Analogical Face Generation based on Feature Points

Andy Kyung-yong Yoon[1*,2], Ki-cheul Park[2], Duck-kyo Oh[2], Hye-young Cho[2], Jung-hyuk Jang[2]

Abstract

There are many ways to perform face recognition. The first step of face recognition is the face detection step. If the face is not found in the first step, the face recognition fails. Face detection research has many difficulties because it can be varied according to face size change, left and right rotation and up and down rotation, side face and front face, facial expression, and light condition. In this study, facial features are extracted and the extracted features are geometrically reconstructed in order to improve face recognition rate in extracted face region. Also, it is aimed to adjust face angle using reconstructed facial feature vector, and to improve recognition rate for each face angle. In the recognition attempt using the result after the geometric reconstruction, both the up and down and the left and right facial angles have improved recognition performance.

Key Words: Face Recognition, Face Detection, Feature Vector, kNN, Blend Shape, 3D Face Generation, Face Depth Calculation.

I. INTRODUCTION

The face recognition method determines the presence or absence of a face in the image, and if the face exists in the image, it means to find the position and size of the face. Existing face detection methods can be classified into the knowledge-based method, the feature-based method, the template-matching method, and the appearance-based method [1].

The knowledge-based method assumes that a person's face consists of two eyes, one nose and mouth, and each face element has a certain distance and position. And it is a method to detect faces considering the relation between these elements. The feature-based methods infer face size and position using face-specific features such as facial features, color, shape, and size [2][3]. It detects faces through inferred data, and also detects faces through distance between facial elements, position on face, and so on. . The template matching method is to create a basic template for the face, and then analyze the inputted face image to create a standard template for the face. Then, the standard template and the input face are compared and detected. The appearance-based method is to detect a face using a model learned by a set of learning images. This method uses statistical numerical values to detect face parts in complex images.

The goal of this study is to improve face recognition rate. Improvement of recognition rate is one of the most important factors. There are many factors that degrade the recognition rate. Changes in skin color due to illumination or strong contrasts, changes in facial expressions, attachments such as glasses, and angle of the face greatly affect the face recognition rate. In this paper, we propose a method to prevent recognition rate degradation due to face angle change. That is, the recognition rate is improved in the turned face which is in up, down, left, and right directions.

II. BASIC MODEL CONFIGURATION

2.1 Face recognition model

The study limit of this paper is to prevent the deterioration of recognition rate, even if the face rotation angle is being increased. Therefore, the facial recognition model as shown in Fig. 1 was created under the assumption that the face is turned up, down, left, and right. The target face recognition rotation angle is 15 degrees up and down and 30 degrees left and right.
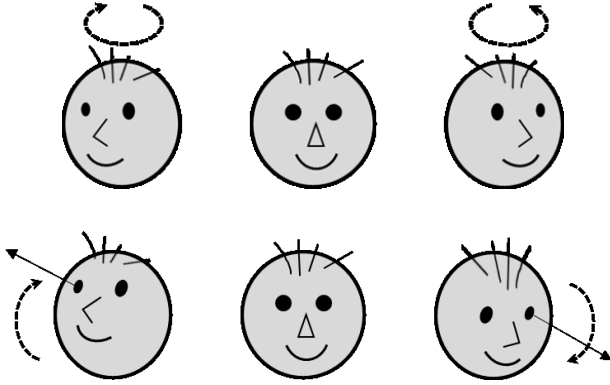
Fig. 1. Face recognition model.

The face source is to generate a three-dimensional face shape using a flat face photograph. The generated 3D face shape is created via information such as the brightness, the contrast, the feature point of the face, the distance between the facial elements, and the bending analogy based on the face rotation angle from the 2D photograph.

The most commonly used algorithms are the kNN algorithm and the blend shape algorithm. The analogy based on the face recognition model is done through complicated calculations.

Therefore, even if they are the same person, different face recognition models can be created depending on the type of photograph. When the algorithm is executed, the face recognition model is accumulated in several layers. In addition, the above-mentioned contrast, feature point, and angle of rotation are different depending on the element. The facial surface, especially the protruding parts of the cheekbones, was not considered in this study.

