Dynamic Hand Gesture Recognition on 3D Virtual Cultural Heritage Ancient Collection Objects Using *k*-Nearest Neighbor

Adri Gabriel Sooai, *Member, IAENG,* Khamid, Kayo Yoshimoto, Hideya Takahashi, Surya Sumpeno, *Member, IAENG,* and Mauridhi Hery Purnomo, *Member, IAENG*

***Abstract*—This paper discusses on how to prepare a specific dynamic hand gesture, modeling and testing it to interact with 3D virtual objects of cultural heritage ancient collection. Those virtual objects prepared to avoid damage on the original one. Several kinds of research work for recreation or reactivating ancient heritage for educational purposes can take place using it. The dynamic hand gesture detected using hand movement sensor. We recorded ten specific dynamic hand gesture that stands for the interaction between museum visitors and the ancient collection chosen for the test. All ten gestures consist of fingers tips coordinates, palm, and wrist movement. A Total of 14474 rows in 30 features forming fingers and palm movements information. Those gestures namely: pick-up, sweep from right to left, sweep from left to right, grab from above, grab from the right side, pinch from above, pointing, scooping, push and picking. We train ten dynamic hand gestures using *K*-NN classifier and using different distance metric namely Cosine, Euclidean and Cubic. The best result of trained model reaches 99.3% accuracy. Later, we use the new hand gesture to test the trained model. It consists of 15000 rows of fingers coordinates in 30 features. The results show that from all ten gestures, there are four gestures reach recognition accuracy more than 92%. One gesture reaches 100%, two gestures on 82% and 89% and three gestures below 64%. The gesture which reaches high accuracy in training and testing consider selected for default model.**

***Index Terms*—dynamic-hand-gesture-recognition, *k*-nearest-neighbor, 3D-virtual-object, cultural-heritage, gaussian-mixture-model.**

I. Introduction

THE museum is a place to get a variety of information about the cultural heritage ancient collection/CHAC. However, due to aging and prone to damage, some of those artifacts are forbid to touch. There are many ways to preserve those collections and give the visitor access to exploring it. 3D virtual reality is one of it. Those artifacts can be preserved by reconstructs into the 3D virtual object[1]. However, there is another problem on how to interact in an immersive way between people and those virtual objects. An earlier research conducted to give an environment consists of 3D virtual objects. The user can explore the environment using a head-mounted display in passive mode[1], [2]. In addition to enhancing the application is by equipping it with a hand gesture sensor. Therefore, the user can interactively explore the environment as in the real environment. However, there is a significant difference between exploring artifacts in the 3D virtual and real environment. There must be a guide on how user treats those virtual artifacts or act freely. A free movement in a virtual environment is possible to equip with a head-mounted display and hand sensor as mention before. Furthermore, the user can walk freely inside it, they can navigate on foot. However, the size of the exploring space limited to the range of movement sensor[3], [4]. Another simple idea is to provide an interactive system that is not so complicated but can give immersive experience to the user, such as touching experience[5]. In a small space or virtual environment, what kind of gesture can suite the interaction? The gestures must have predefined on the first place, so it will give best immersive interaction to the user. In a cultural heritage ancient collection, there are various of objects. Each object with its designation. There are relics, which usually used for religious activities. Various forms of statues that symbolize human civilization. The statues selected in this study consist of imitations of animals, gods, worship monuments and the trinkets to perform daily activities. There are masks for traditional dancing and devices for cooking or storing food, an open water container for washing hands or face, and much more. The various objects need a unique movement of its own. Not all movements in real environments possible to recreate in virtual environments.

This is due to the limitations of motion sensor devices used. In this study, we are using leap motion[6] sensor to detect several gestures. However, there is some limitation still exist in using this sensor. For example, when we need to create an interaction in a fountain, where people can mimic washing their hands virtually. The sensor cannot detect an overlapping finger. Or when a user tries to mimic a scoop from a water bowl or pond using their palm, even it right or left handed, it cannot detect accurately. The problem is when the hand palm rotates by it wrist upside down, right hand facing down can turn to left-hand vice versa. This is a problem that we need to choose which gesture suitable for the immersive interaction. This paper discusses the preparation to set several models of dynamic hand gestures. Some of the best models will selected. Then it prepared for an immersive interaction. Ten basic hand movement prepared for this study.

Manuscript submitted December 5, 2017; revised June 11, 2018. This work was supported by the Postgraduate Scholarship from Indonesian Ministry of Higher Education and the Enhancement International Publication Program.

A. G. Sooai, S. Sumpeno, M. H. Purnomo are with the Department of Electrical Engineering and Computer Engineering, Faculty of Electrical Technology, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia (e-mail: adri14@mhs., surya@, hery@{ee.its.ac.id}).

H. Takahashi, K. Yoshimoto are with the Dept. of Electrical and Information Engineering, Osaka City University, Osaka, Japan (e-mail: hideya@elec., yoshimoto@{eng.osaka-cu.ac.jp}).

Khamid is with the Engineering Faculty, Universitas Wahidiyah, Kediri, Indonesia (e-mail: perumsat@gmail.com).

A. G. Sooai is with the Engineering Faculty, Universitas Katolik Widya Mandira, Kupang, Indonesia (e-mail: adrigabriel@gmail.com).

