

PEMODELAN PROBABILISTIK NEURAL NETWORK UNTUK KONVERSI SUARA GITAR KE CORD

Arviani Rizki⁽¹⁾
arvianirizki@gmail.com

Agus Buono⁽²⁾
pudesha@gmail.com

Abstract

Almost all music genre use guitar as its instrument. To produce a harmonic guitar voice needs guitar chords mastery. However, only few people are able to distinguish guitar chords. This paper is addressed to develop a computational model to convert guitar voice into appropriate cord. In this research, we use Mel Frequency Cepstrum Coefficient (MFCC) as feature extraction because this technique is often used for voice processing and good enough in presenting the characteristics of a signal voice. Probabilistic Neural Network (PNN) is implemented to classify the feature into one out of 24 classes of cord. We record 345 for each cord (totally we have 8640 recording data with WAV format). Experiments are conducted for some number of cepstral coefficients (13, 26, 39 and 52), with 100 millisecond as time frame and 40% overlapping between successive frame. According to the experiment, the maximum accuracy is 94.31% for 52 number of cepstral coefficients.

Keywords: *Chord-identification, Mel Frequency Cepstral Coefficients (MFCC), Probabilistic Neural Network, Speaker Recognition.*

1. Pendahuluan

Gitar merupakan alat musik yang paling umum dimainkan oleh berbagai kalangan usia. Hampir semua jenis musik menggunakan gitar sebagai instrumennya. Untuk menghasilkan suara gitar yang harmonis dibutuhkan penguasaan kunci atau *chord* gitar. Hendro (2004) menjelaskan bahwa *chord* merupakan satuan nada-nada yang dibunyikan secara serentak yang berfungsi sebagai pengiring dalam lagu maupun permainan musik (Wisnuadisastra 2009). Pendengaran manusia memiliki kemampuan dan kepekaan yang beragam dalam mengenali *chord* gitar. Seorang yang terbiasa dengan suara gitar dapat dengan mudah mengenali dan membedakan *chord* gitar, namun bagi orang awam hal tersebut sulit untuk dilakukan.

Penggunaan teknologi dan komputer dapat dilakukan untuk memecahkan masalah tersebut, yaitu pembuatan sistem yang dapat mengenali suara *chord* gitar. Kinerja komputer yang subjektif dan konsisten dalam mengeksekusi setiap perintah dapat memudahkan pengguna untuk mengetahui dan membedakan suara *chord* gitar

¹Departemen Ilmu Komputer, FMIPA Institut Pertanian Bogor

²Departemen Ilmu Komputer, FMIPA Institut Pertanian Bogor

dengan akurat. Pengenalan suara (*speech recognition*) adalah teknologi yang dapat mengubah sinyal suara menjadi sebuah informasi berupa teks. Suara *chord* gitar dapat diidentifikasi dengan mengekstraksi ciri dari setiap *chord* menggunakan teknik *Mel-Frequency Cepstrum Coefficient* (MFCC) untuk kemudian dilakukan pengenalan pola menggunakan *Probabilistic Neural Network* (PNN). PNN sebagai salah satu arsitektur JST telah terbukti menghasilkan akurasi tinggi di beberapa penelitian sebelumnya. Penelitian yang telah dilakukan oleh Fransiswa (2010) tentang pengembangan PNN pada pengenalan kisaran usia dan jenis kelamin berbasis suara memiliki rata-rata akurasi 91.26%.

Penelitian identifikasi suara gitar sebelumnya dilakukan oleh Wisnuadisastra (2009) menggunakan teknik MFCC dan *codebook* sebagai pengenalan pola. Penelitian tersebut menghasilkan rata-rata akurasi sebesar 91% menggunakan 13 koefisien *cepstral* dan 96% menggunakan 26 koefisien *cepstral*. Penelitian tersebut memiliki keterbatasan untuk mengenali satu jenis *chord* gitar sehingga belum dapat mengidentifikasi *chord* gitar yang kontinu. Pada penelitian ini, identifikasi *chord* gitar dilakukan untuk *chord* gitar yang kontinu menggunakan teknik MFCC untuk ekstraksi ciri dan model PNN untuk pengenalan pola.

