MVA2005 IAPR Conference on Machine VIsion Applications, May 16-18, 2005 Tsukuba Science City, Japan
11-2

Robust Face Recognition under various illumination conditions

Atsushi Matsumoto† Nobutaka Shimada‡ Takuro Sakiyama† Jun Miura† Yoshiaki Shirai†

† Graduate School of Engineering
Osaka University
2-1, Yamadaoka, Suita, Osaka, 565-0871, JAPAN.
matumoto@cv.mech.eng.osaka-u.ac.jp
{saki, jun, shirai}@mech.eng.osaka-u.ac.jp

‡ Department of Human and Computer Intelligence
Ritsumeikan University
1-1-1, Noji-Higashi, Kusatsu, Shiga, 525-8577, JAPAN.
shimada@ci.ritsumei.ac.jp

Abstract

We propose a method of face identification under various illumination conditions. Because we use image based method for identification, the accuracy position of the face is required. First, face features are detected, and the face region is determined using the features. Then, by registering the face region to the average face, the horizontal position of the face is adjusted. Finally, the size of the face region is adjusted based on the distance of two eyes determined from all input frames. Because the size of the normalized face image is the same for all faces, the face length feature is lost in the normalized face image. The face is classified into three categories according to the face length, and the subspace is generated in each category so that the face length feature is preserved. We demonstrate the effectiveness of the proposed method by experiments.

1 Introduction

Purpose of this study is to develop a face identification system under various illumination conditions. The input of this system is images of a person taken by the user. Many methods of face identification are proposed, but few of them consider illumination change.

One of methods is the use of PCA [1]. Because only one image is used, the method may fail under various facial expressions. Fukui proposed the constrained mutual subspace method(CMSM) [2], which uses an image sequence.

Because the CMSM requires the accurate face position, Fukui[2] detects pupils and nostrils for face registration. If view point changes, nose detection may fail. In addition, if a face is illuminated from one side, shadows cast around features or in non-feature regions may cause misalignment of the face features or detection error. We adjust the face position by matching the face normalized based on the face features and an average face.

One of factors that affect recognition rate is illumination. Because two-dimensional facial pattern is amenable to illumination, we approximately normalize illumination by image processing.

Another factor is the face length feature. If a face is normalized based on the face features, the face length feature is lost. We classify faces into three categories according to the face length, and match unknown face image to these in each category.

(a)Input (b)Difference image from the skin color (c)Face region

Figure 1: Result of face region extraction

2 Face detection and registration

2.1 Face region extraction

Because the skin color varies depending on illuminations and individuals, we determine the skin color in every frame from the region which is clearly a part of the face. The face must be near the center of the image because the image is taken by the user. In the first frame, the region is set at the center of the image. If the face was detected in the previous frame, the region is set at the center of the detected face. The size of the region is defined as 50 x 50 pixels. The pixel values in the region are projected into the YUV space. The skin color is determined to be (U, V) which gives the peak of the two-dimensional histogram of U and V. The skin pixel is defined to be the pixel whose value in the UV space is closer to the skin color than a threshold (Figure 1(b)). The face region is determined by extracting the largest connected region of those pixels and by filling blanks (Figure 1(c)).

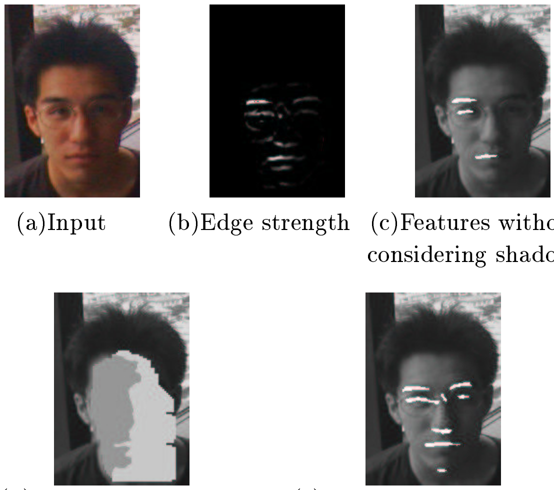
2.2 Feature detection

We detect six face features (eyebrows, eyes, nose and mouth) in the face region. These features have following common characteristics.

