

**SISTEM PENGENALAN *CHORD* PADA *FILE MUSIK DIGITAL*
DENGAN MENGGUNAKAN *PITCH CLASS PROFILES* DAN
*HIDDEN MARKOV MODEL***

Ivanna K. Timotius, Adhi Prayogo

Fakultas Teknik Elektronika dan Komputer, Program Studi Teknik Elektro,
Universitas Kristen Satya Wacana, Salatiga, Indonesia
ivanna_timotius@yahoo.com, adhi.prayogo@gmail.com

Intisari

Chord adalah salah satu elemen utama musik yang dapat memberikan gambaran sederhana dari suatu karya musik. Sebuah sistem yang dapat mengenali *chord* dari suatu rekaman musik secara otomatis akan sangat menarik dan bermanfaat. Tulisan ini mengimplementasikan sebuah algoritma pengenalan *chord* otomatis dengan menggunakan *pitch class profiles* dan *hidden Markov model*. Algoritma yang diimplementasikan menunjukkan tingkat akurasi pengenalan sebesar 100% untuk file lagu progresi *chord* murni dari MIDI, 97,97% untuk file lagu progresi *chord* murni dari rekaman gitar, 85,35% untuk file lagu kompleks dari MIDI, dan 59,80% untuk file lagu kompleks yang mengandung vokal manusia.

Kata kunci: *chord*, *pitch class profiles*, *hidden Markov model*.

1. Pendahuluan

Sistem pendengaran manusia mempunyai kemampuan untuk menyarikan data yang kaya dari sinyal audio yang kompleks. Misalnya manusia dapat memahami percakapan dan menikmati musik. Musik maupun percakapan termasuk sinyal audio yang kompleks. Salah satu elemen utama musik adalah *chord*. Rangkaian *chord* dapat memberikan gambaran sederhana dari sebuah karya musik.

Transkripsi musik didefinisikan sebagai kegiatan mendengarkan sebuah karya musik dan menulis notasi musik untuk setiap nada yang menyusun karya musik tersebut.

sebuah lagu yang berpindah – pindah setiap waktu tertentu berdasarkan probabilitas tertentu. Suatu *chord* dalam sebuah lagu akan mempunyai kecenderungan lebih kuat untuk berpindah ke suatu *chord* tertentu daripada ke *chord* lainnya.

Terdapat beberapa kesulitan yang dihadapi suatu sistem pengenalan *chord* otomatis di antaranya adalah banyaknya jenis alat musik yang digunakan dengan warna suara yang berbeda-beda, adanya alat musik perkusi, adanya *chord* yang nada-nadanya tidak dimainkan secara simultan (*broken chord* atau *arpeggio*), dan adanya melodi lagu.

2. Chord

Chord merupakan sekumpulan nada dengan pola tertentu. Sinyal dari sebuah nada yang berbeda memiliki frekuensi yang berbeda. Karena itu, sinyal dari dua buah *chord* yang berbeda, yang disusun oleh nada yang berbeda, akan memiliki karakteristik frekuensi yang berbeda pula.

Chord dalam musik adalah tiga atau lebih nada – nada yang dibunyikan bersama – sama [3]. *Chord* diklasifikasikan berdasarkan interval di antara nada – nada pembentuknya. Interval dalam musik berarti jarak di antara dua nada musikal. Jenis *chord* yang paling umum dan paling sederhana adalah triad, yaitu yang terdiri dari tiga buah nada. Dua jenis triad yang paling umum adalah major triad dan minor triad. Bila sebuah interval ditambahkan lagi pada sebuah triad, maka akan membentuk *seventh chord* yang terdiri dari empat nada. *Seventh chord* yang paling umum adalah *dominant seventh chord*. Contoh major triad, minor triad, *dominant seventh chord* dengan *root C* yang dituliskan dalam notasi balok ditunjukkan pada Gambar 1



Gambar 1. Notasi Balok untuk (a) C Major Triad (b) C Minor Triad (c) C Dominant Seventh.

sebuah lagu yang berpindah – pindah setiap waktu tertentu berdasarkan probabilitas tertentu. Suatu *chord* dalam sebuah lagu akan mempunyai kecenderungan lebih kuat untuk berpindah ke suatu *chord* tertentu daripada ke *chord* lainnya.

Terdapat beberapa kesulitan yang dihadapi suatu sistem pengenalan *chord* otomatis di antaranya adalah banyaknya jenis alat musik yang digunakan dengan warna suara yang berbeda-beda, adanya alat musik perkusi, adanya *chord* yang nada-nadanya tidak dimainkan secara simultan (*broken chord* atau *arpeggio*), dan adanya melodi lagu.

