

**DANCE MODELLING, LEARNING AND RECOGNITION SYSTEM
UNTUK TARI TRADISIONAL ACEH
BERBASIS HIDDEN MARKOV MODEL**

TESIS

**Karya tulis sebagai salah satu syarat
untuk memperoleh gelar Magister dari
Institut Teknologi Bandung**

Oleh

NURFITRI ANBARSANTI

NIM : 23212071

Program Magister Teknik Elektro



**SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA
INSTITUT TEKNOLOGI BANDUNG**

2014

**DANCE MODELLING, LEARNING AND RECOGNITION SYSTEM
UNTUK TARI TRADISIONAL ACEH
BERBASIS HIDDEN MARKOV MODEL**

**Oleh
Nurfitri Anbarsanti**

Bandung, 19 Agustus 2014

Menyetujui
Pembimbing I,

Dr. techn. Ary Setijadi Prihatmanto

ABSTRAK

DANCE MODELLING, LEARNING AND RECOGNITION SYSTEM UNTUK TARI TRADISIONAL ACEH BERBASIS HIDDEN MARKOV MODEL

Oleh

Nurfitri Anbarsanti

NIM : 23212071

PROGRAM MAGISTER TEKNIK ELEKTRO

Budaya adalah identitas suatu bangsa. Arus globalisasi yang menerpa generasi muda saat ini, menyebabkan generasi muda kehilangan jati diri, kehilangan figur teladan dan kehilangan nilai-nilai kearifan budaya Indonesia yang baik. Jika budaya Indonesia yang baik tidak dipertahankan dan disebarluaskan, maka akan hilang dalam beberapa dekade ke depan. Pengukuhan yang kuat terhadap budaya Indonesia sangatlah diperlukan. Pada penelitian ini dirancang pemodelan, pembelajaran dan klasifikasi tari Likok Pulo dari Aceh, sebagai satu langkah awal untuk ‘mengawetkan’ tari tradisional Indonesia dalam bentuk digital, sehingga dapat diintegrasikan dengan teknologi lain seperti robot penari humanoid atau video game.

Keseluruhan tari Likok Pulo dimodelkan menggunakan hidden markov model. Gesture tari berperan sebagai state diskrit yang tersembunyi, dan frase sebagai rangkaian dari gesture-gesture. Untuk kekokohan sistem dalam menangani input sensor Kinect yang berderau, didesain sebuah representasi skeleton. Sebuah pose tari didefinisikan menggunakan representasi skeleton ini yang telah dikuantisasi berdasarkan range of movement. Gesture tari tunggal didefinisikan sebagai rangkaian pose tari lalu dipelajari dan diklasifikasi menggunakan model HMM.

Sistem diimplementasikan menggunakan pemrograman Matlab dan Simulink. Enam kelas gesture tari dari frase "Assalamualaikum" dilatih dengan ratusan data gesture yang direkam menggunakan sensor Kinect XBOX yang dimainkan oleh tiga orang untuk masing-masing kelas gesture. Sistem mengklasifikasi data gesture menjadi salah satu dari enam kelas gesture terdefinisi atau satu kelas gesture tidak terdefinisi. Sistem klasifikasi memiliki tingkat akurasi 94.87% untuk gesture tunggal.

Kata kunci : representasi sudut skeleton, pemodelan tari, dance recognition, gesture recognition, hidden Markov model, Tari Likok Pulo, Kinect sensor.

ABSTRACT

DANCE MODELLING, LEARNING AND RECOGNITION SYSTEM OF ACEH TRADITIONAL DANCE BASED ON HIDDEN MARKOV MODEL

By

Nurfitri Anbarsanti

NIM : 23212071

ELECTRICAL ENGINEERING MASTER PROGRAM

Culture is the identity of a nation. Globalization that hit today's youth, lead the youth generation lose their identity, lose the good role models, and lose the value of good wisdom Indonesian culture. If the Indonesian culture is not well maintained and disseminated, it will disappear in the next few decades. Strong affirmation of Indonesian culture is required. Modeling, learning and classification system of Likok Pulo Dance from Aceh are designed, as one of the initial steps to 'preserve' Indonesian traditional dance into digital form, so it can be integrated with other technologies such as humanoid robot dancers or video games.

The whole dance of Likok Pulo are modeled by hidden markov model. Dance gesture are cast as hidden discrete states and phrase as a sequence of gestures. For robustness under noisy input of Kinect sensor, an angular representation of the skeleton is designed. A pose of dance is defined by this skeleton representation which has been quantified based on range of movement. One unique gesture of dance is defined by sequence of pose and then learned and classified by HMM model.

The system was implemented using the Matlab and Simulink programming package. Six of dance's gesture classes from the phrase "Assalamualaikum" has been trained with hundreds of gesture instances recorded by the XBOX Kinect sensor which performed by three of subjects for each gesture class. The classifier system classify the input testing gesture into one of six classes of predefined gesture or one class of undefined gesture. The classifier system has an accuracy of 94.87% for single gesture.

Keywords : angular skeletal representation, dance modelling, dance recognition, gesture recognition, hidden Markov model, Likok Pulo dance, sensor Kinect.

PEDOMAN PENGGUNAAN TESIS

Tesis S2 yang tidak dipublikasikan, terdaftar dan tersedia di Perpustakaan Institut Teknologi Bandung, dan terbuka untuk umum dengan ketentuan bahwa hak cipta ada pada pengarang. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau peringkasan hanya dapat dilakukan seizin pengarang dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Memperbanyak atau menerbitkan sebagian atau seluruh tesis haruslah seizin Direktur Program Pascasarjana, Institut Teknologi Bandung.

Perpustakaan yang meminjam tesis ini untuk keperluan anggotanya harus mengisi nama dan tanda tangan peminjam dan tanggal pinjam.

KATA PENGANTAR

Alhamdulillah puji dan syukur penulis panjatkan kepada Allah Subhanahu wa Ta'ala yang telah memberi kesempatan, kesehatan lahir batin, rezeki, inspirasi, dan pertolongan selama proses penggerjaan penelitian ini. Shalawat dan salam semoga tersambung selalu kepada Rasulullah saw, sahabat-sahabat beliau, dan umat Islam yang mengikuti mereka dengan baik. Penulis mengucapkan terima kasih kepada :

1. Bapak Dr. techn. Ary Setijadi P. selaku dosen pembimbing, yang telah banyak memberikan waktu, masukan, arahan, dan saran untuk mengembangkan ide penelitian;
2. Tim Laboratorium Sistem Kendali dan Komputer, juga Institut Teknologi Bandung yang telah memberikan beasiswa kepada penulis;
3. Bapak, Mama, dan adik-adik yang penulis cintai, beserta shahabat-shahabat yang senantiasa memberikan dukungan dan doa restu.
4. teman-teman magister teknik komputer,
5. seluruh staf dan karyawan LSKK dan Tata Usaha, yang telah sigap memberikan banyak bantuan;
6. serta seluruh pihak yang sudah memberikan dukungan, yang tidak dapat penulis sebutkan satu persatu.

Semua karya dan pengetahuan yang sempurna datangnya dari Allah Subhanahu wa Ta'ala, sedangkan kesalahan dan kekurangan pada penelitian ini berasal dari penulis sebagai manusia biasa. Kritik dan saran yang membangun sangat diharapkan untuk pengembangan penelitian selanjutnya. Semoga penelitian ini bisa memberi banyak manfaat. Aamiin.

Bandung, Agustus 2014

Penulis

DAFTAR ISI

ABSTRAK	ii
ABSTRACT	iii
PEDOMAN PENGGUNAAN TESIS	iv
KATA PENGANTAR	v
DAFTAR ISI.....	vi
DAFTAR GAMBAR	viii
DAFTAR TABEL.....	x
BAB I PENDAHULUAN	1
I.1. Latar Belakang	1
I.2. Tujuan Penelitian	3
I.3. Batasan Masalah	3
I.4. Metodologi Penelitian.....	4
I.5. Sistematika Pembahasan.....	5
BAB II STUDI PUSTAKA DAN KARYA BERKAITAN	6
II.1. Pemodelan Sinyal.....	6
II.2. Peluang Diskrit.....	7
II.3. Model Probabilistik Grafis.....	9
II.4. Proses Markov Diskrit	10
II.5. Hidden Markov Model.....	11
II.6. Elemen-Elemen HMM.....	13
II.7. Topologi HMM.....	14
II.8. Algoritma Forward-Backward untuk Evaluasi	15
II.9. Algoritma Viterbi untuk Decoding	18
II.10. Algoritma Baum Welch untuk Learning.....	19
II.11. Maximum Likelihood	21
II.12. Beberapa Penelitian Seputar Pemodelan Tari.....	22
BAB III PEMODELAN TARI DAN GESTURE TARI.....	24
III.1. Pemodelan Tari Likok Pulo	24
III.2. Pemodelan Gesture Tari dalam Hidden Markov Model	30

BAB IV	PEMBELAJARAN DAN KLASIFIKASI GESTURE TARI	32
IV.1.	Model Klasifikasi dengan Threshold	32
IV.2.	Pemilihan Kelas Gesture Tari yang Akan Diuji	33
IV.3.	Perangkat Penginderaan.....	35
IV.4.	Representasi Skeleton	37
	Kerangka PCA Torso	38
	Joint Berderajat Satu	39
	Joint Berderajat Dua	40
	Joint Berderajat Tiga.....	41
	Representasi Pose Tubuh	42
	Range Sudut pada Vektor Pose Berdasarkan Range of Movement.....	43
	Pengujian Kerangka PCA Torso	43
IV.5.	Koleksi Data dan Segmentasi Data.....	45
IV.6.	Kuantisasi Vektor Pose Tari Berdasarkan Range of Movement	46
IV.7.	Pembelajaran Gesture Tari.....	47
IV.8.	Klasifikasi Gesture Tari	48
	Pengujian Pertama	48
	Pengujian Kedua	49
	Pengujian Ketiga.....	50
IV.9.	Rekognisi Frase Tari Berbasis Hidden Markov Model	51
IV.10	Implementasi Realtime	54
BAB V	KESIMPULAN, SARAN DAN TINDAK LANJUT	55
V.1.	Kesimpulan	55
V.2.	Saran dan Tindak Lanjut.....	55
DAFTAR PUSTAKA	xiv
LAMPIRAN A	xvi
LAMPIRAN B	xvii
LAMPIRAN C	xviii
LAMPIRAN D	xxii

DAFTAR GAMBAR

Gambar I.1	Tari Likok Pulo	1
Gambar II.1	Contoh Directed Acyclic Graph.....	9
Gambar II.2	Sebuah Markov chain dengan 5 state dengan transisi state	10
Gambar II.3	Model grafis Hidden Markov Model dengan topologi left-right	12
Gambar II.4	Model 1 koin	12
Gambar II.5	Model 2 koin	12
Gambar II.6	Model ergodic dengan 4 buah state.....	14
Gambar II.7	Model left-right dengan 5 buah state.....	14
Gambar II.8	Ilustrasi rangkaian operasi pada perhitungan variabel maju	16
Gambar II.9	Ilustrasi rangkaian operasi pada perhitungan variabel mundur.....	17
Gambar II.10	Struktur lattice untuk komputasi α ti dan β ti	17
Gambar II.11	Ilustrasi pemilihan θ dari $p_{x1:n\theta}$	21
Gambar II.12	Notasi Laban	22
Gambar III.1	Hierarki Tarian, Frase, Gesture, Pose, dan Observation	24
Gambar III.2.	Beberapa implementasi model L pada sistem praktis.	26
Gambar III.3	Mesin Sekuensial untuk Frase Assalamualaikum.....	26
Gambar III.4	Mesin Sekuensial untuk Frase Hai Aneuk Sen.	27
Gambar III.5	Mesin Sekuensial untuk Frase Lagu Dendang.	27
Gambar III.6	Mesin Sekuensial untuk Frase He Mala Mala.....	28
Gambar III.7	Mesin Sekuensial untuk Frase Hai Laot Sa.....	28
Gambar III.8	Mesin Sekuensial untuk Frase Kisah Hasan Husein.	29
Gambar III.9	Ilustrasi hidden Markov model dari sebuah gesture tari.	30
Gambar IV.1	Diagram Blok Sistem	32
Gambar IV.2	Mesin Sekuensial untuk Frase “Assalamualaikum”.....	34
Gambar IV.3	Sensor Kinect XBOX360	35
Gambar IV.4	Desain blok Simulink untuk akuisisi data sensor Kinect	36
Gambar IV.5	Akuisisi data sensor Kinect menggunakan matlab dan simulink.....	36
Gambar IV.6	Penamaan joint-joint skeleton, beserta hierarkinya.....	37
Gambar IV.7	Kerangka PCA torso, kerangka referensi dan kerangka Kinect.....	38
Gambar IV.8	Sistem koordinat bola pada joint berderajat satu.	39

Gambar IV.9	Sistem koordinat bola pada joint berderajat dua.....	40
Gambar IV.10	Sistem koordinat bola pada joint berderajat tiga.....	41
Gambar IV.11	Hasil aplikasi PCA pada matriks torso.....	44
Gambar IV.12	Rotasi kerangka PCA torso untuk perhitungan joint elbow.....	44
Gambar IV.13	Smartglove dengan sensor tepuk di dalamnya	45
Gambar IV.14	Kurva pelatihan gesture tari dengan HMM.....	47
Gambar IV.15	Penurunan tingkat akurasi dan peningkatan jumlah state.....	48
Gambar IV.16	Diagram batang hasil pengujian setiap satu model.....	50
Gambar IV.17	Diagram batang hasil pengujian pola gesture.	51
Gambar IV.18	Hasil pengujian 20 buah frase Assalamualaikum.	53
Gambar IV.19	Implementasi Realtime : gesture O tidak terdeteksi.....	54
Gambar IV.20	Implementasi Realtime : gesture O terdeteksi.....	54

DAFTAR TABEL

Tabel II.1	Tiga macam kasus directed acyclic graph	10
Tabel IV.1	Daftar enam kelas gesture yang akan diuji	33
Tabel IV.2	Spesifikasi perangkat keras dan perangkat lunak	35
Tabel IV.3	Range of movement untuk vektor pose tubuh bagian atas	43
Tabel IV.4	Hasil koleksi data untuk masing-masing gesture.....	45
Tabel IV.5	Directional codewords berdasarkan range of movement.....	46
Tabel IV.6	Pengujian tanpa menggunakan representasi skeleton	49
Tabel IV.7.	Hasil pengujian setiap satu model oleh seluruh pola gesture.	49
Tabel IV.8	Hasil pengujian setiap pola gesture oleh seluruh model HMM.....	51
Tabel IV.9	Hasil pengujian setiap pola gesture oleh seluruh model HMM (2)	51
Tabel IV.10	Efek penggunaan representasi skeleton dan maximum likelihood.....	52
Tabel IV.11	Hasil pengujian 20 buah frase ‘Assalamualaikum’	53

BAB I PENDAHULUAN

I.1. Latar Belakang

Budaya adalah identitas suatu bangsa. Jatuh bangkitnya suatu bangsa dapat diidentifikasi melalui bagaimana bangsa tersebut menjaga kearifan budayanya. Generasi muda adalah penerus kebangkitan bangsa Indonesia di masa depan. Arus globalisasi yang menerpa generasi muda saat ini menyebabkan generasi muda kehilangan jati diri, kehilangan figur teladan dan kehilangan nilai-nilai kearifan budaya Indonesia yang baik. Jika budaya Indonesia yang baik tidak dipertahankan dan disebarluaskan, maka akan hilang dalam beberapa dekade ke depan. Sehingga pengukuhan yang kuat terhadap budaya Indonesia sangatlah diperlukan. Pada penelitian ini dirancang pemodelan, pembelajaran dan klasifikasi tari Likok Pulo dari Aceh, sebagai salah satu langkah awal untuk “mengawetkan” tari tradisional Indonesia dalam bentuk digital, sehingga dapat diintegrasikan dengan teknologi yang lain seperti robot penari humanoid, atau video game tari Indonesia.



Gambar I.1 Tari Likok Pulo

Tarian adalah rangkaian gerakan manusia yang ekspresif dan memiliki nilai keindahan. Tarian asal Indonesia yang terkenal unik dan menarik adalah tari tradisional Aceh seperti Saman Gayo, Likok Pulo, dan Ratoe Duek. Pada penelitian ini dipilih tari Likok Pulo untuk berkomunikasi dengan komputer, karena tari Likok Pulo dari Aceh memiliki karakteristik-karakteristik berikut :

1. koreografi gerakan yang hampir tidak berubah dari generasi ke generasi,
2. menuntut kesinkronan gerak di antara kelompok penari dengan formasi berbaris,

3. menuntut ketepatan timing gerak tubuh dengan ritme musik yang berubah semakin cepat secara linear,
4. memiliki beberapa gesture tari yang dilakukan berulang-ulang, dan
5. diterima dengan baik di lingkungan internasional, namun tetap memiliki identitas nilai budaya yang sopan dan kuat.

Pada tahun 2010, Microsoft meluncurkan XBOX Kinect System. Kinect merupakan Natural User Interface yang mengkombinasikan kamera stereoskopis dan sensor infrared, sehingga dapat menangkap peta kedalaman dengan kecepatan sekitar 30 frame per detik untuk mengestimasi posisi dari 20 titik joint pada skeleton user. Pemrosesan posisi joint pada skeleton dengan menggunakan algoritma tertentu, dapat digunakan untuk mengenal gerakan tubuh manusia. Sehingga gerakan tubuh manusia, misalnya tarian, dapat diterjemahkan sebagai komando untuk berkomunikasi dengan komputer.

Sinyal trajektori joint-joint skeleton yang ditangkap oleh sensor Kinect sangat mungkin mengalami diskontinuitas, derau, atau parameter yang tidak stabil [7]. Gerakan tarian yang melibatkan banyak artikulasi bagian tubuh akan menghasilkan dimensi input yang sangat besar bagi sistem pemrosesan sinyal trajektori tersebut. Sehingga dibutuhkan representasi skeleton untuk mengurangi entropi sinyal dan mereduksi dimensi data. Representasi skeleton juga harus menangani perubahan posisi dan orientasi penari, relatif terhadap sensor Kinect.

Tarian dapat dideskripsikan sebagai rangkaian seri dari beberapa gesture hingga berbeda yang tersembunyi. Transisi antar gesture tari tidak diatur menggunakan relasi deterministik, namun diatur menggunakan relasi probabilistik untuk memodelkan ketidakideal tari dalam dunia nyata. Gesture tari memiliki dua karakteristik sinyal, yaitu variabilitas ruang-waktu (spatio-temporal variability) dan ambiguitas segmentasi [16]. Variabilitas ruang-waktu adalah fakta bahwa pada gesture yang sama terdapat variasi ukuran, bentuk, dan durasi yang banyak; yang dilakukan oleh orang yang berbeda atau bahkan oleh orang yang sama. Ambiguitas segmentasi adalah bagaimana

menentukan kapan gesture itu dimulai dan kapan gesture itu berakhir dalam trajektori sinyal yang kontinyu.

Beberapa pendekatan yang dapat digunakan untuk menganalisa gesture sebagai pola spasial-temporal adalah Dynamic Time Warping, Artificial Neural Network atau Hidden Markov Model [16]. Pada penelitian ini dipilih HMM untuk memodelkan gesture tari Likok Pulo, karena HMM adalah pemodelan statistik yang dapat diaplikasikan untuk menganalisis deret waktu dengan variabilitas ruang-waktu dan dapat mengatasi pola yang tidak terdefinisi [16]. Pemodelan gesture tari berbasis HMM dapat direalisasikan menjadi sistem praktis yang dapat mempelajari, memprediksi, dan mengklasifikasi gesture-gesture tari Likok Pulo dari Aceh.

I.2. Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut :

1. Membangun model dan hierarki keseluruhan tari Likok Pulo dari Aceh.
2. Membangun representasi skeleton yang mampu menangani diskontinuitas, ketidakstabilan, dan redundansi dimensi pada trajectory data Kinect.
3. Membangun sistem pembelajaran dan klasifikasi gesture tari Likok Pulo dari Aceh.

I.3. Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut :

1. Observasi tarian hanya dilakukan pada gerakan tubuh bagian atas.
2. Representasi skeleton hanya mempertimbangkan aspek-aspek kinematik pada kerangka tubuh manusia, belum mempertimbangkan aspek-aspek dinamis pada kerangka tubuh manusia.
3. Gesture-gesture tari yang dipelajari dan diklasifikasi adalah enam buah gesture yang terangkai pada satu buah frase “Assalamualaikum” tari Likok Pulo.
4. Proses pengujian gesture tari menggunakan sensor Kinect hanya bisa dilakukan dalam ruangan yang terang.

I.4. Metodologi Penelitian

Tahapan-tahapan yang dilakukan dalam penelitian ini sebagai berikut.

1) Penentuan Topik, Tujuan dan Batasan

Pada tahap ini, brainstorming dan konsultasi bersama dosen pembimbing dilakukan untuk memberi gambaran mengenai topik penelitian. Tujuan penelitian, kebutuhan penelitian, dan batasan penelitian ditentukan.

2) Studi Literatur

Pada tahap ini, dilakukan penelusuran pustaka dan penelitian terdahulu yang terkait dengan gesture recognition, gesture classification, dance modelling, kinematic model for human motion, range of movement, skeleton representation, dan hidden markov model.

3) Merancang Model Tari dan Model Gesture Tari

Pada tahap ini, dirancang model tari Likok Pulo keseluruhan; dan model gesture tari berdasarkan hidden markov model. Ditentukan enam buah gesture yang terangkai pada frase “Assalamualaikum” dari tari Likok Pulo untuk implementasi sistem.

4) Merancang Representasi Skeleton

Pada tahap ini, dirancang representasi skeleton yang dapat memenuhi kebutuhan sistem.

5) Implementasi Sistem Pembelajaran dan Klasifikasi Gesture, Pengujian dan Analisis

Pada tahap ini, dilakukan koleksi data dari Kinect, kemudian dipre-process menggunakan representasi skeleton, lalu data tersebut dijadikan baseline untuk ‘melatih’ model HMM. Model HMM yang telah dilatih digunakan untuk klasifikasi gesture-gesture tari Likok Pulo dalam frase “Assalamualaikum”. Pengujian setiap model diamati hasilnya kemudian dianalisis.

I.5. Sistematika Pembahasan

Sistematika pembahasan penelitian ini adalah sebagai berikut

1) BAB I PENDAHULUAN.

Bab ini berisi gambaran umum mengenai penelitian yang akan dilakukan seperti latar belakang, tujuan penelitian, batasan masalah, metodologi penelitian serta sistematika pembahasan tesis.

2) BAB II STUDI PUSTAKA DAN KARYA-KARYA BERKAITAN

Bab ini berisi teori-teori dasar untuk memahami sistem yang akan dibangun. Bab ini juga mengkaji berbagai penelitian yang berkaitan dengan topik penelitian. Wawasan mengenai probabilitas grafis, Hidden Markov Model dan permasalahan terkait, beragam pemodelan; dibahas dalam bab ini.

3) BAB III PEMODELAN TARI DAN GESTURE TARI LIKOK PULO

Bab ini berisi model tari Likok Pulo secara keseluruhan, dan model gesture tari Likok pulo berbasis Hidden Markov Model.

4) BAB IV REPRESENTASI SKELETON

Bab ini berisi rancangan representasi skeleton manusia yang akan digunakan untuk mengolah sinyal trajektori joint-joint skeleton pada sistem : torso PCA frame, joint berderajat satu, joint berderajat dua, joint berderajat tiga, beserta representasinya.

