International Journal of Computer Science and Applications,

Vol. 13, No. 2, pp. 29 – 44, 2016

MINIMUM VARIANCE METHOD TO OBTAIN THE BEST SHOT IN VIDEO AND ITS EFFECTIVENESS FOR FACE RECOGNITION

KAZUO OHZEKI

Graduate School of Engineering and Science,
Shibaura Institute of Technology
3-7-5 Toyosu, Koutou-ku, Tokyo, Japan
ohzeki@shibaura-it.ac.jp

RYOTA AOYAMA

Graduate School of Engineering and Science,
Shibaura Institute of Technology
3-7-5 Toyosu, Koutou-ku, Tokyo, Japan
ma15001@shibaura-it.ac.jp

YUTAKA HIRAKAWA

Graduate School of Engineering and Science,
Shibaura Institute of Technology
3-7-5 Toyosu, Koutou-ku, Tokyo, Japan
hirakawa@ shibaura-it.ac.jp

This paper describes a face recognition algorithm using feature points of face parts, which is classified as a feature-based method. As recognition performance depends on the combination of extracted feature points, we utilize all reliable feature points effectively. From moving video input, well-conditioned face images with a frontal direction and without facial expression are extracted. To select such well-conditioned images, an iteratively minimizing variance method is used with variable input face images. This iteration drastically brings convergence to the minimum variance of 1 for a quarter to an eighth of all data, which proves to take the frontal image in 0.27 second from video at most. The proposed system using six statistic values realizes 98.3% as an authentication rate.

Keywords: face recognition; feature point; distance; distribution.

1. Introduction

There are two major methods for face recognition [Jafri and Arabnia (2009)]; one is the feature-based method, which uses feature points at the endpoints of facial parts. The other is the holistic method, which processes the whole face region without decomposing regions into feature points. The former method has become unpopular because it is difficult to detect accurate feature points automatically. The latter method is now popular. However, Kathryn Bonnen and Anil K. Jain have recently proposed the effectiveness of a component-based method which uses details of facial parts, rather than globally recognizing face information [Bonnen et al. (2013)]. The holistic method conventionally analyzes the whole face region at once to achieve robust recognition. The component-based method utilizes separate regions of the eyebrows, eyes, nose and mouth, and performs dedicated recognition for each separate region, then integrates the results. The component-based method implies that it is now possible to detect facial parts before face recognition. The performance of the component-based method is better than that of the single holistic face recognition method.

The performance of face recognition has improved recently [JCB (2014)], and the results of various contests have been reported [Grother et al. (2010), Ngan and Grother (2014), Grother and Ngan (2014)]. In those reports, recognition of frontal faces was considered as an easy task and more complicated conditions with age changes and with expressions are now targeted. Then, the basic face recognition technology is left away from popular development activity. The recognition rate is now 97.5% for the frontal face without expressions [Gohringer (2012)], and it is still a difficult problem to realize 100% recognition for the frontal face without expressions.

One of the disadvantages of feature-based methods using feature points of face parts is that it is difficult to detect feature points correctly [Jafri and Arabnia (2009)]. On the other hand, many devices are presented in the holistic methods with additional cases, such as faces with a non-frontal direction, under bad lighting conditions, and with facial expressions. However, for all cases, the recognition rate of the face is up to 99.90%, with a false acceptance ratio (FAR) of 1%. It is now difficult to reduce the FAR value since the recognition rate is at the upper limit.

In this paper, we consider the feature-based method because it provides high-precision results if the feature point detection works well and successfully provides digital precision, while the holistic method views a whole face image with rough ambiguity at most. Recognition improves for a well-conditioned image from the frontal direction and without facial expression. We use moving video as the input and automatically extract the best shot from the frontal direction and without expression. We then use the best shot images for registering the face image, and we match the input and registered face images. A NIST's report newly using video in 2015 may diverged at initial status. To obtain the best shot from an input video sequence, we will try two ways using feature points. One involves maximizing the distance between feature points while viewing the input sequence. The other involves minimizing the variance of the distance between feature points. In this paper, the latter method will be described.

One of the methods using feature points of 3D presented by [Drira et al. (2013)], performs with 99.2% accuracy utilizing distances of 3D curves for faces without expressions. Another method using a 3D mesh and the distance function in a wavelet-transformed domain under bad lighting conditions presented by Toderici et al. (2010), outperforms the 2D case. However, the recognition ratio is not so high. Guillaumin et al. (2009) presented an improvement using a learning method utilizing the nearest neighbor method to feature point distances.

2. Previous Methods and Ideas for Improvement

In contests for face recognition, the recognition ratio for still pictures is reported to be 0.92 for rank one [Grother et al. (2010)]. Rank one means that it only adopts the top datum. How to extract frontal face images from video is important. In a previous face recognition vendor test contest using video by NIST [Grother et al. (2012)], they diverged to multi-frame processing and were at the basic evaluation for many increased items. One of previous trials of frontal face recognition from video [Majumdar et al (2008)], used pre-constructed frontal face database for recognition with investigating rates of frontal faces in a head movement video. They concluded two seconds video sequence is needed for authentication. A previous method of really selecting frontal face images from video was proposed by [Zhang Ping (2008)]. The frontal face image in video is detected by adjusting thresholds in the verification stage. The frontal images were selected using recognition results. It will be expected algorithmiclly to narrow down all-posed face images to frontal face images in this paper (Fig.1). To improve the recognition ratio, it is important to enhance the processing methods in the first stage of the system to acquire larger distances of feature values between people, not to incorporate additional conditions of age changes, expression lighting conditions etc.