Selanjutnya, paper ini disajikan dengan susunan sebagai berikut: bagian 2 mendeskripsikan metode percobaan, data yang dipergunakan, serta pemrosesan yang dilakukan. Bagian berikutnya difokuskan pada diskusi mengenai hasil dan pembahasan. Akhirnya, pada bagian 4 akan disajikan kesimpulan dari penelitian ini.

2. Metode Penelitian, Data, dan Pemrosesan

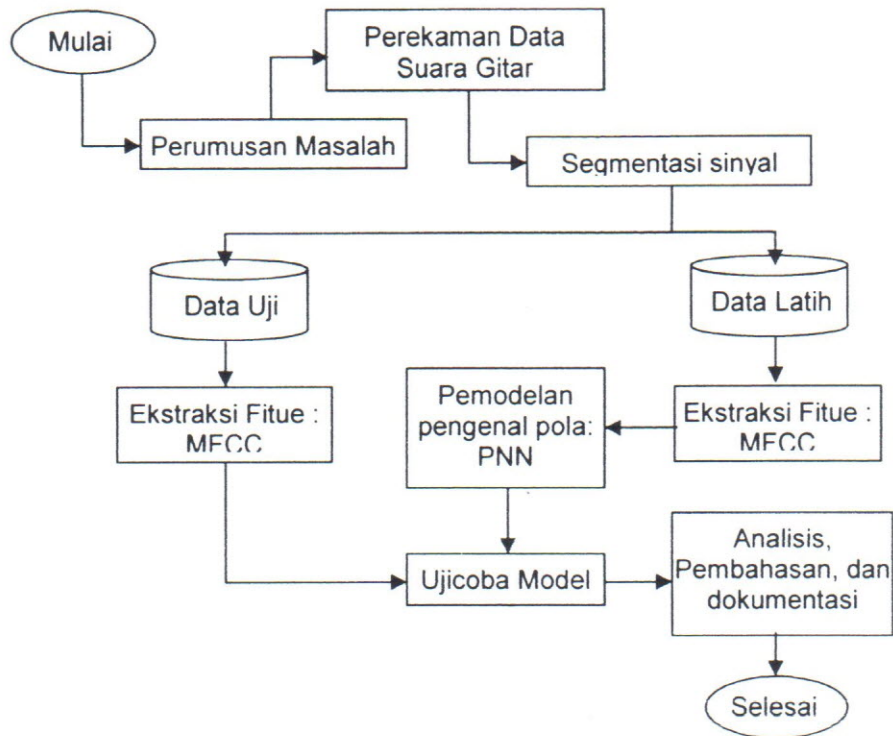
2.1 Metode Penelitian

Sistem identifikasi *chord* gitar diimplementasikan melalui suatu program dengan menggunakan software Matlab R2008b. Pembuatan sistem dibagi menjadi beberapa tahap sesuai dengan diagram alur penelitian yang ditunjukkan pada Gambar 1.

2.2 Data Penelitian

Penelitian ini menggunakan data hasil perekaman suara gitar yang telah dikonversi ke bentuk digital berformat WAV. Data suara gitar merupakan dua campuran dari 12 *chord* mayor dan 12 *chord* minor (Tabel 1). Perekaman suara gitar dilakukan ditempat yang hening untuk meminimalkan adanya *noise*. Perekaman dilakukan 15 kali untuk setiap dua campuran *chord* sehingga dihasilkan 8640 data suara gitar.

Suara gitar direkam menggunakan ukuran waktu tiga detik dengan *sampling rate* sebesar 11.000 Hz dan *bit rate* sebesar 16 bit. Data suara gitar berupa campuran *chord* C ke *chord* C#, *chord* C ke *chord* Cm dan seterusnya masing-masing sebanyak 15 kali. Untuk setiap campuran digunakan sepuluh kali perekaman untuk data latih dan lima kali perekaman untuk data uji, sehingga terdapat 5760 data latih dan 2880 data uji.



Gambar 1. Blok Diagram Penelitian

Tabel 1.