5) BAB V PEMBELAJARAN DAN KLASIFIKASI GESTURE TARI

Bab ini berisi pemilihan gesture apa saja yang akan diuji, kuantisasi vektor pose, dan implementasi HMM untuk pembelajaran dan klasifikasi gesture tari beserta hasil pengujian dan analisis.

6) BAB VI KESIMPULAN, SARAN DAN TINDAK LANJUT.

Bab ini berisi kesimpulan apa saja yang dapat ditarik dari penelitian, beserta saran dan tindak lanjut untuk meningkatkan kualitas dan manfaat penelitian selanjutnya.

BAB II STUDI PUSTAKA DAN KARYA BERKAITAN

II.1. Pemodelan Sinyal

Pembahasan pada subbab ini diadaptasi dari [1].

Proses di dunia nyata menghasilkan keluaran yang dapat diamati dan dikarakterisasi sebagai sinyal. Sinyal dapat dikategorikan menjadi diskrit atau kontinyu. Sumber sinyal dapat dikategorikan menjadi stasioner (properti statistiknya tidak berubah terhadap waktu) atau nonstasioner (properti berubah terhadap waktu). Sinyal tersebut dapat murni (datang secara langsung tepat dari sumber sinyal) atau bisa terusakkan oleh sumber sinyal lain (mis. Derau) atau oleh distorsi transmisi, gaung/gema, dan lain lain.

Pemodelan sinyal merupakan hal yang penting, karena (1) model sinyal memberikan deskripsi teoritis untuk sistem pemrosesan sinyal untuk menghasilkan output yang diinginkan; (2) model sinyal memberikan gambaran mengenai sumber sinyal tanpa harus melihat sumber secara langsung; (3) model sinyal dapat direalisasikan menjadi sistem praktis seperti sistem prediksi, rekognisi, klasifikasi identifikasi, dan lain lain.

Tipe pemodelan sinyal dapat dibagi menjadi dua yaitu (1) model deterministik, memanfaatkan properti tertentu dari sinyal. Misalnya apakah sinyal tersebut gelombang sinus, atau jumlah eksponensial. Spesifikasi model sinyal dilakukan secara langsung untuk menentukan nilai parameter dari sinyal tersebut, misalnya amplitudo, frekuensi, fase gelombang sinus, dll; dan (2) model statistik, mengkarakterisasi properti statistik dari sinyal tersebut. Contohnya proses Gaussian, Poisson, Markov, Markov tersembunyi, dan lain lain. Asumsi yang mendasari model statistik adalah sinyal tersebut dapat dikarakterisasi sebagai proses acak parametrik, dan parameter dari proses stokastik dapat ditentukan secara presisi.

Penelitian ini fokus pada salah satu topologi dari model sinyal stokastik yaitu hidden Markov model. Dalam beberapa literatur, biasa disebut sebagai Markov sources atau fungsi probabilistik dari rantai Markov.

II.2. Peluang Diskrit

Pembahasan pada subbab ini diadaptasi dari [2].

Peluang Terbatas (Finite Probability)

Definisi Laplace : jika S adalah ruang sampel hingga, yang terdiri dari keluaran-keluaran yang memiliki peluang sama untuk muncul, dan E adalah sebuah kejadian yang merupakan subset dari S ; maka peluang dari E adalah $p(E) = \frac{|E|}{|S|}$.

Distribusi Peluang (Probability Distribution)

Misalkan ada n keluaran yang mungkin yaitu x_1, x_2, \dots, x_n , maka dua kondisi yang harus dipenuhi adalah $0 \leq p(x_i) \leq 1$ untuk $i = 1, 2, \dots, n$ dan $\sum_{i=1}^n p(x_i) = 1$.

Fungsi p dari himpunan semua hasil dari ruang sampel S disebut “distribusi peluang” atau *probability distribution*. Peluang munculnya kejadian E adalah jumlah dari peluang semua keluaran di dalam E , yaitu $p(E) = \sum_{s \in E} p(s)$

Kombinasi dari Beberapa Kejadian (Combinations of Events)

E adalah sebuah kejadian dalam ruang sampel S . Peluang dari kejadian \bar{E} , yaitu kejadian komplementer dari E , adalah $p(\bar{E}) = 1 - p(E)$.

Misalkan E_1 dan E_2 adalah kejadian-kejadian dalam ruang sampel S . Maka $p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$.

Jika E_1, E_2, \dots adalah rangkaian dari pasangan kejadian yang terpisah (disjoint) dalam ruang sampel S , maka $p(\bigcup_i E_i) = \sum_i p(E_i)$.

Teorema ini dapat diaplikasikan ketika deret E_1, E_2, \dots terdiri dari jumlah tertentu atau jumlah tidak terbatas namun dapat ditentukan dari pasangan kejadian yang terpisah.

Peluang Bersyarat (Conditional Probability)

Misalkan E dan F adalah kejadian-kejadian dengan $p(F) > 0$. Probabilitas bersyarat munculnya kejadian E , dengan diketahui bahwa F terjadi, dinotasikan dengan $p(E|F)$, dan didefinisikan sebagai $p(E|F) = \frac{p(E \cap F)}{p(F)}$.

Kejadian yang Saling Bebas (Independent events)

Kejadian E dan F disebut saling bebas jika dan hanya jika $p(E \cap F) = p(E)p(F)$.

Teorema Bayes

Misalkan E dan F adalah kejadian dalam ruang sampel S sehingga $p(E) \neq 0$ dan $p(F) \neq 0$. Maka $E = E \cap S = E \cap (F \cup \bar{F}) = (E \cap F) \cup (E \cap \bar{F})$. Karena $E \cap F$ dan $E \cap \bar{F}$ saling lepas, maka

$$p(E) = p(E \cap F) + p(E \cap \bar{F}) \quad \dots (\text{II. 1})$$

Secara umum, karena $E = \bigcup_{i=1}^n (E \cap F_i)$ maka

$$p(E) = \sum_{\forall F} p(E \cap F) \quad \dots (\text{II. 2})$$

Definisi peluang bersyarat untuk (II.1) memberikan

$$p(E) = p(E|F)p(F) + p(E|\bar{F})p(\bar{F}) \quad \dots (\text{II.3})$$

Definisi peluang bersyarat untuk (II.2) memberikan

$$p(E) = \sum_{\forall F} p(E|F)p(F) \quad \dots (\text{II.4})$$

Definisi peluang bersyarat untuk persamaan $p(F|E) = p(E \cap F)/p(E)$ dan $p(E|F) = p(E \cap F)/p(F)$ memberikan $p(F|E)p(E) = p(E|F)p(F)$ sehingga

$$p(F|E) = p(E|F)p(F)/p(E) \quad \dots (\text{II.5})$$

Substitusi $p(E)$ dari (II.3) ke (II.5), diperoleh Teorema Bayes :

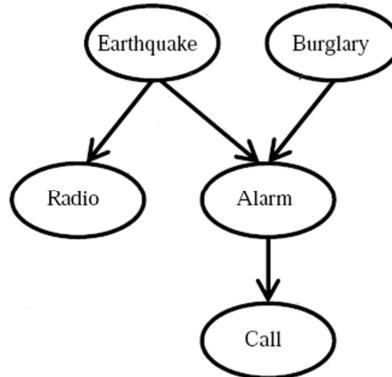
$$p(F|E) = \frac{p(E|F)p(F)}{p(E|F)p(F) + p(E|\bar{F})p(\bar{F})} \quad \dots (\text{II.6})$$

Substitusi $p(E)$ dari (II.4) ke (II.5), diperoleh Generalisasi Teorema Bayes :

$$p(F|E) = \frac{p(E|F)p(F)}{\sum_{\forall F} p(E|F)p(F)} \quad \dots (\text{II.7})$$

II.3. Model Probabilistik Grafis

Pembahasan pada subbab ini diadaptasi dari [3]. Untuk menghitung peluang gabungan (joint probability) dari sejumlah d variabel biner, dibutuhkan 2^d dimensi. Untuk menangani banyaknya dimensi yang muncul, dibutuhkan model probabilistik untuk mempermudah perhitungan peluang gabungan dari sejumlah variabel. Berikut ini akan dipelajari mengenai Directed Acyclic Graph (DAG). DAG terdiri dari node-node yang merepresentasikan variabel acak dan edge-edge yang merepresentasikan pengaruh langsung.



Gambar II.1 Contoh Directed Acyclic Graph

Untuk Gambar II.1, sebuah node X_I saling bebas dengan $X_{\text{ancestors}}|X_{\text{parents}}$ sehingga

$$P(B, E, A, R, C) = P(B)P(E|B)P(A|B, E)P(R|A, B, E)P(C|R, A, B, E)$$

Karena E dan B adalah kejadian yang saling bebas jika diberikan B , maka $P(E|B) = P(E)$. Karena R hanya dipengaruhi E , maka $P(R|A, B, E) = P(R|E)$. Karena C hanya dipengaruhi A , maka $P(C|R, A, B, E) = P(C|A)$. Sehingga persamaan peluang di atas dapat disederhanakan menjadi

$$(B, E, A, R, C) = P(B)P(E)P(A|B, E)P(R|E)P(C|A)$$

Pada DAG, jika dimiliki sejumlah n variabel x_i , peluang gabungan dari variabel ini dapat difaktorisasi sebagai berikut, aturan ini biasa disebut Chain Rule :

$$P(x_{1:n}) = P(x_1, x_2, \dots, x_n) = \prod_{i=1}^n P(x_i | Parents(x_i)) \quad \dots \quad (\text{II.8})$$

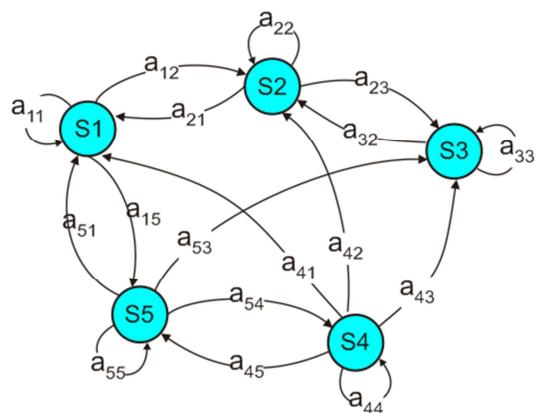
Pada DAG, ada tiga macam kasus untuk mengkaji nilai peluang graph terkait dengan kebebasan (dinotasikan dengan $\perp\!\!\!\perp$) sebuah node dalam graph tersebut :

Kasus Pertama	Kasus Kedua	Kasus Ketiga
$P(C, R, W) = P(C)P(R C)P(W R)$ <p>W dan C saling bebas jika diberikan R:</p> $(W \perp\!\!\!\perp C) R$	$P(R, W, S) = P(R)P(W R)P(S R)$ <p>W dan S saling bebas jika diberikan R:</p> $(W \perp\!\!\!\perp S) R$	$P(S, R, W) = P(R)P(S)P(W R, S)$ <p>R dan S tidak saling bebas jika diberikan W.</p> $(R \perp\!\!\!\perp S) W$

Tabel II.1 Tiga macam kasus directed acyclic graph

II.4. Proses Markov Diskrit

Pembahasan pada subbab ini diadaptasi dari [1]. Misalkan terdapat sebuah sistem yang didefinisikan pada setiap waktu t , yang terdiri dari sejumlah N state berbeda, $S_1, S_2, S_3, \dots, S_N$, yang diilustrasikan pada gambar II.2.



Gambar II.2 Sebuah Markov chain dengan 5 state dengan transisi state

Pada satuan waktu diskrit $t = 1, 2, \dots$ yang berjarak teratur, sistem mengalami perpindahan state (mungkin pindah ke state yang sama), mengikuti himpunan peluang yang berasosiasi dengan state tersebut. State aktual pada saat t dinotasikan sebagai q_t . Deskripsi probabilistik pada sistem pada Gambar II.2 membutuhkan spesifikasi pada state aktual sama seperti semua state-state pendahulu. Untuk rantai Markov diskrit berorde satu, deskripsi probabilistiknya terpotong hanya pada state aktual dengan state sebelumnya (Chain Rule subbab II.3) :

$$P[q_t = S_j | q_{t-1} = S_i, q_{t-2} = S_k, \dots] = P[q_t = S_j | q_{t-1} = S_i] \quad \dots \text{ (II. 9)}$$

Untuk proses selanjutnya hanya digunakan sisi kanan dari (II.9), karena state q_t saling bebas terhadap state $q_{t-2}, q_{t-3}, \dots, q_1$. Proses ini memunculkan himpunan probabilistik a_{ij} dengan bentuk

$$a_{ij} = P[q_t = S_j | q_{t-1} = S_i], \quad 1 \leq i, j \leq N \quad \dots \text{ (II.10)}$$

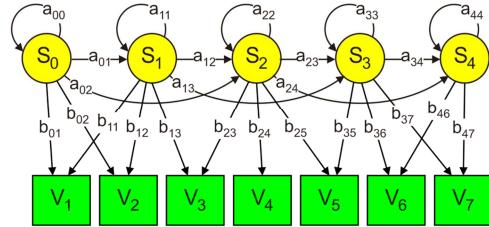
Yang memiliki konstrain stokastik $a_{ij} \geq 0$ dan $\sum_{j=1}^N a_{ij} = 1$.

Proses stokastik di atas dapat disebut sebagai model Markov yang dapat diobservasi (observable Markov model) karena output dari proses tersebut adalah himpunan state, yang di setiap satuan waktu t berkorespondensi dengan kejadian fisik yang terobservasi (physical observable event).

II.5. Hidden Markov Model

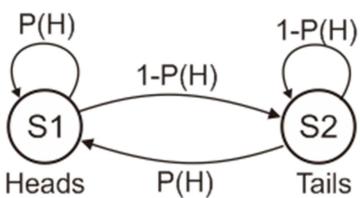
Pembahasan pada subbab ini diadaptasi dari [1]. Hidden Markov Model adalah finite-state automata probabilistik, dimana transisi antara state satu dengan state yang lain diatur menggunakan fungsi peluang. Dalam setiap transisi, state ‘memancarkan’ sebuah output dengan peluang tertentu. Pancaran output tersebut dapat berupa simbol diskrit atau nilai kontinyu multidimensi. Dalam proses markov, peluang transisi diasumsikan hanya bergantung pada transisi sebelumnya (biasanya satu), dan dapat dimodelkan sebagai rantai Markov. Keberadaan transisi dengan probabilitas bernilai nol membatasi jumlah path dan menentukan topologi dari model.

Hidden Markov model adalah Markov model dengan kasus di mana observasinya adalah fungsi probabilistik dari state, terdapat dua buah proses stokastik, salah satunya adalah proses stokastik yang tidak dapat diobservasi, sehingga disebut hidden.



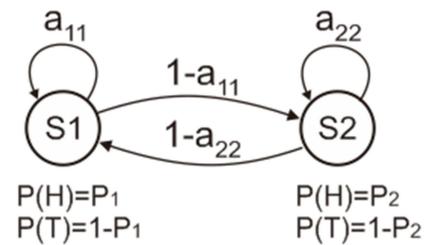
Gambar II.3 Model grafis Hidden Markov Model dengan topologi left-right

Salah satu contoh pemodelan HMM dalam dunia nyata adalah skenario pelemparan koin. Terdapat dua orang yang berada dalam ruangan berbeda. Orang pertama melakukan pelemparan koin, dan melaporkan hasilnya. Orang kedua dapat mengetahui hasil pelemparan koin namun tidak mengetahui apa yang dilakukan orang pertama terhadap koin tersebut. Keluaran dari proses pelemparan koin dapat dinotasikan $O = O_1 O_2 \dots O_T = HHHTHTH \dots H$, H adalah kepala dan T adalah ekor.



Gambar II.5 Model 2 koin

Skenario ini dapat dimodelkan dengan HMM yang memiliki dua buah state yang berkorespondensi terhadap sisi koin (seperti pada Gambar II.5). Pada kasus ini Markov model dapat diobservasi, dan spesifikasi model yang perlu dilengkapi adalah menentukan nilai untuk biasnya (mis. Peluang munculnya kepala). Model ini dapat diringkas menjadi satu state, dimana state berkorespondensi terhadap satu koin. Jika menggunakan dua koin, dapat dimodelkan dengan HMM yang memiliki dua buah state yang masing-masing state berkorespondensi terhadap koin yang berbeda. Masing-masing state dikarakterisasi dengan distribusi peluang dari kepala dan ekor, dan transisi antara state dapat berkorespondensi terhadap ‘himpunan koin yang mana yang dilempar’.



Gambar II.4 Model 1 koin

II.6. Elemen-Elemen HMM

Pembahasan pada subbab ini diadaptasi dari [1].

HMM dikarakterisasi sebagai berikut :

- N , jumlah dari state. State dinotasikan dengan $S = \{S_1, S_2, S_3, \dots, S_N\}$. State pada saat t dinotasikan dengan q_t .
- M , jumlah simbol observasi berbeda per state. Simbol observasi berkorespondensi dengan keluaran fisik pada sistem yang dimodelkan. Simbol observasi dinotasikan dengan $V = \{v_1, v_2, v_3, \dots, v_M\}$. Observasi pada saat t dinotasikan dengan variabel O_t .
- $A = \{a_{ij}\}$, sebuah matriks $N \times N$ yang merepresentasikan distribusi peluang transisi state dimana a_{ij} adalah probabilitas transisi dari S_i ke S_j .

$$a_{ij} = P[q_t = S_j | q_{t-1} = S_i], \quad 1 \leq i, j \leq N$$

- $B = \{b_j(k)\}$, sebuah matriks $N \times M$ yang merepresentasikan distribusi peluang memancarnya observasi simbol, dimana $b_j(k)$ adalah peluang mengeluarkan simbol v_k pada saat t di state S_j .

$$b_j(k) = P(O_t = v_k | q_t = S_j), \quad 1 \leq j \leq N, 1 \leq k \leq M$$

- $\pi = \{\pi_i\}$, distribusi peluang state awal, dimana π_i adalah peluang bahwa state S_i adalah state awal.

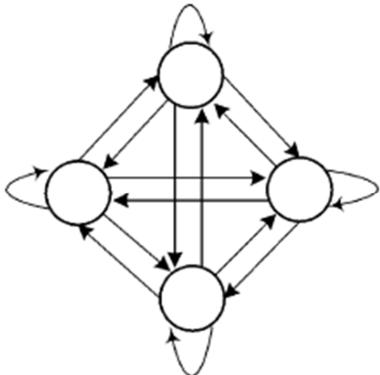
$$\pi_i = P[q_1 = S_i], \quad 1 \leq i \leq N$$

Notasi probabilistik A, B dan π harus memenuhi konstrain probabilistik berikut:

- $\sum_j a_{ij} = 1, \forall i$, dan $a_{ij} \geq 0$.
- $\sum_k b_j(k) = 1, \forall j$, dan $b_j(k) \geq 0$.
- $\sum_i \pi_i = 1$, dan $\pi_i \geq 0$.

Notasi singkat $\lambda = (A, B, \pi)$ digunakan untuk himpunan parameter probabilistik lengkap dari model HMM.

II.7. Topologi HMM

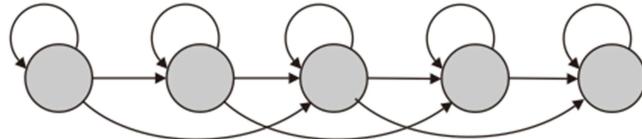


Gambar II.6 Model ergodic dengan 4 buah state

Model HMM yang terkoneksi kuat (strongly connected), disebut model ergodic. Model ergodic dengan state berjumlah $N = 4$ ditunjukkan pada gambar II.6. Model ergodic memiliki properti bahwa setiap koefisien a_{ij} adalah positif. Contohnya untuk model ergodic pada Gambar II.6,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}, a_{ij} > 0, \forall i, j$$

Topologi lain yang dianggap dapat memodelkan properti sinyal dengan lebih baik adalah model left-right. Model left-right HMM tidak memiliki path mundur, state berpindah dari kiri ke kanan atau ke state yang sama, indeks state akan meningkat atau sama seiring dengan berjalanannya waktu. Model left-right dapat memodelkan sinyal yang propertinya berubah terhadap waktu, misalnya pembicaraan (speech).



Gambar II.7 Model left-right dengan 5 buah state

Beberapa properti dari model left right adalah sebagai berikut :

- $a_{ij} = 0, j < i$, state tidak berpindah ke state dengan index yang lebih rendah.
- $\pi_i = \begin{cases} 0, & i \neq 1 \\ 1, & i = 1 \end{cases}$, rangkaian state harus dimulai dari S_1 dan berakhir di S_N .
- $a_{ij} = 0, \quad j > i + \Delta$
- $a_{NN} = 1$, state S_N tidak memiliki akses ke state yang lain
- $a_{Ni} = 0, \quad i < N$

Selain kedua topologi di atas, ada banyak variasi topologi yang mungkin, contohnya model left-right dengan memiliki path paralel [1].

II.8. Algoritma Forward-Backward untuk Evaluasi

Pembahasan pada subbab ini diadaptasi dari [1].

Jika diberikan rangkaian observasi $O_1 O_2 O_3 \dots O_T$, dan sebuah model $\lambda = (A, B, \pi)$, bagaimana menghitung $P(O|\lambda)$, peluang munculnya rangkaian observasi tersebut, jika diberikan model λ ?

Peluang munculnya rangkaian O adalah $P(O|Q, \lambda) = b_{q_1}(O_1) \cdot b_{q_2}(O_2) \cdots b_{q_T}(O_T)$ dan peluang munculnya rangkaian Q adalah $P(Q|\lambda) = \pi_{q_1} a_{q_1 q_2} a_{q_2 q_3} \cdots a_{q_{T-1} q_T}$. Peluang gabungan keduanya adalah $P(O, Q|\lambda) = P(O|Q, \lambda)P(Q|\lambda)$, dan karena $P(O|\lambda) = \sum_{\text{all } Q} P(O|Q, \lambda)P(Q|\lambda)$, maka peluang $P(O|\lambda)$ adalah

$$P(O|\lambda) = \sum_{q_1, q_2, \dots, q_T} \pi_{q_1} b_{q_1}(O_1) a_{q_1 q_2} b_{q_2}(O_2) \cdots a_{q_{T-1} q_T} b_{q_T}(O_T) \quad \dots \quad (\text{II.11})$$

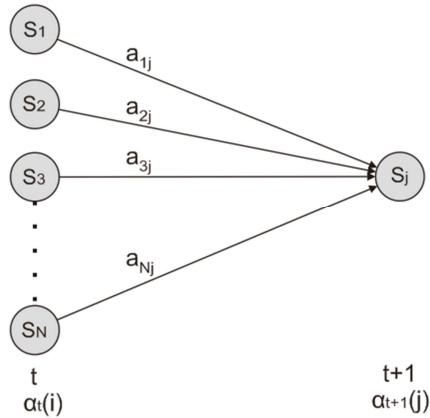
Perhitungan $P(O|\lambda)$ pada persamaan (II.11) melibatkan kalkulasi sebanyak $2T \cdot N^T$, karena di setiap $t = 1, 2, \dots, T$ terdapat N state yang mungkin diraih dan terdapat sekitar $2T$ kalkulasi yang dibutuhkan untuk masing-masing hubungan di persamaan (II.11). (Untuk presisinya, dibutuhkan multiplikasi sebanyak $(2T - 1) \cdot N^T$, dan penambahan $N^T - 1$). Kalkulasi ini tidak mungkin dilakukan secara komputasional, bahkan untuk nilai yang kecil dari N dan T , misalkan untuk $N = 5$ (state), $T = 100$ (observasi), terdapat $2 \cdot 100 \cdot 5^{100} \approx 10^{72}$ komputasi. Sehingga dibutuhkan prosedur yang lebih efisien. Prosedur ini disebut prosedur forward-backward.

