3D Face Recognition without Using the Positional Relation of Facial Elements

Yoshihiro Sato and Yue Bao
Graduate School of Engineering, Tokyo City University, Setagaya, Japan
Email: {g1781807, bao}@tcu.ac.jp

***Abstract*—This paper proposed a three-dimensional face recognition method. It is different from the conventional face recognition methods. The three-dimensional shape of the face is acquired by using a Kinect sensor, and the face is recognized by using point pair feature. Kinect sensor performs skeletal recognition with the depth image to estimate the face region. A robust face recognition becomes possible by the method of using point pair feature. To confirm the effectiveness of the proposal method, at first, an experiment was conducted to determine a threshold for face identification. By the experimental result, a value matching 70% of the 3D point cloud of the face area was set as the threshold for face identification. Next, we registered the face model of 10 subjects (9 males, 1 female), and conducted an authentication experiment based on the determined threshold. As a result of the experiment, the authentication rate of 100% was obtained without an others acceptance and a principal denial (false rejection).**

***Index Terms*—face recognition, image processing, point pair feature, three-dimension, 3D point cloud**

I. Introduction

In recent years, the face recognition systems have been embedded into mobile phone, desktop of personal computer and smart phone. These systems also are well used in various application area such as e-governance, surveillance, room access system etc. Therefore, it is important for the face recognition system to have reliable and robust performance. As the conventional methods, a lot of face recognition systems had been proposed [1]-[9].

In 2D face recognition using 2D image, the face is identified by "shade of face" and "positional relation of facial parts", etc. Therefore, it is vulnerable to changes in brightness and face orientation, and the matching accuracy depends on the surrounding environment. Moreover, since it is possible to impersonate with printed pictures, there is also a security problem. In order to solve these problems of 2D face recognition, 3D face recognition using 3D image data of face was proposed.

3D face recognition using 3D image data acquires the 3D shape of a face using an infrared sensor. Therefore, there is an advantage that it is hard to be affected by changes in the surrounding environment and impersonation acts are also difficult. Recently, 3D face recognition is used in many situations. To the best of our knowledge, there is no experimental study so far published the face recognition using point pair feature.

This paper presents a study on the method of face recognition using point pair feature. Point pair feature is used for bin-picking robots to recognize three-dimensional objects [10]-[12]. There is an advantage that three-dimensional object recognition can be performed using the entire point group without requiring local features such as edges and patterns from the point group acquired from the 3D sensor. Furthermore, there is an advantage that it is possible to perform high-speed, angle independent 3D object recognition by calculating the feature quantities of two points using the local surface information on the three-dimensional object, grouping the feature quantities and comparing them. The following advantages can be obtained by applying a point pair feature to face recognition. Point 1: Facial elements such as nose and mouth are unnecessary. Point 2: Independent on angle of face.

II. Facial Point Pair Feature

Point Pair Feature (hereinafter called "PPF") is adapted for Robotics and Bin-Picking Robots. Therefore, in order to use it for the face recognition, it is necessary to perform noise removal and normalization of 3D point cloud data. This chapter explains the overview of PPF and gives a thorough explanation of noise deletion and normalization.

A. Overview of Point Pair Feature

The PPF is comprised of features between two point points [11]. It is included F1 ~ F4. F1 is the distance between two points, F2 and F3 are the respective normal angles, and F4 is the difference between each normal. Fig. 1 shows a PPF.

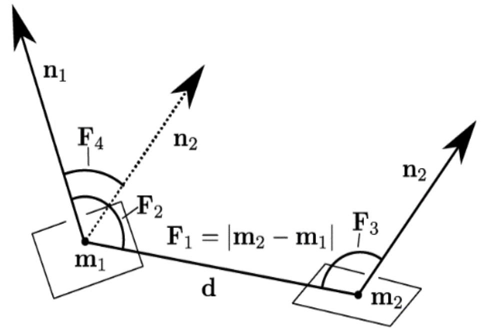


Figure 1. Point pair feature F of two oriented points [11].

