Dance Recognition System of Sigeh Penguten Traditional Dance based on Hidden Markov Model

[1]Aciek Ida Wuryandari, [2]Maria Shusanti Febrianti, [1]Khairurijal Alfathdyanto, [2]Marzuki Syahfirin,
[2]Agus Sukoco, [1]Egi Muhamad Idris Hidayat, [1]Ary Setijadi Prihatmanto and [1]Carmadi Machbub
[1]Department of Electrical Engineering, Institute of Teknologi Bandung,
School of Electrical Engineering and Informatics, Bandung, Indonesia
[2]Department of Informatics, Faculty of Computer Science, Universitas Bandar Lampung,
Lampung, Indonesia
aciek@lskk.ee.itb.ac.id

Abstract: Culture is one of the characteristics created by a group of people inherited from generation to generation. Among various kinds of cultural diversity in Indonesia, cultural arts are one of the cultural riches admired by other countries as well as on dance. Indonesian traditional dance, one of them is Sigeh Penguten dance is a typical dance from Lampung Province to welcome honored guests at certain occasions. Therefore, it is necessary to maintain its sustainability by modeling the dance using the Hidden Markov Model with the representation of the skeleton taken using the Kinect V2 camera. In this research, the phrase modeling using HMM is done on several phrases. The results of the phrase recognition on the dance model were built using Hidden Markov Model (HMM). The results of the dance recognition of the Sigeh Penguten traditional dance will be simulated on the 3D character of the Sigeh Penguten dancer. Motion simulation is generated from the transformation of data in the form of movements from humans to 3D characters.

Key words: Traditional dance, Hidden Markov Model, 3D character, Sigeh Penguten, kinect, Lampung

INTRODUCTION

Culture is one of the characteristics possessed by a group of people who are inheritage from generation to generation. Every country has a diverse culture, one of the countries that has a high variety of cultures is Indonesia. Indonesia has approximately 700 tribes in various archipelago that have varied cultures, from language, cultural arts, traditional houses, traditional clothing, to regional specialties. Among the various kinds of cultural diversity that exist in Indonesia, cultural arts are one of the cultural treasures that are admired such as on dance.

Dances are the work of human creations which are expressed through the medium of motion that has beauty. The dance consists of modern dance and traditional dance. Modern dance is formed from among young people who are only looking for popularity and only temporary. Traditional dance is a dance that comes from the community of an area that has been passed down through generations and has become the culture of the local community. Traditional dance is a cultural heritage that has regional characteristics and strong traditional characteristics, so that, the people involved take part in preserving it through a sense of responsibility and love (Candrawati, 2015) but as modern culture develops today, traditional dances are becoming less attractive to the younger generation. If this is allowed, then the culture that has been inherited by the ancestors will slowly disappear or become extinct. Therefore, the existence of dance in Indonesia must be preserved and maintained properly. One of them is by digitizing traditional dance. In this study, the researcher will make a digitization of traditional Lampung dance, the Sigeh Penguten dance.

Sigeh Penguten dance is one of the new creation dances from Lampung. Sigeh Penguten dance is a dance performed by women (Muli in Lampung language) which is used to welcome honored guests in an event. This dance is a development of the dance of worship which is an indigenous dance tradition of the people of Lampung. Through the regional regulation, the Sigeh Penguten Dance was formalized as a form of respect when welcoming guests. This dance choreography also takes

elements from various Lampung traditional dance to represent the diverse culture of Lampung. One characteristic of the Sigeh Penguten dance which is the original element of the prayer dance is the accessories worn by the dancers. As the name implies, the main accessory used is Siger which is a gold crown that has become the identity of the Lampung Region. Other accessories used are on the fingers of the dancers who wear stairs which are cone-shaped and gold-colored finger covers. In addition to the two accessories, Sigeh Penguten dancers also wear Jajar boards, Jukum fruit necklaces, Canoe bracelets, bird bracelets and sesapur.

In the digitization process, the device used is the Kinect V2 Sensor camera. The Kinect sensor camera is able to detect 25 joints of the human body which have a value in the cartesian coordinate and detect up to 6 skeletons.

