

# Virtual Reality in Puppet Game Using Depth Sensor of Gesture Recognition and Tracking



Ervin Yohannes<sup>1,2\*</sup>, Timothy K. Shih<sup>1</sup>, Fitri Utamingrum<sup>2</sup>

<sup>1</sup> Department of Computer Science and Information Engineering, National Central University, Taoyuan, Taiwan  
ervinyohannes@gmail.com, tshih@g.ncu.edu.tw

<sup>2</sup> Faculty of Computer Science, Brawijaya University, Malang, Indonesia  
f3\_ningrum@ub.ac.id

Received 10 April 2019; Revised 21 September 2019; Accepted 21 November 2019

**Abstract.** Indonesia has many cultures on each island including language, games, food, and another thing. One of the popular game is puppet performance until now still available, especially in the sacred event. The performance still using traditional method which person can play just expert and must be brought many puppets for performance. The puppet game using hand gesture tracking and recognition for playing. Firstly, in-game need two hands to generate or move puppet so need clustering method for cluster right hand and left hand. The clustering method using DBSCAN (Density-based spatial clustering of applications with noise) clustering, which can cluster automatically and efficiently than another method. The right and left hand will search about each median. There is three distance for use in threshold testing of DBSCAN method. From the result, the threshold value by 25 is the best result because can be clustered left and right hand correctly. The last step is gesture recognition using template matching method consist of three templates including down, right, left direction. In the results, we get accuracy above 90% for hand gesture recognition testing both median1 and median2.

**Keywords:** depth sensor, game, gesture, virtual reality

## 1 Introduction

In Southeast Asia, puppets have various form and every place has each story. It consists of some part including values, norms, and beliefs [1]. In these plays, puppeteers take 2D puppet using screen as background and lighting as effect or action so that gives a shadow effect. The puppeteers bring puppet with combining his skill and voice to interesting audience [2]. The focus of audience in puppet game not only form but also character model/texture. Indonesia has many designs and roles in the story. There are about 400 characters in the Indonesia puppet have two big stories. [3]. In Indonesia usually puppet plays together with orchestra or “*gamelan*” so that attract attention many people especially children. Many children of Indonesia feel hard to play puppet since fewer puppeteers in teaching way of puppet play. Besides, the equipment of puppet play is expensive and not easy to bring anywhere. This is a problem which one of traditional game of Indonesia almost extinct. The technology development increasingly stimulates to solve this problem so that Indonesia’s puppet can be preserved.

One of popular application in the multimedia is Kinect. Microsoft Kinect sensor has been widely used in many applications from early launch. At last, Microsoft launching the new version of Kinect is improved the early version. Kinect version 2 much improved the depth measurement accuracy than Kinect version 1. As a result, Kinect version 2 has better quality in capturing depth images [4]. In technology development now, puppeteers can use Kinect to generate puppet. Kinect offers an attractive communication between client and tool with controller-free performance setting. Entertainment, medical,

---

\* Corresponding Author

surveillance, research, and other field are using Kinect in implementation. It is important for the user to control and interacting a tool in natural and efficient in human-robot communication-based. It is using by media in human include head, body, and hand [5]. There are many research methods used by Kinect and it can use clustering and classification method. Clustering is an unsupervised arrangement technique generally utilized for grouping of remote detecting image [6]. In clustering, many methods for implementation and combine such as K-Means [7], ISO-data, fuzzy [8], DBSCAN [9] and more [10]. Unlike another clustering, DBSCAN (Density-based spatial clustering of applications with noise) algorithm has advantages, which is this density-based method can be separate superior identification and cluster that have size and form different when compared to centroid-based technique. This algorithm can arrange data points into separate 'noise' cluster if given point different to rest of the data set, and this algorithm is deterministic [11].

