

Clustering of Human Hand on Depth Image Using DBSCAN Method

Ervin Yohannes¹, Fitri Utamingrum², Timothy K. Shih³

²Faculty of Computer Science, University of Brawijaya, Malang, Indonesia

^{1,3}Department of Computer Science and Information Engineering, National Central University, Taiwan

{ervinyohannes@gmail.com, f3_ningrum@ub.ac.id, timothykshih@gmail.com}

Received 05 August 2019; accepted 29 September 2019

Abstract. In recent years, depth images are popular research in image processing, especially in clustering field. The depth image can capture by depth cameras such as Kinect, Intel Real Sense, Leap Motion, and etc. Many objects and methods can be implemented in clustering field and issues. One of popular object is human hand since has many functions and important parts of human body for daily routines. Besides, the clustering method has been developed for any goal and even combine with another method. One of clustering method is Density-Based Spatial Clustering of Applications with Noise (DBSCAN) which automatic clustering method consists of minimum points and epsilon. Define the epsilon in DBSCAN is important thing since the result depends on those. We want to look for the best epsilon for clustering human hand in the depth images. We selected the epsilon from 5 until 100 for getting the best clustering results. Moreover, those epsilons will be testing in three distance to get accurate results.

Keywords: Human Hand, Clustering, Kinect, DBSCAN

1 Introduction

The hand is required for daily routines and many activities need a hand moving from one to another position such as drinking water, eating, and combing hair so these activities can be called the configuration-to-configuration movements [1]. Today, the immediate input device is using hand and it is the most important and easiest gesture and possibly the most perceptive interface for choice [2]. The human hand is part of human body very complex system and has function for communication and interaction [3]. The result of human hand part including gesture [4][5], fingerprint, motion, sign language, and pose. Many research is using hand in any issues and fields for reaching a goal without notice to its different which right and left hand. These issues can be found in development human-robot for specific activities such as dancing and cooking robot. The implementation of human-robot is using human hand gesture segmentation which the hand gestures in RGB-depth (RGB-D) images took by Kinect [6]. The meaning of pixels in Kinect depth images is the distance from the consistent object to the imaging flat of the Kinect camera [7]. Microsoft Kinect sensor has been widely used

in many applications from early launch. At last, Microsoft launching the new version of Kinect which 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 [8].

Clustering is an unsupervised arrangement technique generally utilized for grouping of remote detecting image [9]. The typical density-based clustering method is Density-Based Spatial Clustering of Applications with Noise (DBSCAN) which is capable to find clusters in various size or form and classify outliers precisely [10]. Commonly, DBSCAN is used in big data clustering and modified by K-Means algorithm so that get the better result since has increased accuracy in every iteration [11]. From those benefit, we propose the DBSCAN method for clustering human hand in-depth images. Our contribution including implementation and analysis of DBSCAN method, capture human hand by using depth camera (i.e. Kinect) and get the best clustering results.

2 Related Work

Our goal is getting the best result from one of the clustering methods since can clustering data precisely and accurately. Besides, Kinect is a popular depth camera and many research discuss its.

2.1 Kinect

The Microsoft Kinect is minimal price sensor that relates many components which consist of traditional RGB camera, a depth sensor including an infrared (IR) camera and projector, a microphone, and built-in motor [12]. One of implementation of Kinect camera is using in detection system [7] and the object is obstacles. It has benefit in the dark place since depth image is produced by infrared camera, so it's tough to illumination and complex scene. Besides, the study about Kinect already proposed and have several important testing including accuracy distribution, depth resolution, depth entropy, edge noise, and structural noise [8]. Those research is verifying that Kinect has many functions, contribution and simulation [3] than another depth camera. Depth sensor of Kinect was successfully applied in breathing and heart rate analysis which get best accuracy especially during various kind of breathing [13]. Moreover, RGB-depth can be applied into human hand gesture for human-robot interaction that improving NAO robot's performance of considering and explanation [6]. The type of sensor inside Kinect including RGB, depth, and infrared. The description of three sensors can be shown in Figure 1.

