International Conference on Information Technology Systems and Innovation (ICITSI) 2014
Bandung-Bali, 24-27 November 2014
ISBN: 978-1-4799-6526-7

Dance Modelling, Learning and Recognition System of Aceh Traditional Dance based on Hidden Markov Model

Nurfitri Anbarsanti[1], Ary S. Prihatmanto[2]
School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung, Indonesia
E-mail: [1]anbarsanti@yahoo.com, [2]asetijadi@lskk.ee.itb.ac.id

***Abstract*—The whole dance of Likok Pulo are modeled by hidden markov model. Dance gestures 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 angular skeleton representation which has been quantified based on range of movement. One unique gesture of dance is defined by sequence of pose and 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.**

***Index Terms*—angular skeletal representation, Kinect sensor, dance modelling, dance recognition, gesture recognition, hidden markov model, Likok Pulo dance.**

I. INTRODUCTION

Culture is the identity of a nation. Globalization that hit today's youth, lead the youth generation lose their identity and the value of good wisdom Indonesian culture. If the Indonesian culture is not well maintained, it will disappear in the next few decades. 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.

Dance is sequence of expressive human body movement and has aesthetic values. Famous Indonesian traditional dance which is known as unique and attractive is Aceh traditional dance such as Likok Pulo. In this study, we choose Likok Pulo dance to communicate with the computer because Likok Pulo dance from Aceh requires synchronous motion among the group of dancers with lined up formation, precision timing of gestures with the rhythm of its music that linearly changed more rapidly. It has several gestures performed repeatedly; and it well accepted in the international environment but still has its strong and decent identity.

Kinect is the Natural User Interface that combines stereoscopic camera and infrared sensors, so it can capture the depth map at a rate about 30 frames per second to estimate the position of the 20 points on the user's skeleton joints. The human body movement such as dance, can be interpreted as a command to communicate with the computer.

Skeleton joint trajectories captured by Kinect sensor is very likely to experience a discontinuity, noise, or instable parameter [5]. Dance movements that involve a lot of body articulation will result in a very large input dimension for the signal trajectories processing systems. So it is necessary to build a representation to reduce the signal entropy and the dimension of data. It must also deal with changes in the dancer's position and orientation relative to the Kinect sensor.

Dance can be defined as sequences of several finite distinct gestures. Gesture transition are not deterministic but probabilistic, due to the unideal dance in real world. Gesture has two aspects of signal characteristics : spatio-temporal variability and segmentation ambiguity [12]. Major approaches for analyzing spatial and temporal patterns include Dynamic Time Warping (DTW), Neural Networks (NNs), dan Hidden Markov Model (HMM) [12]. In this study, HMM-based approach is chosen to model the dance gesture of Likok Pulo dance, because it can be applied to analyzing time-series with spatio-temporal varibilities and can handle undefined patterns [12]. HMM-based dance gesture modelling make us enable to build practical systems that has ability to learn, predict, and classify dance gestures of Likok Pulo Dance from Aceh.

II. BACKGROUND

A. Related Works

Dance choreography has been captured using various formalization approaches, e.g., Laban notation which is initiated in the early 20 th century. Amy Laviers modelled the motion patterns of ballet as a series or event-driven poses that takes the form of a finite automaton [2]. For a system involving two legs without violating the laws of physics or the rules of ballet, it take the Cartesian composition. Amy Laviers also built automatic generation of Ballet phrases using Linear Temporal Logic and Computation Tree Logic as rich motion specification languages for robots' movements [3]. Yaya Heryadi [4] built a syntactical modeling and classification for performance evaluation of Bali traditional dance, adapting the model of skeleton feature descriptor from Michalis Raptis [5]. Dance's pose is represented by spherical coordinate parameter (θ, φ) from several skeleton joints that is clustered as torso frame, first-degree joints, and second-degree joints.