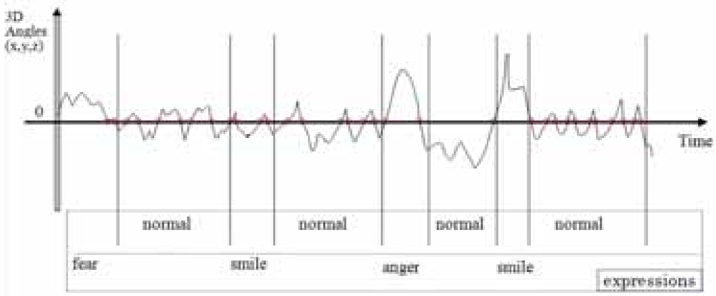

Fig.1 Speaking video sequence and frontal image appearance. The best frontal face image occurs at all angles are zeros and at normal expression.

3. Face Recognition using Feature Point Distance

3.1. *Proposed system[Ohzeki at al. (2014a)(2014b)]*

Fig.2 shows the proposed face recognition system. Using input N frames, the best shot frame is selected by the method to be described in the following section. The best shot frame is registered in a database before recognition. At recognition, the best shot frame is selected and is verified using data in the database. The feature points of face parts are detected from the input face image. The detection of the feature points is performed using software developed by Milborrow et al. [Milborrow and Nicolls (2008)(2014)]. The software is based on the Active Shape Model and detects 77 feature points on the frontal face image. After detection of the feature points, two compensation operations of rotation and scale normalization are performed.

3.1.1 Rotation compensation[Ohzeki at al. (2014b)]

The rotation compensation is to rotate at an angle θ which is a slope between the edges of both eyes. The rotation operation matrix shown in (1) is applied to the (x,y) coordinates of all feature points.

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{1}$$

After this rotation compensation, both eyes are aligned horizontally.

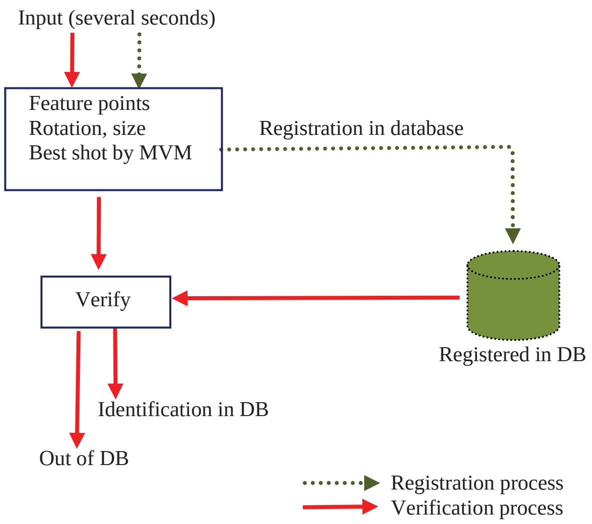

Fig.2 Face recognition system

3.1.2 Size and position normalization

Next, normalization according to the size is performed. This operation is applied to all feature points at the same rate as changing the size of a specified face part to a fixed size. In actual, let NF be

NF=sc/(XLL-XRR) (2)

Where,

XLL= X coordinate of the left edge of the left eye

XRR= X coordinate of the right edge of the right eye,

Then, multiply this NF by all X and Y coordinates.

Position normalization is applied by resizing all points in parallel after the size normalization. Let the input X and Y coordinate values of the first feature point be (In(0),In(1)) and the registered values be (Reg(0), Reg(1)); then the parallel movement value for X is Reg(0)-In(0), and for Y is Reg(1)-In(1)).

3.2. *Best Shot Detection methods from video*

An advantage of using video for face recognition is that it can utilize well-conditioned data and discard poorly conditioned data. Therefore, video provides temporal continuity, so classification information from several frames can be combined to improve recognition performance [Howell and Buxton (1996)]. Also, tracking of detected facial regions is possible and the system can be expanded to carry out facial expression detection [Torres (2004)].

Two methods of obtaining the best shot are considered. One is maximizing the length between two feature points. Another is removing the value with greatest deviation to minimize variance.

3.2.1 Maximizing Length Method

A moving head in an input video shows a three- dimensional rotation pattern. Fig. 3 is a face with a frontal direction and without a facial expression, which is the best shot. We will detect this kind of best shot image from all varying data. As for the Y and X axes rotations, the distances between feature points can be reduced by three-dimensional displacement and there is no data for compensation from the single camera environment. The Maximizing Length Method involves selecting a frame in which the distance between feature points is the maximum in all data[Aoyama et al. (2015)].

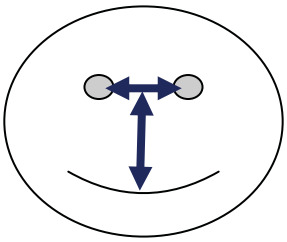

Fig.3 Regular face with frontal direction and without facial expression.