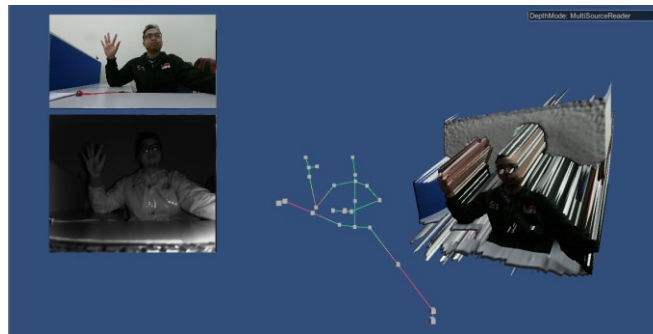


Fig. 1. Description of three sensors of the Kinect

2.1 Clustering

Clustering is a procedure of learning by interpretations and also an unsupervised method which without need training the dataset to produce a model. It is can manage to find of formerly unknown clusters within the data [14]. Moreover, clustering is an unsupervised description of shape development consists of experiments from dataset, data items or feature vector into set which may form clusters. Although, many clustering methods In many conditions, it issues usually being lectured by research academic in various field of study [15]. The amount of user-defined variables in density-based clustering is very great which DBSCAN there are two input parameters and has an important result on the feature of output [16]. The define two parameters epsilon and minimum points are very significant and even sometimes complex to the efficiency of clustering for the decreased set [17]. DBSCAN has successful in location prediction which the results show good deep source of water that used in the water care plant and getting the optimal cluster [18].

3 Research Method

In this part, we will explain two parts of our work. First, about our implementation in the Kinect consist of capture human hand, setting depth scale and testing the tool using Unity tool. Second, about DBSCAN algorithm and how to DBSCAN can clustering human hand which one right and left hand.

3.1 Kinect Implementation

We are using depth camera inside Kinect for capture human hand so that another sensor will hide (i.e. infrared and RGB). Several benefits using depth including not depend into light and depending on the depth scale since we can be setting it's by ourself. Our goal is to capture the human hand and make visible another human body so that we set our Kinect in the depth of 0.1 floats. And also, we are using Unity tool for testing our camera and for developing our work. So, it clear that inside Kinect haven't library or script about DBSCAN method since Kinect just capture our hand and results. Moreover, we must adjust the size of our work by defining width and height screen inside Kinect. The results of our depth image can be shown in Figure 2.



Fig. 2. The result of the depth image in Kinect

3.2 DBSCAN Clustering

DBSCAN has two important parameter epsilon and minimum point which effect to our results. The input is hand point that getting from depth image captured by Kinect. We are calculating each hand point by Euclidian distance as in equation (1).

$$d = \sqrt{\sum(x_i - y_i)^2} \quad , i = 1, 2, \dots, n \quad (1)$$

x and y are denoting as a position of hand point in the Kinect. We are defining epsilon from 5 until 100 and minimum point is zero. Actually, all epsilon will be getting same results but we are selecting the best epsilon which can clustering human hand both right and left hand exactly as match as our distance. Next, each epsilon will be checked which can produce two clusters. Not only epsilon will be checked but also minimum point since in our work define zero so it no need. Then, each cluster is calculated each median so we have median 1 and median 2. The last, each median will be labeled in our system. The block diagram of DBSCAN can be shown in Figure 3.

4 Results and Analysis

Evaluation including epsilon value evaluation in DBSCAN for clustering hand. In epsilon value in DBSCAN evaluation measure for evaluating the position of median 1 or median 2 matches by position or not. Means of position is the position of median 1 or median 2 matches in middle hand or not and so the condition of median 1 or median 2 stables in each frame of performance. Three conditions of the distance can be shown and Figure 4 and the goal is to look for the best epsilon for three distance. Distance is the distance between median 1 and median 2 which both contains x , y , z value so we must use Euclidian distance to get the distance.