(Advance online publication: 28 August 2018)

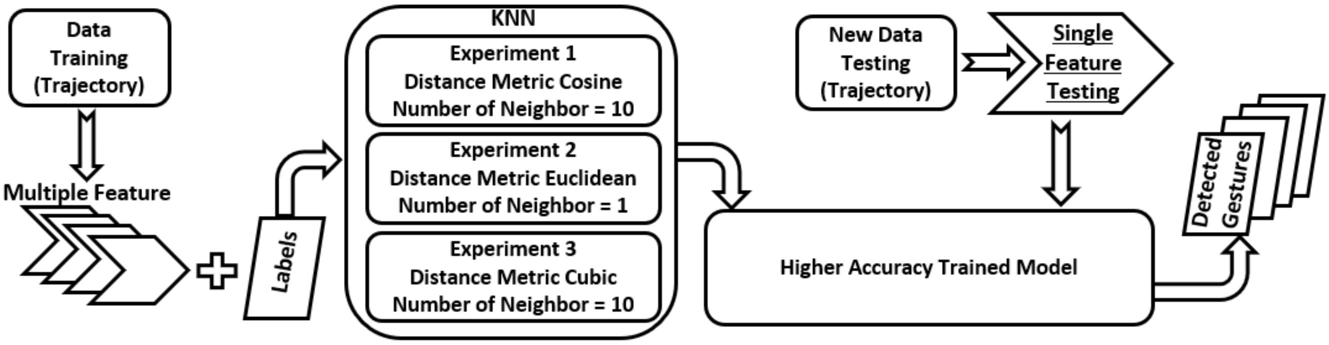

Fig. 1. The dynamic hand gesture recognition model for 3D virtual interaction on Cultural Heritage Ancient Collection. The best detected gesture according to the higher score in percentage, will be selected for application development.

Those gestures namely: pick-up, sweep from right to left, seep from left to right, grab from above, grab from the right side, pinch from above, pointing, scooping, push and picking. These gestures influenced by twenty[7] and ten gestures[8] from previous works. We opt out several gestures and add several new gestures which more reliable to perform immersive interactions.

II. Related Works

A. Preserving Cultural Heritage

The 3D virtual reality technology has been widely used to help preserve the various collections of cultural heritage[9]. The form also varies, ranging from converting the original form into digital. There are also arranged in a story and game[10]. Some researchers even create a virtual environment mimicking the original environment[11]. The goal is to revive the story around the various objects of cultural relics. Special interactions within a virtual environment can enrich it. Therefore, recurring activities is possible in a 3D virtual environment. Detail activities can be various, it is all to interact with virtual objects[12]. These interactions must set up first in an activity theme, and it is done in various games[13].

B. Gesture Recognition

This process aims to decide what gestures are proper in an application. The gesture recognition research using two-dimensional image tried to improve its accuracy. A non-negative matrix factorization combined with compressive sensing was introduced to gain more accuracy[14]. However, it is only to predict two-dimensional image. Several studies have done to test various of gesture models. There are six degrees of freedom models defined previously by researchers[7]. A couple of movements already included in our works, namely: sweep left and right. The left-over gestures cannot have included in this study because of not suitable for the CHAC environment. There is another variation of gesture that has been tested and spread to several types, namely: air writing by finger tips[15], [16], sign language[17], virtual reality object manipulation[18], arm movements in sport[19] and several common gestures for kinds of purposes[20], [21], [22]. The first group of gestures which deal with handwriting in the air and sign language uses hidden Markov model for gesture recognizing. The data characteristic from the recording process suites this algorithm. The second group which focused on sport, record arm movements. They were using Quaternion Dynamic Time Warping[19] to perform all calculation needed for the gesture recognizing. The last group using neural network and pattern matching in processing their calculation[23].

C. Feature Selection

The Basic step before using Machine Learning to recognize a gesture is to perform feature selection[24]. This process, if done with high accuracy will result in decreasing the computational time of machine learning. Several algorithms available to perform feature selection. From some way of performing feature selection, there is a good technique to use to do feature selection. It is known as the normal distribution or Gaussian Mixture Model(GMM). Two steps in preparing GMM are the mean and standard deviation. This will prepare training data for the classification learner. Further, it will determine which features best to be selected using GMM[25], [26], [27].

D. Classifier Learner

After the process for feature selection complete. All data are ready to enter the next step namely classification process. This process can be done using a chosen algorithm such as k-NN[28], [29], SVM, ENSEMBLE, NN and much more[30], [31], [32]. Not all data suits for all algorithm. A good observation of the distribution of data will determine the right algorithm to perform the classification process. The process has done with precision. It takes several repetitions to get the expected result.

E. 3D Stereoscopic Display

There are many options in which 3D Display technology and specification suites to the CHAC environment. It starts from the recent well-known head mounted display combined with virtual reality application. It gives a stereoscopic sight for the user that will make them feel wondering in a realm of another world. The main purposes are only to resemble the real environment. Even though this device should not be using in a relatively long period because it can cause strain in the users eyes[33]. The strain happens when a position of negative parallax[34]occurs too long in the scenery. Therefore, this 3D stereoscopic display is only another option to go with the hand interaction.