3.2.2 Minimizing Variance Method

The minimizing variance method is used to remove irregular data from the total data to obtain concentrated data with smaller variance. The method actually removes iteratively the largest value distant from the temporal average value and makes a new data set. Fig. 4(a)-(d) shows the iterative removal status for the case of a pre-obtained time sequence data set of feature points. The asterisks in Fig. 4(a)-(c) indicate the largest value at each stage. A removed value is replaced by a dotted line. From Fig. 4(a) to (c), two values are removed one after another, and the third largest value is marked in Fig. 4(c). The same processes are repeated to reach Fig. 4(d), which shows a temporally small variance as if the iteration may converge to a fixed average with the smallest variance.

3.3. ***Estimation of the number of distinguishable people***

To distinguish one person from all the other persons using the detected feature points, the values should be clearly different to each other. To reduce the FAR and FRR, the detected feature point values should have meaningful distances. Here, we try to estimate the distance distribution for a larger number of testers from a smaller number of testers.

The distance between two feature points on a face is a feature value. Let us consider a distribution of this feature values that are length data. We assume this distribution is normal with mean m and standard deviation σ.
Fig.5 shows a one-dimensional distribution for a concept of the feature value of a distance between two values. The distribution is made from the data of many testers. When this distribution is normal, the value of a sample taken from the data set randomly is considered to form a random sequence with this distribution.

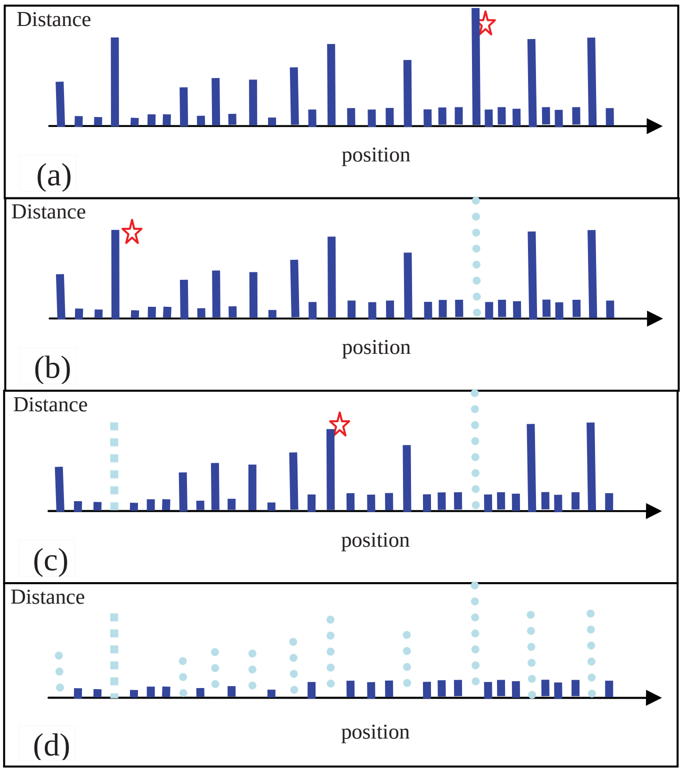

Fig. 4 Variance is reduced by removing the value with the greatest deviation.

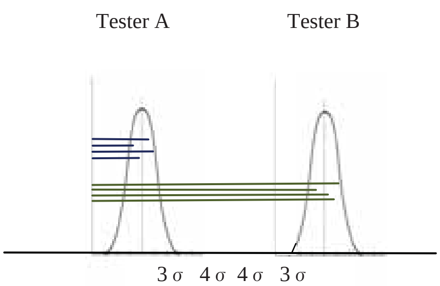

Fig.5 An example of distribution of two different persons.

Fig.6 shows 77 feature points on a face, indicated by black dots, detected by Stasm. (This image is licensed by Datacraft Corp.) The numbers are referred to in later sections.

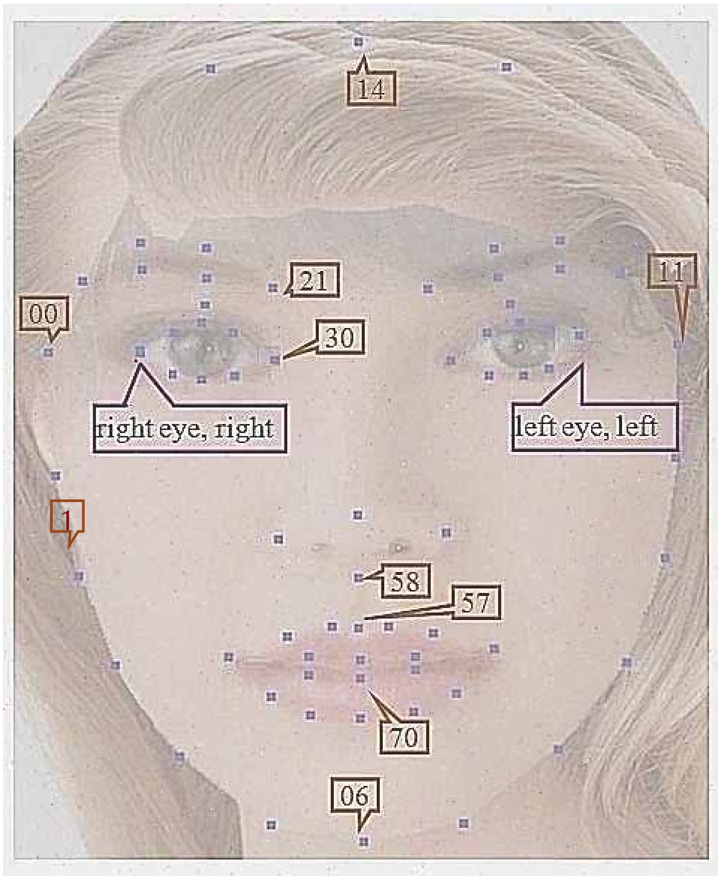

Fig.6 77 feature points on a face, indicated by black dots, detected by Stasm. (This image is licensed by Datacraft Corp.) The numbers are referred to in later sections.