The contribution of the paper including to recognizing one of traditional game in Indonesia, how to playing Indonesia's puppet and recognizing about a story inside, create design Indonesia's puppet from character until the performance, and create animation effect in play for children. So, we propose Indonesia's puppet game changed from traditional to multimedia playing into virtual reality using Kinect version 2. We create about design [1, 12-13], character [14-15], movements [16], and story [3] inside game. The design is changed from 2D to 3D form in each character of puppet. For character, we just create six-character consists of one is bad and rest is good character since to make easier for children to play games. In movement, we using gesture recognition and tracking captured by depth sensor of Kinect version 2. Moreover, we applied DBSCAN method for clustering of left and right hand. And last, we are proposing romance story in our game since this story very popular in Indonesia.

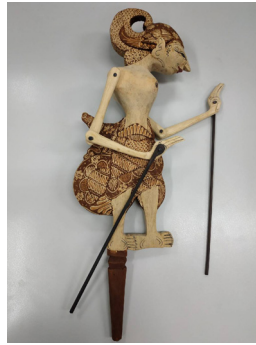
## 2 Related Works

### 2.1 Indonesia Puppet

The design of Indonesia puppet has proposed [1], which is preserve traditional shadow play into a multimedia tool as animate the characters in Indonesia puppet theatre. Jasni et.al. [12] propose a study about Indonesia puppet courseware help to find appeal and satisfaction of the student in the mathematics field. The focus is basic operation such as addition, subtraction, multiply, division combine between number and Indonesia puppet. Many components of Indonesia puppet provides framework that suitable with mathematics such as description, voice, and music into 2D animation. In the [13] explain about graphical model of puppet Kelantan which is directing into four main categorical such as doll, shadow, background, and lighting.

Not only create a design of Indonesia puppet, but also character [14], which is purpose to define a way to combine character of puppet with game in Indonesia. The method using combination between Space-Time-Plane (STP) with Naturalist-perspectives-Momentopname (NPM) which is changing the visual representation of puppet into game character. Moreover, Dahlan [15] propose about study and implementation of framework is Maya MEL Animation Programming language for giving rigging into Indonesia puppet so that it can be move, not only move but also can speak like human activity because it has bones and inverse kinematic for make puppet to move anywhere and anything.

In other hands, there is research about movement of Indonesia puppet [16] first step towards of vision is to take the initiative develop a digital image analysis method. It has aimed to fulfill the vision of Roger Long and but also to discover new and interesting Spatio-temporal cultural patterns which fields of humanities cannot be obtained using qualitative manual observations traditionally. Sigit et al. [3] propose research about relation between Indonesian traditional puppet shapes and function in the story. Therefore, the research focus on recording important information must solve. The first is identifying the form of puppet so that categorization can finish. For these aims, the overlapping lines and patterns between each shape are finding to be obstacles become difficulty in this case. To solve this problem, the way to solve is taking another research report about same way of thinking in understanding the perspective impression for traditional art/painting. The analysis result is comparing puppet part-by-part not whole puppet so that know about relationship with function in the story. One example of Indonesian puppet shown in Fig. 1.



**Fig. 1.** Example of Indonesia puppet

## 2.2 Depth Sensor

The Kinect has successful in much research, including human body movement detection, tracking, and action recognition, and playing games with Xbox. The one sensor of the Kinect is depth sensor, which based on the light coding method that use deformation of projected infrared patterns on object surfaces to measure depth [17]. The Kinect depth sensor has become a versatile tool in various study due to its low cost an easy to use, yet its implementation remains a black box largely [18]. Moreover, Kinect sensor so provides dense, real-time depth measurement of indoor environment at high framerate. The sensor using structured infrared light and based on principle of active stereo. Because of low price, the device has been motivated re-research to exclude the Kinect tool in application, which used for solved problem about non-optical sensor although alternative depth sensors with comparable performance exist [19].

There is some research about the depth sensor in any field and combination by other methods. Kui et al. [20] present about first attempt at fusing data from inertial and vision depth sensor within the framework of hidden Markov model for the application of hand gesture recognition. The data fusion is general purpose in the sense that it can use for recognition of various body movements. Michael et al., [21] propose a high fidelity Kinect IR and depth image predictor and simulator that models the physics of the transmitter/receiver system, unique IR dot pattern, disparity/depth processing technology, and random intensity speckle and IR noise in the detectors. The model accounts for important characteristics of Kinect's stereo triangulation system, consist of depth shadowing, IR dot splitting, spreading, and occlusion, correlation-based disparity estimation between windows of measured and reference IR images, and subpixel refinement. It is capture match by camera Kinect and resulting grayscale image in 3D form. The depth consists of  $x$ ,  $y$ , and  $z$  value in the different range, which is describing based Kinect coordinate system in Fig. 2.