Data generated from the skeleton will be used in the movement recognition process. This is also done by Karavarsamis *et al.* (2016) who do the classification of steps on Salsa dance, Anbarsanti and Prihatmanto (2014) for Aceh traditional dance. They use skeletal poses taken using Kinect sensors for recognition of movement as a feature of the human body. The introduction of movement is a machine learning process in order to recognize the movements carried out by humans. There are many methods used by previous researchers in movement recognition. The Bag of Word (BoW) Model approach for the introduction of the Greek Folk dance movement achieved an accuracy 89.74% (Fotiadou *et al.*, 2016). Introduction of the Baratanatyam dance movement by using automated stick figure generation to visualize dance movement poses reaching an accuracy level of 80% (Jadhav *et al.*, 2014). Movement recognition was also carried out by Himansu for gesture recognition (Kumar, 2014) using active difference signature and Hidden Markov Model as classifier with accuracy rate of 80.7%. Various methods HMM has the advantage of being able to model the human movement well through its stochastic process because the movement of the human body has a random chance. The method used for dance recognition is the Hidden Markov Model (HMM). Hidden Markov Model is a process to predict future events by looking at opportunities to move the current condition/status from the previous conditions that will produce the Markov chain. The Markov chain can be used to calculate a probability order of observable conditions. But there are some cases where sequences of conditions that want to be known but cannot be observed, therefore, the Hidden Markov Model (HMM) is developed to solve this problem (Febrianti *et al.*, 2016).

Sigeh Penguten dance was chosen in this study to introduce one of Lampung's traditional dances to the wider community and make it easier for Lampung people to learn the Sigeh Penguten dance movement.

MATERIALS AND METHODS

Hidden Markov Model: Assume, if you are given a process that produces observable sequences of events. The process is discrete which can be the result of coin throwing experiments, finite alphabet characters and others. The Markov Model is a process for predicting future events by looking at the opportunity to move the current condition/states from the previous conditions which will produce the Markov chain. The Markov chain can be used to calculate a probability sequence of conditions that can be observed. But in some cases there is a sequence of conditions that you want to know but cannot be observed, so, Hidden Markov Model (HMM) was developed to solve this problem (Rabiner and Juang, 1986).

HMM is a Markov Model that has a double stochastic process with a basic process that cannot be observed/hidden but can be observed through a series of other stochastic processes which results in a sequence of observation symbols in the basic process. This model is able to represent data generated from time sequences more efficiently and have characteristics the displacement of a condition of the next condition is regulated by using the opportunity function (Vogler and Metaxas, 1998). HMM denoted by • = (A, B, •) used for the set of probabilistic parameters.

The Markov Model is formed from a system condition that is defined as states. A system can be characterized in several different states. The status of the system can experience changes within a certain time to be another states of the system. This phenomenon of states change is called states transition. This states transition is also a definition of the Markov Model. The relationship between state changes is characteristic of the Markov Model.

States transitions are significantly influenced by many factors that cannot always be stated deterministically. The assumptions taken in making the Markov Model simplify the representation of these state transitions by stating the relationship to the occurrence of the state just by looking at the previous state. As a result, states transitions can be modeled stochastically with the opportunity value of change to states from a previously known states. This opportunity value is called transition opportunity states. The general Markov Model is illustrated as in Fig. 1. States is represented as a states
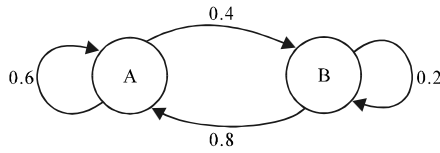
Fig. 1: Model left-right with 2 states

and arrows the states transition along with the transition opportunity value. The model illustrated in Fig. 1 is a Markov Model that has two states, namely A and B. Transitions between states are indicated by arrows and contain the opportunity values of the transition. For example, the status transition from A to B has an opportunity value of 0.4. Transition states can also lead to the states itself. The total value of transition opportunities from each states must be 1 which shows the total value of all events.

HMM is different from the Markov Model in terms of the states you want to know cannot be detected or measured directly. An observation can give clues about the status of the system but it is not enough to be able to determine exactly the desired states of the system.