To reduce the dimension of the data of the skeleton joint signal trajectory from the Kinect sensor before processing it with HMM, there are several methods as follows : mapping joint position in 3D space [7]; grouping joint trajectory relative with K-means clustering [10]; segmenting joint trajectory by the sound of footsteps [9]; using joint angle and joint angular velocity [8]; and using skeleton descriptor [4].

B. Hidden Markov Model

Hidden markov model is Markov model with a case where the observation is a probabilistic function of the state. The resulting model (which is called hidden Markov model) is a doubly embedded stochastic process with an underlying stochastic process that is not observable (it is hidden), but can only be observed through another set of stochastic processes that produce the sequence of observations. [1]

A formal characterization of HMM is as follows :

- $S = \{S_1, S_2, S_3, \dots, S_N\}$ -- A set of N states. The state at time t is denoted by q_t.
- $V = \{v_1, v_2, v_3, \cdots v_M\}$ -- A set of Mdistinct observation symbols. The observation at time t is denoted by the variable O_t. The observation symbols correspond to the physical output of the system being modeled.
- $A = \{a_{ij}\}$ -- An $N \times N$ matrix for the state transition probability distribution where a_{ij} is the probability of making a transition from state S_i to S_j :
  $a_{ij} = P[q_t = S_j | q_{t-1} = S_i], \quad 1 \le i, j \le N$
- $B = \{b_j(k)\}$ -- An $N \times M$ matrix for the observation symbol probability distributions where $b_j(k)$ is the probability of emitting v_k at time t in state S_j :
  $b_j(k) = P(O_t = v_k | q_t = S_j), \ 1 \le j \le N, 1 \le k \le M$
- $\pi = \{\pi_i\}$ -- The initial state distribution where π_i is the probability that the state S_i is the initial state :
  $\pi_i = P[q_1 = S_i], \quad 1 \le i \le N$

Probabilistic notation A, B, and π must satisfy stochastic constraints as follows :

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

An compact notation $\lambda = (A, B, \pi)$ is used which includes only probabilistic parameters.

The left-right model as shown in Fig. 1. It is good for modelling order-constrained time-series whose properties sequentially change over time [12].

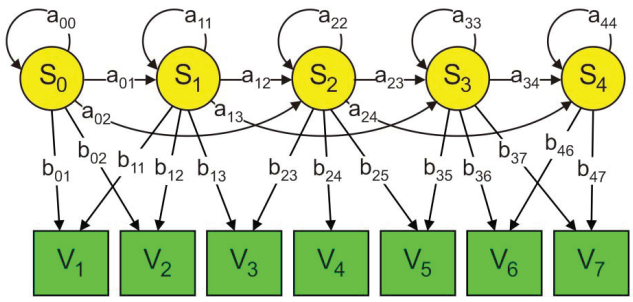


Fig. 1. Graphical model of left-right discrete hidden Markov model

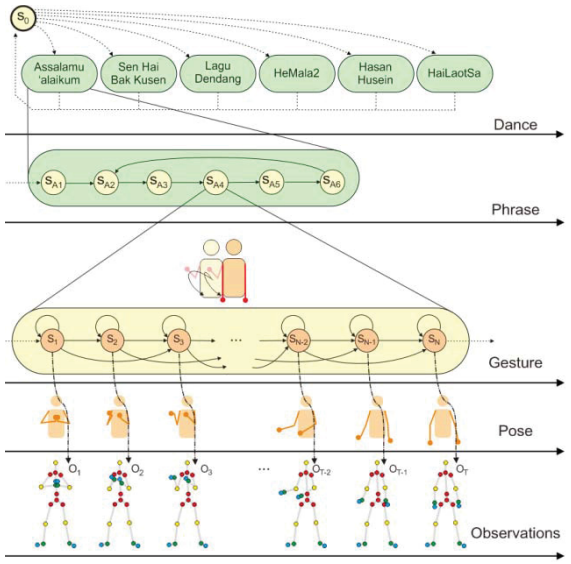


Fig. 2. Hierarchy of the dance

III. Modelling the Whole Dance and the Dance Gesture

A. Modelling The Whole Dance

In this study, some terminologies are used as follows (illustrated in Fig. 2) :

- Pose – Static configuration of human body, without any movement.
- Gesture – Dynamic movement of human body, which is sequence of poses.
- Phrase – Fragment of choreography which consist of sequence of gestures. The same gestures may be repeated.
- Dance – The whole choreography of a dance from the start to the end, which consist of sequence of phrases.