**Fig. 2.** The coordinate system of Kinect version 2

## 2.3 DBSCAN

Image segmentation is one of example computer vision and image processing and it is using the clustering method. Clustering is an important unsupervised learning tool and widely applied in some field knowledge. There are many methods for clustering including k-means, DBSCAN, and NCuts (Normalized Cuts) and another clustering method. DBSCAN is one of the most commonly used density-based clustering algorithms. This algorithm work based on a density-based notion of clusters, and able to generate clusters of arbitrary shapes efficiently. DBSCAN accomplishes the clustering process by extracting cluster sequentially and the algorithm requires as input two parameters, i.e., epsilon and minimum points [13]. Exceeds a given density threshold show region linker a cluster in this model. The

size of the neighborhood and the minimum point in a cluster is a two important parameter defined in DBSCAN that must direct. The process begins from a random point  $x$  and finds all the point which are density-reachable from  $x$  with respect to the neighborhood and minimum points which is  $x$  is main point and return same process. Therefore, there are two-parameter epsilon/neighborhood and minimum points in DBSCAN. It is obvious to note that no points are density-reachable from  $x$  when  $x$  is a border point in case of which DBSCAN begins with an unclassified point to repeat the same process. The function of two-parameter is finding the best value for all cluster so that gives the stable result in DBSCAN. Each point of database visiting by DBSCAN, possibly multiple times so the complexity of this algorithms is  $O(N^2)$  [14]. DBSCAN is the innovator of density-based clustering methods, which based on shape and noise.

The algorithm can extend to large datasets by reducing its time complexity using spatial index structures like R-trees for finding neighbors of a pattern [15]. Unlike another clustering, DBSCAN algorithm has advantages, which this density-based method can be separate superior identification and cluster that have size and form different when compared to centroid-based technique, this algorithm can arrange data points into separate 'noise' cluster if given point different to rest of the data set, and this algorithm is deterministic [16]. The implementation of DBSCAN can explain by Jianbing et al. [17], the research using this algorithm for real-time superpixel segmentation. The detailed algorithm of DBSCAN as follows.

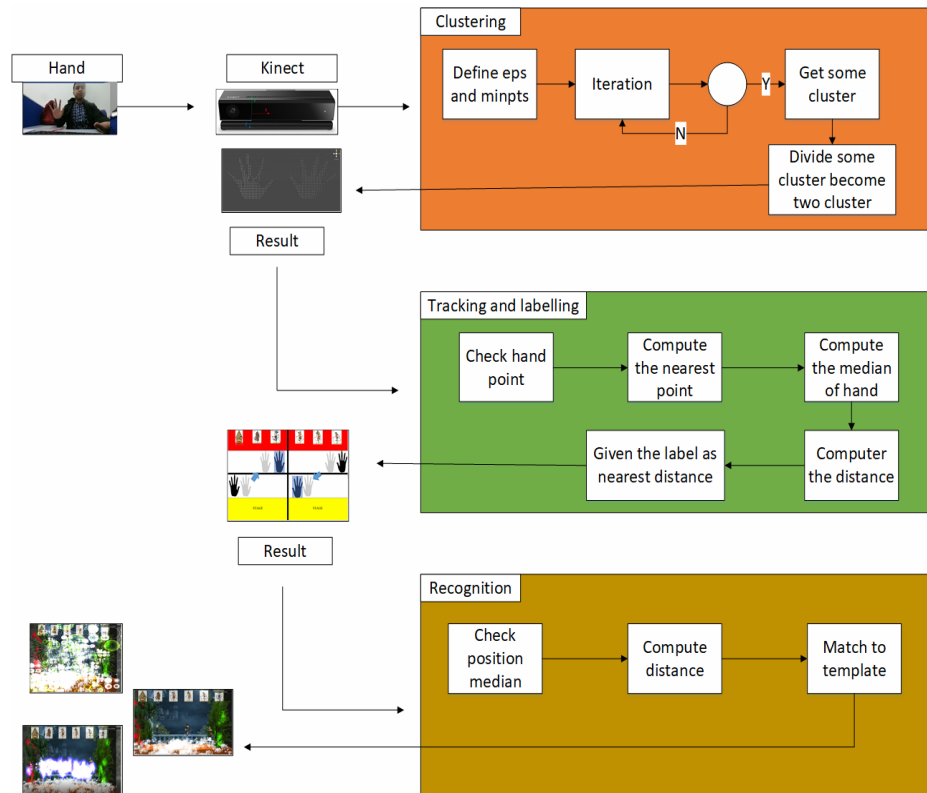
```

For handpoint count
  dist = 0, st = 0;
  if i > 0
    j = i - 1;
    while j >= 0
      dist = Manhattan distance / Euclidian distance of each
handpoint;
      if dist < thresh
        if dist < min
          min = dist;
          minpos = j;
        else
          minpos = j;
          st = 1;
      j--
    if st == 1
      data cluster by array insert into list data cluster;
    else
      identifier++;
      j = 0;
      identifier insert into list data cluster;
  else
    "0" insert into list data cluster;
return list data cluster
end.