In this research, we use HMM left-right topology model. The left-right model has the characteristics of moving status from left right or to the same state and does not have a backward path. Left-right models can model observational data whose properties changed with time. HMM has the following characteristics:

N is the number of state denoted by: $S = \{S_1, S_2, S_3, ..., S_N\}$. State when t is denoted by q_t. M is the number of observation symbols on each state. The observation symbol corresponds to the output on the system that is modeled can be denoted by $V = \{V_1, V_2, V_3, ..., V_M\}$. Observation when t are denoted by variables O_t.

$A = \{a_{ij}\}$, a matrix N×N representing distribution of opportunities where state transition a_{ij} is probability transition from S_i to S_j:

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

With probability constraints that must be fulfilled as follow:
$$\sum_j a_{ij} = 1, \forall i \text{ and } a_{ij} \geq 0$$

$B = \{b_j(k)\}$, a matrix N×M which represents the distribution of probability that produce observation of symbols where, $b_j(k)$ is the probability to produce symbol vk at t in state S_j. $b_j(k) = P(O_t = v_k | q_t = S_j)$, 1•j•N, 1•k•M. With probability constraints that must be fulfilled as follows:

$$\sum_k b_{j(k)} = 1, \forall j \text{ and } b_{j(k)} \geq 0$$

• = {•ᵢ}, initial state probability distribution where, •ᵢ is a probability that state S_i is initial state:

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

with constraint probability that must be fulfilled as follow:

$$\sum_i \pi_i = 1 \text{ and } \pi_i \geq 0$$

HMM on activity recognition is a mathematical modeling (probabilistic) used to represent gesture. There are 3 algorithms used to solve problems in HMM.

Viterbi algorithm: Viterbi algorithm is used to find the optimum value, so that, conclusions can be drawn based on assumptions obtained from the observed probability values obtained in the previous evaluation process.

Forward-backward algorithm: This algorithm is used for evaluation by calculating the probability of the order of observations given by HMM. Difference between forward and backward algorithm is in the observation sequence/the probability value.

Baum-Welch algorithm: This Baum-Welch algorithm works to determine the value of expectations and maximization, so that, it will produce a new model that has maximum value.

Dance transition: Based on data taken from professional dancers, Lampung dance studios through speakers of Lampung traditional leaders, a movement to collect Sigeh Penguten dance was conducted, consisting of dance Phrases (F) consisting of a collection of movements and movements which were a set of poses, the name of all phrases is:

- F1 = Lapah Tebeng
- F2 = Seluang Mudik
- F3 = Merunduk
- F4 = Jong Ippek
- F5 = Jong Silo Khato
- F6 = Sembah
- F7 = Jong Sippuh
- F8 = Samber Melayang
- F9 = Ngerujung Bawah
- F10 = Kilat Mundur
- F11 = Makurancang
- F12 = Kenui Melayang
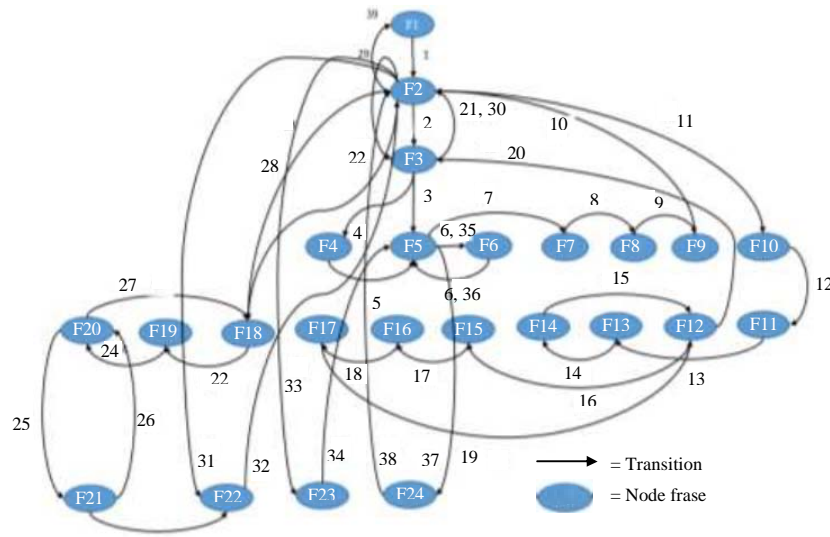- F13 = Gubuh Gakhang
- F14 = Balik Palo

