5 : Hasil pengujian matriks transisi

	Skema 1-2	Skema 3-1
Dataset 1	91.84%	93.38%
Dataset 2	62.39%	61.01%
rata-rata	77.12%	77.20%

Hasil Akurasi Pengujian Perbandingan Matriks Transisi



Gambar 4-11 : Grafik akurasi pada pengujian matriks transisi

Tabel 4-5 merupakan hasil akurasi menggunakan 2 buah skema terhadap 2 dataset dan gambar 4-6 merupakan grafik perbandingan penggunaan matriks transisi antara DNCOF, dan pelatihan anotasi. Penggunaan DNCOF tidak lebih baik dari transisi berbasis anotasi bila sama-sama menggunakan model akor dari metode *trained-HMM* dengan perbandingan terhadap dataset 1 sebesar 91.84% berbanding 93.38%. Sedangkan penggunaan DNCOF pada dataset 2 terlihat lebih baik dibandingkan dengan yang berbasis anotasi dengan perbandingan sebesar 62.39% berbanding 61.01%.

**Analisa :** Pada pengujian ini sama-sama menggunakan *trained*-Model dengan skema 1-2 menggunakan matriks transisi DNCOF sedangkan skema 3-1 menggunakan matriks transisi berbasis anotasi. Terlihat bahwa skema 1-2 dengan matriks transisi DNCOF dapat menghasilkan akurasi yang paling tinggi untuk mengenali dataset sederhana yang yang terdapat pada dataset 1. Ini dikarenakan dataset 1 merupakan dataset yang berisi lagu-lagu yang memiliki komposisi instrument sederhana. Pada dataset 1 instrument didominasi oleh instrument harmonis yaitu instrument yang memainkan akor.

## 5 Kesimpulan dan Saran

### 5.1 Kesimpulan

Berdasarkan hasil pengujian yang dilakukan dapat ditarik kesimpulan bahwa:

1. Penggunaan template akor dengan 6 harmonik pada skema 2-1 terbukti dapat menghasilkan akurasi yang lebih baik dibandingkan dengan yang dibentuk menggunakan pelatihan pada skema 1-1 dengan selisih rata-rata 1,57%.

2. Pada kasus *trained*-model-HMM, penggunaan DNCOF dapat meningkatkan akurasi rata-rata sebesar 4.36%. Sedangkan pada kasus *untrained*-model-HMM, penggunaan DNCOF hanya meningkatkan akurasi rata-rata sebesar 0.01%.
3. Penggunaan DNCOF tidak lebih baik dari transisi berbasis anotasi bila sama-sama menggunakan model akor dari metode *trained*-HMM dengan perbandingan terhadap dataset 1 sebesar 91.84% berbanding 93.38%.

Kombinasi template akor dan DNCOF dapat menghasilkan akurasi terbaik pada dataset 1 yang memiliki komposisi instrument sederhana dengan akurasi sebesar 96.94% dalam mengenali akor yang lebih umum dilihat dari kedua data uji yang digunakan.

#### Daftar Pustaka:

- [1] P. W. Ellis, Daniel. Weller, Adrian. "THE 2010 LABROSA CHORD RECOGNITION SYSTEM". LabROSA, Columbia University
- [2] Uemura, Aiko. Katto, Jiro. "Chord Recognition Using Dolby Nested Circle Of Fifths", IEEE, 2012.
- [3] Perdana, Jati. "Deteksi *Chord* dengan Chroma-Log-Pitch Feature dan Untrained Hidden Markov Model", Informatics Faculty, Telkom Institute of Technology, 2013.
- [4] Syauqi, Riezhan. Intan Nurma, Syahrul., "Chord Segmentation and Recognition using EM-Trained Hidden Markov Models", Informatics Faculty, Telkom Institute of Technology, 2012.
- [5] Ferdian, Adi. "Pengenalan Akor Musik Menggunakan Fuzzy Hidden Markov Model", Informatics Faculty, Telkom Institute of Technology, 2013.
- [6] Harte, C., Sandler, M., Abdallah, S., Gómez, E.: Symbolic representation of musical chords: A proposed syntax for text annotations. In Reiss, J.D., Wiggins, G.A., eds.: Proceedings of the 6th International Conference on Music Information Retrieval, London (2005) 66–71
- [7] P. Bello, Juan. Pickens, Jeremy. "A Robust Mid-level Representation for Harmonic Content in Music Signals". Queen Mary, University of London, 2010
- [8] Papadopoulos, Helene. "Large-Scale Study of Chord Estimation Algorithms Based on Chroma Representation and HMM" Paris, 2007.
- [9] "Audio File Format", [http://en.wikipedia.org/wiki/Audio\\_file\\_format](http://en.wikipedia.org/wiki/Audio_file_format), diakses 28 Oktober 2014
- [10] "Pitch Class", [http://en.wikipedia.org/wiki/Pitch\\_class](http://en.wikipedia.org/wiki/Pitch_class), diakses 28 Oktober 2014
- [11] Oudre, Laurent. Gremier, Yves. Févotte, Cédric. "Chord Recognition by Fitting Rescaled Chroma Vectors to Chord Templates", IEEE, 2011.
- [12] Mauch Matthias, Dixon Simon, "Simultaneous Estimation of Chords and Musical Context From Audio", Music Information Retrieval Evaluation eXchange (MIREX 2010), 2010.
- [13] T. Fujishima, "Realtime chord recognition of musical sound: A system using Common Lisp Music," in Proc. Int. Comput. Music. Conf. (ICMC), Beijing, China, 1999, pp. 464–467.

#### 5.2 Saran

Berdasarkan penelitian yang telah dilakukan, penulis memberikan saran terhadap penelitian lebih lanjut terhadap topik ini:

1. Diharapkan pada penelitian kedepannya lebih fokus terhadap *preprocessing* data yaitu dengan meminimalisasi gangguan (*noise*) yang dapat memperkecil tingkat akurasi sistem serta menaikkan dominasi instrumen yang memainkan akor.

Diharapkan penelitian kedepannya menambah jumlah akor yang dikenali ini dimungkinkan karena dalam dataset 180 lagu *The Beatles* tak hanya terdapat akor mayor dan minor tetapi juga terdapat variasi akor lainnya.