```

### 3 Proposed Method

The system there is three processes including clustering, tracking labeling, and recognition. The architecture system in our research (see in Fig. 3). The clustering is the first step in the process so it is like pre-processing. There are many methods in the clustering but in research only using DBSCAN (Density-based spatial clustering of applications with noise) method. The process begins from point of the hand, and then define optimal epsilon and minimum points. The optimal epsilon is the diameter of the circle in DBSCAN and minimum points are a minimum count of centroid inside the circle. The process will iterate until all data already clustering so the results are some cluster. From some clustering system will be divided two cluster match as a count of the data. Finally, data is clustering into right and left hand. The purpose of clustering for playing puppet more than one for performance by puppeteers.



**Fig. 3.** The system architecture of proposed method

After clustering so hand continues in the next step, which is tracking and labeling. Tracking and labeling are part of pre-processing and there are two hands, right and left in our system. Tracking is processed to track hand exists and this is will follow hand move anywhere and must give a label in each hand. Moreover, tracking and labeling is one part cannot split in our system. In the tracking and labeling processing divide into two parts including look for median for each hand after clustering and check median or labeling. After finding each median so will check whether in the left hand or right hand.

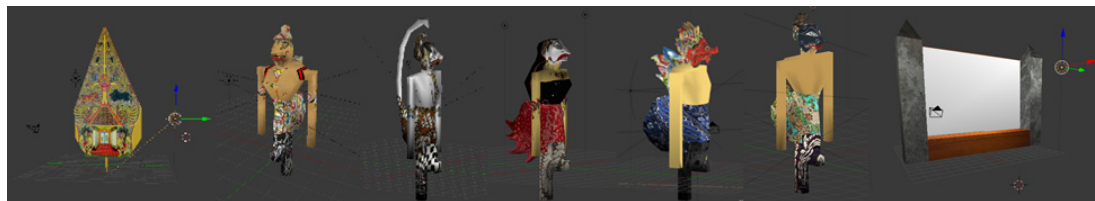
The last, Hand gesture recognition is an important part of the research because using for testing and evaluation of the system. Hand gesture recognition in our research using template matching and it is a clustering because of no need training data. The method only observation two values coordinate  $x$  and  $y$  so that there is three gesture resulted. The three-hand gesture recognition only implemented to two puppets. Thus, the other Indonesia puppet has one hand gesture recognition for show ability. There is three-hand gesture in this research including circle (clockwise and otherwise), up-down, and left-right. Each gesture saves some values to become data training and data testing. The goal is each gesture have each particle system. Hand gesture recognition contains a point of gesture, clustering, tracking, labeling, hand gesture recognition, and show the results of recognition. The overall process in the system using depth data which depth data is one of part from Kinect.