- The direction is horizontal and the shape is ellipsoid.
- The brightness is darker than surrounding neighbors.

Pixels which satisfy the above conditions are extracted. The candidates for the face features are extracted as the connected region of those pixels. The position of each candidate is determined to be the center of corresponding region.

If the face is illuminated from one side(Figure 2(a)), the contrast is different between the bright region and



(a)Input (b)Edge strength (c)Features without considering shadows

(d)Divided two regions (e)Features in our method

Figure 2: Result of feature detection

the dark region(Figure 2(b)). If one threshold is used, feature detection may fail in the dark region(Figure 2(c)). Therefore, we divide the facial region into two regions based on the brightness (Figure 2(d)), and detect face features by thresholding by a different threshold in each region (Figure 2(e)).

Each face feature is determined to be the candidate which has the strongest matching probability (Araki[3]).

2.3 Normalization of the face image

In order to identify individuals, the size and the position of the face are normalized based on the distance of two eyes. By letting (x_{re}, y_{re}) and (x_{le}, y_{le}) denote the position of the right and left eyes, the distance of two eyes D_e is defined as

$$D_e = x_{re} - x_{le} \tag{1}$$

First, we temporarily normalize the size and the position of the face. The face region is approximated by

$$x_{ce} - 0.8D_e \leq x \leq x_{ce} + 0.8D_e \tag{2}$$

$$y_{ce} - 0.5D_e \leq y \leq y_{ce} + 1.1D_e \tag{3}$$

where (x_{ce}, y_{ce}) is the position of the center of two eyes. The temporal face image is generated by normalizing the image in the square region into 20×20 pixels.

Because the temporal face image is normalized based on the face features, the misalignment of the face features causes the misalignment of the face. If shadows are cast on either side of the face features, it is difficult to detect the accurate position of the face features. Furthermore, if pupils move, the position of eyes may be misaligned (Figure 4(a)).

Here, we correct the horizontal misalignment by using the average face (Figure 3). Let $\mathbf{u}(x)$ denote the vector composed of the pixel values of the face image moved x pixels horizontally. The best position is determined based on the similarity between the face image

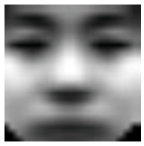
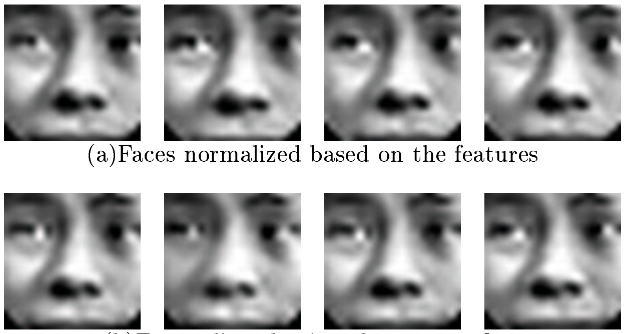


Figure 3: Average face



(a)Faces normalized based on the features

(b)Faces aligned using the average face

Figure 4: Result of registration using the average face

$\mathbf{u}$(x) and the average face $\mathbf{v}$ (Figure 4(b)). The similarity S is defined as

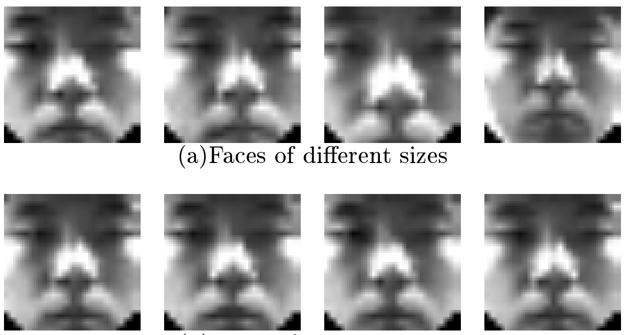
$$S(x) = \frac{\mathbf{u}(x)^T \mathbf{v}}{|\mathbf{u}(x)||\mathbf{v}|} \tag{4}$$

The best face position x satisfies the following condition:

$$arg \min_{-8 \leq x \leq 8} (\mathbf{u}(x), \mathbf{v}) \tag{5}$$

If the size of the face in the face image is different between the input face and the training face, identification by the CMSM may fail. The face size may vary in each frame if the face features are not detected accurately. (Figure 5(a)).