Forward-Backward Procedure

1. Menghitung forward variable atau variabel maju $\alpha_t(i)$

Misalkan terdapat variabel maju, yaitu peluang gabungan munculnya rangkaian observasi parsial O_1, O_2, \dots, O_t dengan state pada waktu t adalah S_i , jika diberikan model λ , yaitu

$$\alpha_t(i) = P(O_1 O_2 \cdots O_t, q_t = S_i | \lambda) \quad \dots \quad (\text{II.12})$$



Gambar II.8 Ilustrasi rangkaian operasi pada perhitungan forward maju

Variabel maju $\alpha_t(i)$ dihitung secara induktif. Untuk $t = 1$, $\alpha_1(i) = \pi_i b_i(O_1)$, $1 \leq i \leq N$. Pada $t = t + 1$, state berpindah dari S_i ke S_j , untuk seluruh N state S_i yang mungkin, $\alpha_t(i)a_{ij}$ dijumlahkan seluruhnya sehingga $\alpha_t(i)$ diinduksi menjadi $\alpha_{t+1}(j) = [\sum_{i=1}^N \alpha_t(i) a_{ij}] b_j(O_{t+1})$. Komputasi dilakukan untuk setiap state S_j , $1 \leq j \leq N$, dengan diiterasi untuk $t = 1, 2, \dots, T - 1$. Akhirnya, pada $t = T$,

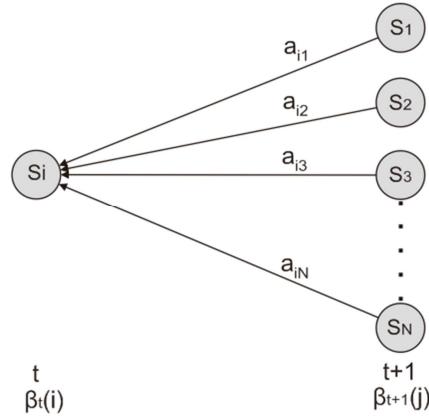
$$P(O|\lambda) = \sum_{i=1}^N \alpha_T(i) \quad \dots \text{ (II.13)}$$

Komputasi $\alpha_t(i)$, $1 \leq t \leq T$, $1 \leq i \leq N$, membutuhkan kalkulasi sebanyak N^2T (Untuk presisinya dibutuhkan $N(N + 1)(T - 1) + N$ kali perkalian dan penambahan $N(N - 1)(T - 1)$).

2. Menghitung backward variable atau variabel mundur $\beta_t(i)$

Misalkan terdapat variabel mundur, yaitu peluang gabungan munculnya rangkaian observasi parsial $O_{t+1}, O_{t+2}, \dots, O_T$ dengan state pada waktu t adalah S_i , jika diberikan model λ , yaitu

$$\beta_t(i) = P(O_{t+1} O_{t+2} \dots O_T | q_t = S_i, \lambda) \quad \dots \text{ (II.14)}$$

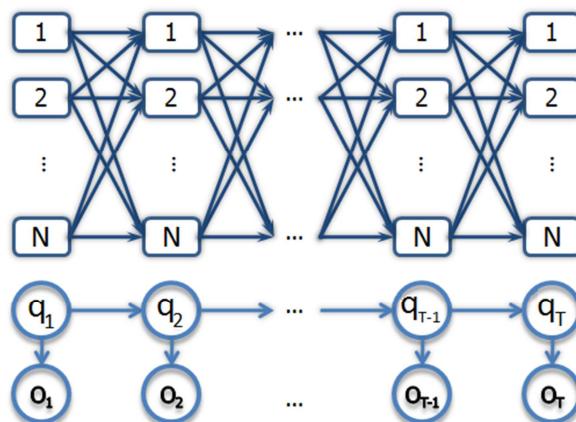


Gambar II.9 Ilustrasi rangkaian operasi pada perhitungan 17ector17k mundur

Variabel mundur $\beta_t(i)$ dihitung secara induktif. Untuk $t = T$, secara acak diset $\beta_T(i) = 1$, $1 \leq i \leq N$. Pada $t = t + 1$, state berpindah dari S_i ke S_j , untuk seluruh N state S_i yang mungkin, $a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)$ dijumlahkan seluruhnya sehingga

$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(O_{t+1}) \beta_{t+1}(j) \quad \dots \text{ (II.15)}$$

Komputasi dilakukan untuk setiap state S_i , $1 \leq i \leq N$, dengan diiterasi pada $t = T - 1, T - 2, \dots, 1$. Komputasi $\beta_t(i)$ membutuhkan kalkulasi sebanyak N^2T , dan dapat dikomputasi menggunakan struktur lattice (Gambar II.10). Variabel mundur ini digunakan pada algoritma Viterbi dan Baum Welch.



Gambar II.10 Struktur lattice untuk komputasi $\alpha_t(i)$ dan $\beta_t(i)$

II.9. Algoritma Viterbi untuk Decoding

Pembahasan pada subbab ini diadaptasi dari [1].

Diberikan rangkaian observasi $O = O_1 O_2 \cdots O_T$, dan model λ , bagaimana “membuka” rangkaian state tersembunyi $Q = q_1 q_2 \cdots q_T$, yang optimal, atau menemukan satu rangkaian state terbaik $Q = q_1 q_2 \cdots q_T$?

Solusi permasalahan ini digunakan untuk mempelajari struktur model, menemukan rangkaian state optimal, atau untuk menemukan titik awal jika diberikan observasi dan titik akhir. Untuk kasus model terdegenerasi, tidak dapat ditemukan rangkaian state yang benar-benar tepat, namun untuk solusi praktis digunakan kriteria optimal. Pemilihan kriteria optimal adalah faktor kuat yang menentukan apakah rangkaian state tersembunyi dapat dibuka.

Algoritma Viterbi.

Didefinisikan sebuah kuantitas

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} P(q_1 q_2 \cdots q_{t-1}, q_t = S_i, O_1 O_2 \cdots O_t | \lambda) \quad \dots \text{ (II.16)}$$

Dengan $\delta_t(i)$ adalah probabilitas terbesar di sepanjang satu path, sampai pada state S_i , pada saat t , dan menghitung untuk yang pertama dari t observasi. Prosedur untuk menemukan rangkaian state terbaik adalah dengan induksi sebagai berikut :

1. $\delta_1(j) = \pi_j b_j(O_1), \quad 1 \leq j \leq N$
2. $\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] \cdot b_j(O_t), \quad 2 \leq t \leq T, 1 \leq j \leq N$
- $\psi_t(j) = \operatorname{argmax}_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] \quad 2 \leq t \leq T, 1 \leq j \leq N$

Untuk informasi backtracking, argumen yang memaksimisasi $\delta_t(j)$ disimpan dalam array $\psi_t(j)$ untuk setiap t dan j .

3. $P^* = \max_{1 \leq i \leq N} [\delta_T(i)]$
- $q_T^* = \operatorname{argmax}_{1 \leq i \leq N} [\delta_T(i)]$
4. Path backtracking, $q_T^* = S_N$,
- $q_t^* = \psi_{t+1}(q_{t+1}^*), \quad t = T - 1, T - 2, \dots, 1$

II.10. Algoritma Baum Welch untuk Learning

Pembahasan pada subbab ini diadaptasi dari [1].

“Bagaimana mengatur parameter model $\lambda = (A, B, \pi)$ untuk mendeskripsikan $P(O|\lambda)$?”

Solusi permasalahan ini digunakan untuk mengoptimasi parameter-parameter model sehingga model memberikan deskripsi terbaik untuk rangkaian observasi yang diberikan. Sejumlah rangkaian data observasi, disebut dengan data pelatihan, digunakan untuk “melatih” model. Parameter-parameter model diatur agar dapat beradaptasi terhadap sejumlah data pelatihan.

Pada faktanya, tidak ada cara yang benar-benar optimal untuk mengestimasi parameter model. Namun parameter model $\lambda = (A, B, \pi)$ dapat dipilih sehingga $P(O|\lambda)$ meraih maksimum lokal dengan prosedur iteratif seperti metode Baum Welch (ekivalen dengan EM (expectation-modification)), atau teknik gradien.

Algoritma Baum Welch

Didefinisikan variabel $\xi_t(i, j) = P(q_t = S_i, q_{t+1} = S_j | O, \lambda)$, yaitu peluang berada di state S_i pada saat t , dan berada di state S_j pada saat $t + 1$, jika diberikan model dan rangkaian observasi. Variabel $\xi_t(i, j)$ ditulis dalam $\alpha_t(i)$ (II.12) dan $\beta_t(j)$ (II.14) sebagai berikut :

$$\xi_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{P(O|\lambda)} = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)} \quad \dots \text{ (II.17)}$$

Didefinisikan variabel $\gamma_t(i) = P(q_t = S_i | O, \lambda)$ sebagai peluang berada di state S_i di waktu t , jika diberikan rangkaian observasi, dan model. Variabel $\gamma_t(i)$ ditulis dalam $\alpha_t(i)$ (II.12) dan $\beta_t(i)$ (II.14) sebagai berikut :

$$\gamma_t(i) = \frac{\alpha_t(i) \beta_t(i)}{P(O|\lambda)} = \frac{\alpha_t(i) \beta_t(i)}{\sum_{i=1}^N \alpha_t(i) \beta_t(i)} \quad \dots \text{ (II.18)}$$

Sehingga relasi variabel $\gamma_t(i)$ dengan $\xi_t(i, j)$ adalah sebagai berikut :

$$\gamma_t(i) = \sum_{j=1}^N \xi_t(i,j) \quad \dots \quad (\text{II.21})$$

Sehingga, jika dijumlahkan $\gamma_t(i), \xi_t(i,j)$, untuk semua t , maka

$\sum_{t=1}^{T-1} \gamma_t(i) =$ frekuensi harapan S_i dikunjungi = transisi datang dari S_i

$\sum_{t=1}^{T-1} \xi_t(i,j) =$ frekuensi harapan transisi datang dari S_i menuju S_j .

Menggunakan formula di atas, didapat metode untuk mere-estimasi parameter model HMM. Himpunan formula re-estimasi untuk π, A , dan B adalah :

$$\bar{\pi}_i = \text{frekuensi harapan di state } S_i \text{ pada saat } t = 1 \text{ adalah } \gamma_1(i) \quad (\text{II.22})$$

$$\bar{a}_{ij} = \frac{\text{freq harapan transisi datang } S_i \text{ menuju } S_j}{\text{freq harapan transisi datang dari } S_i} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \quad (\text{II.23})$$

$$\bar{b}_j(k) = \frac{\text{freq harapan berada di } S_j \text{ dan simbol observasi } v_k}{\text{freq harapan berada di state } S_j} = \frac{\sum_{t=1}^T \gamma_{t O_t=v_k}(j)}{\sum_{t=1}^{T-1} \gamma_t(j)} \quad (\text{II.24})$$

Diinisialisasi model $\lambda = (A, B, \pi)$, lalu dihitung sisi kanan persamaan (II.22)-(II.24). Model yang telah dire-estimasi adalah $\bar{\lambda} = (\bar{A}, \bar{B}, \bar{\pi})$, yang parameternya ditentukan dari sisi kiri persamaan (II.22)-(II.24). Telah dibuktikan oleh Baum bahwa model awal λ menentukan titik kritis dari fungsi likelihood untuk kasus $\bar{\lambda} = \lambda$. Model $\bar{\lambda}$ berpeluang lebih besar daripada model λ dalam kasus $P(O|\bar{\lambda}) > P(O|\lambda)$, sehingga didapat model $\bar{\lambda}$ dimana rangkaian observasi lebih mungkin dihasilkan.

Model $\bar{\lambda}$ digunakan untuk menggantikan λ , lalu kalkulasi re-estimasi diulang terus menerus sampai parameter model meraih titik limit. Hasil akhir dari prosedur re-estimasi disebut estimasi maximum likelihood dari HMM. Karena algoritma forward-backward menuju ke maksimum lokal saja, dan pada banyak kasus, permukaan optimisasi itu sangat kompleks dan mempunyai banyak maksimum lokal.

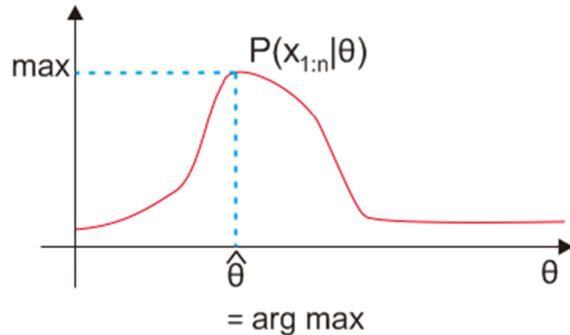
Konstrain stokastik untuk parameter $\bar{\pi}_i, \bar{a}_{ij}$ dan $\bar{b}_j(k)$, secara otomatis dipenuhi oleh masing-masing iterasi.

II.11. Maximum Likelihood

Pembahasan pada subbab ini diadaptasi dari [3].

Proses pembelajaran yang berdasarkan frekuensi data (frequentist learning), mengasumsikan bahwa pasti terdapat sebuah model dengan parameter θ_0 . Estimasi nilai pembelajaran (learned value) dinotasikan dengan $\hat{\theta}$. Jika diberikan sejumlah n data, $x_{1:n} = \{x_1, x_2, \dots, x_n\}$, lalu dipilih nilai θ yang memiliki peluang lebih besar untuk menghasilkan data tersebut, maka

$$\hat{\theta} = \arg \max_{\theta} p(x_{1:n}|\theta)$$



Gambar II.11 Ilustrasi pemilihan $\hat{\theta}$ dari $p(x_{1:n}|\theta)$

Prosedur untuk mencari maximum likelihood adalah sebagai berikut.

1. Jika diberikan sejumlah n data, $x_{1:n} = \{x_1, x_2, \dots, x_n\}$, maka peluang gabungan dari seluruh data adalah :

$$P(x_{1:n}|\theta) = \prod_{i=1}^n P(x_i|\theta)$$

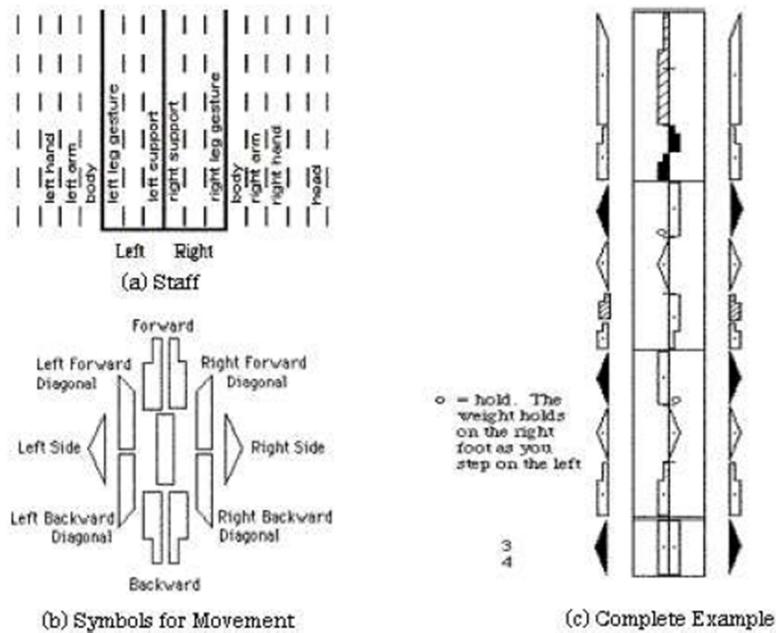
2. Kemudian log-likelihoodnya adalah

$$\mathcal{L}(\theta) = \log P(x_{1:n}|\theta) = \sum_{i=1}^n \log P(x_i|\theta)$$

3. Setelah itu log-likelihood didiferensiasi, disamakan dengan nol untuk menemukan nilai estimasi dari θ .

II.12. Beberapa Penelitian Seputar Pemodelan Tari

Koreografi tari telah dimodelkan/diformalisasi dengan berbagai macam pendekatan, misalnya notasi Laban seperti Gambar II.12, yang diinisiasi oleh Rudolf Laban pada abad ke 20. Salah satu pendekatan yang lain adalah model koreografi pemanasan rutin balet klasik Barre dengan menggunakan deret dari event-driven poses yang mengambil bentuk dari automata hingga [4], dimana state-state merepresentasikan pose dari satu kaki, dan transisi antar state merepresentasikan gerakan atau trajektori antara pose. Satu pose dibangun dari tiga buah sudut dua dimensi pada joint paha, joint lutut, dan joint pergelangan kaki. Digunakan komposisi Cartesian pada dua state machine tersebut untuk memodelkan gerakan pemanasan Barre yang melibatkan dua kaki tanpa bertabrakan. Amy Lavers juga membangun pembangkitan otomatis dari frase tarian balet [5] menggunakan Linear Temporal Logic (LTL) dan Computation Tree Logic (CTL) sebagai bahasa spesifikasi yang kaya untuk gerakan robot.



Gambar II.12 Notasi Laban

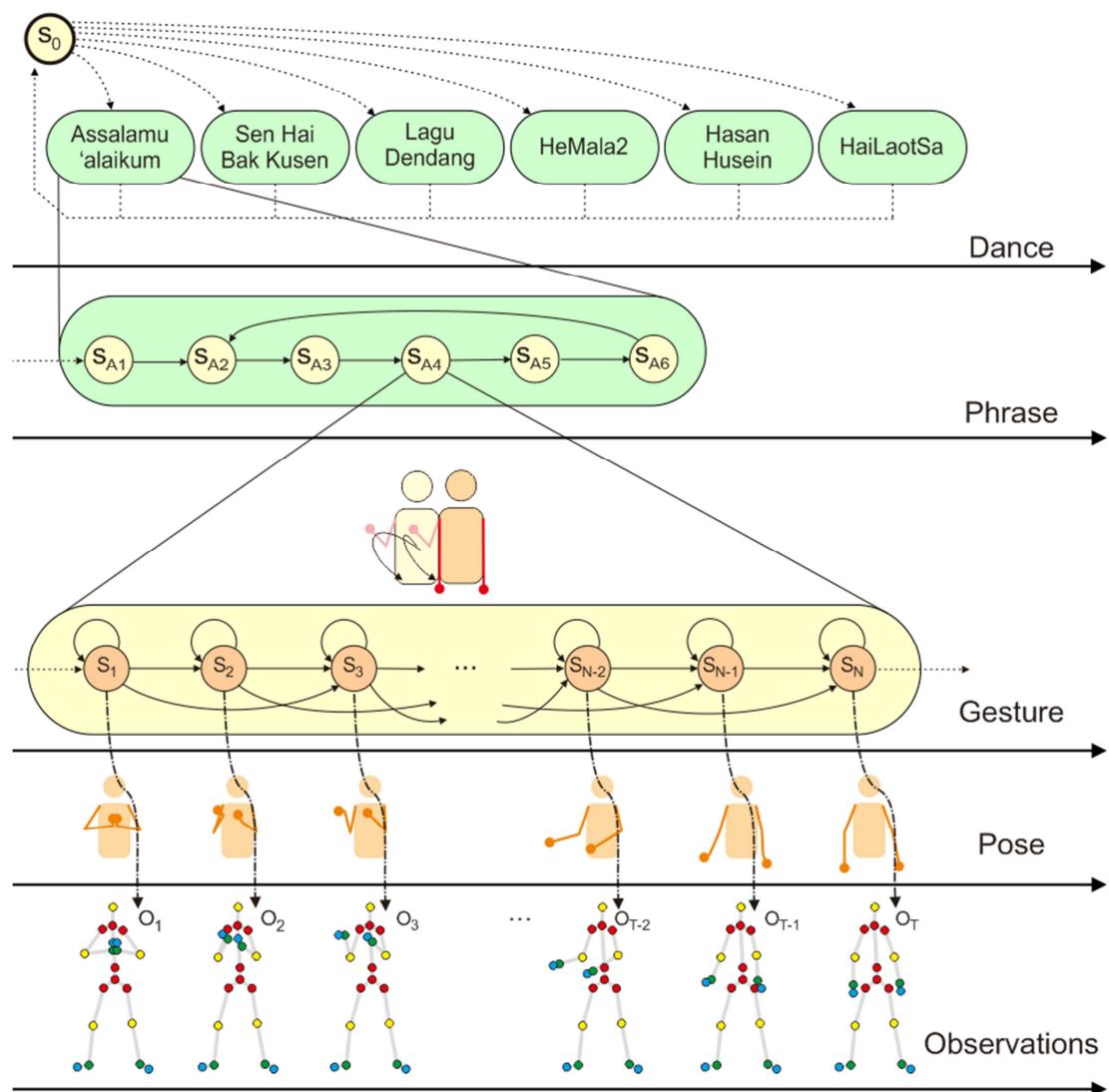
Yaya Heryadi [6] membangun model sintaksis untuk mengklasifikasi dan mengevaluasi pertunjukkan tari tradisional Bali, mengadaptasi model skeleton

feature descriptor yang sebelumnya telah dibangun oleh Michalis Raptis [7], dimana beliau merepresentasikan gesture tari dengan menggunakan parameter koordinat bola (θ, φ) dari beberapa skeleton yang diklasifikasikan menjadi bingkai batang tubuh, joint-joint berderajat satu, dan joint-joint berderajat dua. Digunakan juga matriks rotasi yang menotasikan posisi dan orientasi bingkai tubuh relatif terhadap sistem koordinat kamera yang diparameterisasi dengan sudut Tait-Bryan.

BAB III PEMODELAN TARI DAN GESTURE TARI

III.1. Pemodelan Tari Likok Pulo

Tarian didefinisikan sebagai himpunan transisi dari satu gesture ke gesture yang lain. Jika sebuah gesture dimodelkan dengan sebuah state, maka tarian dapat dimodelkan sebagai mesin sekuensial atau automata hingga yang terdiri dari state-state. Tari Likok Pulo dilakukan dengan banyak user/penari. Keseragaman gerak antar penari dan timing gerak pada ritmik yang berubah menjadi pertimbangan untuk memodelkan gesture tari.



Gambar III.1 Hierarki Tarian, Frase, Gesture, Pose, dan Observation

Berikut ini beberapa terminologi dalam dunia tari yang digunakan dalam penelitian ini :

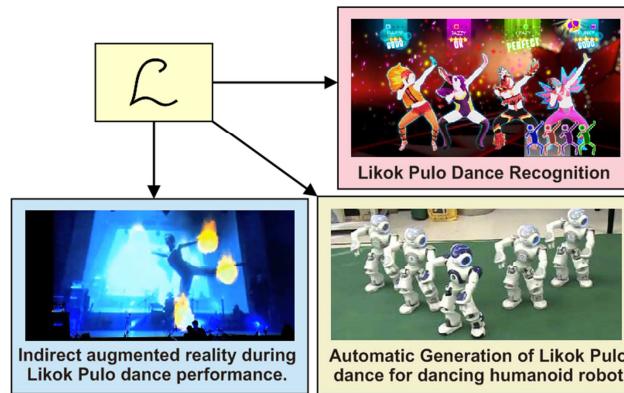
1. Pose, adalah konfigurasi statis dari tubuh, tanpa ada gerakan sama sekali.
2. Gesture, adalah gerakan dinamis yang merupakan rangkaian dari pose-pose.
3. Frase, adalah penggalan koreografi yang terdiri dari rangkaian gesture-gesture yang mungkin berulang.
4. Tarian, adalah koreografi keseluruhan dari gerakan tari dari awal sampai akhir, yang terdiri dari frase-frase yang berderet.