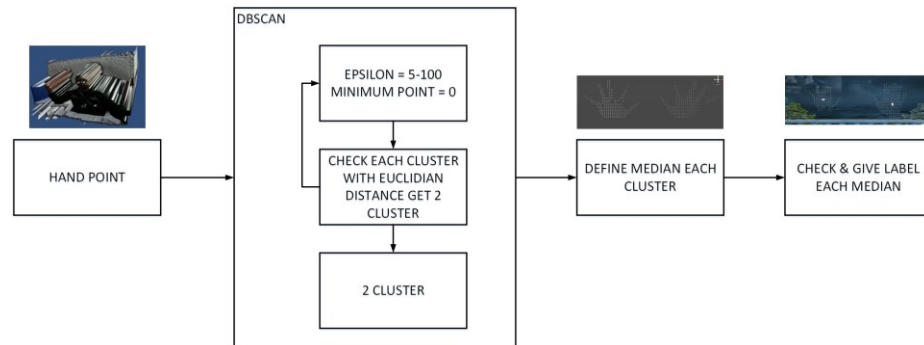
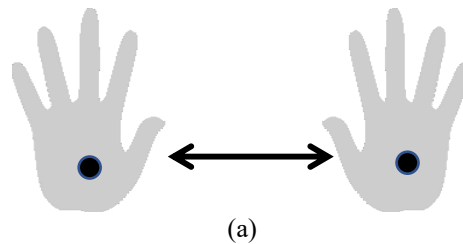


Fig. 3. The block diagram of DBSCAN



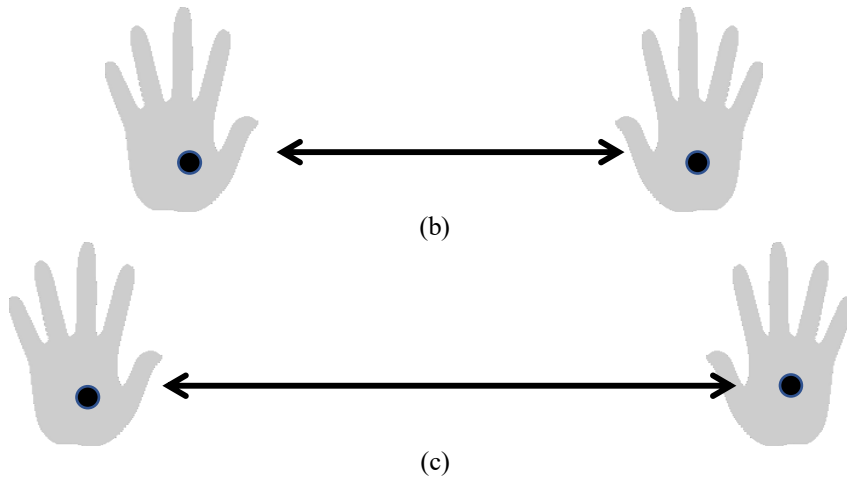


Fig. 4. Three conditions of distance (a) distance between 50 and 100 cm (b) distance between 100 and 200 cm (c) distance more than 200 cm

Table 1. Result of epsilon value in DBSCAN

eps.	Median1 Distance1	Median1 Distance2	Median1 Distance3	Median2 Distance1	Median2 Distance2	Median2 Distance3	acc.
5	false	false	false	false	false	false	0%
10	false	false	false	false	false	false	0%
15	true	true	true	false	false	false	50%
20	false	false	false	false	true	true	33%
25	true	true	true	true	true	true	100%
30	true	false	false	true	true	true	67%
35	false	true	true	false	true	true	67%
40	false	true	true	false	true	true	67%
45	false	true	true	false	true	true	67%
50	false	true	true	false	true	true	67%
55	false	true	true	false	true	true	67%
60	false	true	true	false	true	true	67%
65	false	true	true	false	true	true	67%
70	false	true	true	false	true	true	67%
75	false	true	true	false	true	true	67%
80	false	true	true	false	true	true	67%
85	false	true	true	false	true	true	67%
90	false	true	true	false	true	true	67%
95	false	true	true	false	true	true	67%
100	false	true	true	false	true	true	67%

In this research, we are looking for the best epsilon that stables in three conditions of distance. The epsilon we begin from 5 until 100 and each epsilon contains three distance and two medians value for testing case. The result of the epsilon value in DBSCAN can be shown in Table 1. There are three categories in each epsilon for search the best epsilon. The best epsilon means can different three categories distance. First is a distance with a value between 50 and 100 cm, it denoted “distance 1” in Table 1. Second, the distance by value 100 and 200, it denoted “ distance 2” and the last is

distance by a value more than 200, it denoted “ distance 3” in Table 1. From the result in Table 1, there is 20 kind of epsilons and each epsilon has three conditions which depend on the distance between right and left human hand. Median of two hands can appear in middle or center of each hand as Figure 4.