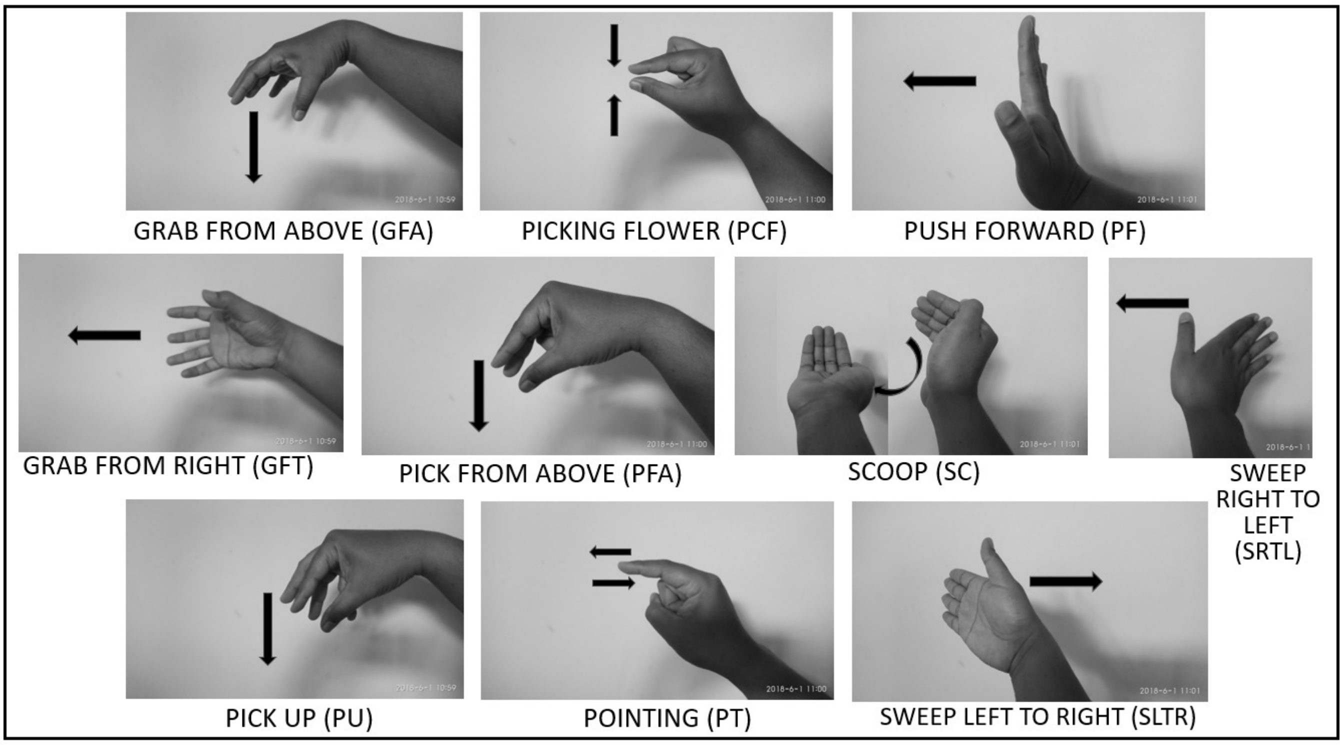

Fig. 2. The Ten gestures prepared for experiments

III. Method

The first stage is to record the data. A total of ten dynamic gestures recorded using a leap-motion sensor device. The result consists of hand gestures, in the form of coordinates of the five fingertips, palms, wrist and rotation. The data will then store in comma separated value (*.csv) format for two groups: training data and test data. Results of data recording cannot directly enter the preprocessing step because it holds some static coordinates. These static coordinates must opt out. This process aims to reduce learning time using the *k*-nearest neighbor classifier. Static data generated by stagnating motion, or when the hand does not make a movement for a while. This data found by examining the time-marking columns with 30 feature columns. In which are similar value for about one to two seconds. After removing static data, each dynamic gesture will have about 2000 row. It consists of 30 dimensions or features. Three experiments prepared in this study, as it can be seen in Fig 1. Each one will be using different distance metric. The first metric is Cosine / vector space distance[35], the second euclidean and third Cubic / minkowski. All experiment using the same steps from feature selection based on gaussian mixture model and classifier learner.

A. Feature selection

Not all features will be using in the classification learning process. Only features that have a standard deviation and a unique average selected for training. The GMM will be used to see which features will be selected as training data. The reason for using GMM is its ability to classify incomplete data[26]. There are steps to use the Gaussian distribution in checking features. The steps are: a) searching for the average value of a feature, b) finding the standard deviation value of a feature, c) finding the normal distribution value, followed by plotting data on a chart using the probability density function (PDF), and cumulative distribution function (CDF). This will make it easier to see the range of each feature checked. The three steps use the equation as follows:

$\bar{X}$= mean, n = dimension, x = data, with the equation as follows:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_1 = \frac{1}{n}(x_1 + x_2 + \ldots + x_n) \tag{1}$$

for $\{x_1, x_2, \ldots, x_n\}$ as the observed value of the sample and $\bar{X}$ is the mean of observations, then the standard deviation is:

$$S_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{2}$$

With μ as the average of the distribution σ, the standard deviation σ^2, as a variant, the PDF is:

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2, \pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma\ 2}} \tag{3}$$

The normal distribution f with mean μ and deviation σ , the CDF is:

$$F(x) = \Phi\Big(\frac{x-\mu}{\sigma}\Big) = \frac{1}{2}\Big[1 + erf\Big(\frac{x-\mu}{\sigma\sqrt{2}}\Big)\Big] \tag{4}$$

B. Classification Learner

The next step is making predictions, we need to calculate the similarity between two examples of available data. This is necessary so that we can find the most similar data samples in the training dataset. The prediction will use cosine, euclidean and cubic distance metric.

1) Experiment One: The first experiment will be using *k*-NN with cosine distance metric to train features. The number of neighbors for this metric is 10, with equal distance weight and calculate angle between two vectors of attributes A and B. Cosine distance is define as:

$$Cosine\ distance = \frac{cos^{-1}\left(\frac{\sum_{i=1}^{n} A_i^2}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}\right)}{\pi} \tag{5}$$

2) Experiment two: The second experiment will be using *k*-NN with euclidean distance metric. The number of neighbors for this metric is 1. This defined as the square root of the sum of squares difference between two numbers arrangement. Multi-dimensional Euclidean distance arranged in the following equation is define as:

$$d(p,q) = d(q,p) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \tag{6}$$

d = distance, p and q = measured values, i = index, and n = dimension. The *k*-NN prediction assumes that adjacent data have potential characteristics similarities. This is possible by forming a set of weights W, one for each adjacent data, defined by the relative proximity of each data neighbor to the reference point p. Determination of distance using two points denoted by x and x_i. The two entry points in the set p_i, it wrote as $D(x,p)$. An improvement for the *k*-NN algorithm done by measure the weight of each k neighbor by the distance between x_p query points, moves to r_j nearest neighbor. The equation algorithm written as follows