The whole dance of Likok Pulo dance is modelled as follows :

$$\mathcal{L} = (S, I, P, O, f, e, s_0, S_t)$$

- S -- the finite nonempty set of hidden states. The states corresspond to gestures. Its segmentations are determined by the dance expert.
- I -- the finite nonempty set of input.
- P -- the vocabulary of all possible discrete pose of dance.
- O – the finite nonempty set of output, where $O = \{o_1, o_2, \dots, o_T\}$, $o_i \in P^*$, $i \in \{1,2,\dots,T\}$. P^* is the Klenee closure of P, the set consisting of concatenations of arbitrarily many string of element from P (pose). Output O corresponds to gesture trajectories, or its features.
- f – state transition function $f : S \times I \to S$. State transition corresponds to gesture transitions, which for $\forall s \in S$ and $\forall x, y \in I$, satisfies $f(s, xy) = f(f(s, x), y)$ and $f(s, \varepsilon) = s$, where ε is empty transition.
- e – the output map $e : S \times I \to O$.
- s_0 -- initial state, $s_0 \in S$. Initial state corresponds to initial pose or initial gesture of all phrases of Likok Pulo.
- S_t – set of final (or accepting) states, $S_t \subseteq S$. Final states correspond to the end of the phrase.

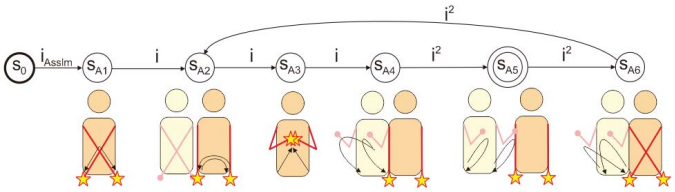


Fig. 3. Model for Assalamualaikum Phrase

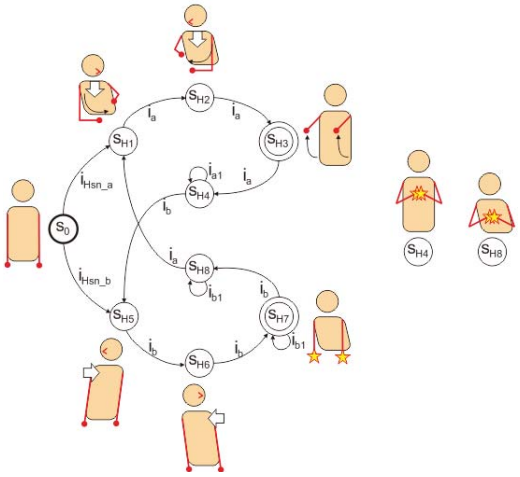


Fig. 4. Model for Kisah Hasan Husein Phrase

The model for "Assalamualaikum" phrase and "Kisah Hasan Husein" phrase are illustrated in Fig.3 and Fig.4. Initial states are indicated by using bold circles. Final states are indicated by using double circles. Actually the whole dance has 6-8 phrases.

B. Modelling The Dance Gesture in HMM

As illustrated in Fig.2, the hidden states $S = \{S_1, S_2, \dots, S_N\}$ correspond to the pose. The observations symbols $V = \{v_1, v_2, \cdots v_M\}$ correspond to physical output at the system, i.e., the discrete pose vector $P_{u,2}$ (will be explained at chapter IV). Matrix $A = \{a_{ij}\}$ corresponds to transition probability distribution between the gestures S_i. Matrix $B = \{b_j(k)\}$ corresponds to observation symbol probability distribution of discrete vector pose v_i. Matrix $\pi = \{\pi_i\}$, corresponds to initial gesture distribution.