2. Chord

Chord merupakan sekumpulan nada dengan pola tertentu. Sinyal dari sebuah nada yang berbeda memiliki frekuensi yang berbeda. Karena itu, sinyal dari dua buah *chord* yang berbeda, yang disusun oleh nada yang berbeda, akan memiliki karakteristik frekuensi yang berbeda pula.

Chord dalam musik adalah tiga atau lebih nada – nada yang dibunyikan bersama – sama [3]. *Chord* diklasifikasikan berdasarkan interval di antara nada – nada pembentuknya. Interval dalam musik berarti jarak di antara dua nada musikal. Jenis *chord* yang paling umum dan paling sederhana adalah triad, yaitu yang terdiri dari tiga buah nada. Dua jenis triad yang paling umum adalah major triad dan minor triad. Bila sebuah interval ditambahkan lagi pada sebuah triad, maka akan membentuk *seventh chord* yang terdiri dari empat nada. *Seventh chord* yang paling umum adalah *dominant seventh chord*. Contoh major triad, minor triad, *dominant seventh chord* dengan *root C* yang dituliskan dalam notasi balok ditunjukkan pada Gambar 1



Gambar 1. Notasi Balok untuk (a) C Major Triad (b) C Minor Triad (c) C Dominant Seventh.

Nilai elemen-elemen pada vektor PCP $x_{pcp}[i]$, $0 \leq i \leq N_c - 1$, didapatkan dengan merata-ratakan magnitudo dari sinyal di ranah frekuensi $X[k]$ yang memiliki kelas nada yang sama:

$$x_{pcp}[i] = \frac{1}{\#\{k \in \{P[k] = i\}\}} \sum_{P[k]=i} |X[k]|^2 \quad (4)$$

4. Hidden Markov Model (HMM)

HMM merupakan sebuah sistem pemodelan yang bertujuan untuk mengenali suatu pola dengan menggunakan proses pembelajaran terlebih dahulu. Sebuah HMM adalah sebuah proses acak ganda (*doubly stochastic process*) [5][6] dengan satu proses acak pokok yang tidak dapat diobservasi (*hidden*), tetapi hanya dapat diobservasi melalui himpunan proses acak lainnya yang menghasilkan urutan simbol yang terobservasi. HMM juga dapat dipandang sebagai sebuah automaton acak terbatas dimana setiap *state*-nya menghasilkan suatu observasi. Transisi antar *state* mematuhi properti Markov, yaitu perpindahan antar *state* pada HMM orde n hanya bergantung pada n *state* sebelumnya.

Salah satu tipe HMM adalah *ergodic* HMM dimana setiap *state* dapat dicapai dari setiap *state* lainnya dalam sejumlah langkah tertentu. Pada penerapan konsep HMM untuk membangun aplikasi pengenalan *chord*, setiap *chord* diwakili oleh sebuah *state* dalam proses acak utama. Proses acak utama adalah proses transisi antar *chord*, yang mematuhi properti Markov. Proses acak utama ini tersembunyi. Pada setiap saat t , *state* aktif akan menghasilkan sebuah vektor PCP, vektor PCP inilah yang merupakan urutan simbol yang dapat diobservasi.

Sebuah HMM λ terdiri dari parameter – parameter A , B , dan π . Pernyataan ini sering ditulis dengan

$$\lambda = (A, B, \pi) \quad (5)$$

Parameter A adalah matriks transisi dari *hidden status*, dengan $A = \{a_{ij}\}$ dan a_{ij} adalah peluang *next state* q_j jika *present state* q_i ,

$$a_{ij} = P[q_j|q_i] \quad (6)$$

Nilai elemen-elemen pada vektor PCP $x_{pcp}[i]$, $0 \leq i \leq N_c - 1$, didapatkan dengan merata-ratakan magnitudo dari sinyal di ranah frekuensi $X[k]$ yang memiliki kelas nada yang sama:

$$x_{pcp}[i] = \frac{1}{\#\{k \in \{1, \dots, N\} : C[k] = i\}} \sum_{k \in C^{-1}(i)} |X[k]|^2 \quad (4)$$

4. Hidden Markov Model (HMM)

HMM merupakan sebuah sistem pemodelan yang bertujuan untuk mengenali suatu pola dengan menggunakan proses pembelajaran terlebih dahulu. Sebuah HMM adalah sebuah proses acak ganda (*doubly stochastic process*) [5][6] dengan satu proses acak pokok yang tidak dapat diobservasi (*hidden*), tetapi hanya dapat diobservasi melalui himpunan proses acak lainnya yang menghasilkan urutan simbol yang terobservasi. HMM juga dapat dipandang sebagai sebuah automaton acak terbatas dimana setiap *state*-nya menghasilkan suatu observasi. Transisi antar *state* mematuhi properti Markov, yaitu perpindahan antar *state* pada HMM orde n hanya bergantung pada n *state* sebelumnya.