$$I(x_p) = \frac{argmax}{veV}\sum_{j=1}^{n} r_j \beta(y, i(x_j)) \tag{7}$$

where the weight is

$$r_j = \frac{1}{d(x_p, x_j)^2} \tag{8}$$

all the training data will be used for weighting calculation

$$F(x_q) = \frac{\sum_{i=1}^{k} w_i f(x_i))}{\sum_{i=1}^{k} w_i} \tag{9}$$

3) Experiment Three: The third experiment using minkowski for it distance metric. The number of neighbors for this metric is 10 with equal distance weight. Minkowski distance defined as:

$$D(X,Y) = \left(\sum_{i=1}^{n}|x_i - y_i|^p\right)^{\frac{1}{p}} \tag{10}$$

IV. Result and Discussion

Ten gestures prepared for the experiments. Starts with Grab from above(GFA), Grab from right(GFR), Pick up(PU), Picking flower(PCF), Pick from above(PFA), Pointing(PT),Push forward(PF), Scoop(SC), Sweep left to right(SLTR) and Sweep right to left(SRTL). Each gesture perform repeatedly five times and recorded using camera to validate the perfect movements. Several gesture seems a little bit similar between each other as it can be seen on Fig 2. The gesture GFA, PU and PFA. However, the width of the opening between the fingers is different for each gesture.

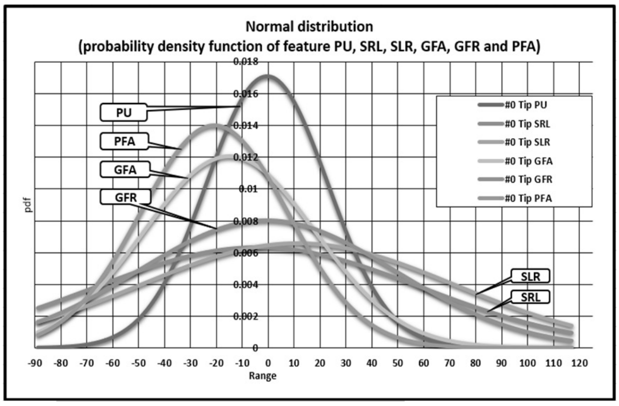

Fig. 3. The gaussian mixture model of six gestures

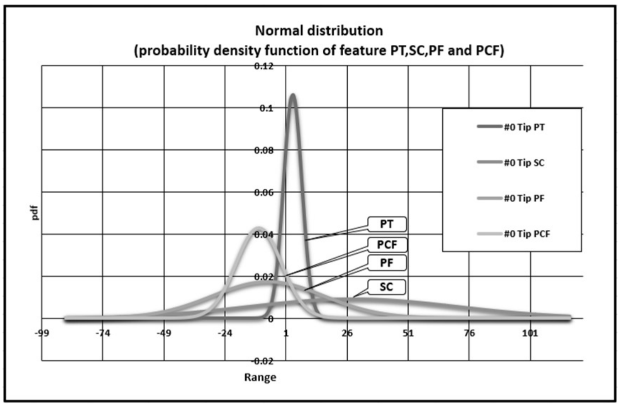

Fig. 4. The gaussian mixture model of four gesture

The preliminary process after gesture recording using hand gesture sensor is feature selection. We arrange all thirty features as predictor plus one response in a matrix shape. The feature selection process begins with calculating the mean and standard deviation of each feature. Plotting on a normal distribution chart. As it seen in Fig 3, there are six gestures shaping the normal distribution curves. It divided into two groups, the first group is gesture PU, PFA and GFA. They plotted on the top of chart area with wide range of its peak and standard deviation. The second is gesture GRF, SLR and SRL. Which almost overlapped and made slight differences between each other compared to the first group.

However, it still a potential feature for further calculation by classifier learner. The next figure as shown in Fig 4 consists of four gestures. The peak of each normal distribution curve spread far one from another. This is a good sign because the classifier learner can easily differ the class. Based on this gaussian mixture model, we continue our recognition process to the next step. Using the *k*-NN algorithm we decided to use the first and second point to compare and find the nearest neighbor between it. The search is using x as the coordinate from thumb and index finger. The results seen as several groups plotted on Fig 5. It consists of ten gestures, forming a specific pattern that distinct one and another. In addition to the three experiments that have been prepared, several tests were also performed and obtained the following results: Several *k*-NN models used to perform the calculation. The recognition accuracy using each classifier learner result are: Medium *k*-NN 98.2%, Coarse *k*-NN 94.0% and Weighted *k*-NN 99.0%. The dataset

TABLE I

THE CONFUSION MATRIX USING COSINE DISTANCE METRIC

		PREDICTED CLASS COSINE KNN									
		GFA	GFR	PU	PCF	PFA	PT	PF	SC	SLTR	SRTL
REAL CLASS	GFA	1754		10					6		
	GFR		934						5	8	2
	PU	46		1564				3			
	PCF				1286		1				
	PFA	2		2		1465	1				
	PT				1		1474				
	PF							1890	2	7	1
	SC		2					15	1596	22	11
	SLTR	1						8	4	1220	13
	SRTL							5	7	62	1040

TABLE II

THE CONFUSION MATRIX USING EUCLIDEAN DISTANCE METRIC

		PREDICTED CLASS									
		GFA	GFR	PU	PCF	PFA	PT	PF	SC	SLTR	SRTL
REAL CLASS	GFA	1753	1	11				1	3	1	
	GFR		939	1						3	6
	PU	5		1606			1	1			
	PCF				1286		1				
	PFA					1470					
	PT						1475				
	PF							1898	2	3	1
	SC	1	1			1		7	1623	12	1
	SLTR	1						5	1	1225	14
	SRTL							1		26	1087

TABLE III

TEH CONFUSION MATRIX USING CUBIC DISTANCE METRIC

		PREDICTED CLASS CUBIC KNN									
		GFA	GFR	PU	PCF	PFA	PT	PF	SC	SLTR	SRTL
REAL CLASS	GFA	1743		14			2		8	3	
	GFR		932	1					4	9	3
	PU	36		1569			3	3		2	
	PCF				1286		1				
	PFA	2		2		1463	3				
	PT				1		1475				
	PF							1891	4	9	
	SC	3	4				2	16	1592	22	7
	SLTR			1				11	1	1223	10
	SRTL		1	1			1	5	6	77	1023

consists of 14474 observations, 30 predictors, 1 response label, 10 response classes with 10-fold cross validation. The first two Medium and Coarse using euclidean distance metric and all using equal distance weight. However, Weighted k-NN using square inverse and failed to reach higher accuracy. Based on the results, we assume that using k-NN classifier is suitable enough to perform gesture recognition for CHAC environment.