#### 4 Experiments and Results

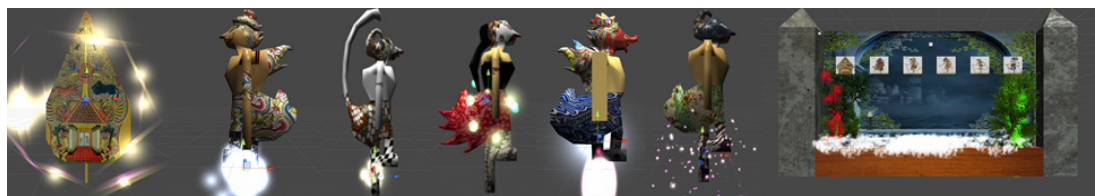
The experiment and results including system implementation, the threshold in DBSCAN, gesture recognition and discussions. The system implementation, we want to describe our system including creating each character before and after use particle system. Moreover, we will show animation effect and our performance results by using Kinect version 2. We want to get the best result in clustering of left and right hand so that we examine threshold from 5 until 100 in DBSCAN. This method will be examined in three condition distance including 50-100 cm, 100-200 cm, and over 200 cm. We also want to check gesture recognition in our performance for animation play (e.g. ability of each puppet) consists of gesture 1, gesture 2, and gesture 3. To check our results, we use precision, recall, and accuracy in each gesture. And last, we want to discuss our work from related works including create design, character, movement, and story.

#### 4.1 System Implementation

The System implementation consists of 3D model and particle system whereas we have six models of Indonesia puppet object including “Gunungan”, “Rama”, “Anoman”, “Sinta”, “Rahwana”, “Laksma” and one model of stage. We also show the game interface from multimedia perspective. The 3D model of implementation there are seven objects and each object has different texture, shape and particle system. For detail of each object and its particle system (see Fig. 4).



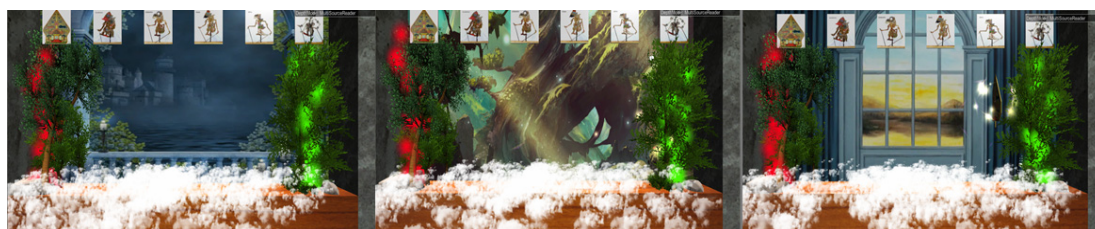
(a) without particle system



(b) with particle system

**Fig. 4.** The character of Indonesia puppet consists of (from left to right) “Gunungan”, “Rama”, “Anoman”, “Sinta”, “Rahwana”, “Laksma”, a stage object

Interface implementation has six interfaces in the system. Interface implementation consists of background change and threshold value. Interface implementation has two-part (see Fig. 5). In Fig. 5(a) is the first condition of performance there is the only background, particle system and Indonesia puppet choose the menu and the background are changed which is when the user takes a “Gunungan” object so the background is changed. Fig. 5(b) when distance value between 50 cm and 100 cm so Indonesia puppet get fight ability as shown above. Next, shown when distance value between 100 cm and 200 cm so Indonesia puppet showed animated as a hand moving, head moving and others. The last, distance more than 200 cm shown ability each Indonesia puppet there are five Indonesia puppet can show ability.