Four experimental items are carried out in the following sections. The minimizing variance method (MVM) described in section 3.2.2 is introduced in the experiments to automatically obtain frontal face images without expressions from input moving video of a person speaking for face recognition.

4. Experiments

4.1. *Reducing variance by MVM*

First, by adapting MVM described in section 3.2.2 to use with a recorded part of a video with a length of about one minute, the convergence status of variances is investigated. The number of pieces of distances between feature points is the combination of two out of 77 feature points, that is ${}_{77}C_2 = 2926$, where the total number of feature points is 77. For all this length data, let an initial value of the number of frames according to the time direction be “n”. The algorithm involves the following three steps:

a. Obtain the average length for time direction with the total number of the lengths of n.
b. Remove the largest value that is distant from the average value.
c. Obtain a new set of length data in which the number of lengths is subtracted by 1. (n=n-1)

This process is repeated until we get n=1.
Fig. 7 shows four data which were randomly chosen from 2926 results with graphs of variances as the process above proceeded. The variances at the position of 1000 times of iteration, which is about half the number of total iterations, are greatly reduced. Also, at the position of 1500 times, which is about a quarter of the total number, the variance is almost below 1-2. Other results show nearly the same tendency, as shown in the graphs.

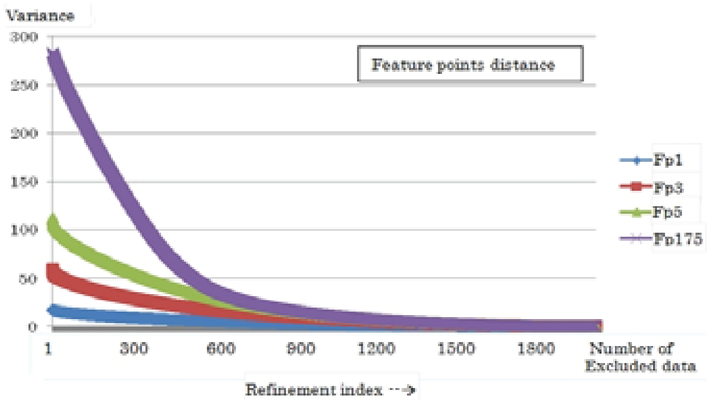

Fig.7 Variances of four distances between feature points converge with deleting iterations.

4.2. *Variation by the positions of feature points*

Fig. 8 shows all 2926 distance values between feature points with well converging status according to the number of iterations. The number of iterations is from 500 to 1935, and their results are displayed overlapped. Fig. 8(a) shows variances with the iteration numbers of 500 and above. Fig. 8(b) shows 1000 and above, Fig. 8(c) shows 1500 and above. These Figures show the variances reduce to small values after 1500 iterations. After 1750 iterations, almost all variance data go below 2, and half of them seem to be below 1. A quarter to an eighth of all data seem to be below 1, where the input data converge to stable values for register and recognition.

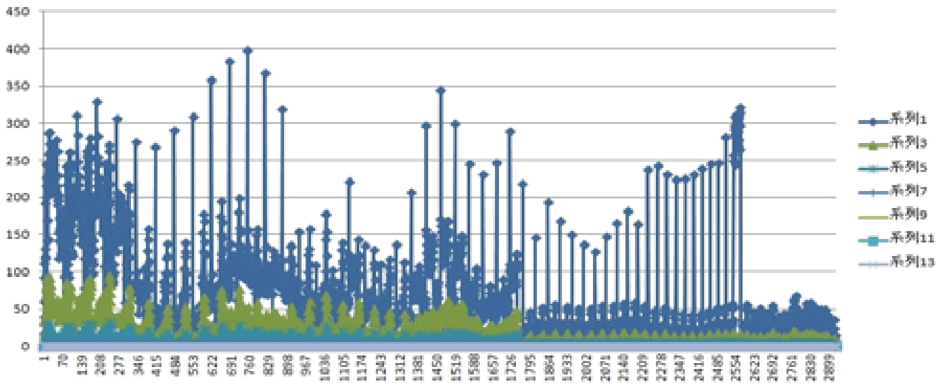

Fig.8 (a) Variances of distances between feature points converge as the number of iterations increases from 500 to 1935.

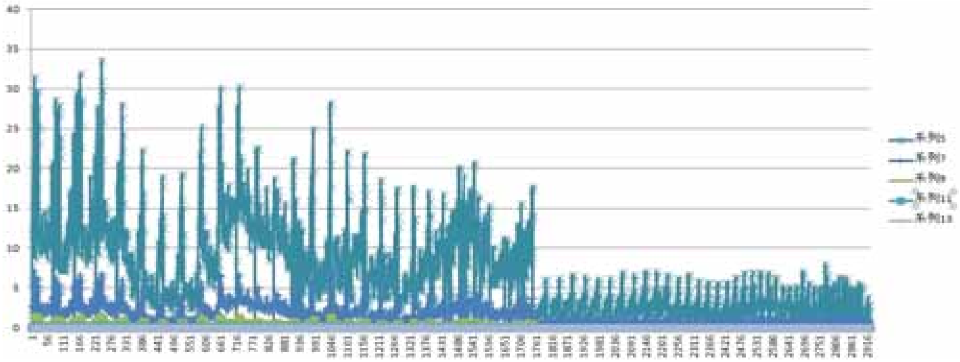

Fig.8 (b) Partial variances of distances between feature points converge as the number of iterations increases 1500, 1750, 1875, 1935. Enhanced in the vertical direction.