Manuscript received February 20, 2018; revised June 21, 2018.

doi: 10.18178/joig.6.1.33-38

PPF has been used in Robotics and Bin-Picking Robots. As for the usage method, the 3D sensor shoots at three-dimensional object flowing in the conveyor from a fixed distance. Then, the object recognition is performed by matching the three-dimensional object to 3D-data using PPF.

It is very difficult to keep the distance between the 3D sensor and the face when trying to recognize faces using PPF because the people moving constantly and the 3D-data of the face need to be shot from a specific position. For this reason, we normalized the 3D-data of the faces and the 3D-data of models so as not to depend on the shooting distance.

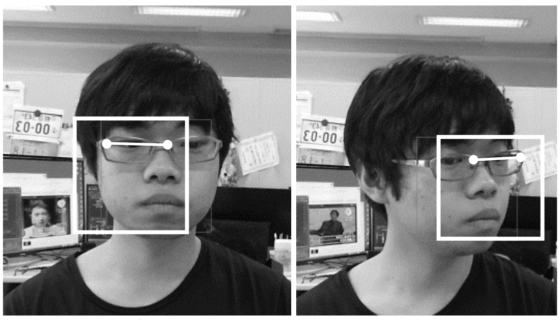
B. Normalization

Among F1 to F4 of PPF, the feature dependent on the shooting distance is F1. The size of the point cloud to compose 3D-data of face also changes when the shooting distance changes. When the size of the point cloud to compose 3D-data of face changes, F1 of PPF formed by the distance between the two points changes.

Therefore, it is necessary to normalize the 3D-data of face by making use of the fact that it is possible to acquire the distance to the face.

There are various methods of normalization. This time, in order to cope with changes in the angle of the face, we use the interocular distance to do the normalization. Both eyes position is detected, and it is conformed the interocular distance of the unknown 3D-data to the interocular distance of the model 3D-data. By resizing the point group according to the interocular distance, it is possible to make F1 of the 3D-data of a face obtained by sensor coincide with the model data. Therefore, face recognition using PPF can cope with changes in shooting distance and angle. A three-dimensional Euclidean distance is used as the interocular distance.

Fig. 2 shows a change in the interocular distance when the angle of the face changes. Fig. 2(a) is the face image when facing the front, and Fig. 2(b) is the face image when facing left to 30 degrees. In these figures, the white spots are the positions of the eyes detected by the kinetic. If turning a face with front direction to 30 degrees left, the interocular distance become shorter.



(a) front (b) 30 degree

Figure 2. The interocular distances when the direction of the front and 30 degree

C. Noise Processing (Remove Background)

Here, we used a rectangle to cut out the face area.

A facial area cannot be represented by a simple shape, and the background may appear in the clipped area in some cases.

If the background is included in the rectangle, the collation accuracy will deteriorate at the time of collation by using PPF.

Therefore, it is necessary to remove the background in the rectangle. We removed the background using neighborhood distance standard deviation of the 3D-data. The equation for calculating the standard deviation is shown in Eq. (1), Eq. (2). Fig. 3 and Fig. 4 shows examples of face images in which the background is included in the point group when the face area point group is cut out with a rectangle.

$$s = \sqrt{\frac{1}{n*n-1}\sum_{i=1}^{n}\sum_{j=1}^{n}(len_{(i,j)} - \overline{len})^2} \quad (i\,!=j) \tag{1}$$

$$len_{(i,j)} = \sqrt{(P_{ix} - P_{jx})^2 + (P_{iy} - P_{jy})^2} \tag{2}$$

s: $standard\ deviation.$ $len_{(i,j)}$: $Euclidean\ distance\ between\ i\ and\ j.$
P_i : $[i]th\ Point(coordinate)$

In Fig. 3, the point group image is viewed from the front, the region where the point group on the right side is the densest is the point group of the face, and the isolated region existing on the left side is the background it is a point cloud.