Hierarki dari pose, gesture, frase, tarian, dapat dilihat pada Gambar III.2. Keseluruhan Tari Likok Pulo dimodelkan sebagai berikut :

$$\mathcal{L} = (\mathbf{S}, \mathbf{I}, \mathbf{P}, \mathbf{O}, \mathbf{f}, \mathbf{e}, \mathbf{s}_0, \mathbf{S}_t)$$

- **S** – himpunan hingga tak kosong dari state tersembunyi yang merepresentasikan satu gesture. Tersembunyi karena tidak bisa diketahui secara pasti. Batas antara satu state dengan state lain pada \mathcal{L} ditentukan oleh pakar tari.
- **I** – himpunan hingga tak kosong dari masukan.
- **P** – perbendaharaan seluruh pose tari yang telah dikuantisasi yang mungkin muncul. Pose dikuantisasi berdasarkan range of movement tubuh manusia, dijelaskan pada Bab IV.
- **O** – himpunan hingga tak kosong dari keluaran, dimana $O = \{o_1, o_2, \dots, o_T\}$, $o_i \in P^*$, $i \in \{1, 2, \dots, T\}$. P^* adalah himpunan dari semua kemungkinan string dengan lebar hingga dari elemen-elemen himpunan P , yaitu pose. Yang dimaksud dengan “string of pose” adalah gesture, sehingga keluaran O berkorespondensi terhadap trajektori gesture, atau fiturnya.
- **f** – relasi transisi state $f : S \times I \rightarrow S$ yang memetakan $S \times I$ ke dalam S . Transisi antar state berkorespondensi terhadap transisi antar gesture, yang memenuhi $\forall s \in S$ dan $\forall x, y \in I$, $f(s, xy) = f(f(s, x), y)$ dan $f(s, \varepsilon) = s$, dimana ε adalah transisi kosong.

- e – sebuah fungsi keluaran $e : S \times I \rightarrow O$, yang memetakan $S \times I$ ke dalam O .
- s_0 – initial state atau state awal, $s_0 \in S$. State awal merepresentasikan pose atau gesture awal dari seluruh frase tari Likok Pulo.
- S_t – himpunan dari terminal state atau himpunan dari state-state akhir, $S_t \subseteq S$. State-state final merepresentasikan akhir dari sebuah frase. State akhir ini akan bermanfaat untuk proses rekognisi dari sebuah frase.



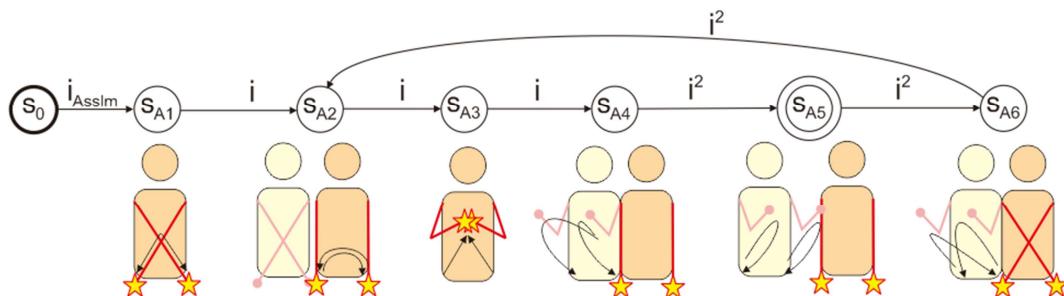
Gambar III.2. Beberapa implementasi model \mathcal{L} pada 26ector praktis.

Jika model diimplementasikan pada sistem praktis, masukan I dapat berkorespondensi terhadap (1) ritmik musik tari, atau (2) input yang ditentukan sebelumnya oleh sistem. Keluaran O dapat berkorespondensi terhadap misal, augmented reality pada dekorasi layar digital dibalik penari Likok Pulo dari hasil rekognisi tari real-time.

Berikut ini model frase-frase dalam Tari Likok Pulo.

1) Frase Assalamualaikum :

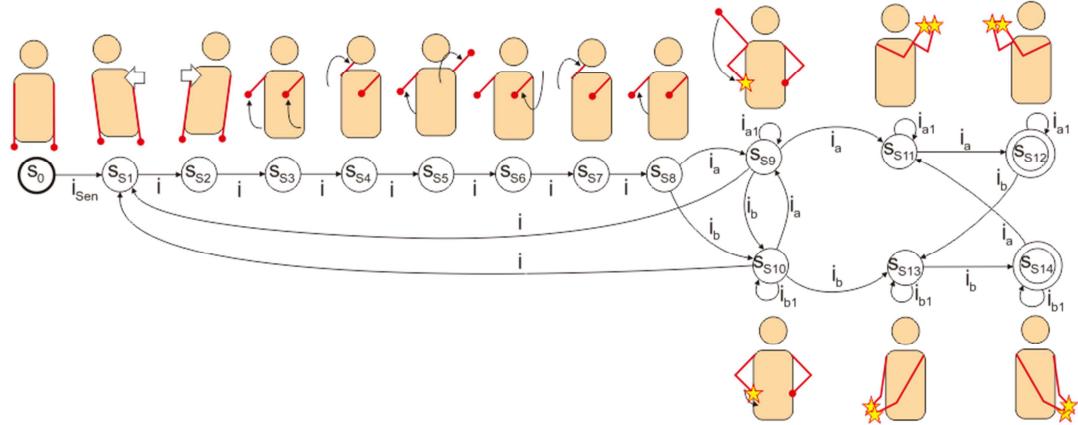
$$\{s_0 s_{A1} (s_{A2} s_{A3} s_{A4} s_{A5} s_{A6})^7 s_{A2} s_{A3} s_{A4} s_{A5}\} .$$



Gambar III.3 Model untuk Frase Assalamualaikum.

2) Frase Hai Aneuk Sen

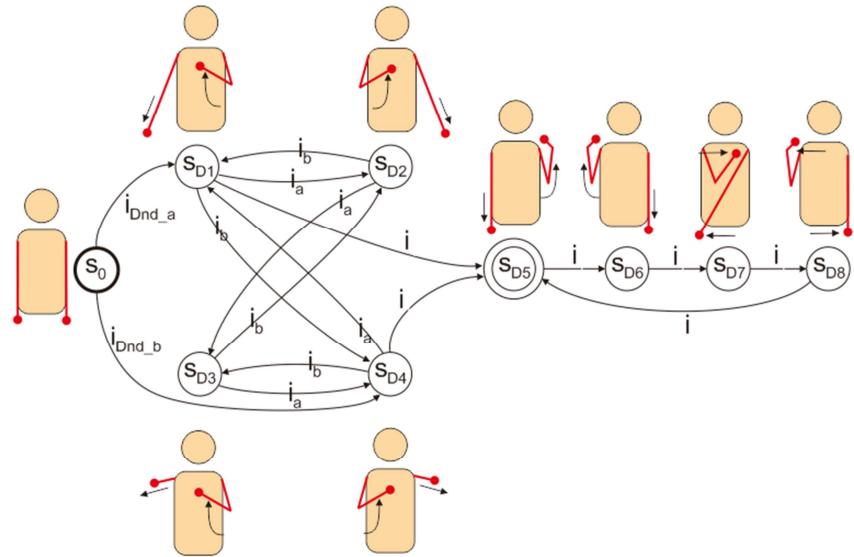
$$\{s_0(s_{S1}s_{S2}s_{S3}s_{S4}s_{S5}s_{S6}s_{S7}s_{S8}(s_{S9}^2s_{S10}^2)^4)^2(s_{S11}s_{S12}s_{S13}s_{S14})^2 \cup \\ s_0(s_{S1}s_{S2}s_{S3}s_{S4}s_{S5}s_{S6}s_{S7}s_{S8}(s_{S10}^2s_{S9}^2)^4)^2(s_{S13}s_{S14}s_{S11}s_{S12})^2\} .$$



Gambar III.4 Model untuk Frase Hai Aneuk Sen.

3) Frase Lagu Dendang

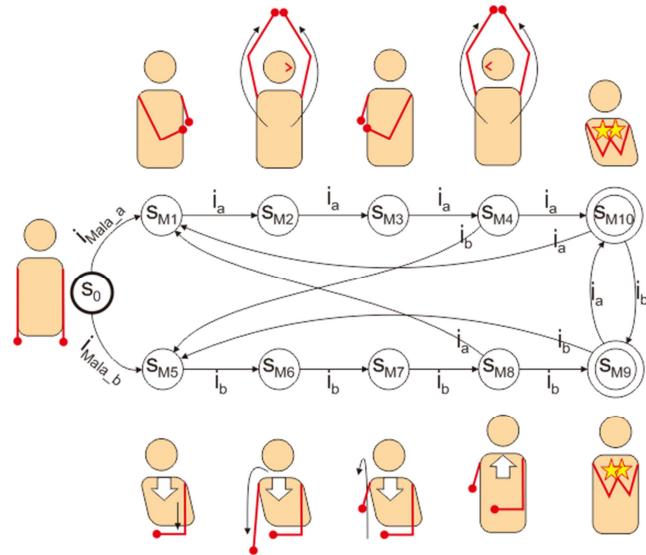
$$\{s_0(s_{D1}s_{D2}s_{D3}s_{D4})^2(s_{D5}s_{D6}s_{D7}s_{D8})^{15}s_{D5} \cup \\ s_0(s_{D4}s_{D3}s_{D2}s_{D1})^2(s_{D5}s_{D6}s_{D7}s_{D8})^{15}s_{D5}\}$$



Gambar III.5 Model untuk Frase Lagu Dendang.

4) Frase He Mala-Mala

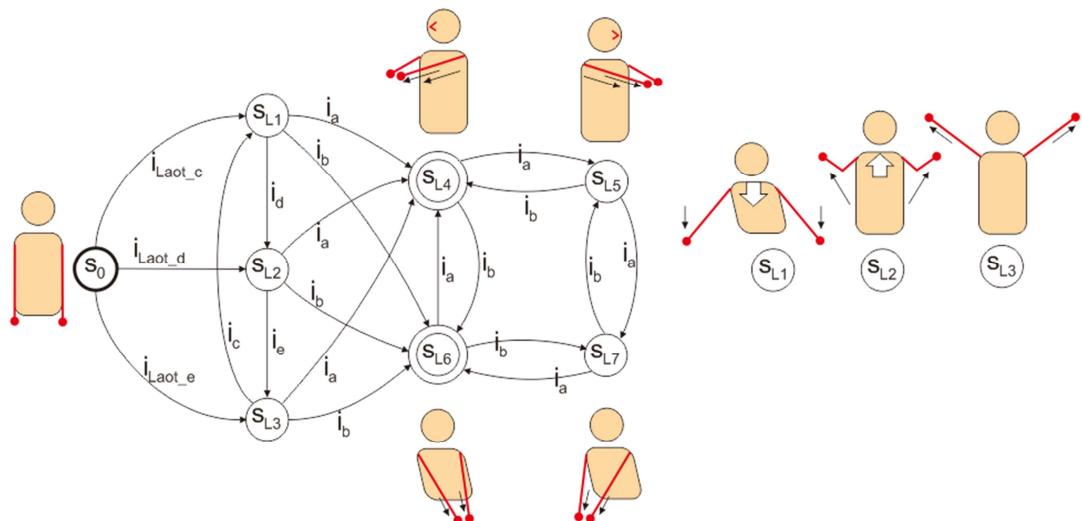
$$\{s_0(s_{M1}s_{M2}s_{M3}s_{M4}s_{M5}s_{M6}s_{M7}s_{M8}(s_{M9}s_{M10})^2)^8 \cup \\ s_0(s_{M5}s_{M6}s_{M7}s_{M8}s_{M1}s_{M2}s_{M3}s_{M4}(s_{M9}s_{M10})^2)^8\}$$



Gambar III.6 Model untuk Frase He Mala Mala.

5) Frase Kisah Hasan Husein

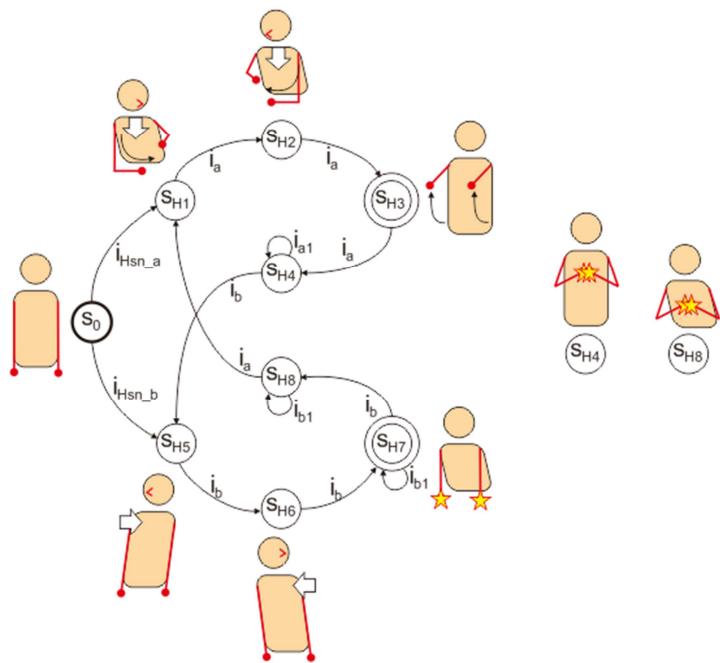
$$\{s_0(s_{H1}s_{H2}s_{H3}s_{H4}^2s_{H5}s_{H6}s_{H7}s_{H8}^2)^7s_{H1}s_{H2}s_{H3}s_{H4}^2s_{H5}s_{H6}s_{H7} \cup \\ s_0(s_{H5}s_{H6}s_{H7}s_{H8}^2s_{H1}s_{H2}s_{H3}s_{H4}^2)^7s_{H5}s_{H6}s_{H7}s_{H8}^2s_{H1}s_{H2}s_{H3}\}$$



Gambar III.7 Model untuk Frase Hai Laot Sa.

6) Frase Kisah Hasan Husein

$$\begin{aligned} & \{s_0(s_{L1}s_{L2}s_{L3})^{10}s_{L1}s_{L2}(s_{L4}s_{L5}s_{L7}s_{L6})^{12} \cup \\ & s_0(s_{L1}s_{L2}s_{L3})^{10}s_{L1}s_{L2}(s_{L6}s_{L7}s_{L4}s_{L5})^{12} \cup \\ & s_0(s_{L2}s_{L3}s_{L1})^{10}s_{L2}s_{L3}(s_{L4}s_{L5}s_{L7}s_{L6})^{12} \cup \\ & s_0(s_{L2}s_{L3}s_{L1})^{10}s_{L2}s_{L3}(s_{L6}s_{L7}s_{L4}s_{L5})^{12} \cup \\ & s_0(s_{L3}s_{L1}s_{L2})^{10}s_{L3}s_{L1}(s_{L4}s_{L5}s_{L7}s_{L6})^{12} \cup s_0(s_{L3}s_{L1}s_{L2})^{10}s_{L3}s_{L1}(s_{L6}s_{L7}s_{L4}s_{L5})^{12}\} \end{aligned}$$

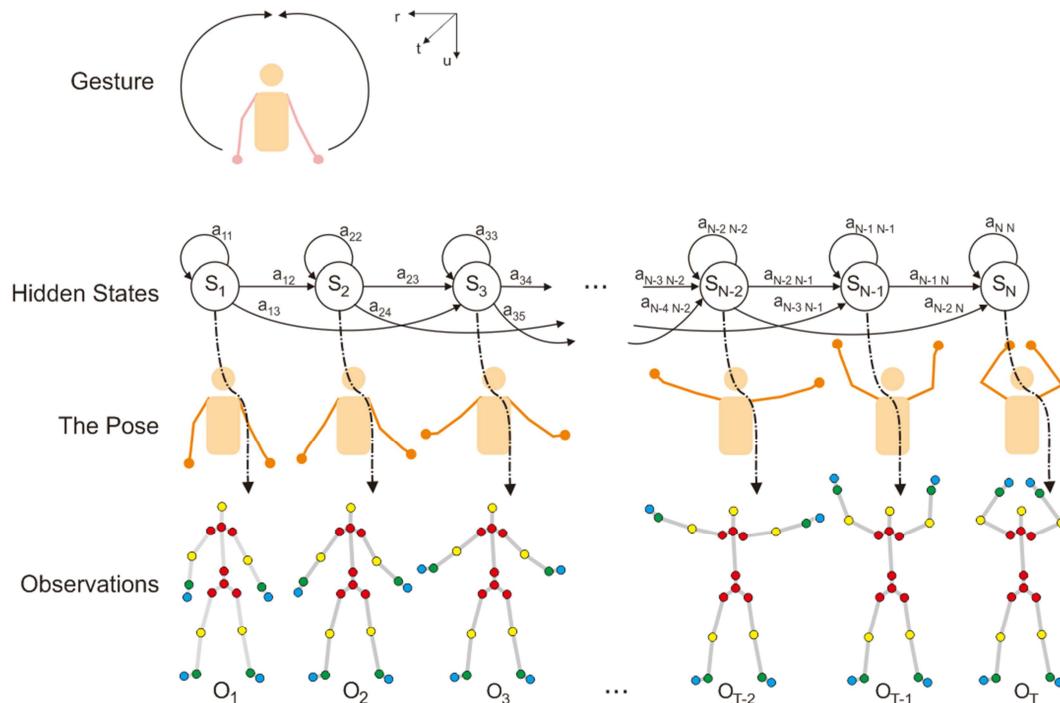


Gambar III.8 Model untuk Frase Kisah Hasan Husein.

III.2. Pemodelan Gesture Tari dalam Hidden Markov Model

Gesture tari dideskripsikan sebagai rangkaian spatio temporal dari vektor pose tari $P_{u,2} = (RS_\varphi, RS_\theta, RE_\varphi, RE_\theta, LS_\varphi, LS_\theta, LE_\varphi, LE_\theta)$ yang terdiri dari sudut-sudut esensial pada skeleton berderajat dua yang merepresentasikan sebuah pose tari dalam ruang 3D. Pose adalah konfigurasi statis dari tubuh tanpa ada gerakan sama sekali. Vektor pose tersebut kemudian dikuantisasi menjadi beberapa codewords berdasarkan range of movement.

Ritmik/beat musik dalam tari Likok Pulo mengalami perubahan semakin cepat secara linear dalam satu frasa. Sehingga setiap gesture, dimodelkan dengan left-right HMM berderajat dua yang memanfaatkan karakteristik temporal dari sinyal gesture. Topologi left-right model dipilih karena model ini baik untuk memodelkan deret waktu yang dibatasi urutan (order-constrained time-series) dimana propertinya berubah secara sekuensial terhadap waktu [16].



Gambar III.9 Ilustrasi hidden Markov model dari sebuah gesture tari.

Karakteristik formal dari HMM untuk sebuah gesture tari adalah sebagai berikut :

- N , adalah jumlah dari state tersembunyi (hidden state). State dinotasikan dengan $S = \{S_1, S_2, S_3, \dots, S_N\}$. State pada saat t dinotasikan dengan q_t . State merepresentasikan pose tari.
- M , adalah jumlah simbol observasi berbeda di setiap state. Simbol observasi dinotasikan dengan $V = \{v_1, v_2, v_3, \dots, v_M\}$. Observasi pada saat t dinotasikan dengan variabel O_t . Simbol observasi merepresentasikan keluaran fisik pada sistem yaitu vektor pose $P_{u,2}$, yang dijelaskan pada Bab IV.
- $A = \{a_{ij}\}$ adalah sebuah matriks $N \times N$ yang merepresentasikan distribusi peluang transisi state, dimana a_{ij} adalah probabilitas transisi dari state S_i ke S_j .

$$a_{ij} = P[q_t = S_j | q_{t-1} = S_i], \quad 1 \leq i, j \leq N$$

- $B = \{b_j(k)\}$, adalah sebuah matriks $N \times M$ yang merepresentasikan distribusi peluang keluarnya simbol observasi pada setiap state, dimana $b_j(k)$ adalah peluang mengeluarkan simbol v_k pada saat t di state S_j .

$$b_j(k) = P(O_t = v_k | q_t = S_j), \quad 1 \leq j \leq N, 1 \leq k \leq M$$

- $\pi = \{\pi_i\}$, adalah distribusi peluang state awal, dimana π_i adalah peluang bahwa state/pose S_i adalah state/pose awal.

$$\pi_i = P[q_1 = S_i], \quad 1 \leq i \leq N$$

Karena A, B dan π adalah probabilistik, maka harus memenuhi konstrain-konstrain probabilistik di bawah ini :

$$\sum_j a_{ij} = 1, \forall i, \text{ dan } a_{ij} \geq 0.$$

$$\sum_k b_j(k) = 1, \forall j, \text{ dan } b_j(k) \geq 0.$$

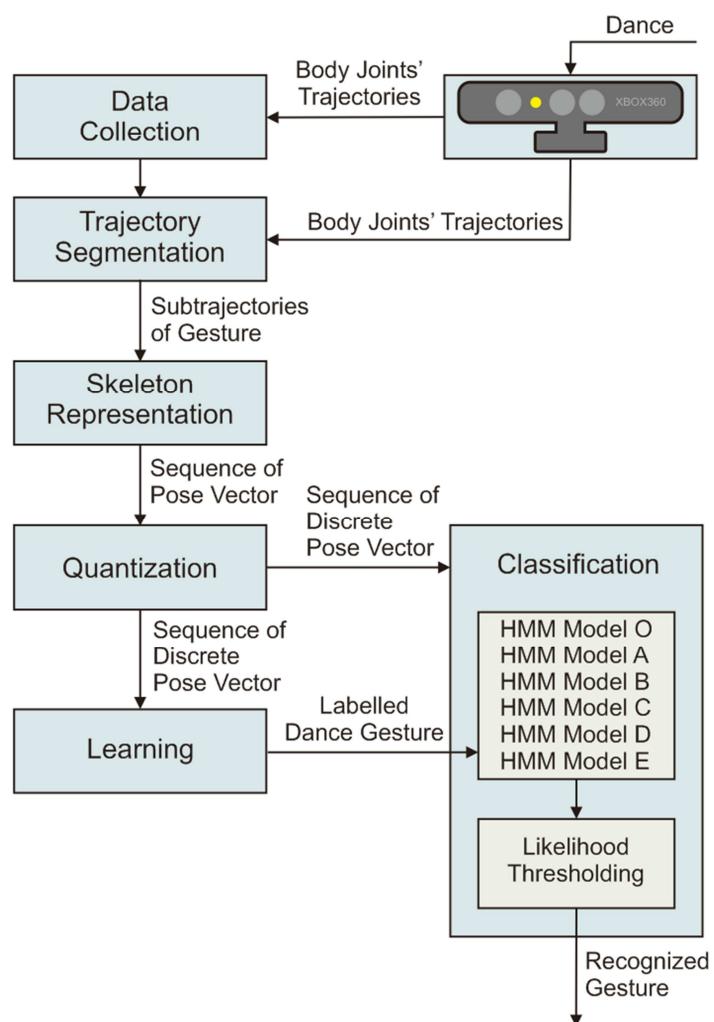
$$\sum_i \pi_i = 1, \text{ dan } \pi_i \geq 0.$$

Notasi singkat $\lambda = (A, B, \pi)$ digunakan untuk himpunan parameter probabilistik lengkap dari model HMM untuk sebuah gesture tari.