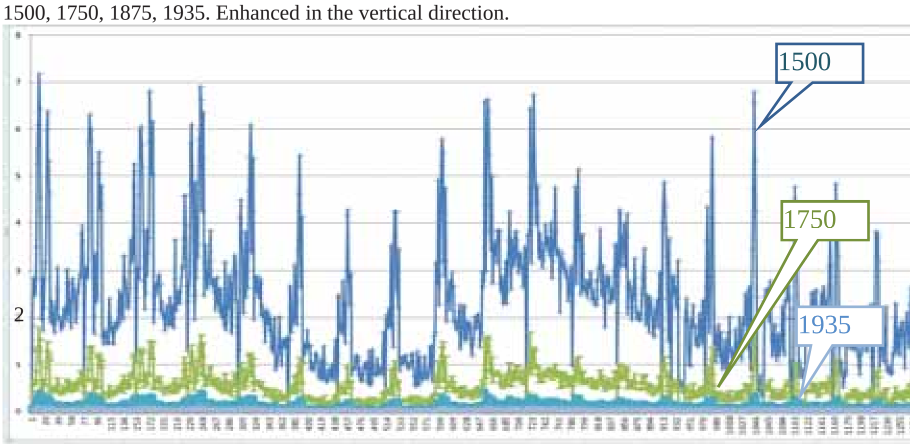

Fig.8 (c) Variances of distances between feature points converge as the number of iterations increases from 1500 to 1935. For the 1750 case, most feature values are below 2 and half of them are below 1.

4.3. ***Maximum interval between minimum variance data***

After adapting the MVM to the length data, a quarter to an eighth of the resulting data form a set with a small variance of 1-2. This quarter means 3.75-7.5 Hertz in the time direction because the original data was 30 Hertz. Face recognition can work at about 5Hz on average in the best conditions. But the value of 5Hz is average and we must consider the possible worst case. Thus, the maximum interval between the converged data with small variances, which shows the worst case for detecting the best shot, is searched for. Fig. 9 shows the maximum interval between the best shot frames vs. thresholds of deviation between the average value and the length value. The average values are the final values obtained by experiment using an MVM, as in section 3.2.2.

4.4. ***The number of distinguishable people***

4.4.1 Theoretical data

For identification by face recognition, the number of distinguishable people will be estimated. The total number of feature values is 2926 as the pieces of distances between feature points. It is important to select feature values whose distance value is large enough to distinguish people. Fig. 10 shows a hundred feature values from the first number. The number of feature values that have a distance larger than five is 45. A distance value larger than 4 can be assumed to be the case for a standard deviation larger than four in former Fig.6 in Ohzeki et al (2015). If we take four more feature values (in total five feature values) we can find the number of distinguishable people is $5^5 = 3125$ [Tanaka et al. (2011)].

Let us take a statistical value of the difference between two line segments consisting of two feature points on each face. Let an average be A and standard deviation σ. If the average A is small, then we cannot discriminate two persons because the distance between two features is small. I the average A is sufficiently large, as A> σ, then we can discriminate two persons as they have different features. For A> σ, at least 70% of normal distribution locates in different area and 30% is in the same. If we take five independent feature statistic values, 0.3^5=0.00243 and for six values 0.3^5=0.000729. We

.

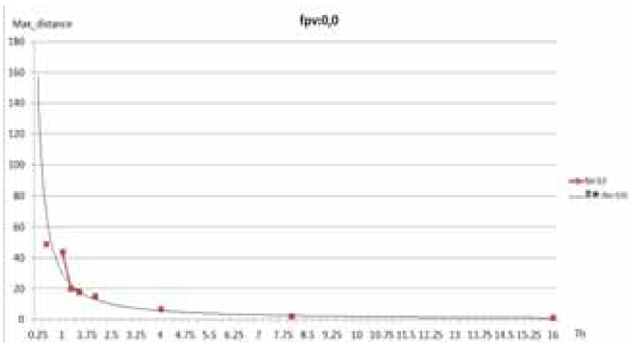

Fig.9 The maximum interval between the best shot frames vs. thresholds of deviation between the average value and length value. Dots are measured intervals. The value belongs to the best shot if it is smaller than the threshold.

would like to use three to six pairs of feature points which are on horizontal, vertical and slant lines and may be independent. The experiments are performed in the later sections.