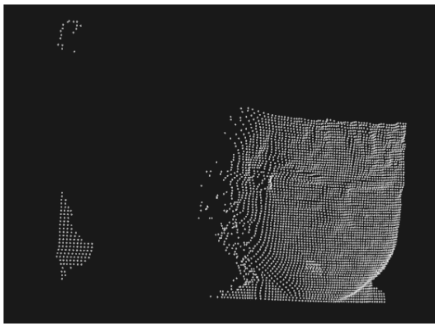


Figure 3. The 3D-data of face and background seen from the front

Fig. 4 shows a view of the point group as viewed from above. In Fig. 4, the depth becomes deeper as going from bottom to top. The point group existing at the bottom of the image is a face, and the small point group existing at the top of the image is the background.

Figure 4. The 3D-data of face and background seen from the top

As shown in Fig. 3 and Fig. 4, since the distance difference between the facial region is obviously far and the number of point clouds is small, it can be eliminated by using the neighborhood distance standard deviation.

III. Face Recognition

The proposed method is consists of process of registration of face data (3D-data of models) and the process of face recognition.

Fig. 5 shows the flowchart of the proposal method.

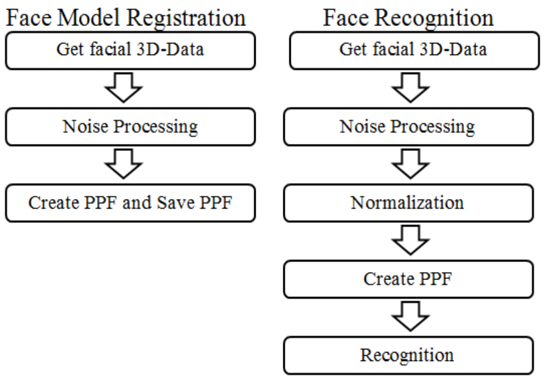


Figure 5. The flowchart of the proposed method

In the registration of face data, face data is first obtained using a 3D sensor, then a noise processing to remove background is performed to the obtained data. After that, the PPF between each point pair is calculated and stored.

Likewise, in the face recognition process, face data is obtained using the 3D sensor, and a noise processing is performed to the acquired data, and a normalization processing is performed according to size of the model data. After that, the PPF of above point group is compared with the PPF of stored model data, so the coincidence rate of the point group is obtained.

IV. Preliminary Experiments for Face Recognition

In this section, we describe preliminary experiments for determining the threshold value for identifying individuals in order to perform face recognition using PPF.

A. *Overview of Preliminary Experiments*

In the preliminary experiment, face recognition is performed under various situations, and a threshold for specifying an individual is determined. Three preliminary experiments are shown in this chapter.

B. *Experiments Environment*

A Kinect v2 (Microsoft) [13], [14] is set on the top of the desktop PC display for obtaining three-dimensional data of the face. In the preliminary experiment, conditions for determining the necessary threshold under the experimental environment are determined. How the conditions are determined are explained in detail in the following subsections.

Fig. 6 and Fig. 7 shows an image of the experimental environment.