(a) changed the background



(b) action of gesture recognition based on distance consists of (from left to right)

**Fig. 5.** Game Interface fighting effect, talking action, and show ability

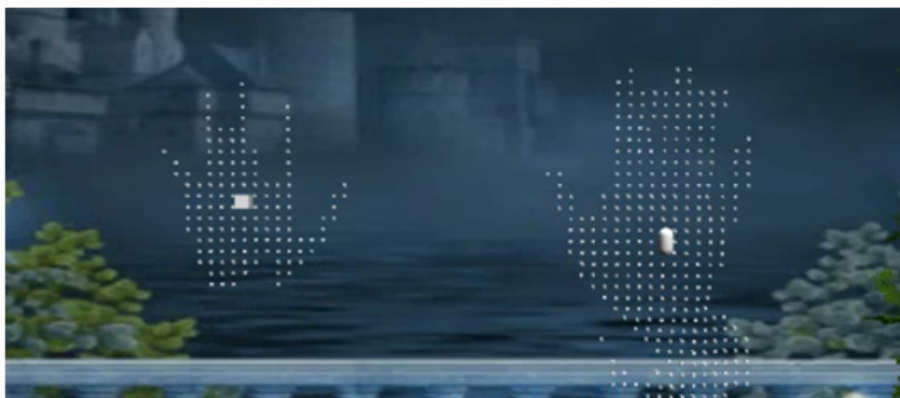
#### 4.2 Threshold in DBSCAN

There are three conditions of the distance and the goal is to look for the best threshold for three conditions. Distance is the distance between median1 and median2 which both contains x, y, z value, so we must use Euclidean distance to get the distance. In this research, we are looking for the best threshold that stables in three conditions of distance. The threshold we begin from 5 until 100 and each threshold contains three conditions. The result of the threshold value in DBSCAN in each median and each distance (see Table 1).

**Table 1.** Result of threshold in each median and each distance

Threshold	Median1 Distance1	Median1 Distance2	Median1 Distance3	Median2 Distance1	Median2 Distance2	Median2 Distance3
5	false	false	false	false	false	false
10	false	false	false	false	false	false
15	true	true	true	false	false	false
20	false	false	false	false	true	true
25	true	true	true	true	true	true
30	true	false	false	true	true	true
35	false	true	true	false	true	true
40	false	true	true	false	true	true
45	false	true	true	false	true	true
50	false	true	true	false	true	true
55	false	true	true	false	true	true
60	false	true	true	false	true	true
65	false	true	true	false	true	true
70	false	true	true	false	true	true
75	false	true	true	false	true	true
80	false	true	true	false	true	true
85	false	true	true	false	true	true
90	false	true	true	false	true	true
95	false	true	true	false	true	true
100	false	true	true	false	true	true

There are three categories in each threshold for search the best threshold. The best threshold means can different three categories distance. First is a distance1 with a value between 50 cm and 100 cm. Second, the distance2 by value 100 cm and 200 cm, and the last is distance3 by a value more than 200 cm. From the result in Table 1, there are 20 kinds of threshold and each threshold has three conditions that depend on the distance between right and left hand. Median of two hands can appear in middle/ center of each hand (see Fig. 6).



**Fig. 6.** Median position of the hand

### 4.3 Gesture Recognition

Each person will be practiced three threshold value and each threshold value has different performance. The first distance effect is fighting effect, the second threshold is playing animation and the third threshold is Indonesia puppet show ability. The setting of the puppet game in virtual reality using Kinect can describe in Fig. 7. We have 20 persons to do hand gesture recognition and each person does five times experiment of hand gesture recognition. The hand gesture recognition has involved 20 respondents or participants in doing this experiment. The participants not only come from computer science and Indonesia but also various knowledge discipline and country. The result of accuracy we are using precision, recall, and accuracy. First, we must define true positive (TP), true negative (TN), false positive (FP), and false negative (FN) value. After finding, so we can describe each value into con-fusion matrix. In our research, TP is cased in which we predicted yes (user do gesture), and system show ability of Indonesia puppet. TN define when we predicted no gesture, and the system does not show the ability of Indonesia puppet. FP is when we predicted yes (user do gesture), but the system not shows the ability of Indonesia puppet. The last, FN is when we predicted no gesture, but the system show ability of Indonesia puppet. The result of precision, recall, and accuracy of gesture recognition in each median and each gesture can show in Table 2.