BAB IV PEMBELAJARAN DAN KLASIFIKASI GESTURE TARI

IV.1. Model Klasifikasi dengan Threshold

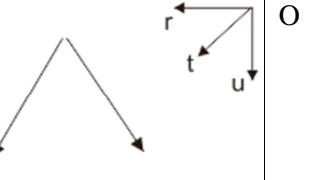
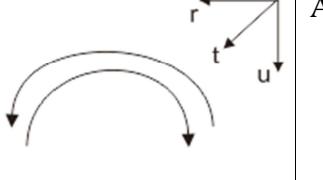
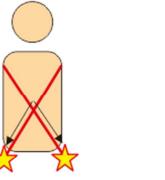
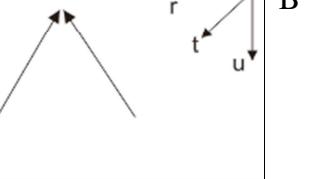
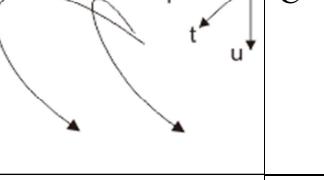
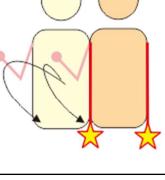
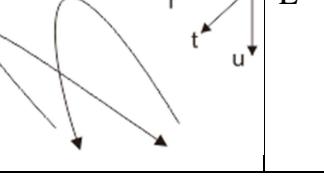
Gambar V.1. menunjukkan diagram blok yang dirancang untuk proses pembelajaran dan klasifikasi gesture tari pada penelitian ini.



Gambar IV.1 Diagram Blok Sistem

IV.2. Pemilihan Kelas Gesture Tari yang Akan Diuji

Dipilih enam buah gesture pada frase “Assalamualaikum”. Enam buah gesture pada frase “Assalamualaikum” disegmentasi oleh tepukan tangan. Gesture-gesture tari terpilih pada penelitian ini dapat dilihat di Tabel IV.1.

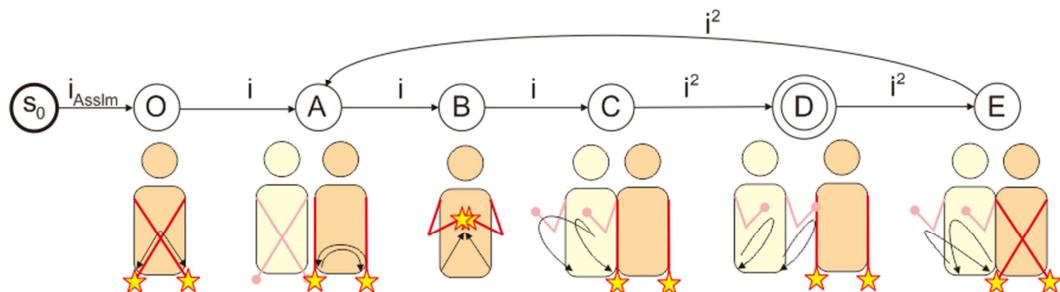
Actual Gesture	Gesture	Wrist Trajectories	Kls
			O
			A
			B
			C
			D
			E

Tabel IV.1 Daftar enam kelas gesture yang akan diuji

Gerakan untuk masing-masing gesture adalah sebagai berikut :

- Gesture O adalah transisi dari pose “tepuk tangan di depan dada” ke pose “tangan bersilangan di atas paha”.
- Gesture A adalah transisi dari pose “tangan bersilangan di atas paha” ke pose “tangan lurus di atas paha”.
- Gesture B adalah transisi dari pose “tangan lurus di atas paha” ke pose “tepuk tangan di depan dada”.
- Gesture C adalah transisi dari pose “tepuk tangan di depan dada” ke pose “kedua tangan berayun ke arah kanan” lalu diakhiri pose “tangan lurus di atas paha”.
- Gesture D adalah transisi dari pose “tangan lurus di atas paha” ke pose “kedua tangan berayun ke arah kiri” lalu diakhiri pose “tangan lurus di atas paha”
- Gesture E adalah transisi dari pose “tangan lurus di atas paha” ke pose “kedua tangan berayun ke arah kanan” lalu diakhiri pose “tangan bersilangan di atas paha”

Penamaan O, A, B, C, D, E adalah penyederhanaan penulisan dari A1, A2, A3, A4, A5, A6 pada Bab III.



Gambar IV.2 Mesin Sekuensial untuk Frase “Assalamualaikum”

Pada frase “Assalamualaikum”, $f(s_0, i) = O$, $f(O, i) = A$, $f(A, i) = B$, $f(B, i) = C$, $f(C, i^2) = D$, $f(D, i^2) = E$, $f(E, i^2) = A$. Rangkaian gesture dikenal sebagai frase “Assalamualaikum” jika $\{s_0 O(ABCDE)^7 ABCD\}$.

IV.3. Perangkat Penginderaan

Sistem pembelajaran dan klasifikasi gesture tari berbasis HMM diimplementasikan pada perangkat lunak dan perangkat keras pada Tabel IV.2.

Perangkat Keras	Perangkat Lunak
Kinect XBOX360	Sistem Operasi Windows 7 32bit
Smartglove (optional)	Realterm
<ul style="list-style-type: none">• ADXL345 (unused)• Gyroscope ITG3200 (unused)• Magnetometer HMC5883 (unused)• DF Bluetooth V3 Module• Clap Sensors• ATMEL Atmega328P-PU	<ul style="list-style-type: none">MATLAB R2013aKinect API<ul style="list-style-type: none">• Microsoft Visual Studio 2012• Kinect for Windows SDK v1.8Simulink<ul style="list-style-type: none">• Computer Vision System Toolbox• Simulink Coder

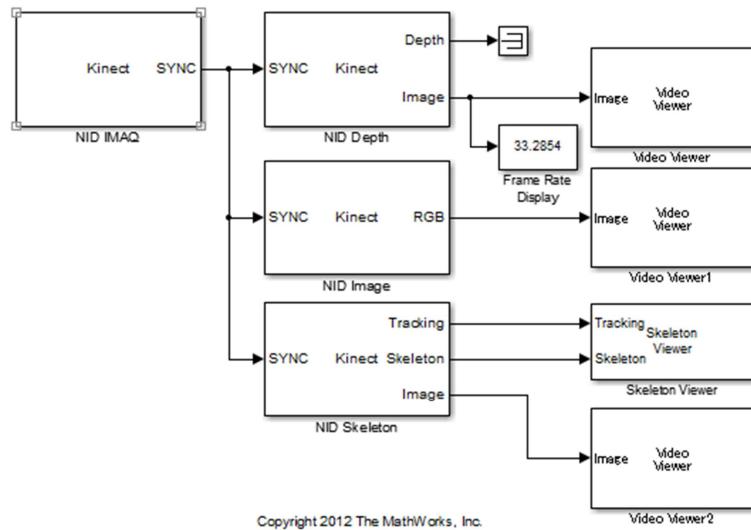
Tabel IV.2 Spesifikasi perangkat keras dan perangkat lunak .

Digunakan sensor Kinect XBOX yang ditunjukkan pada Gambar IV.3. Skeletal tracking algorithm (STA) pada platform Kinect menangkap skeleton dengan kecepatan 30 frame per detik. Skeleton direpresentasikan dengan 20 titik di setiap joint. Setiap titik diberi tanda “tracked” (terdeteksi) atau “inferred” (terduga terdeteksi). Joint diberi tanda “inferred” jika dianggap terganggu oleh derau.

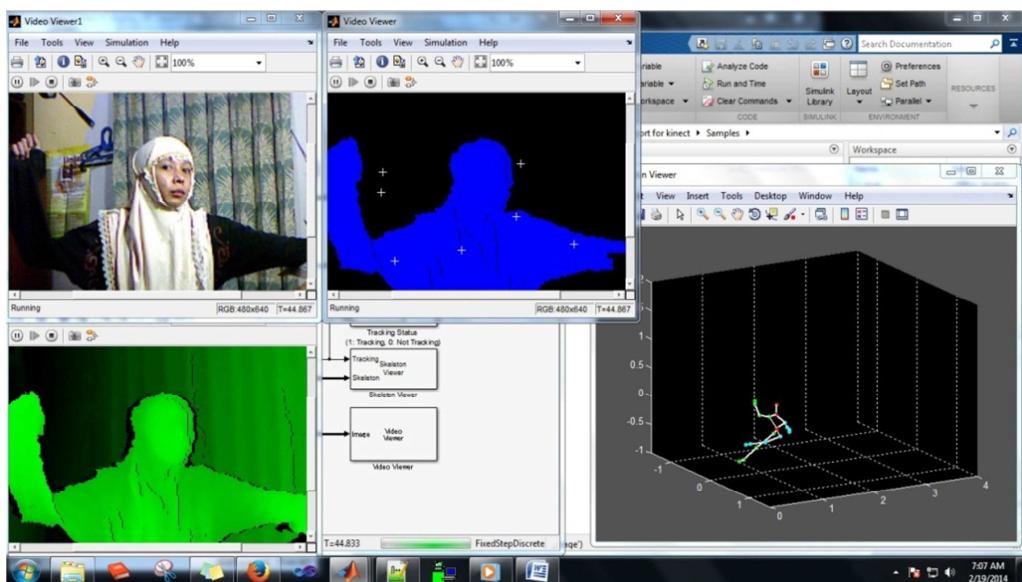


Gambar IV.3 Sensor Kinect XBOX360

Sistem pembelajaran dan klasifikasi diimplementasikan dalam environment MATLAB dan Simulink. Kode untuk mentrigger Kinect dan interfacenya ditulis dalam C++ dan dikompilasi menggunakan mex compiler. Digunakan MATLAB agar penelitian fokus pada algoritma dan hasilnya, bukan pada desain pemrogramannya. Toolbox yang digunakan pada penelitian kali ini adalah Computer Vision System Toolbox. Blok simulink lengkap yang digunakan untuk akuisisi data ditunjukkan pada Gambar IV.4. Akuisisi data sensor Kinect ditunjukkan pada Gambar IV.5.



Gambar IV.4 Desain blok Simulink untuk akuisisi data sensor Kinect



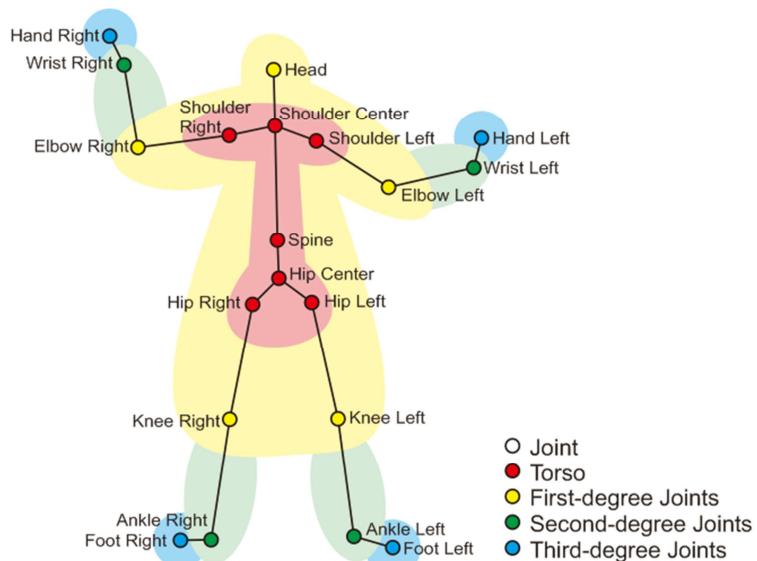
Gambar IV.5 Akuisisi data sensor Kinect menggunakan matlab dan 3dector36k

IV.4. Representasi Skeleton

Penginderaan tarian oleh Kinect melibatkan banyak artikulasi joint skeleton. Tipikal sinyal trajektori data Kinect yang asli kemungkinan besar kehilangan entropi jika dimensinya berkurang. Maka dibutuhkan representasi yang tepat sehingga data dapat dikelompokkan dengan baik oleh classifier dengan tujuan [7]: (1) Sistem koordinat yang kokoh berbasis orientasi tubuh manusia, agar representasi skeleton tidak bergantung pada posisi sensor Kinect. (2) Kontinuitas dan stabilitas sinyal. (3) Mengurangi dimensi sinyal dengan tetap mempertahankan karakter gerak.. [7]

Beberapa pre-processing data trajektori joint-joint skeleton yang diperoleh dari sensor Kinect adalah sbb: posisi joint dipetakan pada ruang gerak 3D [10]; trajektori joint relatif dikelompokkan dengan K-means [13]; trajektori joint disegmentasi dengan suara langkah kaki [11]; menggunakan sudut dan kecepatan sudut joint [12]; dan menggunakan skeleton descriptor [6].

Jika posisi joint xyz, translasi atau akselerasinya dipetakan dalam ruang 3D yang tersegmentasi; untuk beberapa joint yang diamati secara simultan akan menghasilkan dimensi observasi yang redundan bagi model HMM. Pada penelitian ini dipilih representasi sudut bola sebagai representasi yang paling tepat untuk memodelkan pose tari, diadaptasi dari skeleton feature deskriptor [6].



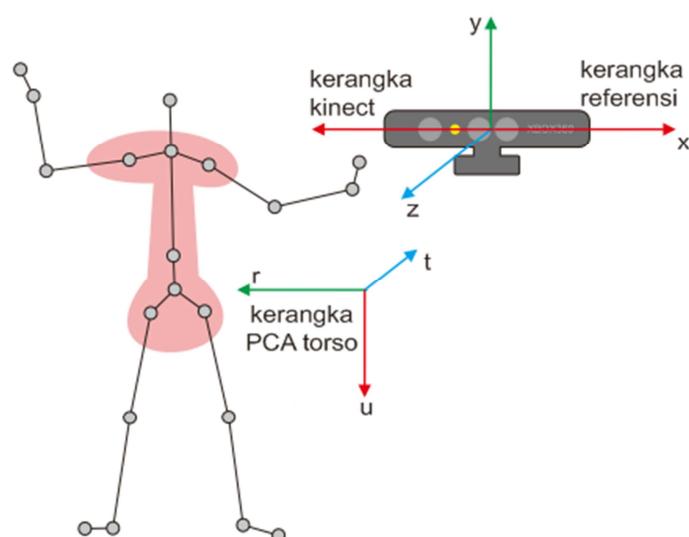
Gambar IV.6 Penamaan joint-joint skeleton, beserta hierarkinya.

Kerangka PCA Torso

Principal Component Analysis (PCA) digunakan untuk mengidentifikasi dan mengekspresikan pola-pola, kesamaan dan perbedaan antara data-data. Biasanya digunakan standar deviasi, variansi, kovariansi, atau vektor/nilai eigen [17].

Joint yang berada pada torso (shoulder right, shoulder left, shoulder center, spine, hip right, hip left) jarang mengeluarkan gerakan independen dengan sudut besar. Derau menyebabkan joint-joint pada torso akan memunculkan gerakan yang tidak realistik jika representasinya tidak dibatasi. Maka torso dapat dianggap sebagai tubuh kaku yang memberikan informasi frame, kemudian frame ini menjadi basis ortonormal 3D bagi joint-joint sisanya. Komponen prinsip hasil dari PCA matriks torso adalah sebagai berikut :

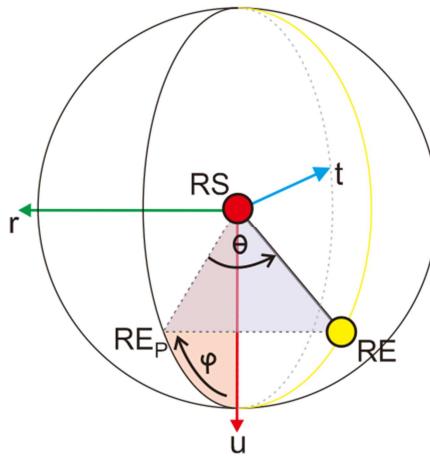
- 1) \vec{u} , vektor dengan arah dari atas tubuh ke bawah tubuh. Pada banyak tarian, torso penari tidak akan berdiri dari atas ke bawah relatif terhadap sensor.
- 2) \vec{r} , vektor dengan arah dari kanan tubuh ke kiri tubuh, (hampir) sejajar dengan vector dari joint right shoulder ke joint left shoulder.
- 3) \vec{t} , merupakan cross product dari dua komponen prinsip, $\vec{t} = \vec{u} \times \vec{r}$. Hasil $\{\vec{u}, \vec{r}, \vec{t}\}$ disebut kerangka PCA torso.



Gambar IV.7 Kerangka PCA torso, kerangka referensi dan kerangka Kinect.

Joint Berderajat Satu

Joint yang berdekatan langsung dengan torso disebut sebagai joint berderajat satu, terdiri dari joint-joint berikut ini : elbow right, elbow left, knee right, knee left, dan head. Joint berderajat satu ditandai dengan warna kuning pada Gambar IV.6.



Gambar IV.8 Sistem koordinat bola pada joint berderajat satu.

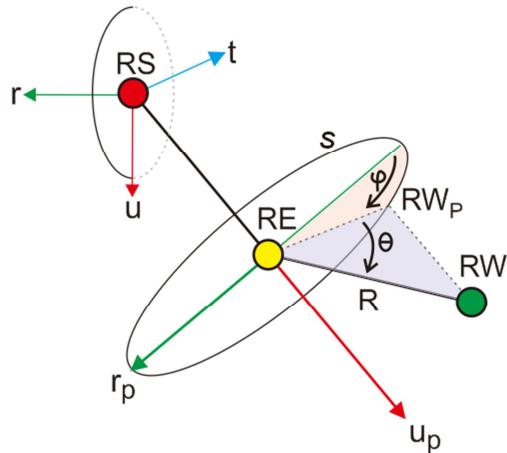
Joint berderajat satu direpresentasikan relatif terhadap torso frame. Jika dilihat pada Gambar IV.8, dimana RE (right elbow) direpresentasikan relatif terhadap RS (right shoulder). Kerangka torso $\{\vec{u}, \vec{r}, \vec{t}\}$ ditranslasikan ke RS dan membangun sistem koordinat bola sehingga titik origin terpusat di RS , sumbu azimuth adalah \vec{u} , sumbu zenith adalah \vec{r} . Sehingga, posisi RE dideskripsikan sebagai berikut :

1. Radius R adalah jarak RE ke titik origin RS . Karena panjang dari tulang lengan atas telah dinormalisasi dan konstan, radius R diabaikan.
2. Azimuth φ adalah sudut antara \vec{u} dengan $\overrightarrow{(RS, RE_P)}$ dimana RE_P adalah proyeksi RE ke bidang dengan vektor normalnya adalah \vec{r} .
3. Elevasi θ adalah sudut antara $\overrightarrow{(RS, RE_P)}$ dengan $\overrightarrow{(RS, RE)}$.

Dengan menggunakan model representasi ini, setiap joint berderajat satu direpresentasikan dengan dua sudut (θ, φ) . Pada joint RS , representasi sudutnya adalah $\{RS_\theta, RS_\varphi\}$.

Joint Berderajat Dua

Joint berderajat dua adalah joint-joint yang merupakan kepanjangan dari torso frame dan berdekatan dengan joint berderajat satu, yaitu wrist right, wrist left, ankle right, ankle left; pada Gambar IV.6 ditunjukkan dengan warna hijau.



Gambar IV.9 Sistem koordinat bola pada joint berderajat dua.

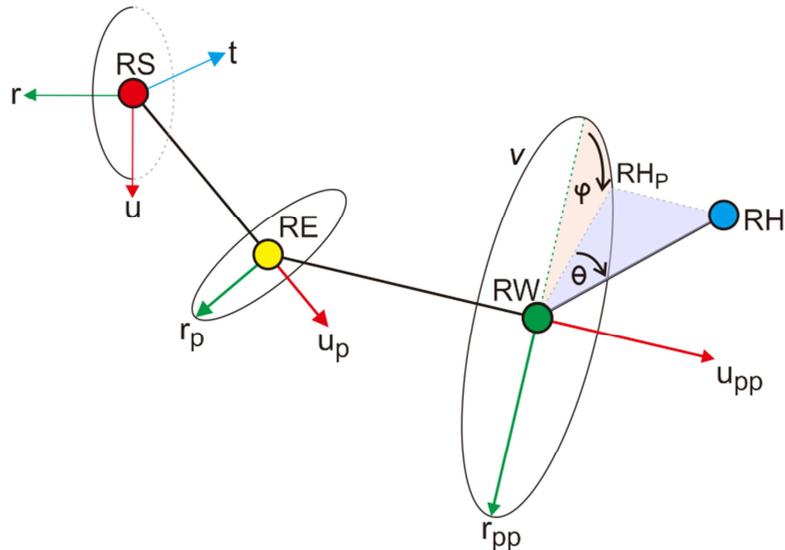
Diilustrasikan pada Gambar IV.9, vector $\vec{u_p}$ menonjol keluar dari RE namun merupakan kepanjangan dari vektor $\overrightarrow{(RS, RE)}$, kemudian dijadikan sebagai sumbu zenith pada sistem koordinat bola dengan titik origin RE . Frame $\{u_p, r_p, t_p\}$ merupakan hasil transformasi kerangka torso $\{\vec{u}, \vec{r}, \vec{t}\}$ yang dirotasi oleh sudut $\{RS_\theta, RS_\varphi\}$. Titik RW adalah notasi bagi joint right wrist. Sehingga posisi RW adalah sebagai berikut :

1. Radius R adalah jarak RW ke titik origin RE , dapat diabaikan.
2. Azimuth φ adalah sudut antara $\vec{r_p}$ dengan $\overrightarrow{(RE, RW_p)}$. Vektor $\vec{r_p}$ adalah hasil transformasi dari vektor \vec{r} ke bidang S dengan vektor normal $\vec{u_p}$, sedangkan RW_p adalah proyeksi RW ke bidang S .
3. Elevasi θ adalah sudut antara $\overrightarrow{(RE, RW_p)}$ dengan $\overrightarrow{(RE, RW)}$

Model ini merepresentasikan setiap joint berderajat dua menggunakan dua sudut (θ, φ) . Pada joint RE , representasi sudutnya adalah $\{RE_\theta, RE_\varphi\}$. Untuk joint knee atau lutut, hanya direpresentasikan dengan sudut elevasi θ saja.

Joint Berderajat Tiga

Joint berderajat tiga adalah joint-joint yang merupakan kepanjangan dari torso frame, dan berdekatan dengan joint berderajat dua, yaitu hand right, hand left, foot right, foot left; pada Gambar IV.6 ditunjukkan dengan warna biru.



Gambar IV.10 Sistem koordinat bola pada joint berderajat tiga.

Dilustrasikan pada Gambar IV.10, vector \vec{u}_{pp} menonjol keluar dari RW dan merupakan kepanjangan dari vektor $\overrightarrow{(RE, RW)}$. Sehingga vektor \vec{u}_{pp} dijadikan sebagai sumbu zenith pada sistem koordinat bola dengan titik origin RW . Titik RW adalah notasi bagi joint right wrist. Sehingga posisi RH adalah sebagai berikut :

1. Radius R adalah jarak RH ke titik origin RW , dapat diabaikan.
2. Azimuth φ adalah sudut antara \vec{r}_{pp} dengan $\overrightarrow{(RW, RH_p)}$. Vektor \vec{r}_{pp} adalah hasil transformasi dari vektor \vec{r} ke bidang V dengan vektor normal \vec{u}_{pp} , sedangkan RH_p adalah proyeksi RH ke bidang V .
3. Elevasi θ adalah sudut antara $\overrightarrow{(RW, RH_p)}$ dengan $\overrightarrow{(RW, RH)}$

Model ini merepresentasikan setiap joint berderajat tiga menggunakan dua sudut (θ, φ) . Pada joint RW , representasi sudutnya adalah $\{RW_\theta, RW_\varphi\}$.

Untuk memperoleh sudut (θ, φ) pada joint berderajat tiga dengan akurat, baik untuk joint-joint pada telapak tangan maupun joint-joint pada telapak kaki, dibutuhkan Wearable Inertial Measurement Units (WIMU) [14], untuk menambah akurasi nilai sudut pembacaan. Berdasarkan pengamatan pada Lampiran A, sensor Kinect XBOX360 tidak mampu mendeteksi orientasi joint berderajat tiga dengan akurat.