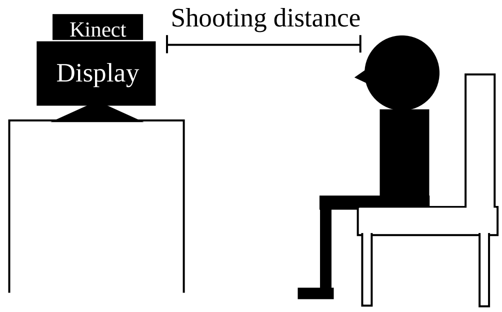


Figure 6. Experiment environment: distance measurement

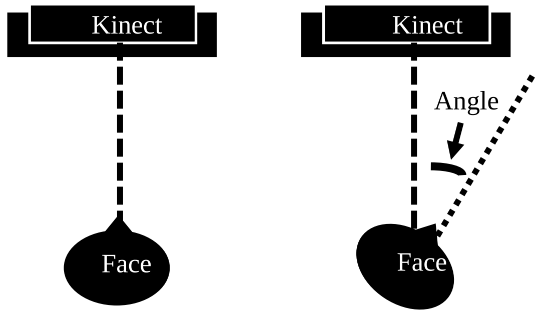


Figure 7. Experiment environment: angle measurement

C. *Preliminary Experiment 1 - Changed the Distance between Face and 3D Sensor*

Preliminary experiment 1 is for confirming whether our system corresponds to the change in distance between the face and the 3D sensor. In this experiment, a front face of the person same as model was used and the angle of the face was not changed. 3D-data were obtained by setting the distance between the face and the 3D sensor to 750mm, 1000mm, 1250mm, and 1500mm. For model 3D-data, 3D-data collected from 1000mm was used. The experimental results are shown in Table I. Table I shows the distance to the sensor, number of points constituting the face, and the percentage of corresponding points.

TABLE I. The Result of Preliminary Experiment 1

Distance to sensor (mm)	Number of points constituting the face (point)	Percentage of corresponding points (%)
750	8200	94.7521
1000	5549	98.4518
1250	4781	81.5418
1500	2847	73.2455

From the results of preliminary experiment 1, even when the distance to the sensor was different, we could confirm that the percentage of corresponding points is high. The reason why the corresponding point decreases as the distance increases is that the number of points constituting the face decreases according to the distance increases, and the number of PPFs that can be calculated also decreases. However, even if it was a distance up to

1500mm, the percentage of corresponding points was over 70%.

D. Preliminary Experiment 2 - Changed the Angle between Face and 3D Sensor

Preliminary experiment 2 is for confirming whether or not it corresponds to the face orientation. 3D-data taken from the same person at 1000mm was used. Three-dimensional data of face facing 0°, 15°, 30°, 45°, and 60° was obtained by the sensor.

For model 3D-data, 3D-data of face facing 0° (Front face) was used. The experimental results are shown in Table II. In Table II, the angle of face, the point number of point cloud obtained, and the percentage of corresponding points are shown.

TABLE II. THE RESULT OF PRELIMINARY EXPERIMENT 2

Face orientation with sensor (°)	Number of points constituting the face (point)	Percentage of corresponding points (%)
0	5549	99.5341
15	5340	98.1944
30	5924	82.8762
45	5551	78.5726
60	5520	52.9629

From the results of preliminary experiment 2, if the orientation of the face to the sensor was within 45°, the high percentage of corresponding points be confirmed. However, with the face orientation of 60° with respect to the sensor, about half of the face area is not displayed on the camera, so the percentage of corresponding points decreases.

E. Preliminary Experiment 3 - Matching with Other People

Preliminary experiment 3 confirms the result of the recognition with others. Face 3D-data of five people in their twenties were obtained from sensors. The face image of subject A was used as the model data. The shooting condition faces the front with respect to the sensor, and the distance to the sensor is about 1300mm. The experimental results are shown in Table III. Table III shows the subject, the point number of point cloud obtained, and the percentage of corresponding points.

TABLE III. THE RESULT OF PRELIMINARY EXPERIMENT 3

Subject	Number of points constituting the face (point)	Percentage of corresponding points (%)
A	3207	97.8546
B	3232	46.1324
C	3180	50.1258
D	3241	48.1024
E	3271	43.2284

From the result of preliminary experiment 3, it was confirmed that the recognition result of the same person is about 97%, and the percentage of corresponding points to different people is about 50% at the maximum. In this case, the distance and angle between the face and the 3D sensor were the same shooting conditions. When the shooting conditions are changed to preliminary experiment 1 or 2, since conditions become more stringent, the percentage of corresponding points will go down further. Therefore, the percentage corresponding points of 50% is the maximum value with other people.

F. Consideration of Preliminary Experiments

From the results of the preliminary experiments 1 and 3, the percentage of corresponding points with other persons is a point rate lower than the percentage of corresponding points with person himself within the distance of 1500mm, and from the results of the preliminary experiments 2 and 3, the percentage of corresponding points with other persons is a point rate lower than the percentage of corresponding points within the angle of 45°.