Fig. 2: Frase transition

- F15 = Makurancang Berbalik
- F16 = Nginyau Bias
- F17 = Ngerujung Tinggi
- F18 = Sabung Melayang
- F19 = Tolak Tebeng
- F20 = Mempam Bias
- F21 = Belah Hui
- F22 = Lippeto
- F23 = Ngerujung Sedang
- F24 = Jong Ippek

Observe the whole dance motion there is a similar movement that is performed more than once in a dance from the beginning to the end of the dance. The F1-F24 are the names of movements that exist in the Sigeh Penguten dance of the and 1-39 are the transition state as a whole in a Sigeh Penguten dance. The results of the identification of all movements in the Sigeh Penguten dance show a total of 24 phrases with 39 transitions. Dance moves consist of a series of phrases in which there are repeated sequences shown in Fig. 2, from the picture it is clear that there are several phrase transitions which are repeated several times from the same phrase in the phrase in the Sigeh Penguten dance.

Dance recognition system: The dance motion recognition system built in this study is depicted in the flowchart (Fig. 3). The process of dance recognition system begins with the data acquisition. Acquisition is data that has been recorded and then processed by the computer for training and evaluation. The sigeh penguten dance movement recorded using the Kinect Version 2 sensor
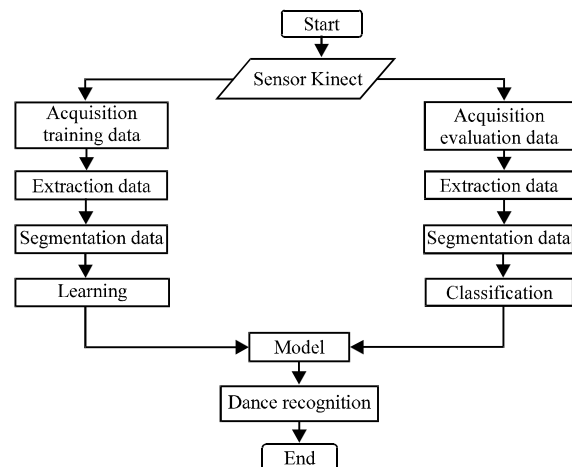

Fig. 3: Flowchart system

produces skeleton data as training data and test data. The extraction feature (or detection) aims to find areas of significant features in the image depending on their intrinsic characteristics and application. The region can be defined in a global or local environment and distinguished by shape, texture, size, intensity, statistical properties and so on. Features are unique characteristics of an object. The dancer's movement is extracted to produce a skeleton from the Kinect V2 sensor. Segmentation data is an operation aims to classify a data into several segments with certain criteria. This type of operation is closely related to pattern recognition. Segmentation data based on feature extraction to classify skeleton data becomes a phrase to be more easily recognized. After segmentation
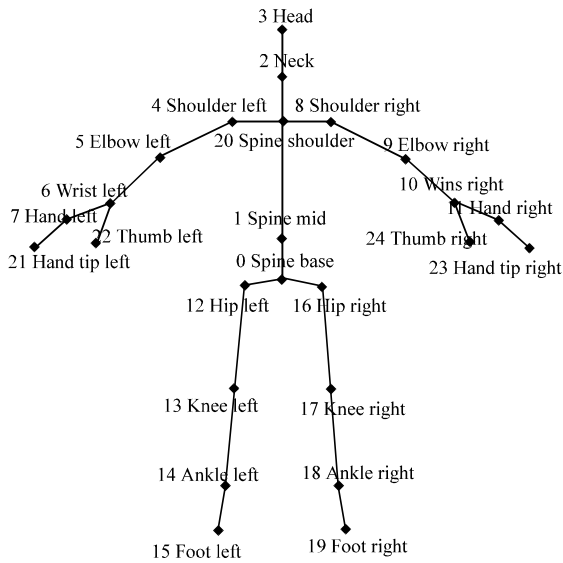
Fig. 4: Skeleton

Fig. 5: Skeleton representation in the human body

data we do the learning, learning is a process to help user learn. From some data on the Sigeh Penguten dance movement that we have used as training data, it is then processed with the Baum-Welch algorithm. And the next classification is needed for each data skeleton is in a series of skeleton datasets compared to each category which determines unknown movements and is most similar to the same category.