Representasi Pose Tubuh

Untuk lingkup pose tubuh yang melibatkan sampai joint berderajat dua,

1. Pose tubuh bagian atas dapat direpresentasikan dengan 8-tuple

$$P_{u,2} = (LE_\varphi, LE_\theta, LS_\varphi, LS_\theta, RS_\theta, RS_\varphi, RE_\theta, RE_\varphi)$$

2. Pose tubuh bagian bawah dapat direpresentasikan dengan 6-tuple

$$P_{l,2} = (LK_\theta, LH_\varphi, LH_\theta, RH_\theta, RH_\varphi, RK_\theta)$$

Untuk lingkup pose tubuh yang melibatkan sampai joint berderajat tiga,

1. Pose tubuh bagian atas dapat direpresentasikan dengan 12-tuple

$$P_{u,3} = (LW_\varphi, LW_\theta, LE_\varphi, LE_\theta, LS_\varphi, LS_\theta, RS_\theta, RS_\varphi, RE_\theta, RE_\varphi, RW_\theta, RW_\varphi)$$

2. Pose tubuh bagian bawah dapat direpresentasikan dengan 12-tuple

$$P_{l,3} = (LA_\varphi, LA_\theta, LK_\theta, LH_\varphi, LH_\theta, RH_\theta, RH_\varphi, RH_\varphi, RK_\theta, RA_\theta, RA_\varphi)$$

3. Pose kepala dapat direpresentasikan dengan 3-tuple

$$H = (H_\varphi, H_\theta, H_\phi)$$

Representasi sudut ini i.A dapat digunakan langsung untuk robot humanoid dengan model kinematika yang setara dengan NAO 24 DOF [15] atau Aaron 28 DOF [16]. Model kinematikanya terdapat pada Lampiran B. Untuk model HMM pada penelitian ini, digunakan $P_{u,2}$ sebagai representasi sebuah pose, 8-tuple representasi gerak tubuh bagian atas yang melibatkan sampai joint berderajat 2 saja.

Range Sudut pada Vektor Pose Berdasarkan Range of Movement

Fleksibilitas sebuah joint didefinisikan sebagai range of movement (ROM) yang mampu dilakukan oleh sebuah joint. Gerakan tipikal pada sebuah joint adalah gerakan angular dengan satuan derajat. ROM dari sebuah joint biasanya diukur dengan jumlah derajat dari posisi awal dari sebuah segmen sampai posisi akhir dari range gerakan penuh, dengan menggunakan geniometer. Berikut ini range of movement dari tubuh manusia untuk vektor pose $P_{u,2}$ yang diperoleh dari [9] dan [8].

Left Arm	Right Arm
$-60 \leq LS_\phi \leq 180$	$-60 \leq RS_\phi \leq 180$
$-75 \leq LS_\theta \leq 180$	$-180 \leq RS_\theta \leq 75$
$20 \leq LE_\phi \leq 180$	$0 \leq RE_\phi \leq 160$
$-60 \leq RE_\theta \leq 90$	$-60 \leq RE_\theta \leq 90$

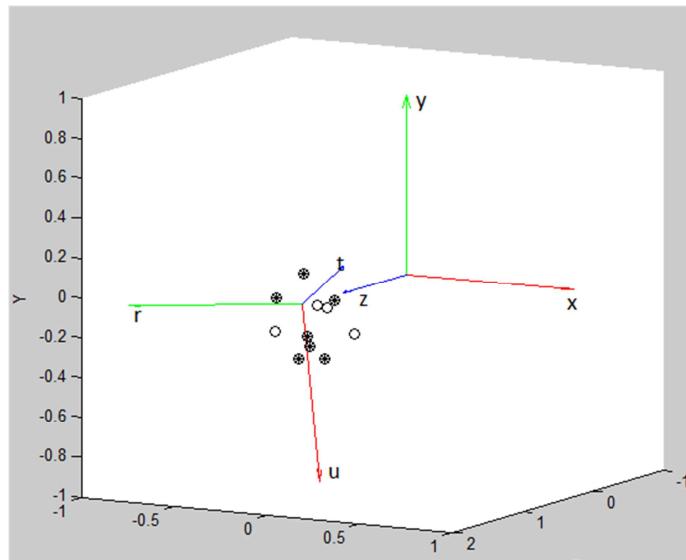
Tabel IV.3 Range of movement untuk 43ector pose tubuh bagian atas

Range of movement lengkap beserta sistem koordinat bola yang digunakan untuk merepresentasikannya, untuk setiap joint pada tubuh manusia yang terdeteksi oleh sensor Kinect, terlampir pada Lampiran C.

Pengujian Kerangka PCA Torso

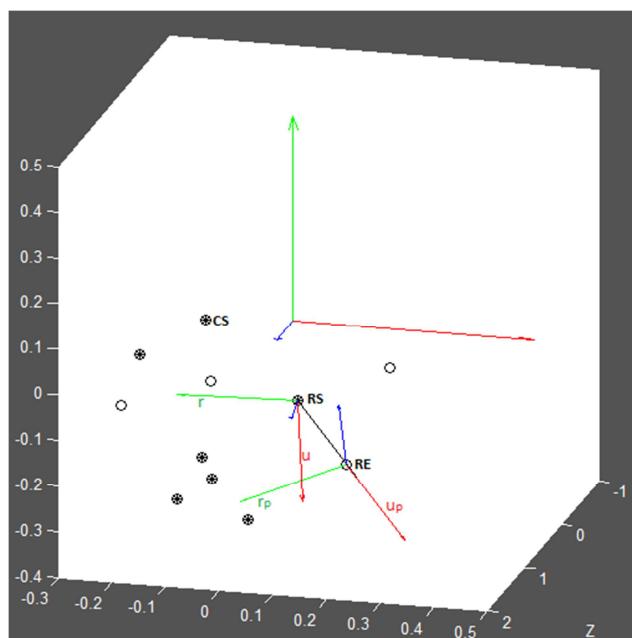
Hasil aplikasi PCA terhadap matriks torso dengan elemen-elemen joint sebagai berikut : neck, spine, center hip, right shoulder, left shoulder, right hip, dan left hip; dijadikan basis ortonormal 3D untuk tubuh user yaitu $\{\vec{u}, \vec{r}, \vec{t}\}$. Perhitungan dilakukan pada sebuah pose yang ditangkap oleh Kinect dengan menggunakan Matlab, dan hasilnya pada Gambar IV.11

Kerangka Kinect yang sebenarnya adalah $\{-\vec{x}, \vec{y}, \vec{z}\}$. Pada gambar IV.6, $\{\vec{x}, \vec{y}, \vec{z}\}$ adalah kerangka referensi yang digunakan, sedangkan $\{\vec{u}, \vec{r}, \vec{t}\}$ adalah kerangka PCA torso user. Joint berwarna buram adalah joint-joint pada matriks torso, sedangkan joint-joint berwarna putih adalah joint-joint elbow dan wrist.



Gambar IV.11 Hasil aplikasi PCA pada matriks torso.

Untuk pemrosesan pada joint berderajat dua pada tubuh bagian atas, kerangka PCA torso user $\{\vec{u}, \vec{r}, \vec{t}\}$ dirotasi oleh sudut-sudut shoulder $\{LS_\theta, LS_\varphi\}$ untuk joint *LE*. Kerangka PCA torso user $\{\vec{u}, \vec{r}, \vec{t}\}$ dirotasi oleh sudut-sudut shoulder $\{RS_\theta, RS_\varphi\}$ untuk joint *RE*, sehingga didapatkan $\{u_p, r_p, t_p\}$ dengan titik originnya adalah joint *RE*. Perhitungan dilakukan pada sebuah pose yang ditangkap oleh Kinect dengan menggunakan Matlab, dan hasilnya pada Gambar IV.12.



Gambar IV.12 Rotasi kerangka PCA torso untuk perhitungan joint elbow.

IV.5. Koleksi Data dan Segmentasi Data

Telah dikoleksi data dari tiga orang subjek yang menarikan frase tari Assalamualaikum pada Tari Likok Pulo secara berulang-ulang. Setelah disegmentasi, diperoleh total 2169 data gesture tunggal dengan rincian pada Tabel IV.4

Gesture	Only Kinect	With Clap Sensors	Total Data	Training Data	Test Data
O	387	0	387	310	77
A	394	71	465	372	93
B	415	92	507	406	101
C	488	127	615	492	123
D	415	131	546	437	109
E	386	113	499	400	99

Tabel IV.4 Hasil koleksi data untuk masing-masing gesture.

Data yang telah dikoleksi kemudian disegmentasi. Segmentasi antara gesture satu dengan yang lainnya dalam rangkaian gesture (frase), dilakukan dengan dua cara :

- 1) Mendeteksi tepukan. Tepukan tangan user dideteksi oleh sensor tepuk di dalam smartglove, kemudian informasi tepukan dikirim ke komputer melalui bluetooth module. Tepukan dijadikan pemisah antara satu gesture dengan gesture yang lain. Proses akuisisi data dengan cara ini tidak dibatasi waktu. (Tabel IV.4 kolom 3)

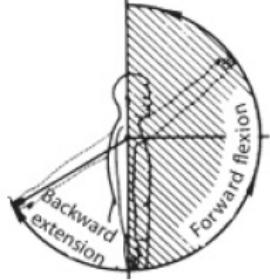
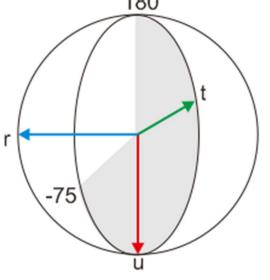
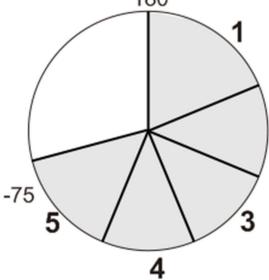
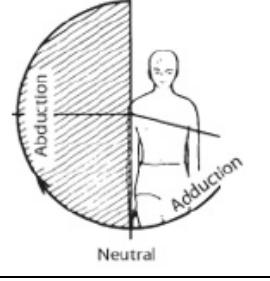
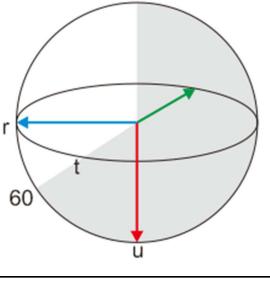
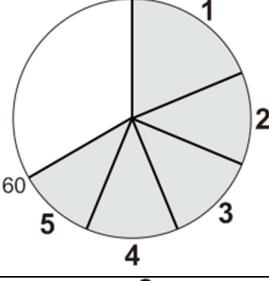
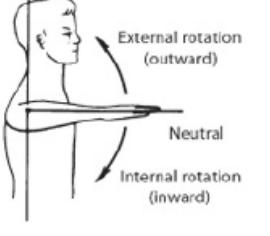
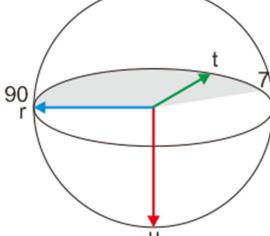
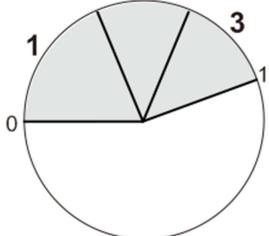
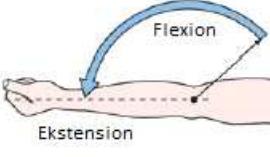
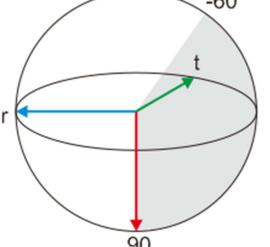
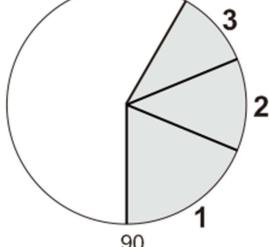


Gambar IV.13 Smartglove

- 2) Menggunakan jendela waktu. Proses dengan sensor tepuk di dalamnya akuisisi dan koleksi data dibatasi waktu, yaitu 1 detik untuk gesture O, A, B dan 2 detik untuk gesture C, D, E. Gesture satu dengan gesture yang lain dipisahkan oleh timeborder yang diidentifikasi melalui timestamps. (Tabel IV.4 kolom 2)

IV.6. Kuantisasi Vektor Pose Tari Berdasarkan Range of Movement

Data yang telah disegmentasi direpresentasikan dengan representasi skeleton menjadi vektor pose tari $P_{u,2} = (RS_\phi, RS_\theta, RE_\phi, RE_\theta, LS_\phi, LS_\theta, LE_\phi, LE_\theta)$. Kombinasi semua kemungkinan nilai-nilai elemen vektor pose $P_{u,2}$ akan menjadi tidak terbatas, sehingga vektor pose $P_{u,2}$ perlu dikuantisasi untuk menghasilkan simbol observasi diskrit untuk model HMM, seperti pada Tabel IV.5.

Element	Physical Representation	Range of Joint Angle	Directional Codewords
RS_ϕ			
RS_θ			
RE_ϕ			
RE_θ			

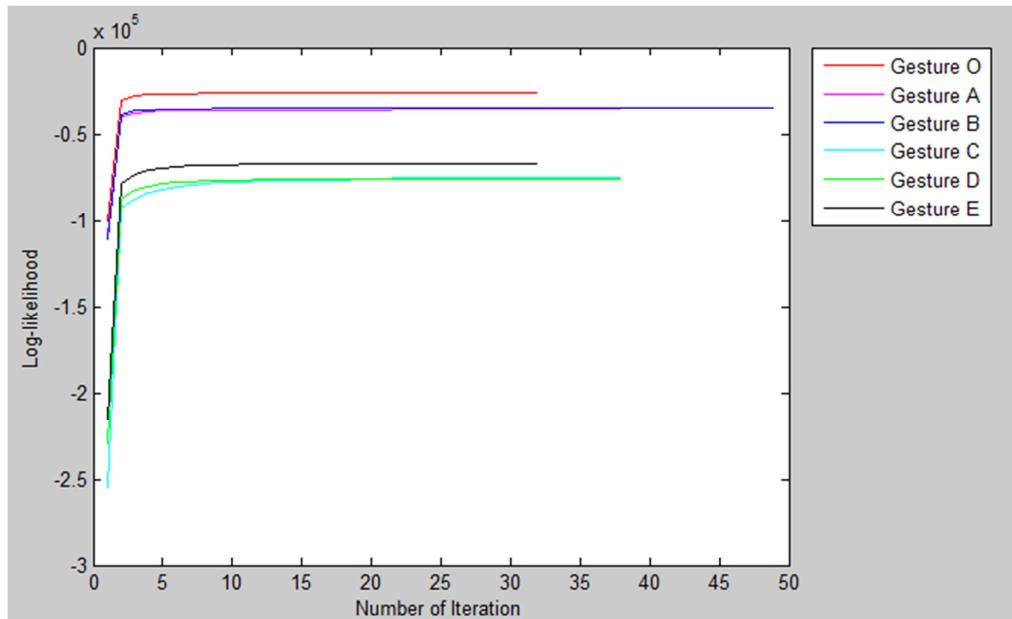
Tabel IV.5 Directional codewords berdasarkan range of movement.

Seluruh konfigurasi pose tari pada joint skeleton berderajat dua yang mungkin muncul pada satu lengan, ada 225 pose. Dan seluruh kombinasi konfigurasi pose tari pada joint skeleton berderajat dua yang mungkin muncul pada kedua lengan ada $225 \times 225 = 50625$ pose = 50625 simbol diskrit.

IV.7. Pembelajaran Gesture Tari

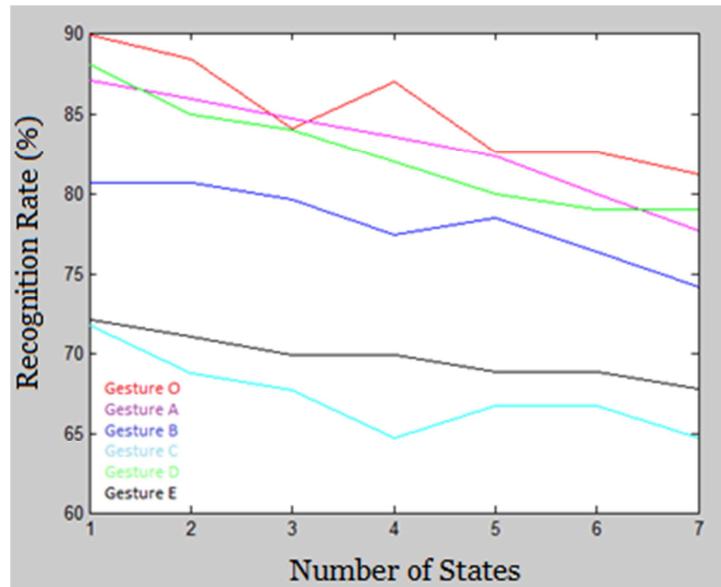
Pada tahap ini, parameter-parameter model HMM ‘dilatih’ dengan menggunakan data training dan algoritma Baum Welch. Digunakan data pelatihan sebesar 80% dari total koleksi data yang diperoleh untuk melatih setiap model HMM. Algoritma Baum Welch mengestimasi parameter-parameter model HMM secara iteratif.

Kurva pelatihan menggunakan algoritma Baum-Welch ditunjukkan pada Gambar IV.14. Kurva pelatihan secara umum mulai stabil setelah 25 iterasi, namun belum benar-benar mengumpul (converge) setelah siklus ke 70. Pelatihan dihentikan setelah 100 siklus.



Gambar IV.14 Kurva pelatihan gesture tari dengan HMM.

Gesture O, A, dan B, dimodelkan dengan 3 buah state. Gesture C, D, E dimodelkan dengan 5 buah state. Penambahan jumlah state yang digunakan berbanding terbalik dengan tingkat akurasi yang diperoleh, ditunjukkan pada Gambar IV.15



Gambar IV.15 Penurunan tingkat akurasi dan peningkatan jumlah state.

IV.8. Klasifikasi Gesture Tari

Model HMM yang telah terlatih digunakan untuk mengklasifikasi data pengujian yang berbeda dengan data pelatihan. Digunakan data pelatihan sebesar 20% dari total koleksi data yang diperoleh untuk menguji setiap model HMM.

Klasifikasi dilakukan dengan menggunakan nilai score $P(O|\lambda)$ untuk menilai tingkat kecocokan antara gesture masukan dengan model gesture. Perhitungan score dilakukan dengan forward probabilities :

$$Scoring = P(O|\lambda) = \sum_{i=1}^N \alpha_T(i)$$

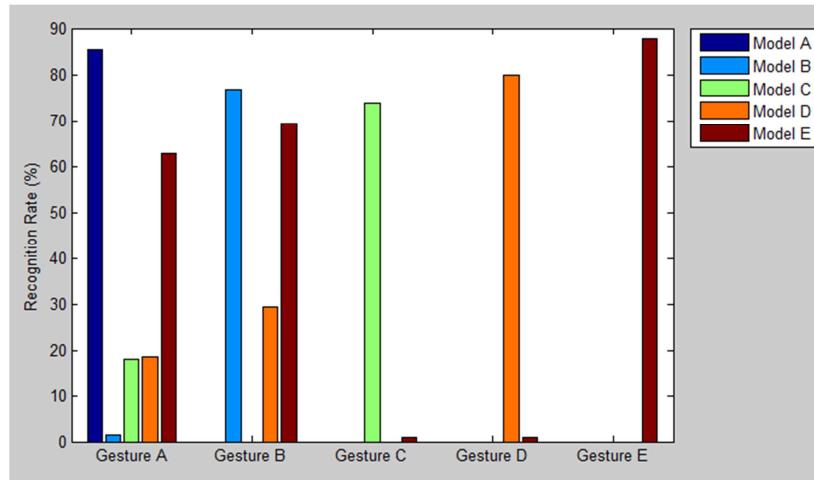
$\alpha_t(i)$ = Forward probabilities.

Pengujian Pertama

Untuk melihat efek dari representasi skeleton, setiap satu model yang telah terlatih diuji dengan seluruh pola gesture dari koleksi data. Representasi skeleton tidak digunakan pada pengujian ini, hasilnya ditunjukkan pada Tabel IV.6 dan Gambar IV.16

Model	Keterdeteksian Model oleh Pola Gesture (%)				
	A	B	C	D	E
A	85.3	1.33	18	18.67	62.67
B	0.00	76.84	0.00	29.47	69.47
C	0.00	0.00	73.91	0.00	0.87
D	0.00	0.00	0.00	79.84	0.88
E	0.00	0.00	0.00	0.00	87.78

Tabel IV.6 Pengujian tanpa menggunakan representasi skeleton



Gambar IV.16 Hasil pengujian sebelum menggunakan representasi skeleton

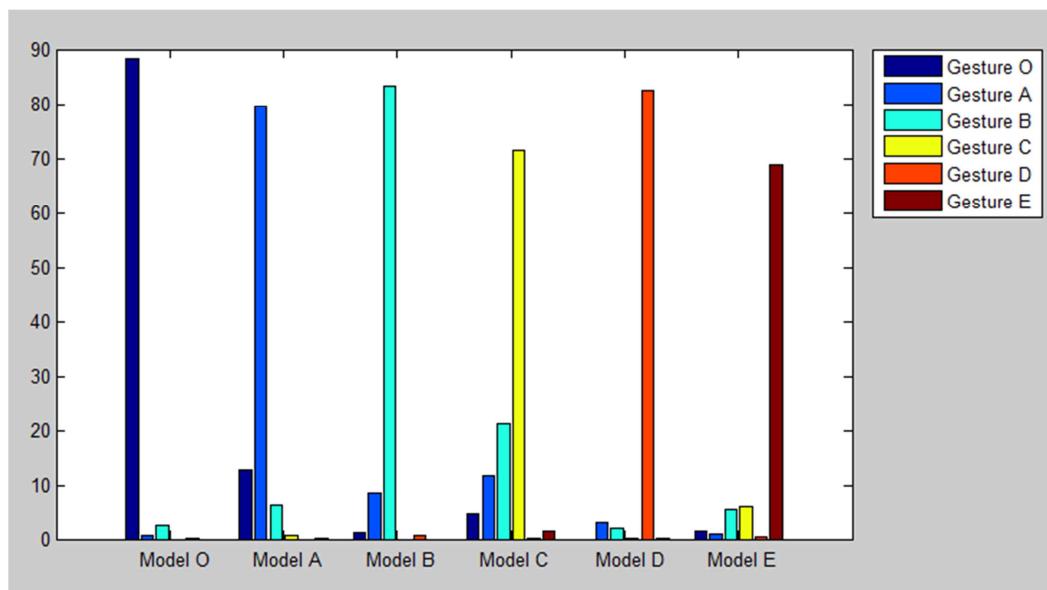
Pengujian Kedua

Digunakan representasi skeleton. Setiap satu model yang telah terlatih diuji dengan seluruh pola gesture dari koleksi data (kecuali 80% data yang digunakan untuk melatih model tersebut). Threshold dari setiap model adalah minimum dari score hasil pengujian data pelatihan. Hasil pengujian ditunjukkan pada Tabel IV.7,Gambar IV.17.

Model	Keterdeteksian Model oleh Pola Gesture (%)					
	O	A	B	C	D	E
O	88.31	0.65	2.56	0.00	0.18	0.00
A	12.66	79.57	6.31	0.65	0.00	0.20
B	1.29	8.60	83.17	0.00	0.55	0.00
C	4.65	11.61	21.10	71.54	0.18	1.40
D	0.00	3.01	1.97	0.16	82.57	0.20
E	1.55	0.86	5.52	6.02	0.37	68.69

Tabel IV.7. Hasil pengujian setiap satu model oleh seluruh pola gesture.

