

SISTEM IDENTIFIKASI CIRI MUSIK UNTUK ROBOT PENARI JAIPONG

Bima Sena Bayu D., Ardik Wijayanto
Politeknik Elektronika Negeri Surabaya, Institut Teknologi Sepuluh Nopember
Kampus ITS Keputih Sukolilo Surabaya 60111
Telp : (+62)-31-5947280, Fax : (+62)-31-5946114
E-mail : bima@eepis-its.edu, ardik@eepis-its.edu

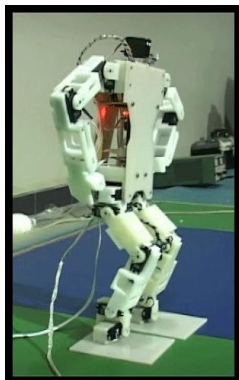
ABSTRAK

Robot penari jaipong adalah sebuah robot humanoid yang dapat bergerak secara serasi dengan irama lagu yang sedang dimainkan. Gerakan yang dimainkan merupakan kombinasi dari 21 derajat kebebasan actuator penyusunnya. Keserasian gerakan robot yang ditampilkan sangat bergantung pada kemampuan robot untuk mengenali irama yang dimainkan dengan menggunakan kombinasi dari beberapa metode, yaitu Filter Bank, Power Sinyal, Thresholding dan Feed Forward Artificial Neural Network (ANN). Dimana struktur dan parameter ANN sebelumnya telah diperoleh dari proses pembelajaran secara offline. Dari kombinasi beberapa metode yang telah digunakan, diperoleh hasil pengenalan pola gerakan offline mencapai 100% selama durasi waktu pemutaran lagu. Dengan kata lain, metode yang digunakan dalam penelitian ini sudah cukup memenuhi persyaratan dalam hal kecepatan proses yang lebih cepat (delay kecil) dan akurasi antara gerakan dengan irama masukan.

Kata kunci : Filter Bank, Power Sinyal, Thresholding, Feedforward ANN

1. PENDAHULUAN

Robot penari jaipong merupakan sebuah robot humanoid yang dibuat dengan tujuan untuk dapat menari mengikuti alunan irama musik tertentu. Robot ini dibuat dalam rangka lomba Kontes Robot Indonesia untuk divisi Kontes Robot Seni Indonesia. Divisi ini merupakan sebuah divisi baru dalam lomba KRI 2009, dimana robot yang dibuat harus mampu mengikuti alunan irama lagu dengan tema musik jaipong dari daerah Jawa Barat. Untuk dapat menari mengikuti alunan irama, maka diperlukan sebuah system identifikasi terhadap ciri-ciri apa saja yang terkandung dalam sebuah musik, yang meliputi tempo, irama, notasi alat instrumentasi dan keserasian.



Gambar 1. Robot penari jaipong

Pembuatan system identifikasi ciri music yang terintegrasi secara hardware merupakan kesulitan

tersendiri dalam penelitian ini, sehingga perlu dilakukan penyederhanaan algoritma agar proses identifikasi dapat berjalan secara real time.

2. TINJAUAN PUSTAKA

Pencarian ciri terhadap irama music telah banyak dilakukan oleh peneliti-peneliti dunia. Namun semuanya masih berbasis pada PC / simulasi. Tetsuro Kitahara dkk [1] melakukan identifikasi terhadap instrument music dengan mengacu pada 3 isu, yaitu : variasi fitur yang disebabkan oleh suara campuran, warna nada (timbre), dan penggunaan konteks musik. Untuk isu pertama, template dari fitur vektor yang mewakili timbres diambil dari suara terisolasi dan juga suara campuran. Menyebabkan beberapa fitur yang tidak kuat dalam campuran, fitur ditimbang menurut ketahanan mereka dengan menggunakan analisis diskriminan linier. Untuk masalah kedua, ia menggunakan F0 multivarian distribusi normal, yang mendekati fungsi ketergantungan dari frekuensi dasar. Untuk masalah ketiga, ketika instrumen dari masing-masing teridentifikasi, probablity apriori dihitung dari probabilitas aposteriori temporer dengan tetangganya. Danennberg [2] menyatakan bahwa pemahaman terhadap musik memerlukan pengenalan dan identifikasi dari struktur dan pola informasi dalam musik. Tiga hal yang dilakukan dalam penelitiannya yaitu : computer accompaniment, metode statistic untuk mengenali lokasi pengimprovisasi dan beat tracking.