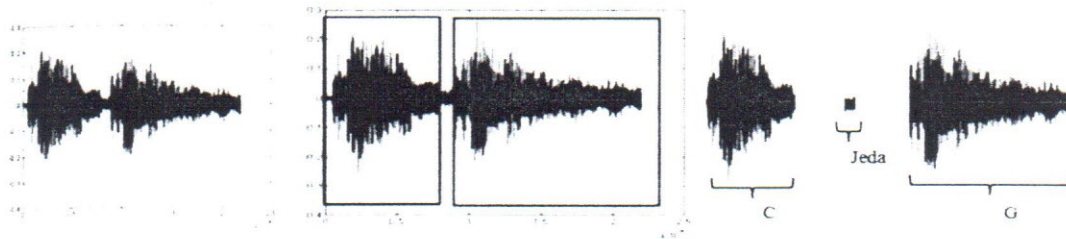
Chord Mayor dan Minor sebagai kelas dalam penelitian ini

Cord Dasar	Mayor	Minor
C	C	Cm
C#	C#	C#m
D	D	Dm
D#	D#	D#m
E	E	Em
F	F	Fm
F#	F#	F#m
G	G	Gm
G#	G#	G#m
A	A	Am
A#	A#	A#m
B	B	Bm

2.3 Pemrosesan Penelitian

Selanjutnya pada setiap data tersebut dilakukan segmentasi sinyal secara manual seperti diperlihatkan pada Gambar 1. Tahap segmentasi merupakan tahap

pemotongan jeda pada setiap sinyal suara. Pendeteksian jeda dilakukan pada sinyal untuk memisahkan dua campuran *chord* menjadi dua sinyal yang berbeda. Dengan demikian sinyal suara akan dipisah secara manual menjadi *chord* pertama, jeda, dan *chord* kedua. Pemisahan dilakukan dengan mencari rentang terbesar dari setiap frame yang nilainya kurang dari rata-ran frame.

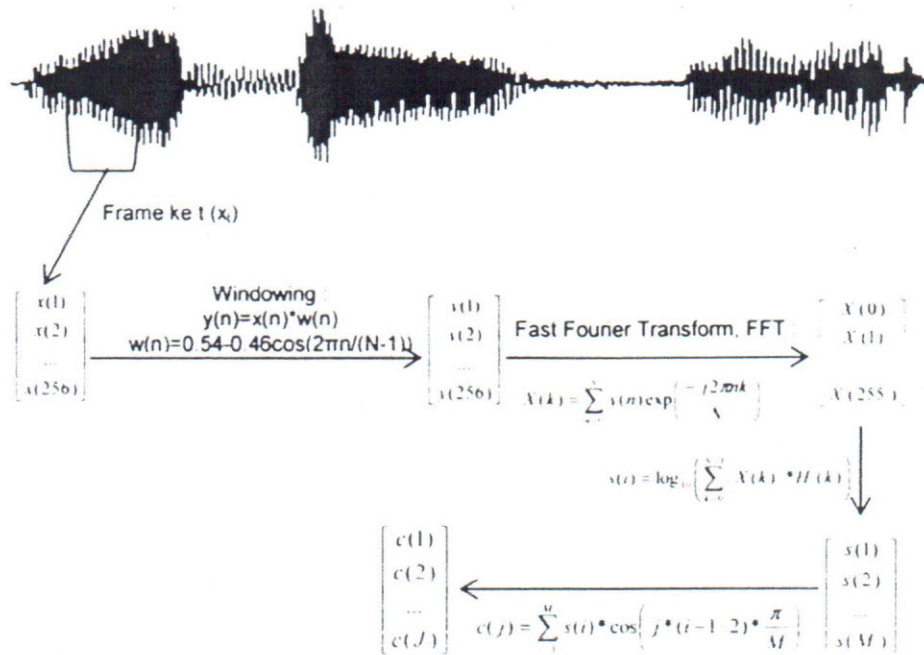


Gambar 2. Segmentasi Sinyal Suara

Pada setiap sinyal yang sudah disegmentasi, selanjutnya dilakukan ekstraksi ciri. Ekstraksi ciri merupakan proses untuk menentukan suatu nilai atau vektor yang dapat dipergunakan sebagai pencari objek atau individu (Buono 2009). Pada Penelitian ini, teknik ekstraksi ciri yang dipergunakan adalah MFCC. Tahapan MFCC adalah seperti disajikan pada Gambar 3. Seperti pada Gambar 3 sinyal suara dibaca dari *frame* ke *frame* dengan lebar tertentu yang saling tumpang tindih (*overlap*), kemudian dilakukan *windowing* pada setiap *frame*. Tahap berikutnya, dilakukan transformasi *Fourier* untuk mengubah suara dari domain waktu ke domain frekuensi. Nilai hasil transformasi *Fourier* ini dihitung *spectrum mel* menggunakan sejumlah filter. Koefisien MFCC merupakan hasil transformasi *Cosinus* dari *spectrum mel* tersebut dan dipilih K koefisien. Transformasi *Cosinus* berfungsi untuk mengembalikan domain, dari domain frekuensi ke domain waktu (Ganchev, 2005).