A. Results of experiment one

The accuracy of first experiment can be seen on Table I explained that the confusion matrix shows that 98.3%. Overall there are none reach 100% prediction. There are three groups of accuracy. The first groups consist of five gestures reach accuracy 99% and more. The second consist of four gestures reach only 97 to 98% accuracy. Only one gesture stay lowest at 93%. The first five gestures are GFA, PF, PFA, PT and PF. The first gesture GFA total missed around 1%. Ten trajectory coordinates misses interpreted as PU and six as SC. A total 1287 trajectory coordinates of Gesture PF only misses one. The PFA gesture, from total 1470, 1465 predict correctly, two predict as GFA, two PU and one PT. The second three gesture are GFR, PU, and SLTR. The last gesture SRL only reach 93% accuracy prediction. It can be seen in the lower right corner of the Table I the number 62 is 3% of total 1114 data from SRTL feature.

B. Results of experiment two

The confusion matrix shows that 99.3% recognition accuracy gained. As it seen in Table II, the first gesture GFA total missed less than 1%, one observation missed interpreting as GFR, PF and SLTR, 3 observations for SC and 11 for PU. The second gesture GFR total missed at exactly 1%, one missed interpret as PU, 3 for SLTR and 6 for SRTL. The third and fourth gesture PU and PCF also missed interpret below 1%, it scattered from PU, PT, and PCF. However, for gesture PFA and PT reached 100% accuracy. The 1470 and 1475 observation successfully interpreted as is. There is a sharp distinction between peaks of two curves which split 20 points shows that it will recognize perfectly correct. The seventh gesture, PF, missed interpret far below 1%, there are 2 observations interpreted as SC, 3 as SLTR and 1 as SRTL. The eight-gesture SC missed interpreted 1.43%. There are 12 observations missed and interpreted as SLTR, 7 observations as PF, the rest 4 observation missed interpreted by 1 as GFA, GFR, PFA and SRTL. The ninth gesture SLTR missed interpreted by 1.71%. There are 14 observations missed interpreted as SRTL, 5 as PF and 1 as GFA and SC. The tenth gesture SRTL missed interpreted by 2.4%. There are 26 observations interpreted as SLTR and 1 as PF. Overall there are 112 missed interpreted observations from total 14474 observations. The recognition accuracy using KNN is 99.3%. This trained model then used to test new data. Each gesture tested against the trained model. The results vary, it spread into three groups as it seen in Fig 6. Gesture 4 until 10 reaches prediction accuracy above 80%. Gestures 6, 8 and 9 reach accuracy above 91%, and gesture 7 reach 100% accuracy. However, for gesture 2 and 3 only reach accuracy above 34% and below 37%. It predicted before using the gaussian mixture model and scattered chart. On the GMM, the curve of both gestures slightly overlapped one and another as it seen also in the scattered plot on cartesian axis on Table 3. The first gesture, Pick-Up only reaches 63.46% test accuracy. This happen because 36.24% detected as Grab from above gesture.

C. Experiment two scattered chart

The scattered chart as shown in Fig 5 explained that The GFA gesture majority plotted on the -x and +y axis, it spread around (-75,55) to (50,250). The GFR plotted evenly on -x to +x axis, and +y, it spread around (-80,75) to (60,155). This ranges shows good separable to make differences between gestures. For the other several gestures namely PU, PFA, PT, PF, and SC each of it made distinct pattern. However, for the last two similar gestures namely SLTR and SRTL, it creates larger part of an overlapped area compared to the other gestures. The scattered plotted coordinates of both gestures ranging from (-100,150) to (100,230) across the x and y axis. Overall, the feature selection performed well to support the classifier learner calculation.

D. Results of Experiment Three

The accuracy of third experiment can be seen on Table III explained that the confusion matrix shows that 98.1%. However, compare to the first experiment, there is one gesture correctly predicted 100% accuracy, it is PT. Besides the PT, there are three groups of accuracy. The first group consist of three gestures reach accuracy 99% and more. The second consists of four gestures reach only 97 to 98% accuracy.