V. EXPRIMENT

Several recognition experiments were conducted to confirm the effectiveness of face recognition using PPF.

A. Overview of Experiments

Three-dimensional face recognition experiments were performed by using the threshold (more than 70%) determined in the preliminary experiment. In the experiments, 3D data was acquired with "posture to use PC properly". Therefore, the position, distance, and angle of the face at the time of authentication was not defined. In this experiment, a person was seated in a chair, the chair was set 2 meters or more away from a background wall. The subjects were 10 subjects (9 males, 1 female) in their twenties.

B. Result

The experimental results are shown in Table IV. As shown in Table IV, the combination of 10 subjects (A ~ J) was collated. When a result equal to or more than the set threshold value is obtained, it is assumed that these data are data of the same person. When a result smaller than the set threshold value is obtained, it is assumed that these data are data of different people. In Table IV, Success or failure of recognition is indicated by "o" or "x" and the correlation rate (%) is indicated.

VI. CONCLUSION

Since the face recognition technology can identify individuals by the photography which is photographed in a noncontact manner from a remote location, unlike other biometric authentication techniques, there is an advantage that authentication can be performed without requiring cooperation from users. In this paper, we have done from a different approach than the conventional authentication method of 3D face recognition, proposed a 3D face recognition method using Point Pair Feature (PPF) which is used to recognize objects in the field of robot vision. The PPF has an advantage that three-dimensional object recognition can be performed using the entire point group without requiring local features such as edges and patterns from the point group acquired from the 3D sensor. Furthermore, there is an advantage that it is possible to perform high-speed, angle independent 3D object recognition by calculating the feature quantities of two points using the local surface information on the three-dimensional object, grouping the feature quantities and comparing them. In the experiments, we used a Kinect sensor to perform a skeletal recognition from the depth image to estimate the face region and acquire data of the three-dimensional shape of the face. As an experimental result, in our experimental environment, the condition (threshold) of identity identification was that the 3D point cloud and the face area 70% or more match. Next, we registered the face model of 10 subjects based on the determined threshold, and conducted an authentication experiment. As the results, 100% authentication rate was got by our proposal method without an others acceptance and a principal denial (authentication errors). The 3D face recognition using point pair feature is effective was confirmed by the experiments. By using our method with the point pair feature, the following advantages are obtained, thus enabling robust face recognition.

1. Facial elements such as nose and mouth are unnecessary.

2. Face recognition is independent on angle of face.

TABLE IV. THE RESULT OF EXPERIMENT 4

(%)	Scene data									
	A	***B***	***C***	***D***	***E***	***F***	***G***	***H***	***I***	***J***
A	**o 91.5668**	x 49.4029	x 44.5695	x 45.2654	x 42.1019	x 35.7142	x 51.5686	x 43.7299	x 40.5603	x 33.1512
B	x 48.2027	**o 89.5522**	x 43.6865	x 42.3893	x 38.9171	x 33.9560	x 49.3627	x 36.8649	x 41.8534	x 33.9915
C	x 46.5898	x 44.4278	**o 95.6512**	x 44.1592	x 47.8343	x 37.4725	x 43.2352	x 39.3821	x 47.2413	x 39.0336
D	x 46.5898	x 43.6815	x 46.1147	**o 89.1592**	x 54.6284	x 37.4725	x 41.7647	x 40.2974	x 45.5172	x 36.7226
E	x 44.5161	x 40.1990	x 48.5430	x 56.1061	**o 90.6581**	x 40.1098	x 40.2941	x 44.1876	x 51.3362	x 41.7647
F	x 35.7603	x 34.7263	x 37.7262	x 38.1858	x 39.1295	**o 91.0989**	x 36.8627	x 41.4416	x 38.1896	x 44.7058
G	x 50.9677	x 48.9054	x 40.8167	x 40.1769	x 40.1910	x 36.8131	**o 86.0294**	x 42.8146	x 42.7155	x 35.4621
H	x 43.1336	x 37.7114	x 40.1545	x 40.8407	x 42.5265	x 39.6703	x 44.4607	**o 81.4645**	x 47.4568	x 35.0420
I	x 41.5207	x 43.9303	x 48.1015	x 49.4690	x 50.1698	x 38.5714	x 45.4411	x 48.9931	**o 88.1465**	x 37.1428
J	x 33.4562	x 34.7263	x 39.7130	x 37.5221	x 41.2526	x 43.6263	x 35.8823	x 35.7208	x 36.6810	**o 80.2521**