Dengan demikian, untuk setiap sinyal hasil segmentasi akan dihasilkan sejumlah frame dan pada setiap frame dikonversi menjadi koefisien MFCC sebagai cirinya. Selanjutnya, vektor ciri setiap sinyal adalah sebuah vektor yang merupakan rata-rata dari semua frame yang berasal dari sinyal tersebut. Tahapan selanjutnya adalah memilah data vektor ciri ini menjadi data latih (untuk membuat model pengenalan pola, yang dalam penelitian ini menggunakan PNN) dan data uji.

Kudang *et al* (2005) mengatakan model PNN merupakan salah satu jenis jaringan saraf tiruan yang memiliki tingkat akurasi yang cukup tinggi dalam pengklasifikasian, yaitu hampir mendekati 100% (Fransiswa 2010). Dengan menggunakan pengklasifikasi Bayesian dapat ditentukan bagaimana sebuah data masukan diklasifikasi sebagai anggota suatu kelas dari beberapa kelas yang ada, yaitu yang mempunyai nilai maksimum pada kelas tersebut. PNN memiliki struktur yang terdiri atas empat *layer* sepertipada Gambar 4.



Gambar 3. Tahapan Ekstraksi Sinyal Suara dengan Teknik MFCC. (dimodifikasi dari Nilsson dan Egnarsson, 2002)

Input layer merupakan layer untuk input data yang akan diuji.

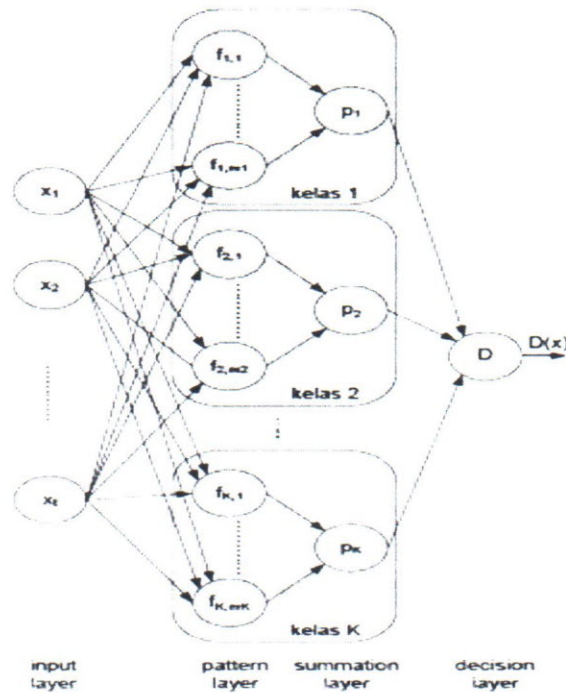
1. *Pattern layer* berfungsi menghitung jarak antara nilai input data suara dengan nilai pola dari tiap anggota kelas yang dirumuskan sesuai formula berikut:

$$f_i(\mathbf{x}) = \prod_{j=1}^d k(z) \left(\frac{x_j - x_{ij}}{h_j} \right)$$

2. *Summation Layer*. Layer ini menghasilkan peluang untuk satu kelas. Peluang tersebut didapat dari penjumlahan *pattern layer* pada kelas tersebut dan hasilnya dibagi dengan $(2\pi)^{d/2} h_1 h_2 \dots h_n$. Nilai h_1, h_2, \dots, h_n adalah nilai *smoothing* dari kelas tersebut.
3. *Decision layer* digunakan untuk membandingkan hasil peluang pada setiap kelas. Selanjutnya, input data dimasukkan dalam kelas yang memiliki nilai peluang terbesar.