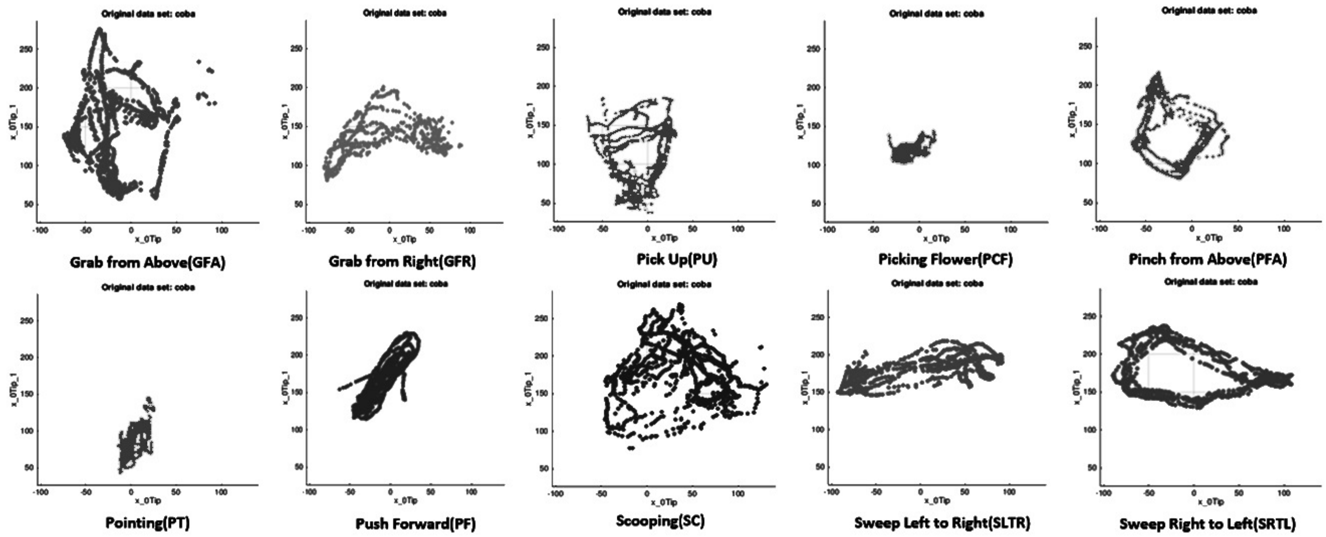

Fig. 5. The Scattered chart shows for each feature

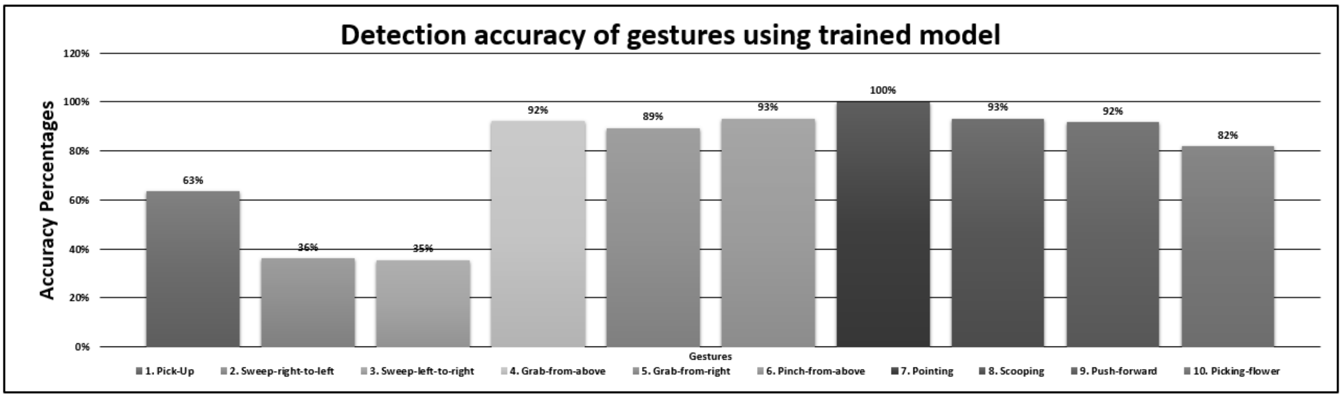

Fig. 6. The accuracy of tested model using new dataset

Only one gesture stay at the lowest at 92%. The first three gestures are PF, PFA, and PF. The first gesture PF total missed less then 1%. A Total Seven trajectories coordinates misinterpreted as GFA, PU and PT. A total 1287 trajectory coordinates of Gesture PF only miss one. The PFA gesture, from total 1470, 1465 predict correctly, two predict as GFA, two PU and one PT. The second three gesture are GFR, PU, and SLTR. The last gesture SRL only reach 93% accuracy prediction. It can be seen in the lower right corner of the Table III the number 77 is 7% of total 1114 data from SRTL feature.

V. Conclusion

Gestures for a specific task like in CHAC environment need to model accurately. A straightforward way to perform feature selection can be done using gaussian mixture model. As described on those three experiments, only experiment two reaches the highest accuracy in predicting dynamic gestures. The first and third experiments reach almost the same value. Both obtained a 98% accuracy rate with a very thin margin. However, from all three trials, the lowest value in accuracy was obtained in two overlapping motions. Both movements are SLTR and SRTL has been predicted before by gaussian mixture model in Fig 3. Overall, as it can be seen on Fig 6 from ten gestures tested using the trained model, there are eight gestures can perfectly recognize using the trained model. One gesture perfectly recognizes +and gain 100% accuracy. There are four gestures gain above 91% accuracy. Two gestures gain above 80% accuracy, one above 60% and two gestures failed the test. There must be an evaluation to replace similar gesture for the trained model. Two gestures show in the test that cannot recognize optimally and only get test accuracy below 50%. One gesture is also lower than 70% need to replace. Fine-KNN classifier learner shown the best performance in performing the training and testing process. It outnumbered the other distance metric on several KNN method and gain score at 99,3% with 10-fold cross validation.

Acknowledgment

The authors would like to thank Mr Ir. I Gusti Ngurah Made Soetedja,MM. for His contribution in supporting this research with the introduction of cultural heritage ancient collection. High appreciation was also presented to Ms.Phillepina Theedens, Mr Ir. Erwin "Carlo" Dasanov Tibuludji and Mr Atyanta Nika Rumaksari,M.Eng. for all their supports.

References

[1] A. G. Sooai, S. Sumpeno, and M. H. Purnomo, "User Perception on 3d Stereoscopic Cultural Heritage Ancient Collection," in *Proceedings*

of the The 2nd International Conference in HCI and UX on Indonesia 2016, ser. CHIuXiD '16. New York, NY, USA: ACM, 2016, pp. 112–119. [Online]. Available: http://doi.acm.org/10.1145/2898459.2898476

[2] A. G. Sooai, A. Nugroho, M. N. A. Azam, S. Sumpeno, and M. H. Purnomo, "Virtual artifact: Enhancing museum exhibit using 3d virtual reality," in *2017 TRON Symposium (TRONSHOW)*. Tokyo Midtown Hall [Midtown East B1F]: IEEE, Dec. 2017, pp. 1–5.