**Fig. 7.** Setting of the puppet game in virtual reality using Kinect

**Table 2.** Result of precision, recall, and accuracy in each median and each gesture

Median	Gesture	Precision (%)	Recall (%)	Accuracy (%)
Median1	Gesture1	97%	98%	97%
	Gesture2	94%	97%	93%
	Gesture3	95%	94%	91%
Median2	Gesture1	100%	100%	100%
	Gesture2	99%	99%	98%
	Gesture3	96%	96%	96%

### 4.4 Discussion

The design in our work including character, stage, and performance. The character of the puppet is important thing since one with another character must be different. Dahlan et al. [1], propose a puppet called “wayang kulit” from Malaysia. It is little different with our puppet from Indonesia especially its form. In Malaysia, the puppet is looked human face [12-13] but Indonesia’s puppet-like abstract face. We are learning about 3D form and inverse kinematic (IK) from Malaysia’s puppet and implementing into Indonesia’s puppet which is inverse kinematic (IK) method for make hand of our puppet can move. Our stages based on puppet game in Indonesia including prolog, dialog, and epilog. The creation of character based on [14-15] since one of popular puppet and we also add skeleton inside for animation effect and particle system. In our work, we creating six characters based on story and designing different from each character. In the movement [16], we are implementing based on hand tracking and clustering of hand



using DBSCAN method and hand recognition for release ability of each puppet. The DBSCAN method will give the best clustering result since be affected threshold value. The story of puppet [3] is taken from popular story in Indonesia and the genre is romance.

## 5 Conclusion

The implementation of the system consists of some main step. Creating a 3D object about Indonesia puppet design, movement, storytelling, and character from starting development the program. Moreover, for multimedia purposed and people interesting to see the performance so each Indonesia puppet character added particle system inside game. In other hands, the process from hand recognition begin from capture the hand from Kinect and clustering between right hand and left hand using DBSCAN method so that movement of Indonesia puppet no cause ambiguous when puppeteers play it. The DBSCAN method work based on neighborhood values one centroid with other centroid and magnitude of the circle. From result of evaluation, the best threshold for DBSCAN is 25 because cluster two hand correctly. After clustering of hand, hand will track and the object will follow the hand to any position and will be given a label for each hand. Lastly, the gesture of hand will recognize into system based on template matching method. In the hand gesture recognition, system there is three results consist of near distance is 50-100 cm so Indonesia puppet will fight effect, medium distance is 100-200 cm so Indonesia puppet with does conversation, and long-distance is more than 200 cm so Indonesia puppet will show each ability [22]. In the recognition system, we get accuracy above 90% for median1 and median2. For future works, add more Indonesia puppet characters and change the story because in Indonesia have many characters and story about Indonesia puppet. Besides that, we need professional player advice to develop next version of Indonesia puppet performance.

## Acknowledgments

I say thank you to the Bureau of Foreign Cooperation Planning (BPCLN) give me the challenge to study in foreign and support. This work is done by join degree collaboration between Department of Computer Science Brawijaya University with Multimedia Information Networking (MINE) Lab Department of Computer Science and Information Engineering National Central University, Taiwan that lead by Professor Timothy K. Shih.

## References

- [1] D.A. Ghani, M.S.B.A. Ishak, Preserving wayang kulit for future generations, *IEEE Multimedia* 18(4)(2011) 70-73.
- [2] D.B.A. Ghani, Seri Rama: Converting a shadow play puppet to street fighter, *IEEE Computer Graphics and Applications* 32(1)(2012) 8-11.
- [3] S. Kusumanugraha, A. Ito, K. Mikami, K. Kondo, An analysis of Indonesian traditional “Wayang Kulit” puppet 3D shapes based on their roles in the story, in: *Proc. 2nd International Conference on Culture and Computing, Culture and Computing, 2011.*
- [4] L. Yang, L. Zhang, H. Dong, A. Alelaiwi, A. ElSaddik, Evaluating and improving the depth accuracy of Kinect for Windows v2, *IEEE Sensors Journal* 15(8)(2015) 4275-4285.
- [5] H.M. Jais, Z.R. Mahayuddin, H. Arshad, A review on gesture recognition using Kinect, in: *Proc. 5th International Conference on Electrical Engineering and Informatics: Bridging the Knowledge between Academic, Industry, and Community, ICEEI 2015.*
- [6] E. Yohannes, F. Utamingrum, Building segmentation of satellite image based on area and perimeter using region growing, *Indonesian Journal of Electrical Engineering and Computer Science* 3(3)(2016) 579-585.