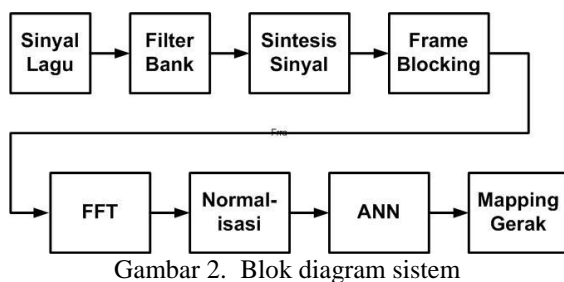
Masataka Goto dkk [3] membuat sebuah system real time yang dapat mengenali struktur ritme dari sinyal audio tanpa suara drum. Ia membagi pengamatan menjadi beberapa level not, untuk mempermudah pengamatan ritme/tempo. Hasil percobaan menunjukkan bahwa system yang dibuat cukup handal untuk memproses sinyal audio yang berasal dari CD musik populer.

Olivier Lartillot dkk [4] memperkenalkan sebuah toolbox di matlab untuk mengekstraksi ciri musik dari audio. Ia melakukan percobaan untuk mendapatkan rhythm / tempo dari musik.

Dari beberapa referensi yang telah dirujuk diatas, kebanyakan menggunakan ritme/tempo sebagai ciri musik yang cukup diandalkan. Sehingga teknik pencarian dan tracking beat sangat diperlukan. Namun, perbedaan di system yang dibuat dalam penelitian ini adalah terletak pada bagaimana cara memahami isi musik (lebih condong pada irama yang sangat dipengaruhi oleh kultur) karena irama ini digunakan sebagai pengiring gerak sehingga kesesuaian dan keserasian dengan konten dan konteks kultural mutlak diperlukan.

3. METODOLOGI

Dalam penelitian ini, metodologi yang digunakan menjadi sangat disederhanakan dimana tujuannya adalah agar dapat diterapkan pada perangkat keras sehingga dapat berjalan secara real time. Sistem dapat digambarkan pada blok diagram sebagai berikut :



Gambar 2. Blok diagram sistem

3.1 Sinyal / Stimulant

Sinyal / stimulant berupa sebuah lagu dengan irama khas Jawa Barat yaitu berjudul Bubuy Bulan yang dimainkan dalam waktu ± 3 menit (2:56). Gambar 2 menunjukkan potongan sinyal lagu Bubuy Bulan.

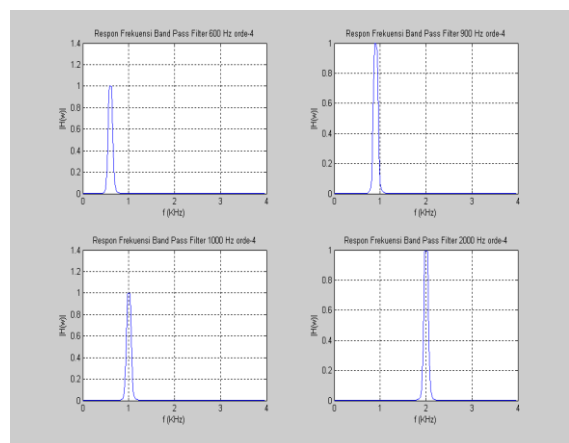


Gambar 3. Sinyal/stimulant lagu Bubuy Bulan

Lagu disampling dengan frekuensi sampling (fs) sebesar 8000 Hz, sehingga keseluruhan sinyal lagu akan memiliki 176 detik x 8000 titik = 1.408.000 titik sampling.