Data acquisition using Kinect V2 camera with skeleton representation. Kinect is hardware made by Microsoft that can record 3D images of an object. Kinect is actually a sensor for human interaction on the XBOX console but is also often used for research. Kinect can also be used for the purposes of motion recording, motion detection and human-computer interaction. Most of these capabilities can be realized with Microsoft SDK (Software Development Tools). There are two types of Kinect namely: Kinect XBOX 360 and XBOX one. Kinect XBOX 360 is also called Kinect V1.0 while Kinect XBOX one is also called Kinect V2.0. This research uses Kinect for XBOX one because it can produce more detailed observers than the previous version. The change that was considered by the researcher was that many of the joints that could be detected increased from 20-25, so that, more accurate movement details were obtained. The details of the joint of skeleton that Kinect XBOX one can acquire are detailed in Fig. 4. As illustrated in Fig. 4 and the skeleton represented on the human body using the Kinect sensor is shown in Fig. 5.

Data acquisition is divide into 2 parts, for learning and for evaluation. Data learning is taken from the movements performed by professional dancers and the

Fig. 6: Lapah Tebeng

data evaluation was taken from the the studio dancers consisting of 10 different people in each phrase. Some of the movements used in this study consisted of Lapah Tebeng, Seluang Mudik, Merunduk, Jong Ippek and Jong Silo Khato. Here's the explanation:

- Lapah Tebeng: moves place motion. This motion is used when entering and exiting the show area shown in Fig. 6
- Seluang Mudik: transition motion from a standing position toward bend (Fig. 7)

Fig. 7: Seluang Mudik

Fig. 9: Jong Ippek

Fig. 8: Merunduk

Fig. 10: Jong Silo Khato

- Merunduk: the transition movement from a sitting position to Jong Ippek (Fig. 8)
- Jong Ippek: the transition movement from a sitting position to Jong Silo Khato (Fig. 9)
- Jong Silo Khato: the movements of both hands are placed on the knees (Fig. 10)

The data of these movements are extracted in the form of a skeleton using the Kinect V2 sensor which produces the features of the joint-joint value in the cartesian coordinate plane. Next, segmentation is done, if necessary on certain body parts features that will be used for modeling, for example, right and left hands, right and left legs or other body parts. Modeling using the Hidden Markov Model is a Markov Model that has a double

stochastic process with a basic process that cannot be observed/hidden but can be observed through a series of other stochastic processes that produce a sequence of observation symbols in the basic process (Rabiner and Juang, 1986). This model is able to represent data generated from time sequences more efficiently and has the characteristics of moving a condition to the next condition set using the opportunity function.

RESULTS AND DISCUSSION

The introduction of this dance movement uses professional dancer data used for learning. Professional dancer movements on 5 phrases (Table 1 and Fig. 11):

Table 1: Frase evaluation using the resulting HMM

Takes	Lapah Tebeng	Seluang Mudik	Merunduk	Jong Ippek	Jong Silo Khato
Dancer 1	•	•	•	•	Merunduk
Dancer 2	Merunduk	•	•	•	•
Dancer 3	•	•	•	•	Jong Ippek
Dancer 4	-	-	-	-	•
Dancer 5	-	-	Lapah Tebeng	•	•
Dancer 6	-	-	•	•	•
Dancer 7	•	-	•	•	Lapah Tebeng
Dancer 8	-	•	Merunduk	•	•
Dancer 9	Jong Silo Khato	•	•	Merunduk	-
Dancer 10	•	•	•	-	Merunduk
Recognize	4	6	6	6	5
Un-recognize	4	4	1	2	1
Incorrect recognize	2	0	2	1	4

Fig. 11: The 3D character Sigeh Penguten dancer

- Tebeng Lapah phrases
- Seluang Mudik phrases
- Merunduk phrases
- Jong Ippek phrases
- Jong Silo Khato phrases taken 3 times in each phrase

The learning process is carried out in each phrase with the Baum-Welch method which uses each of the 3 professional dancer data to produce a convergent model. Convergent models are generated through several iterations. For classification, forward-backward algorithm is applied to compute the likelihood of certain validation data to the five frase models. A dataset is annotated as correctly recognized, if the likelihood is above the threshold value. Validation results are given in Table 1.