The advantages of this face recognition model are: First, it is not necessary to store all faces. As a result, storage space and recognition time can be drastically reduced. It can also be a powerful force in the surveillance system. A single camera can take the role of dozens because it can simultaneously recognize the faces of many people with different face angles [4].

Second, it is possible to solve the problem that the same person is mistakenly recognized as another person or failed in recognition depending on the facial expression or angle. For example, although the face model in Fig. 1 is the same person, there is a problem in that it can be recognized as a different person by using the existing system [4].

The feature of this method is that it can internally generate and recognize faces of various angles with only one front face without having to store various faces of the same person. Thus, internally generated faces of various angles are stored as angle vectors.

When a new face image is searched, first, the coordinates of the minutiae of the eyes, nose, and mouth are extracted and the minutiae points of the left eye and the right mouth are connected. Then, the direction and angle of the face are determined by connecting the feature points of the right eye and the left mouth and the feature points of the nose.

2.2 Structural features of face

Feature points on the face are relatively easy to distinguish from the face, and each person has a different shape. In order to compare each feature point, it will be used the correlation value of the eye, the nose, and the mouth, and the linear relation of the chin, mouth, and face area.

Fig. 2 schematically shows each feature point of the face and the structural position value of each point. And the structural position value of the corresponding component is obtained from the coordinate of the candidate area of each extracted feature point [8].

A is the distance between the eyes, B is the distance between the eye and the nose, and C is the distance between the nose and mouth. In addition, the distance between the eyebrow and the eye, and the distance between the two eyebrows can be obtained.

In this study, the error of the ratio was reduced by using the average of the positions of the eyes, nose, and mouth and the minimum / maximum values between them. The ratio of distance between feature points is calculated by the Euclidean distance equation. Recognition by Euclidean distance is normalized to the standard deviation of each person.

The height of the male face of Korean is about 199m, the female is about 190mm and the male is about 9mm big. Each person's face has a different distance for each element, and when the face is divided into three, the ratio of the height of each of the Upper, Middle, and Lower faces is also different. As shown in Table 1, the percentage of the height of the Upper, Middle, and Lower faces of the male was about 40%, 37%, and 21%, respectively. The female ratio was about 42%, 37% and 21% [14]. In a simpler way, the ratio of Upper face height : Middle face height : Lower face height was about 1.0 : 1.0 : 0.6 for males and 1.0 : 0.9 : 0.5 for females.
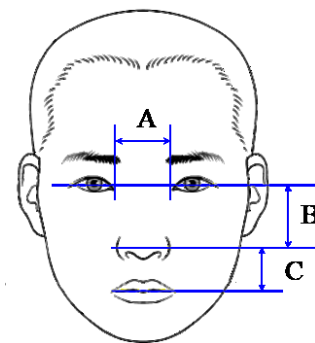

Fig. 2. Structural position value of facial feature points.

Ratio	Male	Male	Female
UHF / THF	39.8±2.6	41.9±2.0	41.1±2.5
MFH / TFH	38.3±1.9	37.1±1.7	37.5±1.9
LFH / TFH	22.0±1.8	21.0±1.6	21.4±1.7
LFH / UFH	55.7±7.8	50.3±5.7	52.5±7.1

Table 1. Measurement data matched total face height.
Male samples are selected within female mean±SD of total face height (n=84). Statistically significant difference between the genders ($p<0.05$).

Table 2. Ratios among face heights.

	Male	Female
Total face height	191.9±4.0	190.2±7.9
Upper face height	74.3±5.8	79.8±5.8
Mid face height	74.8±3.4	70.4±3.6
Lower face height	42.9±3.8	40.0±3.3

Statistically significant difference between the genders ($p<0.05$).
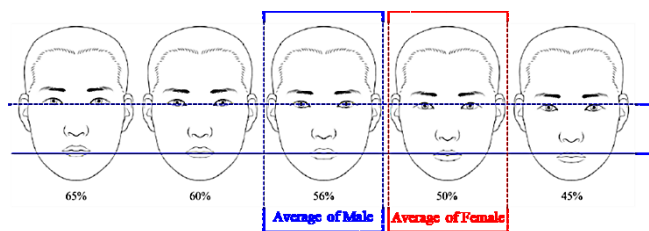Total face height, TFH; Upper face height, UFH; Mid face height, MFH; Lower face height, LFH..