We can see in Table 1, median 1 have true at epsilon 15 but median 2 if false. Moreover, at epsilon 25 and 30 that both median 1 and 2 have true results. Median 2 many have values false than median 1 in the epsilon 1 result because of the distance of two hands so near. It means higher epsilon results so will more difficult to recognize. The number of false is higher than true results in the epsilon 1 results which count of false is 17 contains median 1 and median 2.

We can see in Figure 5, our results get zero accuracies in epsilon 5 and 10, which it can't cluster human hand. Then, we are reaching 50% in epsilon 15 and decrease to 33% in epsilon 20. The optimal results got when epsilon 25, which the accuracy reaches 100%. Next, the accuracy is static in epsilon 30 until 100 by accuracy 67%. The comparison of our work with another clustering method can be shown in [4], which the optimal accuracy by using K-means method reaches about 80 % and using SVM method reaches about 93.5%. Moreover, hybrid between K-Means and Particle Swarm Optimization (PSO) can reaches accuracy 86% [19]. In other hand, it is proved that DBSCAN method better than another method depending on the two-parameter including epsilon and minimum points.

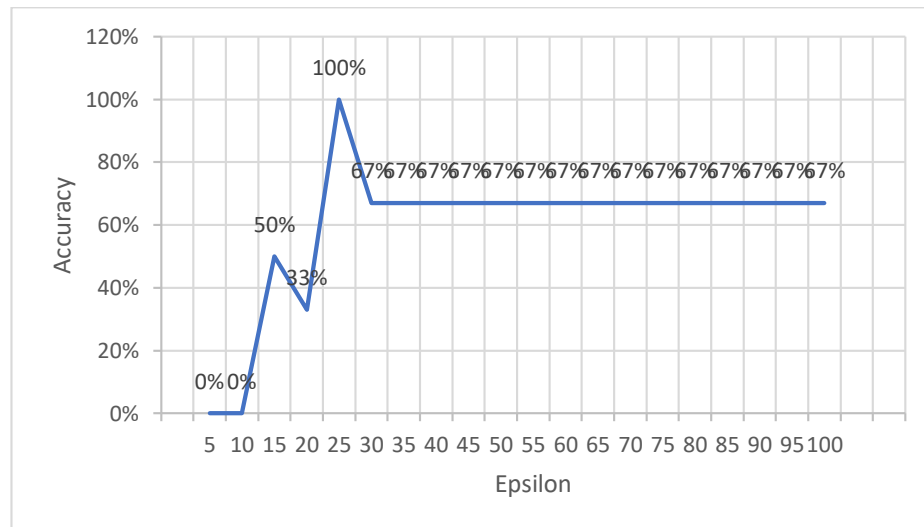


Fig. 5. Accuracy results in each epsilon

5 Conclusion

There is a sensor in the Kinect version 2 including NIR, Skeleton and depth sensor. Especially, in the newest version the depth sensor improvement result from the previous version. Moreover, the depth sensor is most popular in any field because of it more sensitive and effective than another sensor in the Kinect. We combine the depth sensor and DBSCAN method for clustering the left and right human hand. The DBSCAN can be a clustering of left and right hand at the epsilon value 25 because three conditions can work properly than another. Three conditions are the distance between the left and right hand. In the first distance, 50 – 100 cm, the epsilon value less than 25

and more than 30 cannot cluster the object. Second distance 100 – 200 cm, the epsilon more than 25 will be correct than less than 25. The last distance more than 200 cm, the epsilon more than 25 will be correct. It means DBSCAN can be clustered correctly by the big epsilon value and for small epsilon value, it will difficult. For future work, we want to develop this system into recognition system or another research in human-computer interaction field. In another hand, we can change another method for clustering or classification for a different human hand.