Satu gesture terdeteksi oleh satu model jika gesture tersebut memiliki nilai kecocokan yang lebih tinggi dari threshold yang ditentukan untuk model tersebut. Nilai threshold ditentukan berbeda-beda untuk masing-masing model gesture. Namun demikian satu model dapat mengenal beberapa gesture atau satu gesture dapat diklasifikasikan ke dalam beberapa model. Untuk menangani hal ini, dilakukan pengujian kedua.



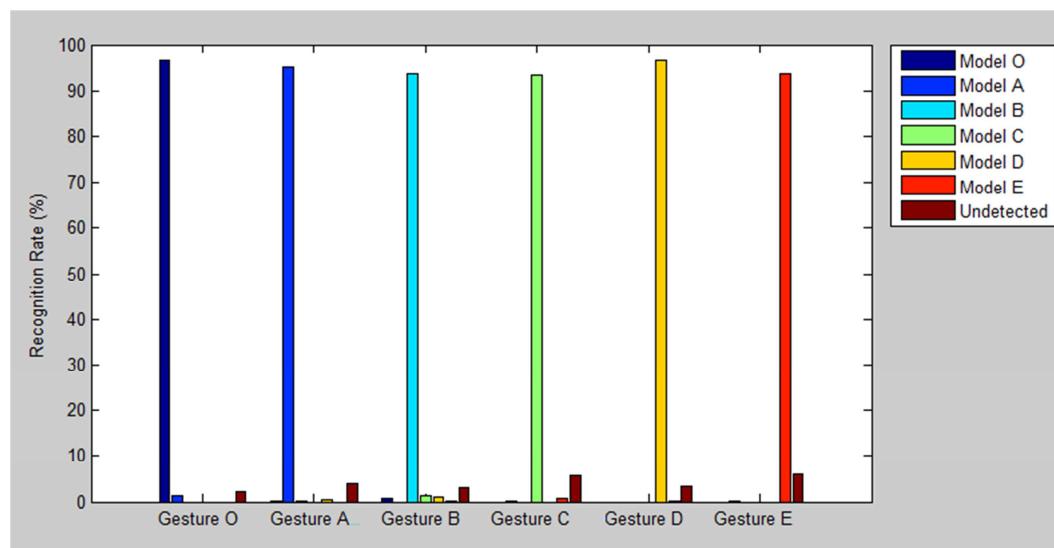
Gambar IV.17 Diagram batang hasil pengujian setiap satu model.

Pengujian Ketiga

Digunakan representasi skeleton. Setiap satu masukan gesture X diuji oleh seluruh model untuk menemukan satu model Y yang memiliki nilai kecocokan tertinggi (maximum log-likelihood). Jika nilai maksimum log-likelihood yang diperoleh lebih tinggi dari threshold yang ditentukan untuk model Y, maka gesture X terklasifikasi sebagai gesture Y. Jika log-likelihoodnya minus infinity (tidak memiliki kemiripan sedikitpun) saat diuji oleh semua model, maka gesture tersebut tidak terdeteksi oleh model manapun. Threshold dari setiap model adalah minimum dari score hasil pengujian data pelatihan. Hasil pengujian ditunjukkan pada Tabel IV.8, Gambar IV.18, dan Lampiran D.

Gesture	Keterdeteksian Gesture oleh Model (%)						Tidak terdeteksi
	O	A	B	C	D	E	
O	96.64	1.29	0.00	0.00	0.00	0.00	2.07
A	0.22	95.27	0.22	0.00	0.43	0.00	3.87
B	0.79	0.00	93.69	1.18	0.99	0.20	3.16
C	0.00	0.16	0.00	93.33	0.00	0.81	5.69
D	0.00	0.00	0.00	0.00	96.52	0.18	3.30
E	0.00	0.20	0.00	0.00	0.00	93.79	6.01

Tabel IV.8 Hasil pengujian setiap pola gesture oleh seluruh model HMM.



Gambar IV.18 Diagram batang hasil pengujian pola gesture.

Gesture	Detected as True (%)	Detected as False Positive (%)	Detected as False Negative (%)
O	96.64	1.01	1.29
A	95.27	1.49	0.87
B	93.69	0.22	3.16
C	93.33	1.18	0.97
D	96.52	1.33	0.18
E	93.79	1.19	0.20

Tabel IV.9 Hasil pengujian setiap pola gesture oleh seluruh model HMM (2)

	Sebelum	Sesudah
Data Dimension	33-tuple	8-tuple
Detected as False	10 %	0.22 %
Detected as True	80.7 %	94.87 %

Tabel IV.10 Efek penggunaan representasi skeleton dan maximum likelihood

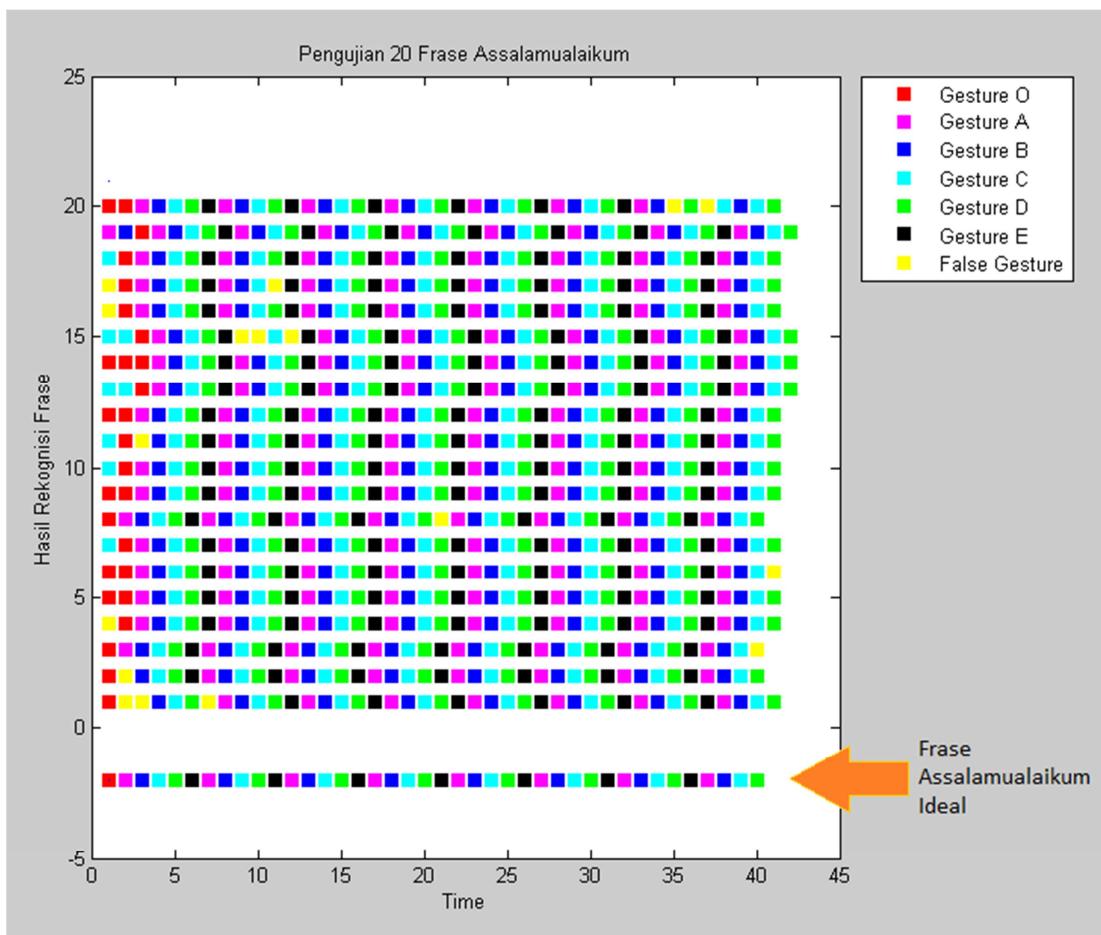
IV.9. Rekognisi Frase Tari Berbasis Hidden Markov Model

Dikoleksi 20 data rekaman sensor Kinect untuk frase ‘Assalamualaikum’ lengkap. Setiap gesture pada frase Assalamualaikum disegmentasi menggunakan jendela waktu, kemudian diuji dengan oleh setiap model HMM yang terlatih. Hasil pengujian 20 buah frase ‘Assalamualaikum’ ditunjukkan pada Tabel IV.9 dan Gambar V.18.

Phrase	True Gesture	False Gesture	Recognition Rate (%)	Is Phrase Detected
1	37	3	92.50	No
2	39	1	97.50	No
3	39	1	97.50	No
4	40	0	100.00	Yes
5	40	0	100.00	Yes
6	39	1	97.50	No
7	40	0	100.00	Yes
8	39	1	97.50	No
9	40	0	100.00	Yes
10	40	0	100.00	Yes
11	39	1	97.50	No
12	40	0	100.00	Yes
13	40	0	100.00	Yes
14	40	0	100.00	Yes
15	37	3	92.50	No
16	40	0	100.00	Yes

17	39	1	97.50	No
18	40	0	100.00	Yes
19	40	0	100.00	Yes
20	37	3	92.50	No
Mean			98.125 %	55 %

Tabel IV.11 Hasil pengujian 20 buah frase ‘Assalamualaikum’

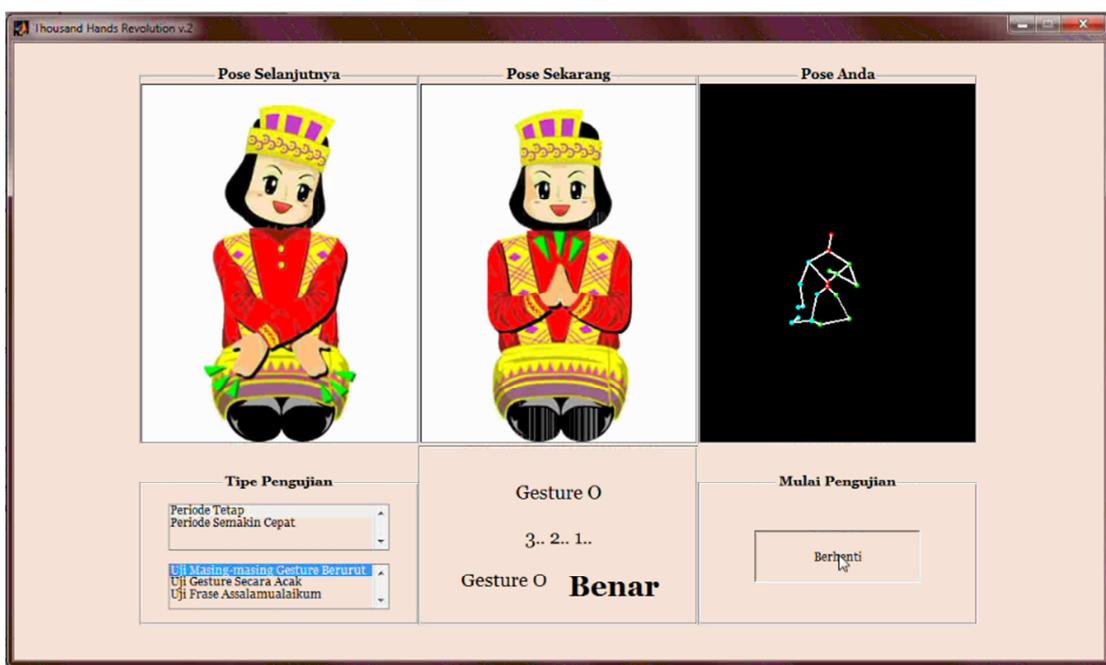


Gambar IV.19 Hasil pengujian 20 buah frase Assalamualaikum.

IV.10. Implementasi Realtime



Gambar IV.20 Implementasi Realtime : gesture O tidak terdeteksi.



Gambar IV.21 Implementasi Realtime : gesture O terdeteksi.

BAB V KESIMPULAN, SARAN DAN TINDAK LANJUT

V.1. Kesimpulan

Keseluruhan tari Likok Pulo dimodelkan dengan $\mathcal{L} = (\mathbf{S}, \mathbf{I}, \mathbf{P}, \mathbf{O}, \mathbf{f}, \mathbf{e}, \mathbf{s}_0, \mathbf{S}_t)$. Gesture tari berperan sebagai state diskrit tersembunyi dan rangkaian antar state membentuk frase tari.

Representasi skeleton $P_{u,2} = (RS_\varphi, RS_\theta, RE_\varphi, RE_\theta, LS_\varphi, LS_\theta, LE_\varphi, LE_\theta)$ yang dikuantisasi berdasarkan range of movement cukup efektif untuk menangani input sensor Kinect berderau, mengurangi dimensi data, dan menangani perubahan posisi dan orientasi penari relatif terhadap sensor Kinect.

Enam buah gesture tari pada frase “Assalamualaikum” dipelajari dan diklasifikasi menggunakan hidden Markov model pada environment Matlab dan Simulink. Sistem mengklasifikasi data gesture menjadi salah satu dari enam kelas gesture terdefinisi atau satu kelas gesture tidak terdefinisi. Sistem klasifikasi memiliki tingkat akurasi 94.87% untuk gesture tunggal dan 55% untuk frase yang terdiri dari 40 gestures

V.2. Saran dan Tindak Lanjut

Observasi tarian dapat diperluas sampai tubuh bagian bawah, dan/atau diperluas sampai joint berderajat tiga. Diperlukan sensor inersial tambahan untuk observasi joint berderajat tiga karena sensor Kinect memiliki kelemahan dalam mendeteksi posisi dan orientasi joint berderajat tiga (telapak tangan dan telapak kaki).

Representasi skeleton dapat diperluas untuk juga mempertimbangkan aspek-aspek dinamis pada kerangka tubuh manusia untuk membangun robot humanoid penari dinamis yang mampu bangkit dari kejatuhan. Implementasi sistem dapat diperluas untuk pelatihan tari robot humanoid, atau video game untuk seluruh tari Likok Pulo atau tari tradisional Indonesia lainnya.

Pada pengujian frase, data pengujian dan data pelatihan ada kemungkinan tercampur. Sebaiknya data pelatihan dan data pengujian tidak tercampur.

Untuk gerakan-gerakan tari yang lebih luas, diperlukan segmentasi antar-gesture yang tidak tergantung pada wearable sensor. Metode segmentasi dapat menggunakan metode gesture spotting [16] dan algoritma Viterbi.

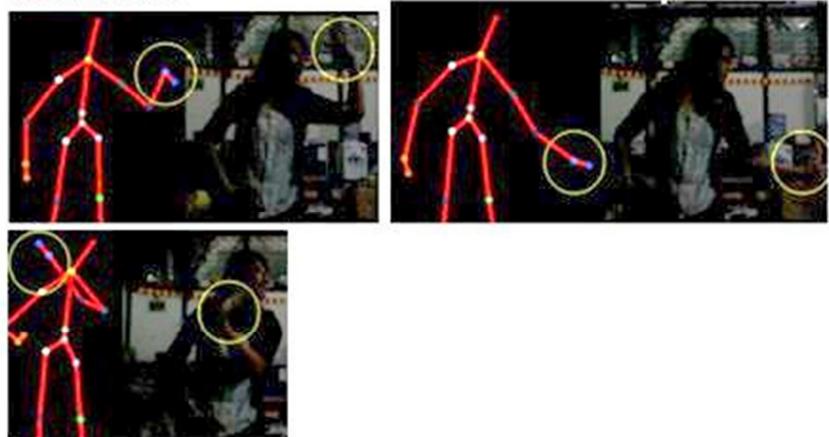
DAFTAR PUSTAKA

- [1] R. Rabiner, Lawrence. *A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition*. Proceedings of the IEEE, vol 77 no.2, Februari 1989.
- [2] Kenneth H. Rosen. *Discrete Mathematics and Its Application, Sixth Edition*. Mcgraw Hill. International Edition 2007
- [3] Nando de Freitas. *Undergraduate Machine Learning*. Youtube Video. UBC Course. Available :
http://www.youtube.com/watch?v=pid0IUH467o&list=PLE6Wd9FR--Ecf_5nCbnSQMHqORpiChfJf
- [4] Amy LaViers and Magnus Egerstedt. *The Ballet Automaton : A Formal Model for Human Motion*. 2011
- [5] Amy LaViers, Yushan Chen, Calin Belta, and Magnus Egerstedt. *Automatic Generation of Ballet Phrases*. 2011
- [6] Yaya Heryadi, Mohammad Ivan Fanany, Aniati Murni Arymurthy. *A Syntactical Modeling and Classification for Performance Evaluation of Bali Traditional Dance*. ICACSIS 2012.
- [7] Michalis Raptis, Darko Kirovski, Hugues Hoppe. *Real-Time Classification of Dance Gestures from Skeleton Animation*. ACM SIGGRAPH Symposium on Computer Animation, 2011.
- [8] Alan Graham Apley. *Apley's Sistem of Orthopaedics and Fractures, Ninth edition*. Hodder Arnold, 2010.
- [9] Warren Jahn, DC. *Musculoskeletal examination – range of motion*. The Journal of the CCA/Volume 23 No.2/ June 1979
- [10] J.Huang, Chun-wei Lee, Junji Ma. *Gesture Recognition and Classification using the Microsoft Kinect*. CS229 Final Project, 2012.

- [11] Aymeric Masurelle, Slim Essid, Gael Richard. *Multimodal Classification of Dance Movements using Body Joint Trajectories and Step Sounds.*
- [12] Hao Zhang, WenXiao Du, Haoran Li. *Kinect Gesture Recognition for Interactive System.* CS229 Final Project
- [13] J.C. Hall. *How to Do Gesture Recognition with Kinect Using Hidden Markov Models (HMMs).* [online].
Available : <http://www.creativedistraction.com/demos/gesture-recognition-kinect-with-hidden-markov-models-hmms/>
- [14] Marc Gowing, Cyril Concolato, Ebroul Izquierdo. *Enhanced Visualisation of Dance Performance from Automatically Synchronised Multimodal Recordings.* 2011.
- [15] Nikos Kofinas, Emmanouil Orfanoudakis, Michail G Lagoudakis. *Complete Analytical Inverse Kinematics for NAO.* IEEE Paper 2013.
- [16] H. K. Lee, J. H. Kim. *An HMM-Based Threshold Model Approach for Gesture Recognition.* IEEE Transactions on Pattern Analysis and machine Intelligence, Vol 21. No. 10. October 1999.
- [17] Lindsay I Smith. *A Tutorial on Principal Components Analysis.* 2002.

LAMPIRAN A

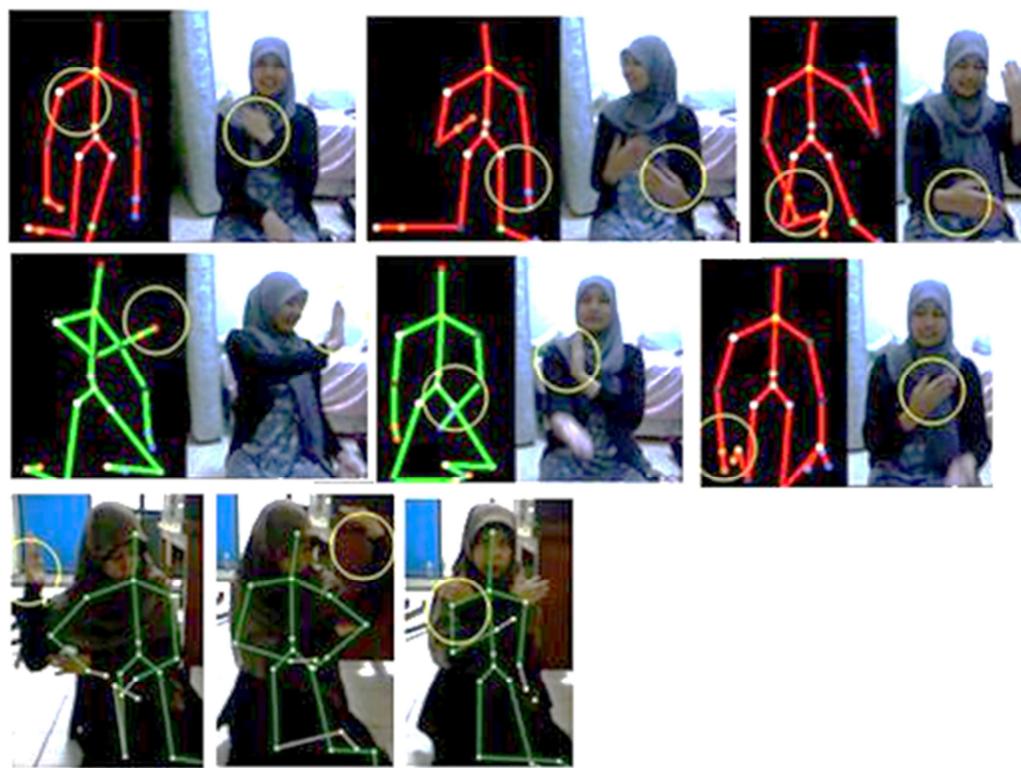
Gambyong Dance



Mamayang Dance

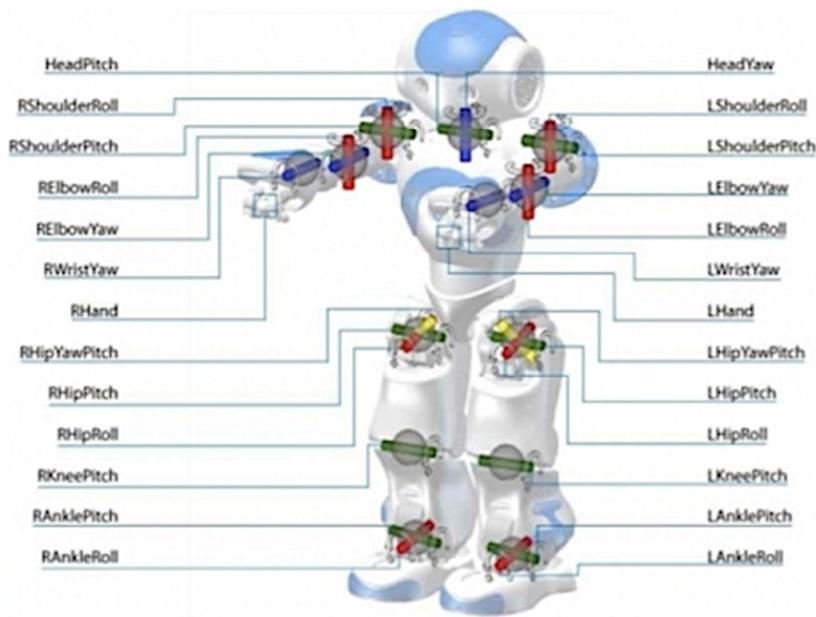


Saman Dance

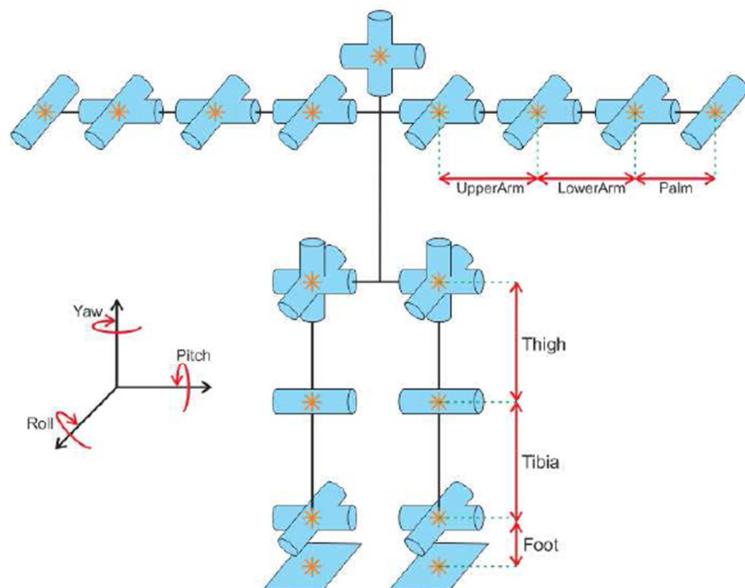


LAMPIRAN B

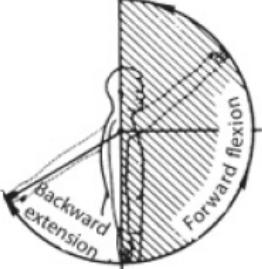
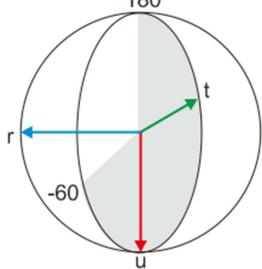
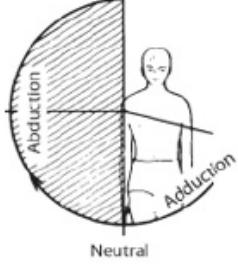
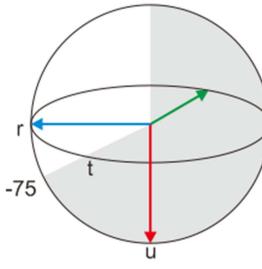
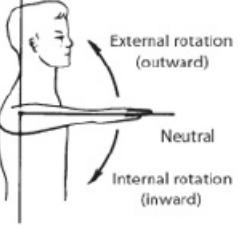
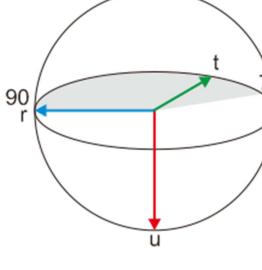
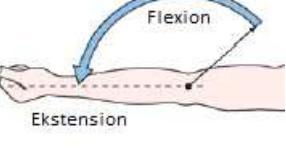
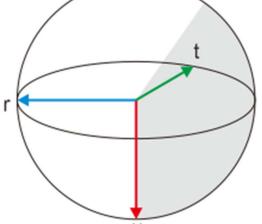
Model Kinematika Robot NAO 24 DOF