Fig. 3. Ratio of lower and upper face height.

As shown in Table 2, the ratio of the Lower face height to the Upper face height was about 56% for male and about 50% for female [14]. As a result of classification according to the height of the face structure in the face, male and female showed different patterns. The Middle face type, which the percentage of the Lower face to the Upper face (50 ~ 60%) was slightly higher than half of the male face (54.1%). Upper face type and low face type were similar to 24.3% and 21.6%, respectively. However, in female, the rate of Low facial type was more than half (53.5%), Middle facial type is 40.8% and Upper facial type is 5.7%. Figure 3 shows the various face ratios of Koreans [14].

III. EXTRACTING FEATURE POINT

3.1 Extracting depth data of 2D face

In this section, a method of extracting depth information from a 2D face is described. Generally, 2D face recognition performs detection by brightness and contrast feature values. However, since there is no depth information, that's why this is an obvious limitation.

Therefore, in order to recognize various angular faces from a 2D face, it is necessary to calculate the depth information that does not exist in the 2D face. The Point Signature method is used to extract depth information. The Point Signature is a method of expressing distance information as a one - dimensional spatial signal for angle, and Chua and Ho [6][7] introduced the concept of Point Signature to face recognition [9].

This method extracts depth information of face corresponding to each feature point based on one point and extracts curvature information of face structure. The procedure of the Point Signature method is as follows.

(1) Set angle θ and radius r for depth data extraction
(2) Generate the normal vector after setting the reference point
(3) Project a circle of radius r centered on the normal vector onto the face shape
(4) Extraction of depth data corresponding to angle θ

3.2 3D Face generation

The cause of performance degradation of the 2D recognition model is the recognition error due to the change of the external environment, that is, the generation of the lost data due to the change of the facial expression of the recognition face, and the illumination change.

These problems often cause false detection of face recognition. In order to overcome this problem, 3D face recognition should be performed using depth information of a face shape that does not exist in a 2D face.

The 2D face information can obtain a certain amount of face depth information by analyzing the reflection pattern according to the brightness and contrast ratio of the photograph. When the face depth information is acquired, it becomes basic data for making a 3D face shape. The obtained 3D face shape obscures the face shape of the point cloud shape. It is also robust to illumination changes because it has depth information [6][9].

Moreover, it is possible to compensate poses freely, which can compensate for the disadvantages of 2D face recognition. Since the generated 3D face is an imperfect face, the face shape is reconstructed by performing the pose compensation by the preprocessing step.
In order to perform face recognition, it is necessary to normalize the coordinates of the generated face shape through the reference point. 2D faces differ in size and position of facial features acquired according to the environmental factors at the time of shooting and personal facial characteristics [10].

Therefore, it is difficult to extract certain facial feature data if the normalization process is not performed. Therefore, pose compensation is performed in front to extract accurate feature data.

In the case of 2D face recognition, the rotated face has data loss in the rotated part. However, the 3D face shape has the advantage that the respective points are posing compensation in all axes rotatable since it has the coordinates of the X, Y, Z in the form of a point cloud possible.

3.3 kNN Algorithm

The kNN algorithm is very useful for extracting the feature points of face recognition. It serves to rationally classify the feature points that are difficult to classify by the nearest neighbors.

kNN is a method of predicting new data with information of the nearest k neighbors of existing data when new data is given. As shown in the Fig. 4, spotted ball category information can be inferred with neighbors. If k is 1, it will be classified as white, if k is 3, it will be classified as black. If it is a regression problem, then the average of the neighbors dependent variable is the predicted value.

In fact, kNN does not have a procedure that is called "Learning". That's also because when new data comes in, it draws neighbors from the distance between existing data. So some people call it a lazy model, meaning that it does not build a model for kNN separately. It's also called Instance-based Learning. This contrasts with Model-based learning, where models are created from data to perform tasks. Therefore, it performs tasks such as classification/regression using only each instance without generating a separate model [15].