Keterangan:

- d = banyaknya data pada satu *pattern layer*
- $k(z)$ = $e^{-0.5 \cdot z^2}$
- x_i = input data uji ke- j
- x_{ij} = *pattern* ke- i data ke- j
- h_j = *smoothing* parameter ($\alpha \times$ simpangan baku ke- $j \times n^{1/5}$)
- n = banyaknya *pattern* pada satu kelas



Gambar 4. Struktur Probabilistik Neural Network (Ganchev, 2005)

3. Hasil dan Pembahasan

3.1 Percobaan dengan 13 Koefisien Cepstral

Hasil percobaan menggunakan 13 koefisien *cepstral* dapat dilihat pada Tabel 2. Akurasi terkecil terdapat pada percobaan *chord* E yaitu 13.33% sedangkan akurasi terbesar terdapat pada percobaan *chord* Cm yaitu 86.67%.

Tabel 2.
Akurasi dengan 13 Koefisien Cepstral

Chord	Akurasi	Chord	Akurasi	Chord	Akurasi
A	59.17	Cis	62.50	F	14.17
Am	33.33	Cism	64.17	Fm	49.17
Ais	52.50	D	29.17	Fis	67.50
Aism	52.50	Dm	70.83	Fism	67.50
B	40.00	Dis	75.00	G	60.00
Bm	70.83	Dism	48.33	Gm	42.50
C	44.17	E	13.33	Gis	28.33
Cm	86.67	Em	60.83	Gism	71.67
Total					52.67

3.2 Percobaan dengan 26 Koefisien Cepstral

Hasil percobaan menggunakan 26 koefisien *cepstral* dapat dilihat pada Tabel 3. Akurasi terkecil terdapat pada percobaan *chord* F yaitu 45% sedangkan akurasi terbesar

terdapat pada percobaan *chord* A yaitu 97.5%.

Tabel 3.
Akurasi dengan 26 Koefisien Cepstral

Chord	Akurasi	Chord	Akurasi	Chord	Akurasi
A	97.50	Cis	84.17	F	45.00
Am	82.50	Cism	83.33	Fm	68.33
Ais	85.00	D	51.67	Fis	90.83
Aism	89.17	Dm	95.83	Fism	88.33
B	86.67	Dis	97.50	G	88.33
Bm	88.33	Dism	70.00	Gm	75.83
C	71.67	E	80.00	Gis	90.00
Cm	95.83	Em	91.67	Gism	86.67
Total					82.67

3.3 Percobaan dengan 39 Koefisien Cepstral

Hasil percobaan menggunakan 39 koefisien *cepstral* dapat dilihat pada Tabel 4.

Tabel 4.
Akurasi dengan 39 Koefisien Cepstral

Chord	Akurasi	Chord	Akurasi	Chord	Akurasi
A	99.17	Cis	90	F	80.83
Am	90.83	Cism	87.5	Fm	81.67
Ais	89.17	D	60.83	Fis	96.67
Aism	92.5	Dm	97.5	Fism	90.83
B	91.67	Dis	99.17	G	93.33
Bm	90	Dism	82.5	Gm	99.17
C	79.17	E	90.83	Gis	95
Cm	98.33	Em	95.83	Gism	95
Total					90.31

Akurasi terkecil terdapat pada percobaan *chord* D yaitu 60.83% sedangkan akurasi terbesar terdapat pada percobaan *chord* A yaitu 99.17%.

3.4 Percobaan dengan 52 Koefisien Cepstral

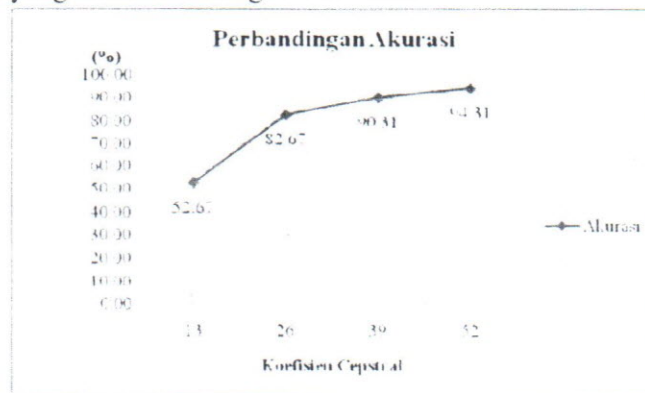
Hasil percobaan menggunakan 52 koefisien *cepstral* dapat dilihat pada Tabel 5. Akurasi terkecil terdapat pada percobaan *chord* D yaitu 65.83% sedangkan akurasi terbesar terdapat pada percobaan *chord* A dan *chord* Gm yaitu 100%.