Model Kinematika Robot AARON 28 DOF

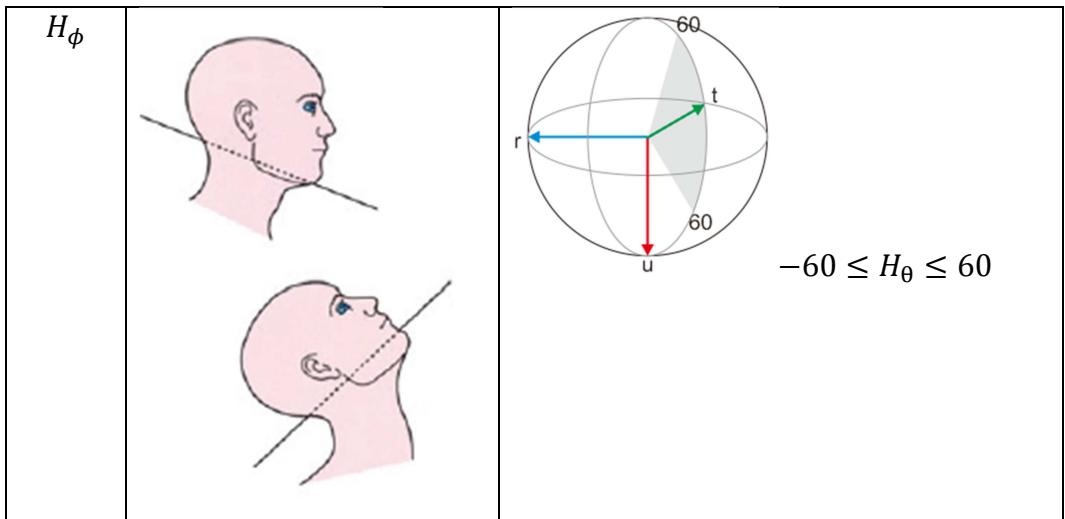


LAMPIRAN C

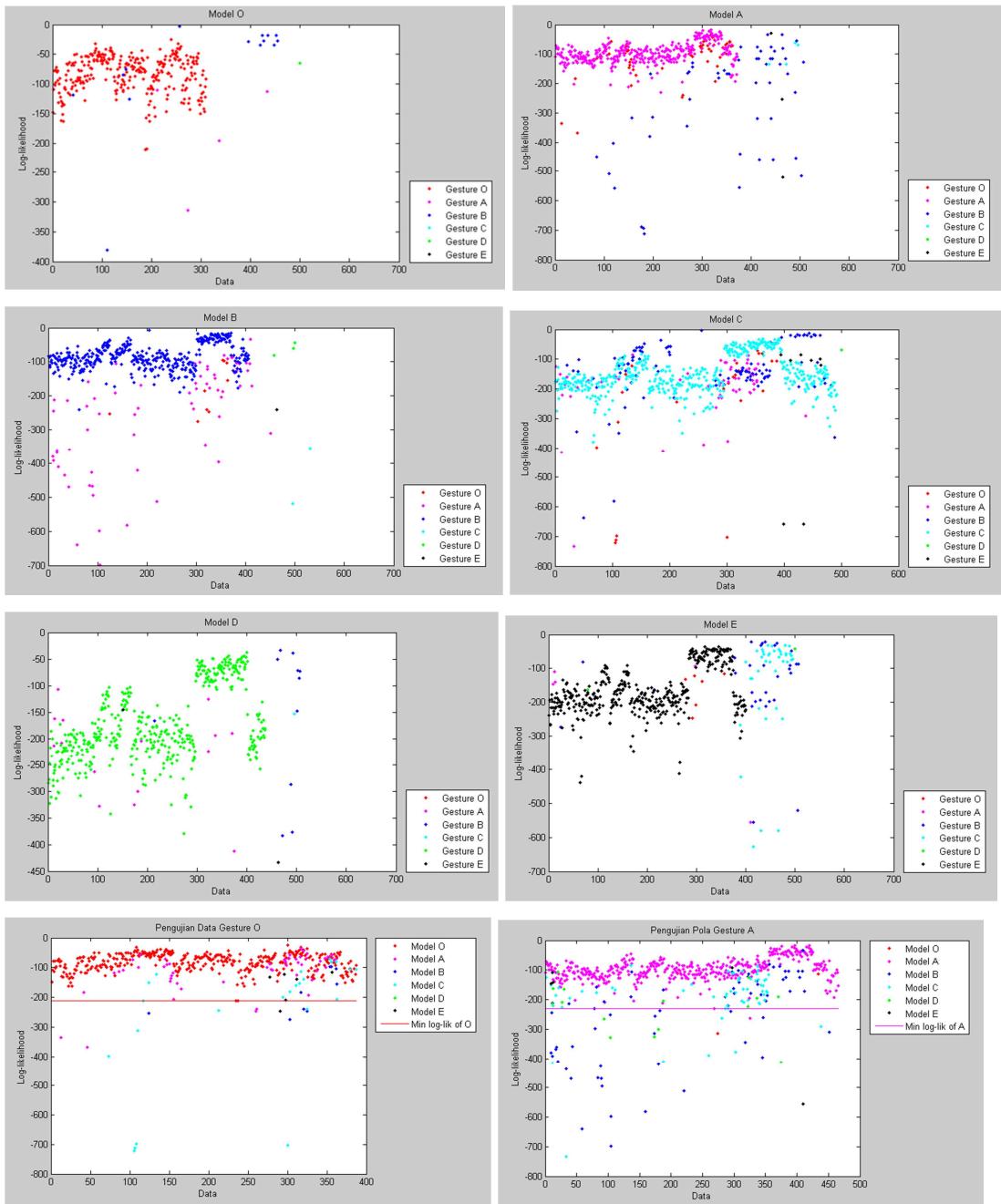
Sudut	Representasi Fisik	Range Sudut
RS_φ		 $-60 \leq RS_\varphi \leq 180$
RS_θ		 $-75 \leq RS_\theta \leq 180$
RE_φ		 $0 \leq RE_\varphi \leq 160$
RE_θ		 $0 \leq RE_\theta \leq 150$

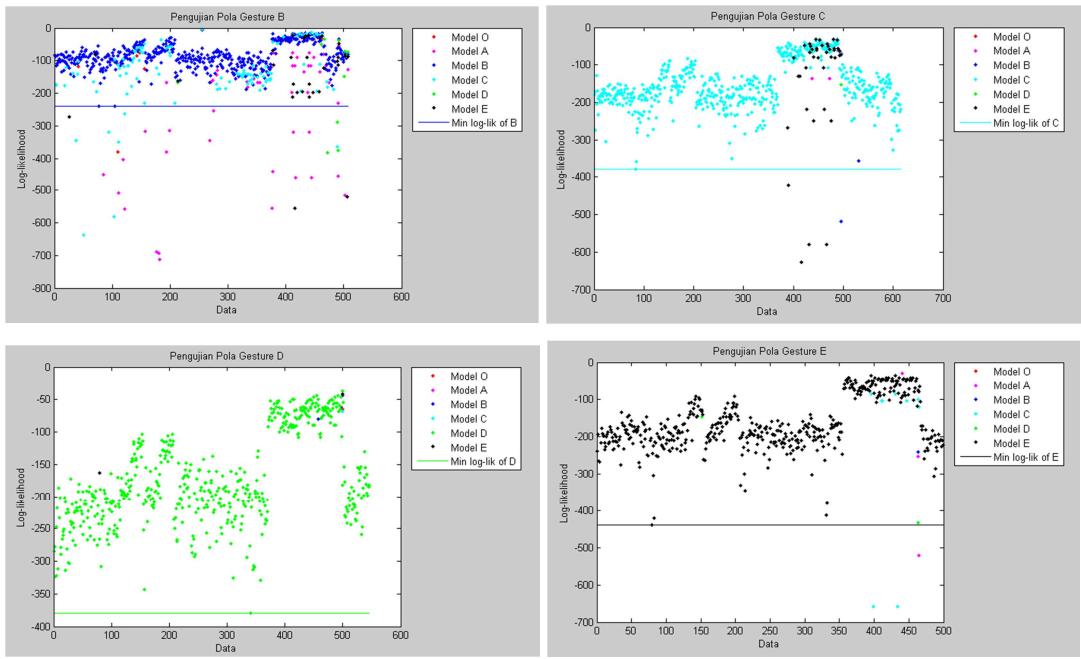
RW_φ			$-30 \leq RW_\varphi \leq 20$
RW_θ			$-80 \leq RW_\theta \leq 70$
RH_φ			$-35 \leq RH_\varphi \leq 45$
RH_θ			$-30 \leq RH_\theta \leq 120$
RH_ϕ			$-30 \leq RH_\phi \leq 50$

RK_θ		<p>$0 \leq RK_\theta \leq 135$</p>
RA_θ		<p>$-50 \leq RA_\theta \leq 20$</p>
RA_ϕ		<p>$-35 \leq RA_\phi \leq 15$</p>
H_φ		<p>$-80 \leq H_\varphi \leq 80$</p>
H_θ		<p>$-45 \leq H_\theta \leq 45$</p>



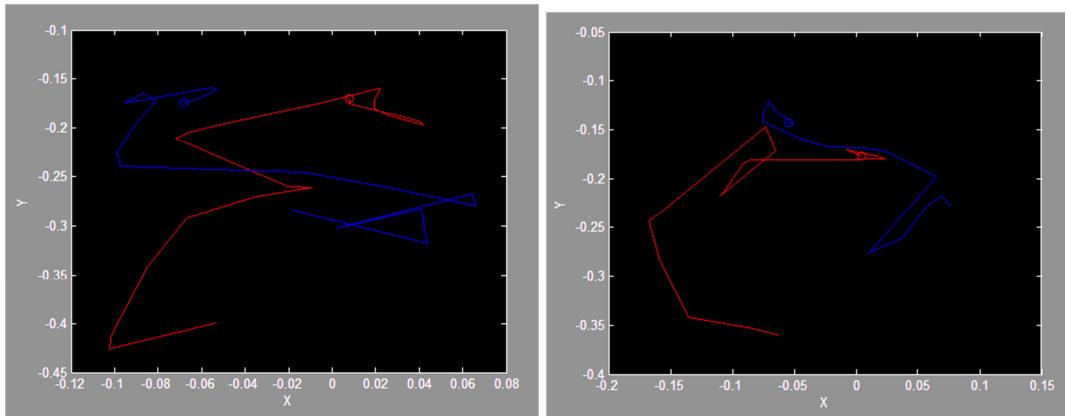
LAMPIRAN D



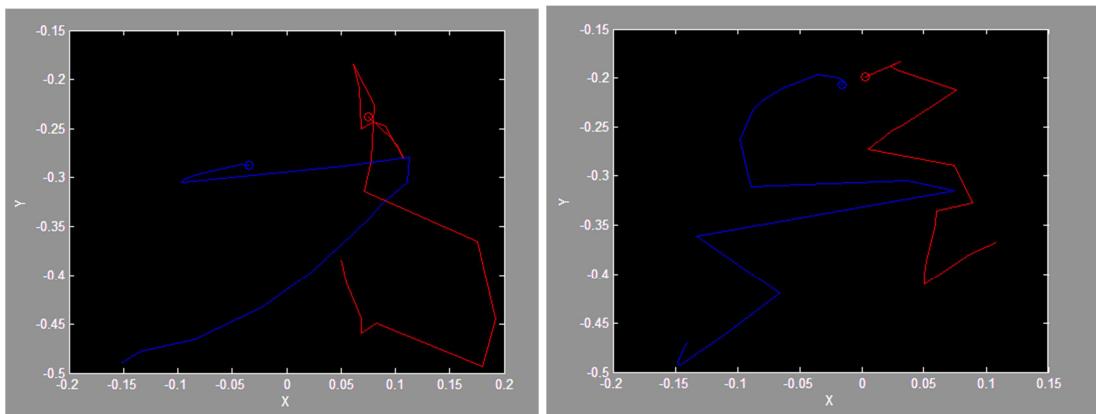


LAMPIRAN E

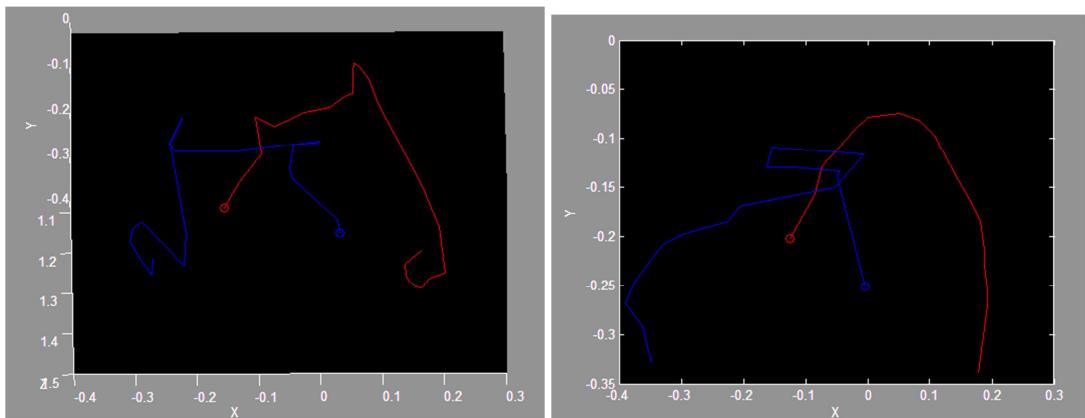
Trajektori Gesture O -- benar



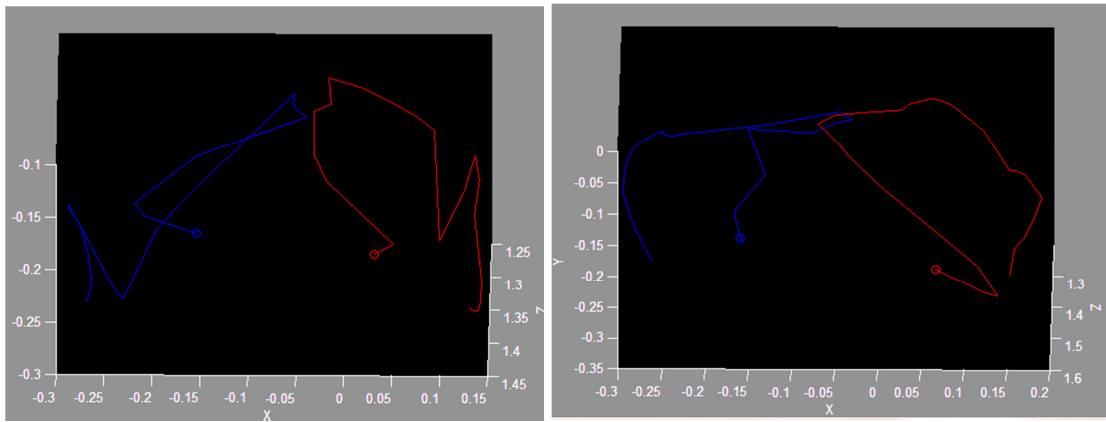
Trajektori Gesture O -- salah



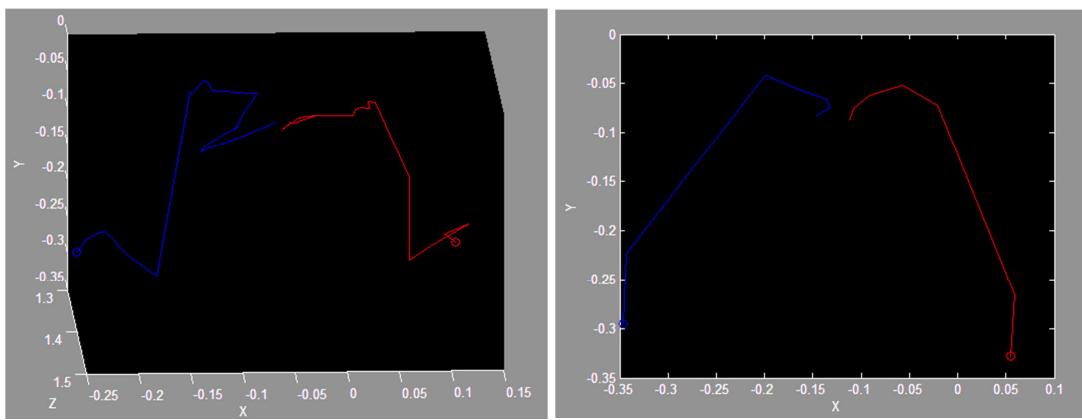
Trajektori Gesture A -- benar



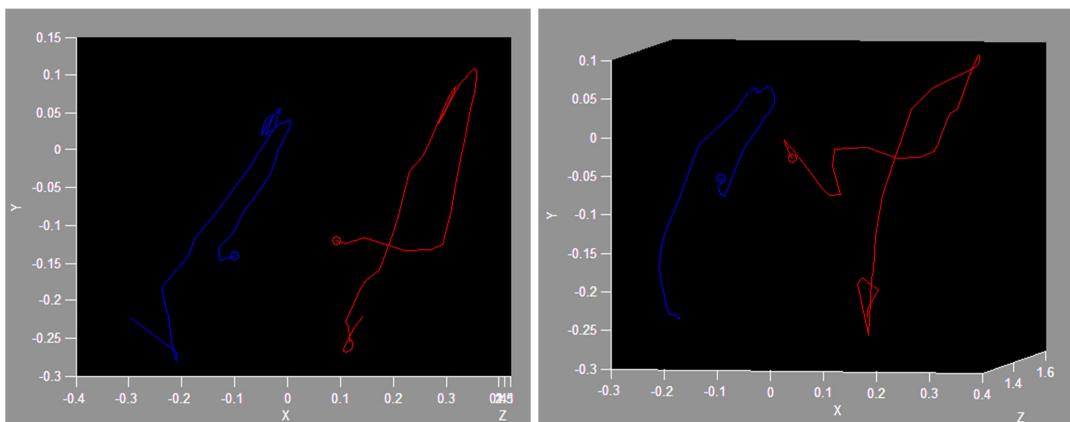
Trajektori Gesture A -- salah



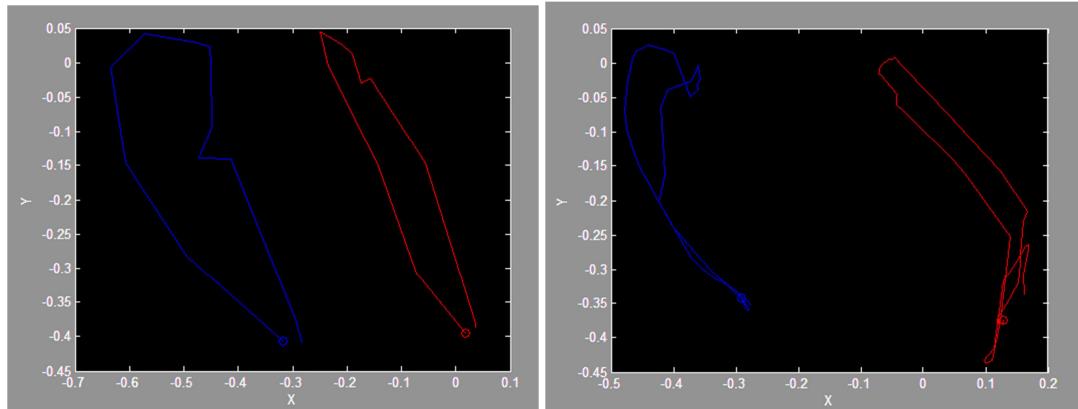
Trajektori Gesture B -- benar



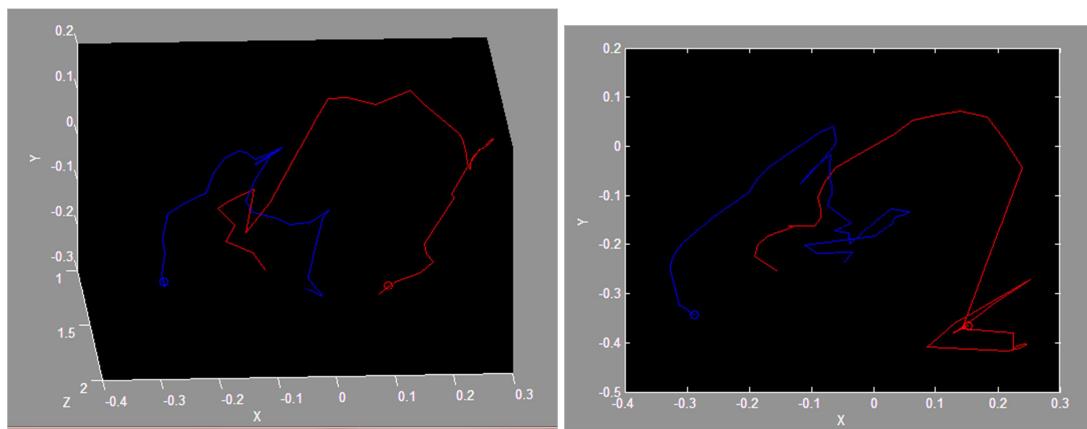
Trajektori Gesture C – benar



Trajektori Gesture D – benar



Trajektori Gesture E – benar



Trajektori Gesture E – salah