Thus, the minimum distance classification rule according to the basic principle is referred to as kNN (k-Nearest Neighbor) classification rule. In order to do this, a standard pattern must be selected in advance for each class. That is, it is examined how many k of an arbitrary pattern x belongs to which class of the nearest neighbors, and the class of the largest number belonging to is determined as the class of x.
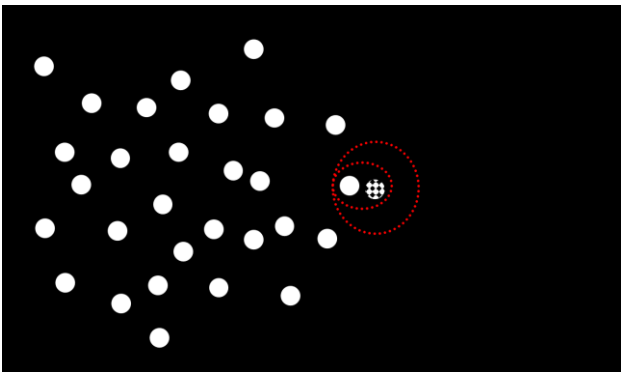

Fig. 4. Concept of 3-Nearest Neighbors method.

The order of execution of the kNN algorithm will be briefly described below with reference to Fig. 4.

In the first step, k nearest neighbors are searched for given unknown data. The second step determines the class to which the unknown data given in the voting method among k-nearest neighbors belongs [15].

In addition, several forms can be used depending on the distance metric of Equation 1.

$$d_p(x) = \left(\sum_i |x_i|^p\right)^{1/p}, \tag{1}$$

where $p = 1$, As Manhattan Distance, this is used as Equation 2.

$$d_1(x) = \sum_i |x_1|. \tag{2}$$

When $p = 2$, As Euclidean Distance, this is presented as Equation 2.

$$d_2(x) = \sqrt{\sum_i |x_i|^2}. \tag{3}$$

When $p = \infty$, As the maximum distance metric, this is presented as Equation 4.

$$d_\infty(x) = max_i |x_i|. \tag{4}$$

The kNN algorithm in this study uses the Euclidean Distance of $p = 2$.

3.4 Blend Shape Algorithm

Blended Shapes are a nearly standard way of expressing facial expressions in computer animation, and are also used in 3D authoring tools such as Maya, and are also used in movies such as Lord of the Rings, KingKong, and Final Fantasy.

The blend shape is represented by a linear combination of Basis vectors representing each facial expression. When the blend shape with n vertices is expressed as a $3n \times 1$ vector, the blend shape is expressed as Equation 5 [2][3].

$$f = \sum_{k=0}^{n} w_k b_k, \tag{5}$$

where b represents the Neutral Expression. A commercial program such as Maya uses Delta Blendshape, which can be expressed as: In this case, the weights are normalized between 0 and 1.

$$f = b_0 + \sum_{k=1}^{n} w_k (b_k - b_0). \tag{6}$$

In addition, there are Intermediate Blend shapes and Combination Blend shapes. In the case of CG movie production, more than 100 blend shapes are made, so Combination Blend shape is often used.

Blend Shapes have advantages over other facial modeling methods. The biggest advantage is that each blend shape has its meaning, and by adjusting the weights of these blend shapes, you can create the desired look directly [11].

This contrasts with a similar method, PCA (Principal Component Analysis), in that the Basis vectors of the PCA produce facial expressions that are not intuitively understood by humans [13]. Blend shape, however, also has its disadvantages, and the biggest problem is that it requires a significant number of Blend shapes to create a natural look.

IV. EXPERIMENT

4.1 Extracting geometric feature information at 2D

The geometric feature information of a face collectively refers to eye, nose, mouth, eyebrow size, relative position relation, distance between both ears, and shape information of chin line.

In this study, a candidate region in which an eye, a nose, and a mouth exist in a face image is determined in advance, and then a plurality feature information is extracted by projecting on a X-axis and a Y-axis. And a relative distance is obtained based on the feature information.

For face recognition based on geometric feature information extraction, FMG (face model graph) was generated for training face image samples [5][6].

Then, the feature points of the face were searched by the method of finding the face graph and FG (Face Graph) according to the arbitrary facial image. FG was obtained from the face photographs and the positions of the feature points of the face were automatically found and Gaber and LBP features were extracted for each feature point [5].