IV. HMM-BASED DANCE GESTURE LEARNING AND CLASSIFICATION SYSTEM

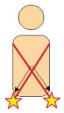
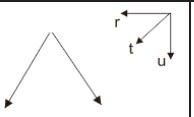
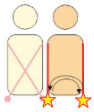
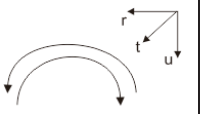
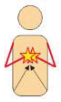
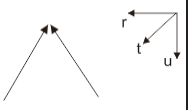
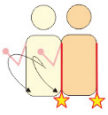
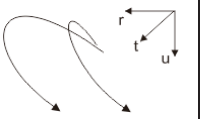
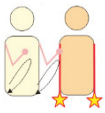
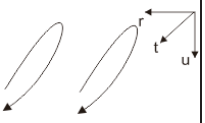
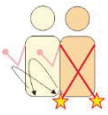
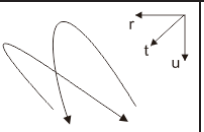
A. Definition of the Dance Gesture Classes

The gesture classes used in this study are 6 dance gesture elements of Likok Pulo traditional dance on "Assalamualaikum" phrase. $f(s_0, i) = O$, $f(\mathrm{O}, i) = \mathrm{A}$, $f(\mathrm{A}, i) = \mathrm{B}$, $f(\mathrm{B}, i) = \mathrm{C}$, $f(\mathrm{C}, i^2) = \mathrm{D}$, $f(\mathrm{D}, i^2) = \mathrm{E}$, $f(\mathrm{E}, i^2) = \mathrm{A}$. Sequence of dance gestures is recognized as "Assalamualaikum" phrase if it is $\{s_0 O(\mathrm{ABCDE})^7 \mathrm{ABCD}\}$.

B. Sensing System Environment

XBOX Kinect sensor is used to capture the skeleton joints at about 30 fps. The system is implemented in MATLAB and Simulink Environment. The code to trigger the Kinect and its interface are written in C++ and compiled using mex compiler.

Table 1. Dance gestures used in this study

Gesture	Wrist Trajectories	Class	Description
		O	"Clapping the hand in front of the chest" to "crossing the hand over the thigh".
		A	"Crossing hand over the thigh" to "straightening the hand over the thigh".
		B	"Straightening the hand over the thigh" to "clapping the hand in front of the chest".
		C	"Clapping the hand in front of the chest" to "swinging the hand to the rightside" to "straightening the hand over the thigh".
		D	"Straightening the hand over the thigh" to "swinging the hand to the leftside" to "straightening the hand over the thigh".
		E	"Straightening the hand over the thigh" to "swinging the hand to the rightside" to "crossing the hand over the thigh".

C. Skeleton Representation

The skeleton representation must satisfy these objectives [5]: (1) Robust coordinate system based on human body orientation, so that the skeleton representation does not depend to the position of the Kinect sensor. (2) Continuity and stability of the signal. (3) Reduce the dimension of the signal while maintaining the character of the motion.

Torso PCA Frame

The joints of the human torso rarely exhibit strong independent motion with large angle. Due to the strong noise in the depth sensing system, individual torso points, in particular shoulder and hips, may exhibit unrealistic motion that it would like to be limited.

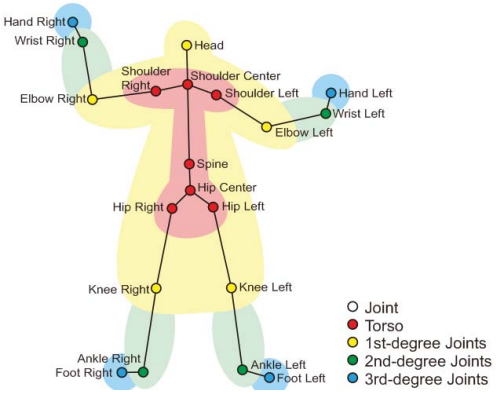


Fig. 5. Hierarchy of skeleton joints

Therefore, the torso can be considered as a rigid body which provides 3D orthonormal basis will be used as reference frame for the remaining joints.