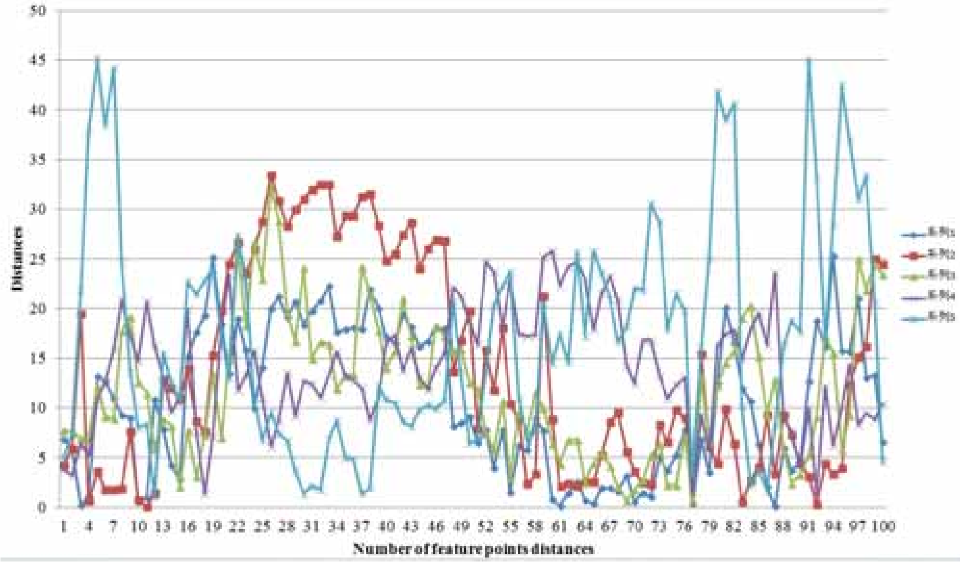

Fig.10 Five distance values obtained by six testers. The horizontal axis is the numbering of 100 feature values out of 2926. The feature values that have a distance of more than 4 are good distinguishable features for recognition.

4.4.2 Experiments using real data

Based on the considerations above, we now conduct experiments by sampling statistical data using uniform distribution characteristics. We assume sampled feature point data, and subtracted data follow normal distributions. Video specifications for these experiments are as follows. First, speaking persons are recorded for three minutes, then 2000 frames from the video sequence which represent about a one-minute period are used for the experiments. On registration of each person's face, using MVM as described in section 3.2.2, the frame after 1750 iterations out of 2000 frames which can be considered as well-converged, is selected for recognition. The number of persons involved is six.

To recognize other persons apart from oneself, it is used to compare the difference in the distance between two feature points.

(1) For the case of using a difference between two feature point distances

Six pairs of the two feature point distances are used as the difference. The total number of feature points is 77. The number of combinations to take two from 77 is 2926. The six important pairs are presented in Ohzeki(2014b) representing vertical height, horizontal width of a face and other major distances between eyes and mouth edge points. The six pairs are represented by the number of feature points as (1,11), (6,14), (14,30), (21,58), (57,58), (57,70). The points are indicated in Fig.6. Out of six persons, we get 15 combinations in choosing two persons. Then, there are 15 combinations for statistical values of line segments consisting two feature point distances for six persons.

Depending on the length of the differences of feature points the variances are different. Fig. 11 shows distributions of differences of feature point distances and variances for different pairs.

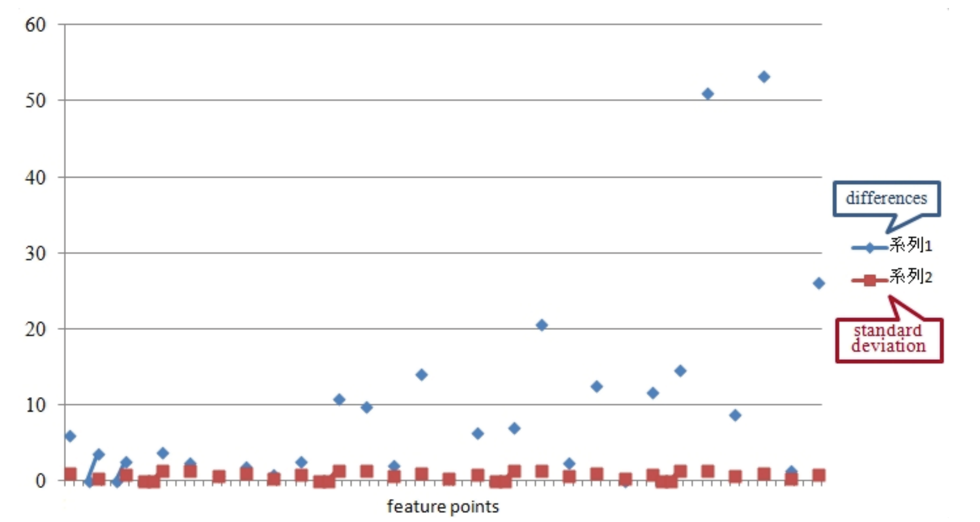

Fig.11 Differences and standard deviations for many feature points in the horizontal axis. The vertical axis represents values of the differences and standard deviations.

As this graph shows sampled data statistics, it represents differences of feature point distances and standard deviations. The differences are absolute values of the two distances of the two feature points in two dimensional images. This graph shows an intermediate result because this is a sampled distribution from many fixed pairs. This graph is only effective for considering how much the trial data distribute when many trials of taking differences of the fixed pairs of feature points distances for the many same pairs of two persons. For the recognition stage, we should obtain the averages and variances (or standard deviations) and evaluate the ratio of the average and the standard deviation to see how near or far away the ratio is from zero distance, at which we cannot distinguish the two persons. To evaluate statistics for recognition, we assume that the differences of the two feature point distances for different persons is independent in statistics and the distribution follows the normal distribution. In these experiments, we assume the differences follow a normal distribution with the average μ , and the standard deviation σ .