The face has a fixed spatial arrangement of eyes, nose, and mouth. For example, the position of eyes and nose is not changed. It is arranged in the order of eyes, nose, and mouth from the top, and two eyes are arranged. Since the spatial arrangement information is fixed, it is possible to easily find an expected candidate space of eyes, nose, and mouth.

After finding the face space in which the facial elements are located, Gaber and LBP features are extracted from each feature point of the FG. Then, each of the Gaber and LBP features is represented as a set, and the similarity is calculated by comparing these facial images with the inputted face image.

Because of the variety of faces, FMG is generated to represent facial images in one representative form. It is possible to obtain FMG by sex, race, and age by collecting facial images that can reflect various kinds of man, woman, age, and race, and then displaying the characteristic points by handwork and then finding the average position of each

feature. Fig. 2 shows the concept of feature extraction for the better understanding [5].
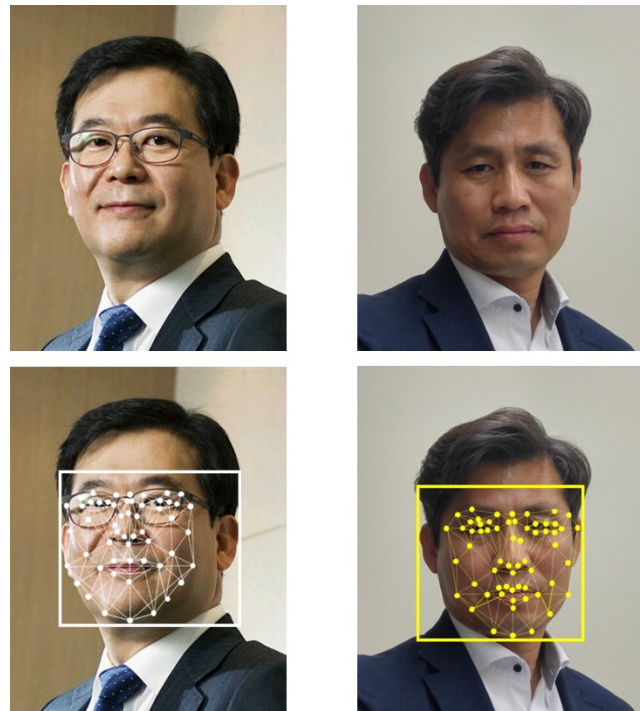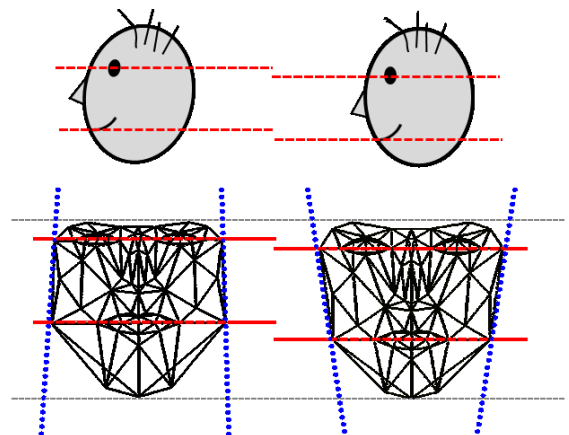

Fig. 5. FMG and feature points extracting.

Fig. 6. Comparison of extracted FMG.

The FMG for the pictures in Fig. 5 is shown in Fig. 6. Depending on the face angle in the photo, it can be seen that the FMG is generated slightly differently. Accordingly, the feature points also have different vector values.

If the feature points are not found or the featureless face area information is considered noise. These results show that facial feature extraction can obtain much information from face elements rather than skin. Based on these results, we applied the kNN algorithm and the blend shape algorithm to FMG to generate a 3D rotated face.

Fig. 7 shows the feature points extracted from the 2D image. Because it was extracted from the frontal face, both sides are balanced distribution around the nose. In Fig. 7 (a), FMG is

constructed by connecting extracted feature points with wires. Fig. 7 (b) shows extracted feature points.