Its principal components as follows : $\vec{u}$, the vector with the direction out of the upper to the lower (in most dancing, the player's torso will never stand upside-down relative to the sensor); $\vec{r}$, the vector with the direction out of the right body to the left side of the body; $\vec{t}$, is the cross product of two principal components, $\vec{t} = \vec{u} \times \vec{r}$.

First-Degree Joints

These joints are represented relative to the adjacent joint in the torso in a coordinate system derived from torso PCA frame as illustrated in Fig. 6 (a). The torso PCA frame is translated to RS (right shoulder) and construct spherical coordinate system such that the origin is RS, its azimuth axis is $\vec{u}$ and its zenith axis is $\vec{r}$.

Azimuth φ is the angle between $\vec{u}$ and $\overrightarrow{(RS, RE_P)}$ where RE_P is the projection of RE onto the plane whose normal is $\vec{r}$. Elevation θ is the angle between $\overrightarrow{(RS, RE_P)}$ and $\overrightarrow{(RS, RE)}$. Then each first-degree joint is represented with two angles (θ, φ). Angular representation for RS is $\{RS_\theta, RS_\varphi\}$.

Second-Degree Joints

These joints are represented relative to the adjacent joint in the first-degree joints in a coordinate system $\{u_p, r_p, t_p\}$ which is derived from rotationed torso PCA frame $\{\vec{u}, \vec{r}, \vec{t}\}$ by angle $\{RS_\theta, RS_\varphi\}$ as illustrated in Fig. 6 (b). The vector $\overrightarrow{u_p}$ protuding out of the vector $\overrightarrow{(RS, RE)}$. The vector $\overrightarrow{u_p}$ be a zenith axis of the spherical coordinate system with origin RE. The azimuth axis is $\overrightarrow{r_p}$ and the zenith axis is $\overrightarrow{u_p}$. Each second-degree joint is represented with two angles (θ, φ). Angular representation for RE is $\{RE_\theta, RE_\varphi\}$. Knee joint is represented by one angle θ.

Third-Degree Joints

These joints are represented relative to the adjacent joint in the second-degree joints in a coordinate system $\{u_{pp}, r_{pp}, t_{pp}\}$

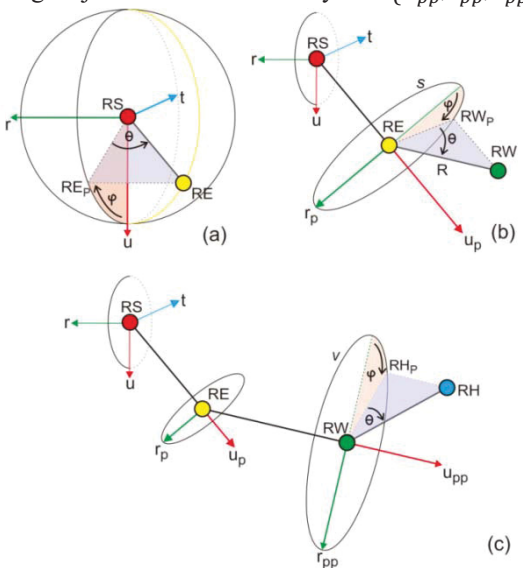


Fig. 6. Spherical coordinate system for (a) first-degree joints, (b) second-degree joints, (c) third-degree joints.

which is derived from rotationed frame $\{u_p, r_p, t_p\}$ by angle $\{RE_\theta, RE_\varphi\}$ as illustrated in Fig.8(c). The vector $\overrightarrow{u_{pp}}$ protuding out of the vector $\overrightarrow{(RE, RW)}$. The vector $\overrightarrow{u_{pp}}$ be a zenith axis of the spherical coordinate system with origin RW. The azimuth axis is $\overrightarrow{r_{pp}}$ and the zenith axis is $\overrightarrow{u_{pp}}$. Each third-degree joint is represented with two angles (θ, φ). Angular representation for RW is $\{RW_\theta, RW_\varphi\}$.