Table 1 shows a result for all obtained difference data by a pair of a two feature point distance case. The average of the sampled data is 18.5, and the standard deviation is 13.2, while the ratio of the average to standard deviation is 1.40. The ratio corresponds to 1.4 σ if the sampled data really follows the normal distribution, The recognition rate for this distribution is limited to 85% and we can only distinguish 85 persons at most without error.

Table 1. Minimum values of differences for combinations of one to five persons.

Number of data	The average of the sampled data	The standard deviation	Ratio
15	18.548	13.220	1.403

(2) For the case of using three and six differences between two feature point distances

Next, we increased the number of differences of two feature point distance up to three and six. In this experiment, we consider sums of several pairs of differences between feature point distances for the same two persons. For the case using three pairs of differences, there are 20 combinations by ${}_6C_3=6\times5\times4/(3\times2\times1)=20$. The selection of these three out of six is also made according to the importance of features, that is one is the horizontally longest feature pair, the other is the vertically longest pair, and another is a pair of an eye and mouth edge points. The numbers of feature points of these three pairs are (1,11), (6,14), (21,58). Fig. 12 shows distributions of differences of feature point distances and variances for different pairs. This graph is only for confirmation of the distribution. Next, the fixed three pairs of differences of feature point distances for pairs of two persons are measured. The number of pairs of two persons from six persons is 15. Table 2 shows a result for all obtained difference data by pairs of two persons with a distance function of three differences. The average of the sampled data is 61.65, and the standard deviation is

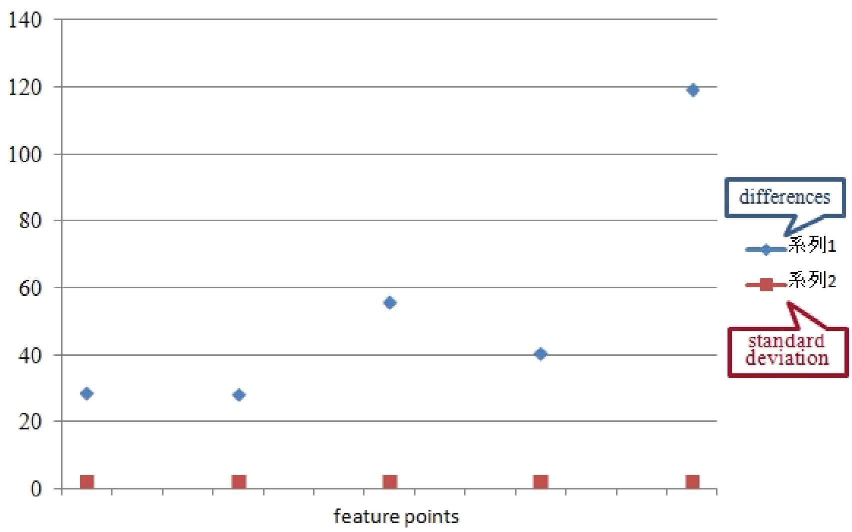

Fig. 12 Differences and standard deviations for five feature points in the horizontal axis. The vertical axis represents values of the differences and standard deviations. Three differences are used for distance measurements.

20.06 with the ratio of the average to standard deviation is 2.20. The ratio corresponds to 2.20 σ if the sampled data really follows the normal distribution, The. recognition rate for this distribution is limited to 97.2% and we can only distinguish 972 persons out of 1000 at most without error.

Table 2. Minimum values of differences for combinations of one to five persons. Three differences are used for distance measurements.

Number of data	The average of the sampled data	The standard deviation	Ratio
15	61.65	28.057	2.197

(3) For the case of using six differences between two feature point distances
Finally, we increased the number of differences between two feature point distances to six which was already introduced in the above (2). For using the six pairs of the differences of feature point distances out of six pairs, there is only one combination as to feature points. The six differences are summed up to a single evaluation value. As the number of cases is small, three minor variations are tried. They are (a) the minimum distance value only adopted from five distance values for different persons, (b) the absolute sum of five distance values for different persons. (c) the maximum values among absolute values of five difference pairs of different persons. The results are shown in Table 3. Variations (a) seem to be slightly better than others in performance. The result of (a) is 2.39 σ which shows that the recognition rate is 98.3%, and 983 persons out of 1000 can be identified.

Table 3. Differences for three versions of distance measures (a), (b) and (c). Six differences are used for distance measurements.

Version	Number of data	The average of the sampled data	The standard deviation	Ratio
(a)	15	52.81	22.08	2.39
(b)	15	79.74	36.64	2.18
(c)	15	34.71	15.50	2.24

4.5. *Results*

The proposed minimum variance method (MVM), which shakes off non-frontal face images from input video sequence, extracts eighth frontal face images out of 2000 facial images. The variance of the feature points is small enough for the extracted images. As for statistic values of the feature point distances, the standard deviation locates around one σ for a single statistic value, exceeds two σ for a set of three statistic values, and reaches 2.39 σ for a set of six statistic values. The last case corresponds to 98.3 percent of authentication rate.