Fig. 8 shows the calculation of the face depth by calculating the distance between the feature points when the front face is turned to the left. Also, through the kNN algorithm, feature points can be moved, or they cannot be used and can be erased. The depth of the face was changed by excessive calculation. Fig. 8 (a) is the 2D image and (b) is the feature point of the face turned 30 degrees to the left. In Fig. 9 (b), FMG is constructed by connecting the feature points of the rotated face with wires. The facial form appears relatively intact, but the facial depth of the side is excessive.
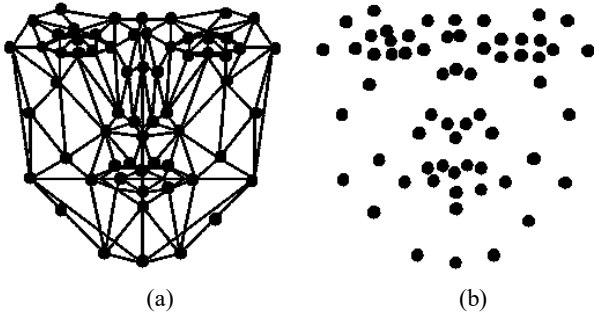

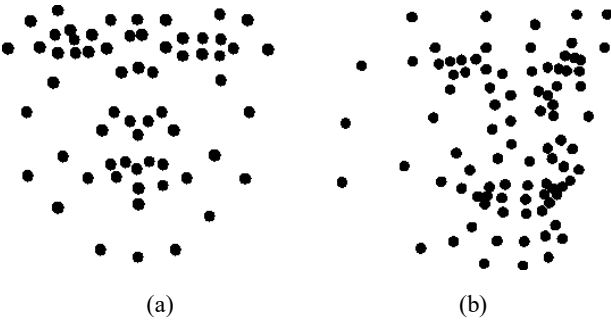(a)   (b)

Fig. 7. Feature points of 2D image.

(a)   (b)

Fig. 8. Rotated 3D image creation.
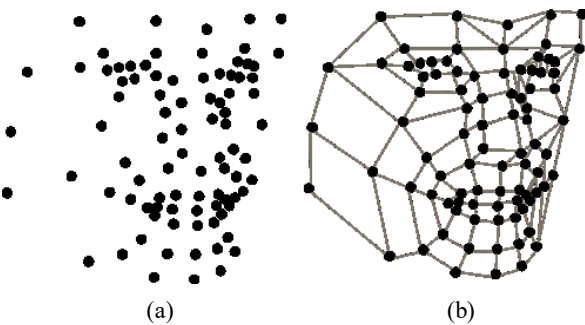

(a)   (b)

Fig. 9. Obtained FMG.

4.2 Face Tracking method in Video

Particle Filter (PF) have been widely used to track objects in video images. Particle Filter has a unique position in the processing of similarity with Kalman Filter.

The PF can predict the dynamic state of an object through the Monte Carlo method given in the Bayesian framework. The posterior probability is estimated using the prior probability and the measured likelihood function by repeating Equation 7[12].

The construction of a general PF has the following four steps: [Selection], selects a new sample proportional to the weight of the sample obtained in the previous step. [Propagation], propagates the selected sample to a new location. [Observation], and weights are measured and weighed in each sample. [Estimation], Estimates the object as an average of the positions of the weighted samples. Then repeat the above procedure to track the object [13].

The color-based PF tracking method is a typical tracking algorithm that calculates the similarity of the color distribution with respect to the tracking object at each sample position, measures the probability that the object exists, and tracks the object. However, when the pose of the face to be tracked changes, such as the face tracking, the shape and color distribution of the tracking object changes according to the pose change.

In this case, the IVT tracking algorithm can improve the performance degradation when the shape of the tracking target changes. The IVT algorithm can extract the features from the previous images using the PCA and track them using the PF technique to effectively cope with the object whose shape changes. In IVT, the following Bayesian probabilities are calculated to track the object [13].

$$p(X_t|Z_{1:t}) \propto p(Z_t|X_t) \int_{X_{t-1}} p(X_t|X_{t-1})p(X_{t-1}|Z_{t-1}), \quad (7)$$

where X_t and Z_t are the state of the object at time t and the input image frame, respectively. And assumes that the initial X_0 is known. Parameters of the state X_t of the sample include the object's center coordinates, the size relative to the reference image, and the rotation angle with respect to the horizontal axis.