It is needed to use Wearable Inertial Measurement Units (WIMU) [11] to obtain accurate angles at third-degree joints because Kinect sensor can not detect third-degree joints orientation and position accurately.

Human Pose Representation

For the scope of body poses which involves up to second-degree joints,

- Upper body poses are represented by an 8-tuple
$$P_{u,2} = (LE_\varphi, LE_\theta, LS_\varphi, LS_\theta, RS_\theta, RS_\varphi, RE_\theta, RE_\varphi).$$
- Lower body poses are represented by the 6-tuple
$$P_{l,2} = (LK_\theta, LH_\varphi, LH_\theta, RH_\theta, RH_\varphi, RK_\theta).$$

For the scope of body poses which involves up to third-degree joints,

- Upper body poses are represented by an 12-tuple
$$P_{u,3} = \begin{pmatrix} LW_\varphi, LW_\theta, LE_\varphi, LE_\theta, LS_\varphi, LS_\theta, \dots \\ RS_\theta, RS_\varphi, RE_\theta, RE_\varphi, RW_\theta, RW_\varphi \end{pmatrix}$$
- Lower body poses are represented by the 12-tuple
$$P_{l,3} = \begin{pmatrix} LA_\phi, LA_\theta, LK_\theta, LH_\phi, LH_\varphi, LH_\theta, \dots \\ RH_\theta, RH_\varphi, RH_\phi, RK_\theta, RA_\theta, RA_\phi \end{pmatrix}$$
- Head poses are represented by the 3-tuple
$$H = (H_\varphi, H_\theta, H_\phi)$$

$P_{u,2}$ will be used for implementing HMM-based dance learning and classification.

D. The System Block Diagram

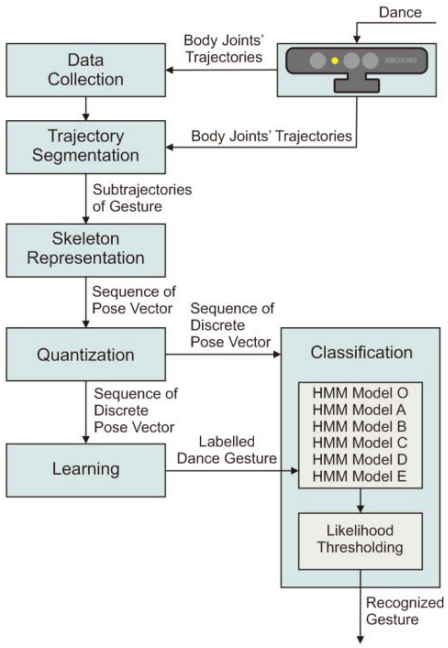


Fig. 7. System block diagram with major components.

Table 2. Isolated dance gesture pattern for the HMM Training and Testing

Gesture	Only Kinect	W/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

E. Data Collection and Segmentation

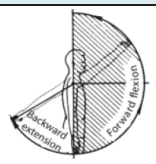
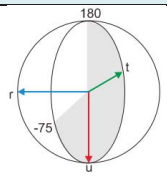
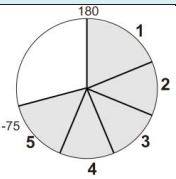
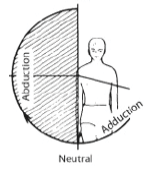
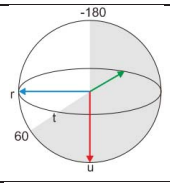
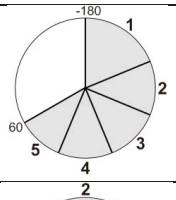
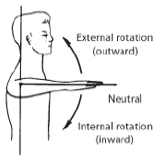
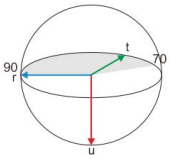
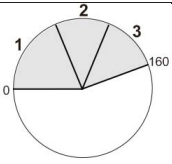
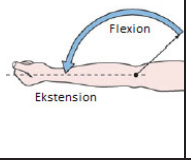
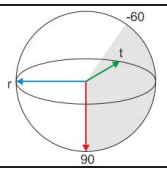
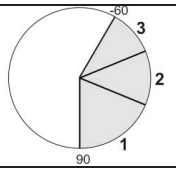
It has been collected 2169 isolated dance gestures data from three subjects which are classified to 6 sets of data for each of dance gesture classes. Total isolated data are partitioned into 80% training data and 20% test data.