[3] I. G. A. Dharmayasa, S. Sumpeno, I. K. E. Purnama, and A. G. Sooai, "Exploration of prayer tools in 3d virtual museum using leap motion for hand motion sensor," in *2017 TRON Symposium (TRONSHOW)*. Tokyo Midtown Hall [Midtown East B1F]: IEEE, Dec. 2017, pp. 1–8.

[4] Y. Zheng, M. McCaleb, C. Strachan, and B. Williams, "Exploring a Virtual Environment by Walking in Place Using the Microsoft Kinect," in *Proceedings of the ACM Symposium on Applied Perception*, ser. SAP '12. New York, NY, USA: ACM, 2012, pp. 131–131. [Online]. Available: http://doi.acm.org/10.1145/2338676.2338713

[5] N. Rosa, W. Hrst, W. Vos, and P. Werkhoven, "The Influence of Visual Cues on Passive Tactile Sensations in a Multimodal Immersive Virtual Environment," in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, ser. ICMI '15. New York, NY, USA: ACM, 2015, pp. 327–334. [Online]. Available: http://doi.acm.org/10.1145/2818346.2820744

[6] Khamid, A. D. Wibawa, and S. Sumpeno, "Gesture Recognition for Indonesian Sign Language Systems (ISLS) Using Multimodal Sensor Leap Motion and Myo Armband Controllers Based-on Naive Bayes Classifier," in *2017 International Conference on Soft Computing, Intelligent System and Information Technology (ICSIIT)*. Denpasar Bali Indonesia: IEEE, Sep. 2017, pp. 1–6. [Online]. Available: https://doi.org/10.1109/ICSIIT.2017.42

[7] M. Chen, G. AlRegib, and B. H. Juang, "Feature Processing and Modeling for 6d Motion Gesture Recognition," *IEEE Transactions on Multimedia*, vol. 15, no. 3, pp. 561–571, Apr. 2013.

[8] W. Lu, Z. Tong, and J. Chu, "Dynamic Hand Gesture Recognition With Leap Motion Controller," *IEEE Signal Processing Letters*, vol. 23, no. 9, pp. 1188–1192, Sep. 2016.

[9] S. Rizvic, D. Pletinckx, and V. Okanovi, "Enhancing museum exhibitions with interactive digital content: Sarajevo city model interactive," in *2015 XXV International Conference on Information, Communication and Automation Technologies (ICAT)*. IEEE, Oct. 2015, pp. 1–5.

[10] G. Kontogianni and A. Georgopoulos, "A realistic Gamification attempt for the Ancient Agora of Athens," in *2015 Digital Heritage*, vol. 1. IEEE, Sep. 2015, pp. 377–380.

[11] F. Gabellone, I. Ferrari, M. T. Giannotta, and A. Dell'Aglio, "From museum to original site: A 3d environment for virtual visits to finds re-contextualized in their original setting," in *Digital Heritage International Congress (DigitalHeritage), 2013*, vol. 2. IEEE, Oct. 2013, pp. 215–222.

[12] M. T. Yang and W. C. Liao, "Computer-Assisted Culture Learning in an Online Augmented Reality Environment Based on Free-Hand Gesture Interaction," *IEEE Transactions on Learning Technologies*, vol. 7, no. 2, pp. 107–117, Apr. 2014.

[13] L. Chittaro and F. Buttussi, "Assessing Knowledge Retention of an Immersive Serious Game vs. a Traditional Education Method in Aviation Safety," *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, no. 4, pp. 529–538, Apr. 2015.

[14] Huiwei Zhuang, Mingqiang Yang, Zhenxing Cui, and Qinghe Zheng, "A Method for Static Hand Gesture Recognition Based on Non-Negative Matrix Factorization and Compressive Sensing," *IAENG International Journal of Computer Science*, vol. 44, no. 1, pp. 52–59, 2017.

[15] M. Chen, G. AlRegib, and B. H. Juang, "Air-Writing Recognition-Part I: Modeling and Recognition of Characters, Words, and Connecting Motions," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 3, pp. 403–413, Jun. 2016.

[16] ——, "Air-Writing Recognition-Part II: Detection and Recognition of Writing Activity in Continuous Stream of Motion Data," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 3, pp. 436–444, Jun. 2016.

[17] J. Gaka, M. Msior, M. Zaborski, and K. Barczewska, "Inertial Motion Sensing Glove for Sign Language Gesture Acquisition and Recognition," *IEEE Sensors Journal*, vol. 16, no. 16, pp. 6310–6316, Aug. 2016.

[18] Y. Jang, I. Jeon, T. K. Kim, and W. Woo, "Metaphoric Hand Gestures for Orientation-Aware VR Object Manipulation With an Egocentric Viewpoint," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 1, pp. 113–127, Feb. 2017.

[19] R. Srivastava and P. Sinha, "Hand Movements and Gestures Characterization Using Quaternion Dynamic Time Warping Technique," *IEEE Sensors Journal*, vol. 16, no. 5, pp. 1333–1341, Mar. 2016.

[20] N. Neverova, C. Wolf, G. Taylor, and F. Nebout, "ModDrop: Adaptive Multi-Modal Gesture Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 8, pp. 1692–1706, Aug. 2016.

[21] N. Rossol, I. Cheng, and A. Basu, "A Multisensor Technique for Gesture Recognition Through Intelligent Skeletal Pose Analysis," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 3, pp. 350–359, Jun. 2016.

[22] R. Xie and J. Cao, "Accelerometer-Based Hand Gesture Recognition by Neural Network and Similarity Matching," *IEEE Sensors Journal*, vol. 16, no. 11, pp. 4537–4545, Jun. 2016.

[23] F. Abedan Kondori, S. Yousefi, J.-P. Kouma, L. Liu, and H. Li, "Direct hand pose estimation for immersive gestural interaction," *Pattern Recognition Letters*, vol. 66, pp. 91–99, Nov. 2015.