Salah satu tipe HMM adalah *ergodic* HMM dimana setiap *state* dapat dicapai dari setiap *state* lainnya dalam sejumlah langkah tertentu. Pada penerapan konsep HMM untuk membangun aplikasi pengenalan *chord*, setiap *chord* diwakili oleh sebuah *state* dalam proses acak utama. Proses acak utama adalah proses transisi antar *chord*, yang mematuhi properti Markov. Proses acak utama ini tersembunyi. Pada setiap saat t , *state* aktif akan menghasilkan sebuah vektor PCP, vektor PCP inilah yang merupakan urutan simbol yang dapat diobservasi.

Sebuah HMM λ terdiri dari parameter – parameter A , B , dan π . Pernyataan ini sering ditulis dengan

$$\lambda = (A, B, \pi) \quad (5)$$

Parameter A adalah matriks transisi dari *hidden status*, dengan $A = \{a_{ij}\}$ dan a_{ij} adalah peluang *next state* q_j jika *present state* q_i ,

$$a_{ij} = P[q_j|q_i] \quad (6)$$

dipilih sedemikian sehingga dapat mewakili seluruh jenis *chord* yang ingin diklasifikasikan.

2. 14 *file* lagu berisi rekaman progresi *chord* murni yang direkam dari permainan gitar dengan progresi *chord* yang sama dengan kelompok sebelumnya.
3. 5 *file* lagu berisi lagu kompleks yang direkam dari MIDI. Lagu kompleks adalah lagu umum yang sudah mengandung melodi dan diiringi oleh beberapa alat musik, lengkap dengan alat musik perkusi. Lagu didapatkan dari *file* – *file* MIDI yang suaranya direkam menjadi *file* berformat wav.
4. 3 *file* lagu berisi lagu kompleks komersial yang mengandung suara vokal manusia dan diiringi oleh beberapa alat musik.

Pada sistem yang dikembangkan ini dipakai *frame* yang saling *overlap* dengan panjang *frame* $N = 4096$ dan data *ter-overlap* 50%. Dengan frekuensi cuplik dari sinyal masukan sebesar 11025 Hz, akan menghasilkan *frame* sinyal dengan kerapatan sekitar 5 *frame* per detik. Panjang *frame* ini dipilih karena lagu – lagu pada umumnya, *chord* baru akan berganti setelah beberapa ketukan. Jika sebuah sinyal lagu masukan mempunyai tempo 120 ketukan per menit dan ada bagian lagu yang *chord*-nya berganti setiap ketukan, maka dibutuhkan *frame* sinyal untuk setiap ketukan, yang dapat dicapai cukup dengan kerapatan $120/60 = 2$ *frame* per detik saja. Karena itu kerapatan 5 *frame* per detik yang diimplementasikan sudah memadai untuk menangani lagu – lagu pada umumnya.

Jumlah kelas nada N_c pada PCP adalah 24 ditentukan secara empirik sehingga menghasilkan akurasi yang terbesar. Gambar 2 menunjukkan hasil vektor PCP terhadap sumbu waktu dalam detik.

dipilih sedemikian sehingga dapat mewakili seluruh jenis *chord* yang ingin diklasifikasikan.

2. 14 *file* lagu berisi rekaman progresi *chord* murni yang direkam dari permainan gitar dengan progresi *chord* yang sama dengan kelompok sebelumnya.
3. 5 *file* lagu berisi lagu kompleks yang direkam dari MIDI. Lagu kompleks adalah lagu umum yang sudah mengandung melodi dan diiringi oleh beberapa alat musik, lengkap dengan alat musik perkusi. Lagu didapatkan dari *file* – *file* MIDI yang suaranya direkam menjadi *file* berformat wav.
4. 3 *file* lagu berisi lagu kompleks komersial yang mengandung suara vokal manusia dan diiringi oleh beberapa alat musik.

Pada sistem yang dikembangkan ini dipakai *frame* yang saling *overlap* dengan panjang *frame* $N = 4096$ dan data *ter-overlap* 50%. Dengan frekuensi cuplik dari sinyal masukan sebesar 11025 Hz, akan menghasilkan *frame* sinyal dengan kerapatan sekitar 5 *frame* per detik. Panjang *frame* ini dipilih karena lagu – lagu pada umumnya, *chord* baru akan berganti setelah beberapa ketukan. Jika sebuah sinyal lagu masukan mempunyai tempo 120 ketukan per menit dan ada bagian lagu yang *chord*-nya berganti setiap ketukan, maka dibutuhkan *frame* sinyal untuk setiap ketukan, yang dapat dicapai cukup dengan kerapatan $120/60 = 2$ *frame* per detik saja. Karena itu kerapatan 5 *frame* per detik yang diimplementasikan sudah memadai untuk menangani lagu – lagu pada umumnya.