3.2 Filter Bank

Filter bank adalah sebuah array pass filter yang memisahkan sinyal masukan menjadi beberapa komponen, masing-masing membawa satu subband frekuensi sinyal yang asli. Filter bank merupakan cara untuk dapat kembali pada sinyal yang asli. Proses filter bank ini meliputi proses yang disebut *analisa* dan *sintesis*. Filter bank digunakan untuk mengisolasi komponen frekuensi dalam sinyal jika perbedaan kecil pada frekuensi dianggap signifikan. Filter Bank didesain untuk 4 buah filter pass band (BPF). Masing-masing Filter berorde 4 dimana lebar pita yang sempit diharapkan dapat mengisolasi frekuensi seruling. Frekuensi yang dibentuk dari seruling dapat dibagi menjadi 4 unsur frekuensi yaitu 600Hz, 900Hz, 1000Hz dan 2000Hz. Dengan bantuan program Matlab, maka koefisien IIR filter dapat diperoleh.



Gambar 4. Desain digital filter bank

3.3 Sintesis Sinyal

Sintesis sinyal dilakukan untuk mendapatkan kembali gabungan frekuensi seruling yang terisolasi hasil filter bank.

$$y = \sum_{n=0}^N x_1(n) + x_2(n) + x_3(n) + x_4(n) \quad (1)$$

3.4 Frame Blocking

Pada langkah ini, sinyal yang telah disintesis akan dibagi-bagi menjadi beberapa frame, dimana lebar setiap frame adalah sebesar 10ms atau terdapat 80 titik per frame. Sehingga jumlah frame total dalam satu detik adalah 100 frame dan jumlah frame total satu lagu adalah 176 detik x 100 frame = 17600 frame.

3.5 Fast Fourier Transform (FFT)

Sebuah DFT menguraikan urutan nilai ke dalam komponen frekuensi yang berbeda. Operasi ini berguna dalam berbagai bidang tetapi komputasi secara langsung terlalu lambat untuk segera mendapatkan hasilnya. FFT adalah sebuah cara untuk menghitung hasil yang sama dengan lebih

cepat; komputasi DFT untuk N buah titik membutuhkan N^2 operasi aritmatika, sedangkan FFT dapat menghitung hasil yang sama hanya dalam $N \cdot \log N$ operasi. Perbedaan kecepatan sangat besar, terutama untuk kumpulan data dimana N dalam ribuan atau jutaan. Perhitungan waktu dapat dikurangi beberapa kali lipat dalam kasus seperti itu, dan perbaikan secara kasar sebanding dengan $N \cdot \log(N)$. Perbaikan besar ini membuat banyak DFT sebagai algoritma berbasis praktis; sedangkan FFTs adalah sangat sesuai untuk berbagai jenis aplikasi pemrosesan sinyal digital yang menuntut hasil yang cepat dengan jumlah data yang besar.

Secara formulasi, persamaan DFT / FFT sebagai berikut :

$$X_l = \frac{1}{N} \sum_{k=0}^{N-1} x_k \cdot e^{-il\Delta\omega k} \quad (2)$$

3.6 Normalisasi

Normalisasi merupakan proses untuk menyeragamkan atau memberikan batasan nilai tertentu untuk amplitudo maksimum sinyal sehingga dapat digunakan langsung pada proses ANN.