As in the conventional face tracking PF technique, IVT uses a simple Gaussian distribution model instead of using a complex dynamic model to propagate samples from X_{t-1} to Z_t.

In other words,

$$p(X_t|X_{t-1}) = N(X_t; X_{t-1}, \Sigma) \quad (8)$$

where Σ is a properly selected diagonal covariance matrix. In the IVT, PCA is performed on past traced images X_{t-1} to the similarity $p(Z_t|X_t)$ of the image Z_t obtained from the state X_t of one sample in the Observation step. At this time, the current image Z_t can be obtained in a subspace represented by the mean (μ) and eigenvector obtained.
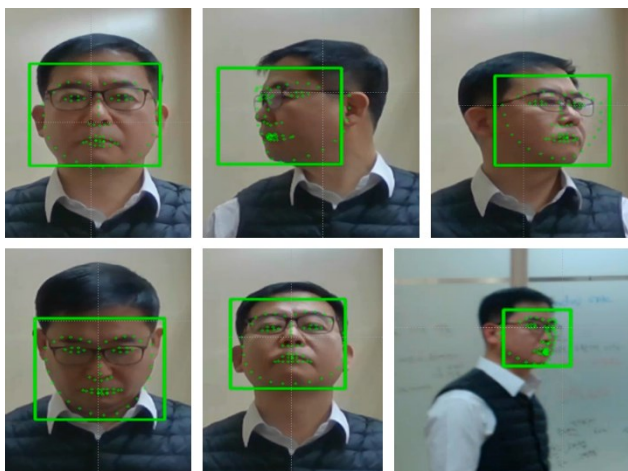
Fig. 10. Tracking in video.

This probability is inversely proportional to the distance between the image Z_t and the center point in the subspace, so we can obtain the following two distances. That is, the similarity is calculated based on the distance from the Z_t to the projected point and the distance from the projected point to the center point in the subspace when projecting Z_t to the subspace.

Fig. 10 shows feature points extracted from video images and feature points are obtained by adapting to face movements and changes. Despite face changes, occlusions, and scene changes, IVT and PF are interlocked by continuously tracking changes in the face continuously.

Detector and tracker operate independently for every frame, and the situation of detection and tracking failure can be compensated by exchanging information between them. If the similarity value of the detection results is lower than the specific value, it is determined that the target to be tracked has not been detected correctly, and the detection is considered as failure.

V. CONCLUSIONS

The subject of this paper is to extract the feature points from the face and rotate the extracted feature points in 3D to generate the side face. In this process, the kNN algorithm and the blend shape algorithm are applied. As the angle of the face is increased, the accuracy of the face is deteriorated. Creating a whole 3D face through computation using a 2D face is a near-impossible task at present. However, if the contrast ratio of the face photograph is large and the sharpness is excellent, it can be upgraded considerably.

Also, having the face location information of extracted feature points together to improve the face recognition performance can be a solution for promoting the enhancement. In addition, when the division method using the ASM (Active Shape Model), which is a method for

improving the background component, is applied, the face recognition performance is relatively improved.

It is a solution that can extract the features of the facial component or extract the components from the facial feature points, or both methods can contribute to enhance the face recognition.

In this paper, we created left and right faces, but did not create up and down faces. It is necessary to implement a more difficult algorithm that is inferior to the left and right. Therefore, it is a reality that the vertical angle is difficult to exceed 15 degrees.

Face recognition technology will play a central role in modern biometrics technology. Although there are various biometrics technologies, it is difficult to make them universal due to having many fatal weaknesses.

Relatively, face recognition technology has no fatal weaknesses, so it is expected to lead the biometric market for a while.

Therefore, it is necessary to study in many fields such as 3D face generation and facial expression generation through algorithms such as Blend shape, kNN, and PF.

REFERENCES

[1] M.Akhil jabbar, B.L Deekshatulua, Priti Chandra, "Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm," *Procedia Technology* 10, pp. 85 – 94, 2013.

[2] Mohammed Hazim Alkawaz, Dzulkifli Mohamad, Ahmad Hoirul Basori, Tanzila Saba, "Blend Shape Interpolation and FACS for Realistic Avatar," Springer *3D Res*., Vol. 6, No.6, 2015

[3] Henry A. Rowley, Shumeet Baluja, Takeo Kanade, "Neural Network Based Face Detection," *Computer Vision and Pattern Recognition*, Carnegie Mellon University, 1996.