5. Conclusions and Future Work

A new method of extracting the best shot from a moving video input for face recognition is proposed. The best shot is obtained at the average probability of 1/4 to 1/8, which means 3.75-7.5 Hz in the time direction on average. The best shot can be obtained for all 2926 combinations of feature points. The maximum interval between the best shots is 0.8 second, while the average time between the best shots is in 0.13-0.27 second. The recognition rate was estimated using the normal distribution assumption and independency of data. Using distance measurements of six difference pairs of feature point distances, nearly 1000 persons can be distinguished. The minimum variance method takes much time to obtain the converged frame and a faster method will be studied in the future.

References

Aoyama R.; Hirakawa, Y; Ohzeki, K (2015), “Face recognition by selecting the best frontal image from video sequence”, IMPS I-1-3 Nov. (in Japanese)

Bonnen, K.; Klare, B.F.; Jain, A.K. (2013) "Component- Based Representation in Automated Face Recognition", IEEE Transactions on Information Forensics and Security, Vol.8, No.1 pp.239-253, Jan.

Drira, H.; Ben Amor, B. ; Srivastava, A. ; Daoudi, M. ; Slama, R. (2013) ” 3D Face Recognition under Expressions, Occlusions, and Pose Variations”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 35, ,Issue: 9, pp.2270 – 2283 Feb.

Gohringer, C. (2012) “Advances in Face Recognition Technology and its Application in Airports”, Allevate Limited. pp.1-10., 17 Jul. http://allevate.com/index.php/2012/07/17/

Grother, P. J.; Quinn, G. W.; Phillips, P J. (2010) “Report on the Evaluation of 2D Still-Image Face Recognition Algorithms”, NIST Interagency/Internal Report (NISTIR) – 7709 June.

Grother, P. J.; Quinn, G. W.; Ngan, M. (2012) “Still Face Image and Video - Concept, Evaluation Plan and API Version 0.5" Face Recognition Vendor Test(FRVT), National Institute of Standards and Technology (NIST) April.

Grother, P.; Ngan, M. (2014) “Face Recognition Vendor Test (FRVT) Performance of Face Identification Algorithms” NIST Interagency Report 8009 May.

Guillaumin, M. ; Verbeek, J. ; Schmid, C.(2009) “Is that you? Metric Learning Approaches for Face Identification”, Proc. of IEEE 12th International Conference on Computer Vision, 2009 pp. 498–505, Sept.

Howell, A.J.; Buxton, H. (1996) "Towards unconstrained face recognition from image sequences," in Proceedings of the Second IEEE International Conference on Automatic Face and Gesture Recognition, pp.224-229.

Jafri, R.; Arabnia, H. R (2009) "A Survey of Face Recognition Techniques", Journal of information Processing Systems Volume: 5, No: 2, pp. 41-68.

JCB 2014 Conference report http://www.ijcb2014.org/IJCB%20Conference/IJCB14_Conference_Report.pdf

Majumdar,; A.; Nasiopoulos, P. (2008) "Frontal Face Recognition from Video", Proceedings of 4th International Symposium, ISVC, Part II, Volume 5359, Lecture Notes in Computer Science Springer Advances in Visual Computing, pp 297-306, Dec.

Milborrow, S.; Nicolls, F. (2008) "Locating Facial Features with an Extended Active Shape Model", Proceeding of ECCV Part IV pp.504-513, Springer-Verlag Berlin, Heidelberg.

Milborrow, S; Nicolls, F. (2014) "Active Shape Models with SIFT Descriptors and MARS", International Conference on Computer Vision Theory and Applications (VISAPP) pp.380-387.

Ngan, M.; Grother, P. (2014) “Face Recognition Vendor Test (FRVT) Performance of Automated Age Estimation Algorithms”, NIST Interagency Report 7995 Mar.

Ohzeki, K.; Wei, Y. Y.; Hirakawa, Y.; Sugimoto, T. (2014) “Authentication System using Encrypted Discrete Biometrics Data”, Proceedings of TRUST 2014 Greece Springer LNCS 8564 pp.210-211 June 30-July2.

Ohzeki, K,; Takatsuka, M.; Kajihara, M.; Hirakawa, Y.; Sato, K. (2014b) "On the False Rejection Ratio of Face Recognition Based on Automatic Detected Feature Points", Proc. international workshops on “Pattern Recognition and Image Understanding”OGRW-9 Mo.3-1, ogrw2014_024_Ohzeki.pdf, Dec.

Ohzeki, K.; Aoyama, R.; Hirakawa, Y. (2015), “Minimum Variance Method to Obtain the Best Shot in Video for Face Recognition”, Proceedings of the 8th International Symposium on Multimedia – Applications and Processing, (MMAP 2015) in Fedcsis, pp.881-886,Sept.

Tanaka, Y.; Miyazaki, E.; Ohzeki, K. (2011), “Feature Point Analysis Using Facial Parts for Face Recognition”, National Convention, D-12-36 Institute of Electronics, Information, and Communication Engineers Mar. (in Japanese)

Toderici, G.; Passalis, G.; Zafeiriou, S. ; Tzimiropoulos, G. (2010) “Bidirectional relighting for 3D-aided 2D face recognition”, Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2721-2728, 2010 June.

Torres, L. (2004) "Is there any hope for face recognition?" in Proc. of the 5th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2004). Lisboa, Portugal.

Zhang P. (2008) "A video-based Face detection and recognition System using Cascade Face Verification Modules”, Proc. of 37th IEEE Applied Imagery Pattern Recognition Workshop, pp. 1-8, Oct.