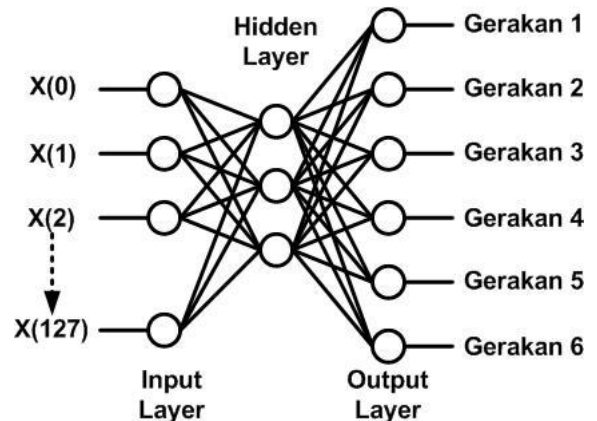
3.7 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) atau Jaringan saraf tiruan (JST) adalah jaringan dari sekelompok unit pemroses kecil yang dimodelkan berdasarkan jaringan saraf manusia. ANN merupakan sistem adaptif yang dapat merubah strukturnya untuk memecahkan masalah berdasarkan informasi eksternal maupun internal yang mengalir melalui jaringan tersebut.

Secara sederhana, ANN adalah sebuah alat pemodelan data statistik non-linier. ANN dapat digunakan untuk memodelkan hubungan yang kompleks antara input dan output untuk menemukan pola-pola pada data.

Terdapat dua tahapan dalam penggunaan ANN, yaitu tahap pembelajaran dan tahap pengujian. Pada tahap pembelajaran, sebuah system struktur ANN yang telah didesain sebelumnya akan diberi set masukan berupa data input dan target, dimana data input berupa keluaran data thresholding dari keempat filter dan targetnya adalah indeks pola gerakan tari. Hasil yang diharapkan dari proses pembelajaran adalah didapatkannya parameter bobot dan bias yang akan berlaku untuk semua set masukan.

Gambar 5 berikut ini adalah desain struktur ANN yang digunakan dalam penelitian :



Gambar 5. Struktur jaringan ANN

Keterangan :

Jumlah layer = 3 (input, hidden dan output)

Jumlah node input = 128 (Hasil FFT)

Jumlah node hidden = 3 - 10

Jumlah node output = 6 (6 kombinasi gerakan selama durasi lagu)

Learning rate = 0.25

Kriteria pemberhentian : toleransi error/iterasi

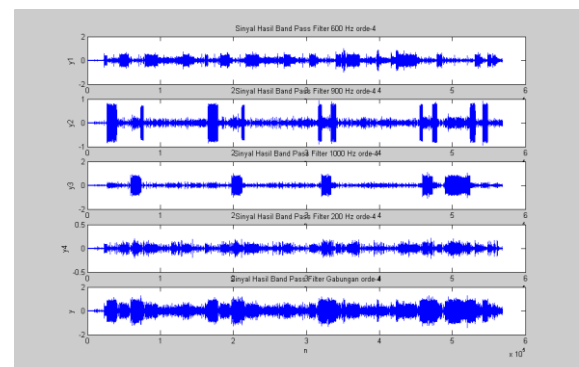
Jumlah pola = 5.867

4. HASIL DAN ANALISA

Berikut ini adalah hasil pengujian terhadap beberapa tahapan yang telah dilakukan :

4.1 Pengujian Filter Bank dan Sintesis Sinyal

Hasil filter bank pada frekuensi 600 Hz, 900 Hz, 1000Hz dan 2000 Hz.



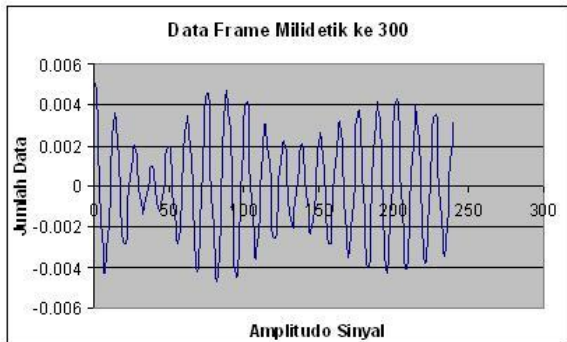
Gambar 6. Hasil digital filter bank

Dari hasil diatas tampak bahwa filter bank yang dibuat telah berhasil memisahkan bunyi alat musik seruling dari bunyi alat musik lain dalam campuran. Dan pada saat semua hasil filter digabungkan kembali (disintesis), maka dominasi suara seruling sangat jelas terdengar.