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

References

- [1] W. Chen, C. Xiong, and S. Yue, “On Configuration Trajectory Formation in Spatiotemporal Profile for Reproducing Human Hand Reaching Movement,” *IEEE Trans. Cybern.*, vol. 46, no. 3, pp. 804–816, 2016.
- [2] S. M. S. Shajideen and V. H. Preetha, “Human-computer interaction system using 2D and 3D hand gestures,” *2018 Int. Conf. Emerg. Trends Innov. Eng. Technol. Res. ICETIETR 2018*, pp. 1–4, 2018.
- [3] M. J. Landau, B. Y. Choo, and P. A. Beling, “Simulating Kinect Infrared and Depth Images,” *IEEE Trans. Cybern.*, vol. 46, no. 12, pp. 3018–3031, 2016.
- [4] D. A. Maharani, H. Fakhruroja, Riyanto, and C. Machbub, “Hand gesture recognition using K-means clustering and Support Vector Machine,” *ISCAIE 2018 - 2018 IEEE Symp. Comput. Appl. Ind. Electron.*, pp. 1–6, 2018.
- [5] F. A. Mufarroha and F. Utaminigrum, “Hand gesture recognition using adaptive network based fuzzy inference system and K-nearest neighbor,” *Int. J. Technol.*, vol. 8, no. 3, pp. 559–567, 2017.
- [6] Z. Ju, X. Ji, J. Li, and H. Liu, “An integrative framework of human hand gesture segmentation for human-robot interaction,” *IEEE Syst. J.*, vol. 11, no. 3, pp. 1326–1336, 2017.
- [7] T. Wang, L. Bu, and Z. Huang, “A new method for obstacle detection based on Kinect depth image,” *Proc. - 2015 Chinese Autom. Congr. CAC 2015*, pp. 537–541, 2016.
- [8] L. Yang, L. Zhang, H. Dong, A. Alelaiwi, and A. El Saddik, “Evaluating and improving the depth accuracy of Kinect for Windows v2,” *IEEE Sens. J.*, vol. 15, no. 8, pp. 4275–4285, 2015.
- [9] E. Yohannes and F. Utaminigrum, “Building segmentation of satellite image based on area and perimeter using region growing,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 3, no. 3, pp. 579–585, 2016.
- [10] M. Li, D. Meng, S. Gu, and S. Liu, “Research and improvement of DBSCAN cluster algorithm,” *Proc. - 2015 7th Int. Conf. Inf. Technol. Med. Educ. ITME 2015*, pp. 537–540, 2016.
- [11] D. Sengupta, S. S. Roy, S. Ghosh, and R. Dasgupta, “Modified K-Means Algorithm for Big Data Clustering,” *Proc. - 2017 Int. Conf. Comput. Sci. Comput. Intell. CSCI 2017*, pp. 1443–1448, 2018.
- [12] N. M. DiFilippo and M. K. Jouaneh, “Characterization of Different Microsoft Kinect Sensor Models,” *IEEE Sens. J.*, vol. 15, no. 8, pp. 4554–4564, 2015.
- [13] A. Procházka, M. Schätz, O. Vyšata, and M. Vališ, “Microsoft Kinect visual

- and depth sensors for breathing and heart rate analysis,” *Sensors (Switzerland)*, vol. 16, no. 7, pp. 1–11, 2016.
- [14] D. Jain, M. Singh, and A. K. Sharma, “Performance enhancement of DBSCAN density based clustering algorithm in data mining,” *2017 Int. Conf. Energy, Commun. Data Anal. Soft Comput. ICECDS 2017*, pp. 1559–1564, 2018.
- [15] G. . Nandana, S. Mala, and A. Rawat, “Hotspot Detection of Dengue Fever Outbreaks Using DBSCAN Algorithm,” *2019 9th Int. Conf. Cloud Comput. Data Sci. Eng.*, pp. 158–161, 2019.
- [16] A. Sharma and A. Sharma, “KNN-DBSCAN: Using k-nearest neighbor information for parameter-free density based clustering,” *2017 Int. Conf. Intell. Comput. Instrum. Control Technol. ICICICT 2017*, vol. 2018-Janua, pp. 787–792, 2018.
- [17] X. Lin and W. Han, “Opinion leaders discovering in social networks based on complex network and DBSCAN Cluster,” *Proc. - 14th Int. Symp. Distrib. Comput. Appl. Business, Eng. Sci. DCABES 2015*, pp. 292–295, 2016.
- [18] M. Perumal and B. Velumani, “Design and development of a Spatial DBSCAN Clustering framework for location prediction- An optimization approach,” *Proc. 3rd Int. Conf. Commun. Electron. Syst. ICCES 2018*, no. Icces, pp. 942–947, 2018.
- [19] Y. A. Auliya and W. F. Mahmudy, “Land Clustering for Potato Plants Using Hybrid Particle Swarm Optimization and K-Means Improved by Random Injection,” vol. 4, no. 1, pp. 42–56, 2019.