- [7] T. Jyothirmayi, K.S. Rao, P.S. Rao, C. Satyanarayana, Image segmentation based on doubly truncated generalized Laplace mixture model and K means clustering, *International Journal of Electrical and Computer Engineering* 6(5)(2016) 2188-2196.
- [8] F.A. Mufarroha, F. Utamingrum, Hand gesture recognition using adaptive network based fuzzy inference system and K-nearest neighbor, *International Journal of Technology* 8(3)(2017) 559-567.
- [9] M. Kazemi-Beydokhti, R. Ali Abbaspour, M. Mojarab, Spatio-temporal modeling of seismic provinces of Iran using DBSCAN algorithm, *Pure and Applied Geophysics* 174(5)(2017) 1937-1952.
- [10] B.S. Chandana, K. Srinivas, R.K. Kumar, Clustering algorithm combined with hill climbing for classification of remote sensing image, *International Journal of Electrical and Computer Engineering* 4(6)(2014) 923-930.
- [11] J.M. Dudik, A. Kurosu, J.L. Coyle, E. Sejdić, A comparative analysis of DBSCAN, K-means, and quadratic variation algorithms for automatic identification of swallows from swallowing accelerometry signals, *Computers in Biology and Medicine* 59(2015) 10-18.
- [12] J. Ahmad, Z. Jamaludin, Embedding interaction design in Wayang Kulit mathematics courseware, in: *Proc. 3rd International Conference on User Science and Engineering: Experience*, 2014.
- [13] K.K. Kia, Y.M. Chan, A study on the visual styles of wayang kulit Kelantan and its capturing methods, in: *Proc. 6th International Conference on Computer Graphics, Imaging and Visualization*, 2009.
- [14] A. Kurnianto, F. Limano, Visual representation of character of wayang kulit purwa in the wayang-based games: Case studies of Kurusetra and Mahabarat warrior games, in: *Proc. 1st International Conference on Game, Game Art, and Gamification*, 2016.
- [15] D.A. Ghani, Wayang kulit: digital puppetry character rigging using Maya MEL language, in: *Proc. 4th International Conference on Modeling, Simulation and Applied Optimization*, 2011.
- [16] I. Kusumaningtyas, G.O.F. Parikesit, Digital image analysis of wayang Kulit movements, in: *Proc. 8th International Conference on Information Technology and Electrical Engineering: Empowering Technology for Better Future*, 2016.
- [17] X. Ning, G. Guo, Assessing spinal loading using the Kinect depth sensor: A feasibility study, *IEEE Sensors Journal* 13(4)(2013) 1139-1140.
- [18] J. Wang, Z. Xiong, Z. Wang, Y. Zhang, F. Wu, FPGA design and implementation of Kinect-like depth sensing, *IEEE Transactions on Circuits and Systems for Video Technology* 26(6)(2016) 1175-1186.
- [19] M. Stommel, M. Beetz, W. Xu, inpainting of missing values in the Kinect sensor's depth maps based on background estimates, *IEEE Sensors Journal* 14(4)(2014) 1107-1116.
- [20] K. Liu, C. Chen, R. Jafari, N. Kehtarnavaz, Fusion of inertial and depth sensor data for robust hand gesture recognition, *IEEE Sensors Journal* 14(6)(2014) 1898-1903.
- [21] M.J. Landau, B.Y. Choo, P.A. Beling, Simulating Kinect infrared and depth images, *IEEE Transactions on Cybernetics* 46(12)(2016) 3018-3031.
- [22] Indonesia puppet system demo video. <<https://youtu.be/kM62ZY0pCA/>>, 2018 (accessed 18.03.02).