Segmentation between gestures in a continuous joint trajectory signal is done in two ways : (1) By detecting hand's clap by clap sensors in the smartgloves and sent by bluetooth module. (2) By using a time window. The system inform the performing subject to do the gesture with limited time. The border between the gestures are identified through timestamps.

F. Pose Vector Quantization based on Range of Movement

Isolated joint trajectory signals are represented by pose vector $P_{u,2} = (RS_\varphi, RS_\theta, RE_\varphi, RE_\theta, LS_\varphi, LS_\theta, LE_\varphi, LE_\theta)$. The combination of all possible values of its elements is infinite. For discrete HMM-based approach, each element of pose vector should be converted to one of the 3 or 5 directional codewords, based on ROM (range of movement) [6].

Table 3. Directional codewords for each joint angle based on ROM

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

G. Dance Gesture Learning

The parameters of each HMM models estimated using the Baum Welch algorithm iteratively. Training likelihood curves generally appeared to be stable after 25 cycles, but has not really converge until approximately 70 cycles. The training stops after 100 cycles. The number of states in gesture models ranges from three to five, depending on the complexity of the gesture shape. Increasing the number of hidden states may lower down the recognition rate.

H. Dance Gesture Classification

Classification is done by using a score value $P(O|\lambda)$ to assess the likelihood (degree of match) between the input test gesture and gesture models. Score computation is done with the forward probabilities:

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

$\alpha_t(i)$ = Forward probabilities.

One input test datum of dance gesture tested by six trained HMM model to find the one model that reflects the highest likelihood. If the value of its maximum likelihood pass the predefined threshold value for that model, then the test datum is classified as that model. If the log-likelihood is minus infinity (has no likelihood at all), then the gesture is not classified to any gestures. Threshold of each model is the minimum of the scores of tested training data.

Table 4.The classification results of dance gesture

Gesture	Detected gesture pattern by the model (%)						Un-detected
	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

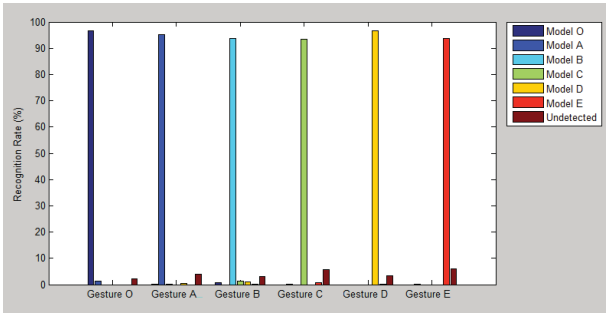


Fig. 8. Bar chart of classification results of dance gesture

Table 5. Effect of Using Skeleton Representation and Maximum Likelihood

	Before	After
Data Dimension	33-tuple	8-tuple
Detected as False	10 %	0.22 %
Detected as True	80.7 %	94.87 %

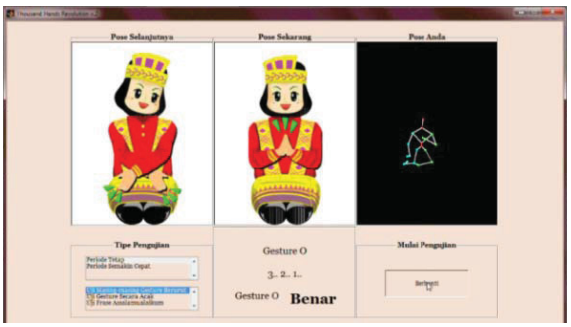


Fig.9. Realtime implementation, gesture O detected

I. *Dance Phrase Recognition*

Twenty of complete recording data of "Assalamualaikum" dance phrase are collected. The results are shown in Table 6 and Fig. 10.