4.2 Frame Blocking

Frame blocking dilakukan setiap 30ms dengan jumlah data sample sebanyak 240 titik data.

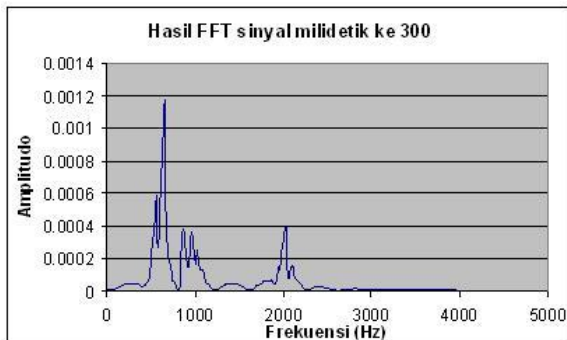
Gambar 7 menunjukkan hasil frame blocking pada milidetik ke 300



Gambar 7. Cuplikan dari data sampel sinyal milidetik ke-300

4.3 Fast Fourier Transform (FFT)

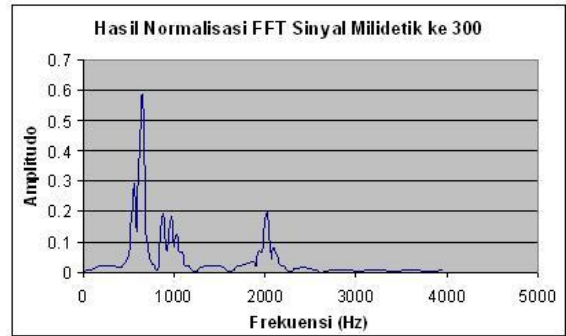
Pada proses ini sinyal yang sebelumnya berada dalam domain waktu akan dirubah dalam domain frekuensi. Setiap sinyal yang berasal dari alam merupakan sinyal analog yang bila diolah harus dirubah dalam bentuk sinyal digital. Dan pengolahan dalam digital merupakan pengolahan dalam bentuk diskrit. Pada proyek akhir ini sinyal dalam domain waktu akan dirubah dalam domain frekuensi dengan 256 titik. Karena hasil yang diperoleh berupa hasil dari fungsi konvolusi maka hanya akan diambil 128 titik saja yang akan diolah dalam proses selanjutnya. Sedangkan 128 sisanya tidak dipergunakan karena berupa pencerminan saja.



Gambar 8. Cuplikan dari data FFT sinyal milidetik ke-300

4.4 Normalisasi

Pada proses ini, data FFT sinyal akan diseragamkan ke level tertentu yang dapat diproses oleh ANN sehingga nilai sinyal harus dibawa kedalam range 0 – 1.

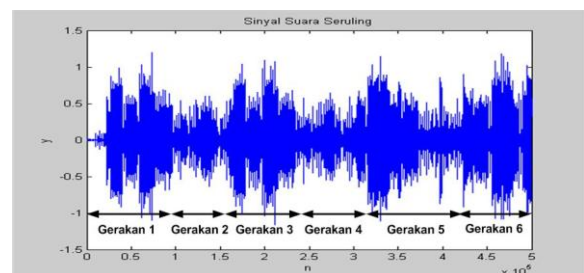


Gambar 9. Cuplikan dari data Normalisasi FFT sinyal milidetik ke-300

4.5 Artificial Neural Network (ANN)

Proses Artificial Neural Network yang digunakan disini adalah proses forward saja dimana yang diperlukan adalah kecepatan proses. Sedangkan proses pembelajarannya dilakukan secara offline pada PC untuk mendapatkan parameter bobot dan bias.

Gambar 10 menunjukkan proses penentuan target berupa pola gerakan tari yang diinginkan terhadap masukan sinyal lagu.