REFERENCES

[1] H. Nakajima, S. Nagashima and T. Aoki, "A hybrid face recognition technology using a two-dimensional image and a three-dimensional structure," *AZBIL Technical Review*, vol. 53, pp. 18-24, 2012.

[2] M. Turk and A. Pentland, "Eigenfaces for recognition," *IEEE Trans Pattern Anal Mach Intell*, vol. 3, no. 1, pp. 71-86, 1991.

[3] X. Lu and A. K. Jain, "Matching 2.5D face scans to 3D models," *Journal of Cognitive Neuroscience*, vol. 28, no. 1, pp. 31-43, 2006.

[4] A. A. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D face recognition: A survey," *Pattern Recognition Letters*, vol. 28, no. 14, pp. 1885-1906, 2007.

[5] K. I. Chang, K. W. Bowyer, and P. J. Flynn, "Face recognition using 2D and 3D facial data," in *Proc. Workshop on Multimodal User Authentication*, 2003.

[6] W. Haririab, H. Tabiaa, N. Farahb, A. Benouarethb, and D. Declercqa, "3D face recognition using covariance based descriptors," *Pattern Recognition Letters*, vol. 78, pp. 1-7, 2016.

[7] Y. Guo, Y. Lei, L. Liu, Y. Wang, M. Bennamoun, and F. Sohelf, "EI3D: Expression-invariant 3D face recognition based on feature and shape matching," *Pattern Recognition Letters*, vol. 83, part 3, pp. 403-412, 2016.

[8] B. A. Echeagaray-Patron, V. I. Kober, V. N. Karnaukhov, and V. V. Kuznetsov, "A method of face recognition using 3D facial surfaces," *Journal of Communications Technology and Electronics*, vol. 62, no. 6, pp. 648-652, 2017.

[9] I. A. Kakadiaris, *et al*., "3D-2D face recognition with pose and illumination normalization," *Computer Vision and Image Understanding*, vol. 154, pp. 137–151, 2017.

[10] S. Akizuki and M. Hashimoto, "High-speed object pose recognition using distinctive 3-D vector pairs," *IEEJ Transactions on Electronics, Information and Systems*, vol. 133, no. 9, pp. 1853-1854, 2013.

[11] B. Drost, M. Ulrich, N. Navab, and S. Ilic, "Model globally, match locally: Efficient and robust 3D object recognition," *IEEE Computer Vision and Pattern Recognition*, pp. 998-1005, 2010.

[12] S. Akizuki and M. Hashimoto, "High-speed and reliable object recognition using distinctive 3-D vector pairs in a range image," in *Proc. International Symposium on Optmechatronic Technologies*, 2012, pp. 1-6.

[13] T. Ringbeck, "A 3D time of flight camera for object detection," in *Proc. 8th Conf. on Optical 3D Measurement Techniques*, 2007.

[14] T. Nishibayashi, "Kinect and natural user interface," *The Journal of the Institute of Image Information and Television Engineers*, vol. 66, no. 9, pp. 755-759, 2012.

Yoshihiro Sato received the B. E degree in computer science engineering from Tokyo City University, Tokyo Japan, in 2016. He is currently a graduate student at Tokyo City University. His research interest includes face recognition, defect detection and image processing.

Yue Bao received Ph.D. degrees from Kanazawa University, Japan, in 1996. He has been a professor at Tokyo City University since 2004. His research interest includes 3D display, image processing and mobile robot.