Jumlah kelas nada N_c pada PCP adalah 24 ditentukan secara empirik sehingga menghasilkan akurasi yang terbesar. Gambar 2 menunjukkan hasil vektor PCP terhadap sumbu waktu dalam detik.

MIDI dan 3 lagu yang mengandung suara vokal manusia. Kepada peserta tes lagu diperdengarkan sebanyak dua kali, hal ini dilakukan untuk memberi kesempatan berpikir dan menulis. Nada dasar dari lagu yang diujikan juga diberitahukan terlebih dahulu kepada para peserta tes. Hal ini dilakukan karena orang yang mampu mengenali *chord* dari lagu tanpa diberitahukan nada dasar lagu terlebih dahulu adalah orang yang memiliki talenta istimewa dan sangat jarang ditemui. Tes diberikan kepada 10 orang yang terdiri dari 4 orang baru berlatih musik di bawah 5 tahun, 3 orang sudah berlatih musik antara 5 hingga kurang dari 10 tahun, dan 3 orang sudah berlatih musik selama 10 tahun atau lebih. Gambar 3 menunjukkan grafik perbandingan tingkat akurasi rata – rata pengenalan *chord* oleh manusia dari tiap kelompok orang dengan pengenalan *chord* otomatis yang berbasis HMM

Dari grafik dapat dilihat bahwa tingkat akurasi sistem dalam mengenali *chord* kira – kira sebanding dengan orang yang berlatih musik selama 10 tahun. Dengan demikian sistem pengenalan *chord* otomatis yang telah direalisasikan dapat bermanfaat untuk membantu orang – orang yang belum begitu terlatih musik untuk mendapatkan *chord* yang benar dari suatu rekaman musik.

Tabel 1. Akurasi Sistem Pengenalan *Chord*.

Kelompok File Lagu	Akurasi (%)
Progresi <i>chord</i> murni dari MIDI	100
Progresi <i>chord</i> murni dari gitar	97,97%
Lagu kompleks dari MIDI	85,35%
Lagu kompleks dengan vokal manusia	59,80%

MIDI dan 3 lagu yang mengandung suara vokal manusia. Kepada peserta tes lagu diperdengarkan sebanyak dua kali, hal ini dilakukan untuk memberi kesempatan berpikir dan menulis. Nada dasar dari lagu yang diujikan juga diberitahukan terlebih dahulu kepada para peserta tes. Hal ini dilakukan karena orang yang mampu mengenali *chord* dari lagu tanpa diberitahukan nada dasar lagu terlebih dahulu adalah orang yang memiliki talenta istimewa dan sangat jarang ditemui. Tes diberikan kepada 10 orang yang terdiri dari 4 orang baru berlatih musik di bawah 5 tahun, 3 orang sudah berlatih musik antara 5 ~~bingga kurang dari 10 tahun, dan 3 orang sudah berlatih musik selama 10 tahun atau lebih.~~ Gambar 3 menunjukkan grafik perbandingan tingkat akurasi rata – rata pengenalan *chord* oleh manusia dari tiap kelompok orang dengan pengenalan *chord* otomatis yang berbasis HMM

Dari grafik dapat dilihat bahwa tingkat akurasi sistem dalam mengenali *chord* kira – kira sebanding dengan orang yang berlatih musik selama 10 tahun. Dengan demikian sistem pengenalan *chord* otomatis yang telah direalisasikan dapat bermanfaat untuk membantu orang – orang yang belum begitu terlatih musik untuk mendapatkan *chord* yang benar dari suatu rekaman musik.

Tabel 1. Akurasi Sistem Pengenalan *Chord*.

Kelompok File Lagu	Akurasi (%)
Progresi <i>chord</i> murni dari MIDI	100
Progresi <i>chord</i> murni dari gitar	97,97%
Lagu kompleks dari MIDI	85,35%
Lagu kompleks dengan vokal manusia	59,80%

**SISTEM PENGENALAN CHORD PADA FILE MUSIK DIGITAL DENGAN
MENGUNAKAN PITCH CLASS PROFILES DAN HIDDEN MARKOV MODEL**

Iyanna K. Timotius, Adhi Prayogo

6. L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceeding of IEEE*, vol. 77, no. 2, pp. 257-286, Februari 1989.



**SISTEM PENGENALAN CHORD PADA FILE MUSIK DIGITAL DENGAN
MENGUNAKAN PITCH CLASS PROFILES DAN HIDDEN MARKOV MODEL**

Iyanna K. Timotius, Adhi Prayogo

6. L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceeding of IEEE*, vol. 77, no. 2, pp. 257-286, Februari 1989.