The experiment result show that the obtained HMM can be used to recognize the five frase. However, the result is still unreliable, in particular for frase Lapah Tebeng, only 40% correct recognition is achieved. The best result is obtained by frase Seluang Mudik, Merunduk and Jong Ippek 60% which is still low to be considered as a valid model. This poor evaluation performance is to be expected as the measurement observations are based only from three observation data from main dancer for each frase. Furthermore, the phrase movements that are successfully known will be simulated on the 3D characters of the Pengeh Penguten dancer as shown in Fig. 11 is an example of the Tebeng Lapah movement that is driven in the 3D character Sigeh Penguten.

CONCLUSION

The dance recognition framework based on HMM using Kinect sensor has been implemented and tested for five frase recognition of Sigeh Penguten traditional dance. The best result is obtained by frase Seluang Mudik, Merunduk and Jong Ippek 60%. The results show that there are possible aspects for future research work to achieve better results.

ACKNOWLEDGEMENT

REFERENCES

Anbarsanti, N. and A.S. Prihatmanto, 2014. Dance modelling, learning and recognition system of ACEH traditional dance based on hidden Markov model. Proceedings of the International Conference on Information Technology Systems and Innovation (ICITSI), November 24-27, 2014, IEEE, Bandung, Indonesia, ISBN:978-1-4799-6526-7, pp: 86-92.

Candrawati, L., 2015. [Functions of Traditional Dance in Indonesia]. PPPPTK Seni Dan Budaya Yogyakarta, Indonesia,.

Febrianti, M.S., E. Hidayat, A.I. Wuryandari, A.S. Prihatmanto and C. Machbub, 2016. Preliminary result on gesture recognition of Sigeh Penguten Dance using Hidden Markov Model. Proceedings of the 6th International Conference on System Engineering and Technology (ICSET), October 3-4, 2016, IEEE, Bandung, Indonesia, ISBN: 978-1-5090-5090-1, pp: 169-172.

Fotiadou, E., I. Kapsouras, N. Nikolaidis and A. Tefas, 2016. A bag of words approach for recognition of Greek folk dances. Proceedings of the 9th Hellenic Conference on Artificial Intelligence (SETN '16), May 18-20, 2016, ACM, New York, USA., ISBN: 978-1-4503-3734-2, pp: 1-4.

Jadhav, S., A. Aras, M. Joshi and J. Pawar, 2014. An automated stick figure generation for BharataNatyam dance visualization. Proceedings of the International Conference on Interdisciplinary Advances in Applied Computing (ICONIAAC '14), October 10-11, 2014, ACM, New York, USA., ISBN: 978-1-4503-2908-8, pp: 12:1-12:8.

Karavarsamis, S., D. Ververidis, G. Chantas, S. Nikolopoulos and Y. Kompatsiaris, 2016. Classifying salsa dance steps from skeletal poses. Proceedings of the 14th International Workshop on Content-Based Multimedia Indexing (CBMI), June 15-17, 2016, IEEE, Bucharest, Romania, pp: 1-6.

Kumar, H., 2014. Gesture recognition using hidden markov models augmented with active difference signatures. MCS Thesis, Rochester Institute of Technology, Rochester, New York, USA.

Rabiner, L.R. and B.H. Juang, 1986. An introduction to hidden Markov models. IEEE ASSP Mag., 3: 4-16.

Vogler, C. and D. Metaxas, 1998. ASL recognition based on a coupling between HMMs and 3D motion analysis. Proceedings of the 6th International Conference on Computer Vision (IEEE Cat. No. 98CH36271), January 7, 1998, IEEE, Bombay, India, pp: 363-369.