Gambar 10 proses penentuan target

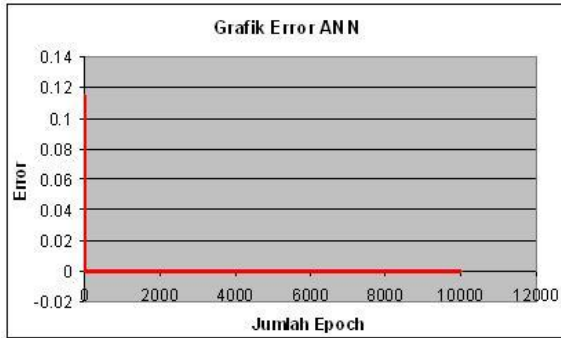
Target diberikan langsung pada node keluaran ANN yang mewakili 6 macam gerakan. Dimana jika gerakan pertama yang diinginkan untuk dapat diidentifikasi maka keluaran ANN harus 1 0 0 0 0 0, jika gerakan kedua yang diinginkan sebagai keluaran, maka pola keluaran harus 0 1 0 0 0 0, dst.

Tabel 1 menunjukkan akurasi algoritma artificial neural network dalam menggambarkan pola gerak yang diinginkan.

Tabel 1 Tingkat akurasi algoritma ANN

No	Data Milidetik ke-	Target	Keluaran Uji	Akurasi
1	6250	100000	100000	100%
2	15000	010000	010000	100%
3	25000	001000	001000	100%
4	37500	000100	000100	100%
5	43750	000010	000010	100%
6	56250	000001	000001	100%

Gambar 11 menunjukkan grafik error pembelajaran artificial neural network untuk keseluruhan sinyal lagu.



Gambar 11 Grafik error pembelajaran ANN

Berdasarkan grafik error ANN diatas, maka dapat dikatakan bahwa algoritma neural network backpropagasi mampu menyelesaikan permasalahan pola sinyal yang cukup kompleks.

5. KESIMPULAN

1. Sebuah lagu memiliki beberapa potensi fitur yaitu : ritme/tempo, warna nada (timbre) dan konten lagu.
2. Permasalahan mendasar dalam penelitian ini adalah terletak pada bagaimana menjadikan irama lagu sebagai factor pembentuk keserasian dengan gerakan yang sudah dibuat. Sehingga yang perlu dipertimbangkan adalah konten dari irama itu sendiri, karena irama yang dimainkan mengandung unsur kultur kebudayaan.
3. Penggunaan kombinasi beberapa metode pengolahan sinyal dengan ANN sudah cukup baik untuk menghasilkan keputusan gerakan tari.
4. Berdasarkan pengujian pengolahan sinyal secara real time pada perangkat keras dan pengujian secara offline data real time pada ANN di PC, diperoleh hasil yang cukup baik dengan prosentase keberhasilan / akurasi 100%.

6. REFERENSI

- [1] Tetsuro Kitahara, Masataka Goto, Kazunori Komatani, Tetsuya Ogata and Hiroshi G. Okuno, "Instrument Identification in Polyphonic Music : Feature Weighting with Mixed Sounds, Pitch Dependent Timbre Modelling and Use of Musical Context"
- [2] Dannenberg, "Recent Work In Real-Time Music Understanding By Computer" Music, Language, Speech, and Brain, Wenner-Gren International Symposium Series, Sundberg, Nord, and Carlson, ed., Macmillan, 1991, pp. 194-202.
- [3] Masataka Goto, Yoichi Muraoka, "Real Time Rhythm Tracking For Drumless Audio Signals – Chord Change Detection For Musical Decisions", IJCAI-97 Workshop on Computational Auditory Scene Analysis.

[4] Olivier Lartillot, Petri Toiviainen, "MIR in Matlab (II) : A Toolbox For Musical Feature Extraction From Audio", 2007 Austrian Computer Society (OCG)