Tabel 5.
Akurasi dengan 39 Koefisien Cepstral

Chord	Akurasi	Chord	Akurasi	Chord	Akurasi
A	100.00	Cis	91.67	F	95.00
Am	94.17	Cism	94.17	Fm	87.50
Ais	93.33	D	65.83	Fis	97.50
Aism	97.50	Dm	99.17	Fism	95.00
B	95.83	Dis	99.17	G	96.67
Bm	94.17	Dism	91.67	Gm	100.00
C	85.83	E	96.67	Gis	98.33
Cm	98.33	Em	98.33	Gism	97.50
Total					94.31

Pada Gambar 5 terlihat grafik perbandingan hasil akurasi menggunakan beberapa koefisien *cepstral*. Pada penelitian ini menunjukkan bahwa tingkat akurasi sangat bergantung pada jumlah koefisien *cepstral* yang digunakan. Berdasarkan penelitian ini semakin besar koefisien *cepstral* yang digunakan, semakin tinggi akurasi yang dihasilkan. Hasil akurasi tertinggi terdapat pada percobaan yang menggunakan 52 koefisien *cepstral*, sehingga didapatkan rata-rata akurasi sebesar 79.99%.

Berdasarkan rata-rata akurasi tiap *chord*, akurasi tertinggi terdapat pada *chord* Cm yaitu sebesar 94,4% sedangkan rata-rata akurasi terendah terdapat pada *chord* D. Setelah diamati, rendahnya rata-rata akurasi *chord* D diakibatkan banyaknya kesalahan deteksi *chord* D yang terdeteksi sebagai *chord* Dm.



Gambar 5 Perbandingan Hasil Akurasi dengan 13, 26, 39 dan 52 koefisien *cepstral*

4. Kesimpulan

Penelitian ini telah berhasil dalam mengkonversi suara gitar ke tangga nada dengan mengimplementasikan metode PNN dan teknik ekstraksi ciri MFCC. Hasil akurasi tertinggi yang dihasilkan sebesar 94.31% dengan penggunaan *time frame* 100 ms, *overlap* 40% dan 52 koefisien *cepstral*. Penelitian ini mampu mengidentifikasi campuran dua *chord* gitar menggunakan model PNN.

Konversi suara gitar ke tangga nada dalam penelitian ini dilakukan dengan menggunakan empat koefisien *cepstral* yang berbeda, yaitu 13, 26, 39, dan 52. Hasil akurasi dengan menggunakan empat koefisien *cepstral* tersebut secara berurutan ialah 52.67%, 82.67%, 90.31%, dan 94.31%. Hal ini menunjukkan bahwa semakin besar nilai koefisien *cepstral* semakin tinggi akurasi yang didapatkan.

Daftar Pustaka

- Buono A. (2009). Representasi Nilai HOS dan Model MFCC Sebagai Ekstraksi Ciri Pada Sistem Identifikasi Pembicara di Lingkungan Ber-Noise Menggunakan HMM.[Disertasi]. Depok: Program Pascasarjana, Universitas Indonesia.
- Fransiswa RR. (2010). Pengembangan Model *Probabilistic Neural Network* (PNN) pada pengenalan kisaran usia dan jenis kelamin berbasis suara.[Skripsi]. Jurusan Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor.
- Ganchev T. (2009). Speaker recognition [disertasi]. Patras (GR): University of Patras.
- M. Nilsson & M. Ejnarsson. (2002). *Speech Recognition using Hidden Markov Model: Performance Evaluation in Noisy Environment*. Master Thesis, Departement of Telecommunications and Signal Processing, Blekinge Institute of Technology.
- Wisnudisastra E, Buono A. (2010). Pengenalan *Chord* pada Alat Musik Gitar Menggunakan *CodeBook* dengan Teknik Ekstraksi Ciri MFCC. Jurnal Ilmiah Ilmu Komputer, Vol 14 No. 1, Mei 2010 : 16 – 21.