Table 6. Testing results of 20 complete phrases

Phrase	True Gesture	False Gesture	Gesture Recognition Rate (%)	Complete 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 %

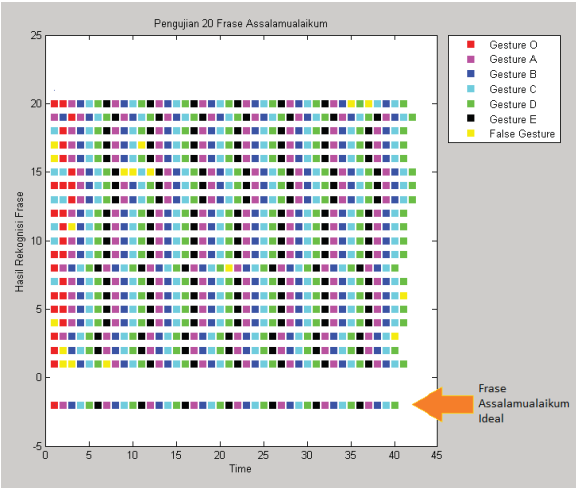


Fig. 10. Testing results of 20 complete phrases compared to ideal phrase

V. Conclusion

Hidden Markov model can be used to model the whole dance; dance gestures cast as hidden discrete states and phrase as a sequence of gestures.

Skeleton representation that is quantized based on range of movement can effectively handle noisy joint trajectory data, reduce the data dimension, and handle the change of position and orientation of user relative to the Kinect sensor.

HMM are an effective and efficient method of both learning and classifying dance gestures involving several joints.

Observation of the dance can be expanded up to lower body, and/or expanded to third-degree joints. It is required additional inertial sensors for capturing position and orientation of third-degree joints (palm hands and foots) due to Kinect sensor can not detect it. Skeleton representation can be deepened to also consider the dynamic aspects of the human body.

Acknowledgment

The author greatly appreciate the contributions of Mr. Ary Setijadi Prihatmanto for the guidance and the teaching. The author is also appreciate the help of Sayid Tarmizi and Sundari Mega for their help related to Likok Pulo dance.

References

[1] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proceedings of the IEEE, vol 77 no.2, Februari 1989.

[2] A. LaViers, M. Egerstedt, "The Ballet Automaton : A Formal Model for Human Motion", 2011.

[3] A. LaViers, Y. Chen, C. Belta, M. Egerstedt, "Automatic Generation of Ballet Phrases", 2011.

[4] Y. Heryadi, M. I. Fanany, A. M. Arymurthy, "A Syntactical Modeling and Classification for Performance Evaluation of Bali Traditional Dance", in ICACSIS (2012).

[5] M. Raptis, D. Kirovski, H.Hoppe, "Real-Time Classification of Dance Gestures from Skeleton Animation". In ACM SIGGRAPH Symposium on Computer Animation (2011).

[6] A. G. Apley. "Apley's Sistem of Orthopaedics and Fractures", Ninth edition. Hodder Arnold, 2010.

[7] J.Huang, C. Lee, J. Ma. "Gesture Recognition and Classification using the Microsoft Kinect". 2012.

[8] H. Zhang, W. X. Du, H. Li. "Kinect Gesture Recognition for Interactive System." 2012.

[9] A. Masurelle, S. Essid. "Multimodal Classification of Dance Movements using Body Joint Trajectories and Step Sounds.".

[10] 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/

[11] M. Gowing, C. Concolato, E. Izquierdo. "Enhanced Visualisation of Dance Performance from Automatically Synchronised Multimodal Recordings". 2011

[12] 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