[4] Wooky Lee, Jongtae Baek, Hwaki Lee, Young-Mo Kim, "Face Recognition Model based on Angled Feature Vectors", *Journal of KIISE*. Vol.39, No.11, pp. 871 - 875 , 2012.

[5] Jin-Ho Kim,, "Face Recognition by Fiducial Points Based Gabor and LBP Features," *The Journal of the Korea Contents Association,* Vol.13 No.1, pp. 1 - 8 , 2013.

[6] Kaveh, A., *Particle Swarm Optimization." Advances in Metaheuristic Algorithms for Optimal Design of Structures*, Springer International Publishing, pp 9-40. Mar. 2014.

[7] Karaboga, Dervis, et al. "A comprehensive survey: artificial bee colony (ABC) algorithm and applications," *Artificial Intelligence Review*, Vol. 42, No. 1, pp. 21-57, June 2014.

[8] Young-Il, Eung-Joo Lee, "Real-Time Automatic Human Face Detection and Recognition System Using Skin Colors of Face, Face Feature Vectors and Facial Angle

Informations," *The KIPS transactions. Software and Data Engineering*, pp. 491 - 500 , 2002.

[9] Chan-Jun Park, Sung-Kwun Oh, Jin-Yul Kim, "A Study On Three-dimensional Optimized Face Recognition Model: Comparative Studies and Analysis of Model Architectures," *The Transactions of the Korean Institute of Electrical Engineers,* Vol. 64, No. 6, pp. 900- 911, 2015.

[10] Hae Min Moon, Sung Bum Pan, "Long Distance Face Recognition System using the Automatic Face Image Creation by Distance", *Journal of The Institute of Electronics and Information Engineering,* Vol. 51, NO. 11, November 2014.

[11] Jae Mo Chung, Hyeon Bae, Sung Shin Kim, "Realtime Face Recognition by Analysis of Feature Information," *Journal of fuzzy logic and intelligent systems*, Vol.11 No.9, pp.822 – 826, 2001.

[12] Jeany Son, Ilchae Jung, Bohyung Han, "Real-Time Human Tracking with Detection Feedback", *Journal of KIISE. Software and application,* Vol.40 No.12, pp.859 – 868, 2013.

[13] Jin Yul Kim, Yong-Seok Kim, "Face Tracking and Recognition in Video with PCA-based Pose-Classification and $(2D)^2$PCA recognition algorithm," *Journal of Korean Institute of Intelligent Systems*, Vol. 23, No. 5, pp. 423-430, October 2013.

[14] Woo-Cheol Song, Sung-Ho Kim, Ki-seok Ko, "High-Set or Low-Set of Korean Face," *Korean J. Phys Anthropol.,* Vol. 30, No. 1, pp. 1~6, 2017.

[15] Park, Moon Kyu, "Face Recognition Using Principal Component Analysis and KNN method," Thesis of M.S., Chungbuk National University, 2003.

Duck-kyo Oh graduated from Kwangwoon University and a bachelor of computer science and engineering. He served as General Manager of Beijing Inca Co., Ltd., and is currently the director of development at Neosecu Co., Ltd. in Korea.

Hye-young Cho graduated from Dongmyung University and a bachelor of computer science and engineering. She is currently head of development at Neosecu Co., Ltd in Korea.

Jung-hyuk Jang graduated from Semyung University and a bachelor of electronics engineering. He is currently a researcher at the Neosecu Co., Ltd. in Korea

Authors

Andy Kyung-yong Yoon completed his MS and Ph.D. degree from the Yonsei University, Korea. He is currently a Professor in the Professional school of Electronic Engineering at San Martin University, Peru, also a CTO of NEOSECU, YAP Company, Korea where his main activities include research, undergraduate and postgraduate training. His research interests include Mobile Agent Systems and AI related O2O application, and Indoor Positioning System.

Ki-cheul Park graduated from Kwangwoon University and a bachelor of electronics engineering. He participated in the TV development project of Samsung Electronics and is currently CEO of Neosecu Co., Ltd. in Korea.