[24] E. E. Bron, M. Smits, W. J. Niessen, and S. Klein, "Feature Selection Based on the SVM Weight Vector for Classification of Dementia," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 5, pp. 1617–1626, Sep. 2015.

[25] D. Wu, L. Pigou, P. J. Kindermans, N. D. H. Le, L. Shao, J. Dambre, and J. M. Odobez, "Deep Dynamic Neural Networks for Multimodal Gesture Segmentation and Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 8, pp. 1583–1597, Aug. 2016.

[26] Pierre Lorrentz, "Classification of Incomplete Data by Observation," *Engineering Letters*, vol. 18, no. 4, pp. 1–10, 2010.

[27] A. N. Rumaksari, S. Sumpeno, and A. D. Wibawa, "Background subtraction using spatial mixture of Gaussian model with dynamic shadow filtering," in *2017 International Seminar on Intelligent Technology and Its Applications (ISITIA)*. IEEE, Aug. 2017, pp. 296–301. [Online]. Available: https://doi.org/10.1109/ISITIA.2017.8124098

[28] Eiji Uchino, Kazuhiro Tokunaga, Hiroki Tanaka, and Noriaki Suetake, "IVUS-Based Coronary Plaque Tissue Characterization Using Weighted Multiple k-Nearest Neighbor," *Engineering Letters*, vol. 20, no. 3, pp. 1–6, 2012.

[29] Yun He and De-chang Pi, "Improving KNN Method Based on Reduced Relational Grade for Microarray Missing Values Imputation," *IAENG International Journal of Computer Science*, vol. 43, no. 3, pp. 1–7, 2016.

[30] A. Baldominos, P. Isasi, and Y. Saez, "Feature selection for physical activity recognition using genetic algorithms," in *2017 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, Jun. 2017, pp. 2185–2192.

[31] S. Fong, R. Wong, and A. V. Vasilakos, "Accelerated PSO Swarm Search Feature Selection for Data Stream Mining Big Data," *IEEE Transactions on Services Computing*, vol. 9, no. 1, pp. 33–45, Jan. 2016.

[32] J. Gui, Z. Sun, S. Ji, D. Tao, and T. Tan, "Feature Selection Based on Structured Sparsity: A Comprehensive Study," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 7, pp. 1490–1507, Jul. 2017.

[33] B. Keshavarz and H. Hecht, "Stereoscopic Viewing Enhances Visually Induced Motion Sickness but Sound Does Not," *Presence: Teleoperators and Virtual Environments*, vol. 21, no. 2, pp. 213–228, Apr. 2012.

[34] K. R. Martin Hasmanda, "The Modelling of Stereoscopic 3d Scene Acquisition," *Radioengineering*, vol. 21, no. 1, 2012.

[35] Dat Huynh, Dat Tran, and Wanli Ma, "Contextual Analysis for the Representation of Words," *IAENG International Journal of Computer Science*, vol. 41, no. 2, pp. 1–5, 2014.

Adri Gabriel Sooai received the B.E, M.E degree in Computer Science from Sekolah Tinggi Teknologi Indonesia in 2001 and masters in civil engineering with specialty in Remote Sensing and Satellite Digital Imaging from Institut Teknologi Sepuluh Nopember in 2006. He is pursuing His Ph.D. degree in Electrical Engineering Institut Teknologi Sepuluh Nopember Indonesia since 2015. His current research interest includes Human Computer Interaction, 3D-Stereoscopic with Virtual and Augmented Reality. He is a member of IAENG, and Student member in IEEE and ACM.

Khamid received the B.E, M.E degree in electrical engineering from Institut Teknologi Sepuluh Nopember in 1998, respectively, and completed the master program at Institut Teknologi Sepuluh Nopember in 2017. He became a research associate of the Dept. of Computer Science at the Universitas Wahidiyah of Kediri 2011. He is pursuing His Ph.D. degree in Electrical Engineering Institut Teknologi Sepuluh Nopember Indonesia since 2018. His current research interests include Pattern Recognition and Human Computer Interaction.

Kayo Yoshimoto received the B.E and M.E. degrees in mechanical engineering from Osaka University in 2009 and 2011, respectively, and completed the doctoral program at Osaka University (Graduate School of Medicine) in 2014. She became a research associate of the Dept. of Electric and Information Engineering at the Graduate School of Engineering of Osaka University since 2014. Her current research interests include Medical Engineering and Nursing Engineering.

Hideya Takahashi received the B.E., M.E., and Ph.D. degrees in electrical engineering from Osaka City University in 1982, 1984, and 1992, respectively. Since 1987, he has been a Faculty Member with Osaka City University, and since 2011, he has been a Professor with the Department of Electric and Information Engineering. His current research interests include interactive 3-D display, retinal projection display, wearable computers, and medical engineering. He is a member of SPIE and OSA.

Surya Sumpeno earned his bachelors degree in electrical engineering from ITS, Surabaya-Indonesia in 1996, and M.Sc. degree from the Graduate School of Information Science, Tohoku University, Japan in 2007 (email: surya@ee.its.ac.id). He earned doctor degree in Electrical Engineering from ITS, Surabaya, in 2011. He is with the Electrical Engineering Department, Institut Teknologi Sepuluh Nopember (ITS) Surabaya, Indonesia. His research interests include natural language processing, human computer interaction, and artificial intelligence. He is IAENG, IEEE, and ACM SIGCHI (Special Interest Group on Computer-Human Interaction) member.

Mauridhi Hery Purnomo earned his bachelor degree from Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia, in 1985, then his M.Eng., and Ph. D degrees from Osaka City University, Osaka, Japan in 1995 and 1997 respectively. He joined ITS in 1985 and has been a Professor since 2004. His current interests include intelligent system applications, electric power systems operation, control and management. He is an IAENG and IEEE Member.