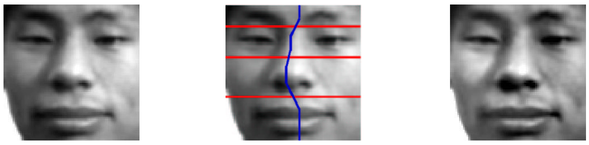
Because the face size must not change in the input images of two seconds, the size of all faces is determined from the peak of the histogram of distance D_e. The positions of eyes are adjusted in each frame so that the distance between two eyes corresponds to the one which gives the peak value (Figure 5(b)).



(a)Faces of different sizes

(b)Faces of the same size

Figure 5: Result of registration using the distance between eyes



(a)Face containing illumination effect (b)Devided eight regions (c)Face normalized illumination

Figure 6: Result of nromalization of illumination

3 Face identification

We use the CMSM for identification. First, the face image is converted to the grayscale image. Because the CMSM uses the vector composed of the pixel values of the face image, the effect of the variation of the illumination condition needs to be removed. In addition, if the face is normalized into one size, the length of the face is lost.

First we normalize illumination, and then solve the problem of the face length.

3.1 Normalization of illumination

We assume that the front face is approximated by a part of an ellipsoid except a nose. If the face is illuminated from one side, this side of the face become brighter. We approximately normalize illumination in such a case.

The face is horizontally divided into two regions by the brightness, and eight regions are generated by vertically dividing into four regions according to the face features (Figure 6(b)). Assuming that the pixel values are distributed as normal distribution, the average and the variance of the right and left region divided according to the face features are adjusted. Near the boundary lines, pixel values are smoothed. Figure 6(c) shows an example of normalization of illumination.

3.2 Subspaces depending on the face length

Because the face is normalized based on two eyes, the position of the mouth is not normalized. Therefore, the mouth may be out of the face image (Figure 7(b)). There are two ways of solving the problem of the mouth. One is to normalize the face based on two eyes and the mouth, and the other is to change the size of the face. If the face is normalized based on the eyes and the mouth, the face length information may be lost (Figure 7(c)). Therefore, we select the way to change the size.

First, we classify faces into three categories according to the ratio, which is defined as

$$R = \frac{y_m - y_{ce}}{D_e} \tag{6}$$

where y_m is the y coordinate of the mouth. The classification C of the face with R is expressed as

(a)Input (b)Face normalized based on eyes (c)Face normalized based on eyes and the mouth

Figure 7: Loss of the face length information

(a)Face in the category C=1 (b)Face in the category C=2 (c)Face in the category C=3

Figure 8: Examples of training face

$$\begin{cases} C = 1 & if(R \leq 1.050) \\ C = 2 & if(1.050 < R \leq 1.172) \\ C = 3 & if(1.172 < R) \end{cases} \tag{7}$$

The height/width ratio in each category is defined as

$$\begin{cases} 1.0 & (for\ C = 1) \\ 1.1 & (for\ C = 2) \\ 1.2 & (for\ C = 3) \end{cases}$$

In the training phase, the face is classified according to the ratio of the distance of two eyes to the vertical distance between the center of two eyes and the mouth. Each distance is determined to be the peak of the histogram of the distance. The subspace is generated in the category.

Because the face length of a person may vary due to the change of the view point, the face in the category C may become category $C \pm 1$. Therefore, the subspace is generated in the category $C \pm 1 (1 \leq C \leq 3)$. Figure 8 shows examples of the face images generated in the training phase.

In the identification phase, the input face is classified into a category in the same way and identified in the corresponding subspace.

Table 1: Identification rate

illumination condition	*the original CMSM*	*our method*
condition 2	86.7%	90.0%
condition 3	80.0%	90.0%
total	83.3%	90.0%

4 Experiment

We consider the following three illumination conditions:

1: Upper fluorescents.
2: Upper fluorescents and a lamp from left.
3: Upper fluorescents and a lamp from front.

We collect faces of 30 individuals under the above conditions. The faces under condition 1 are used for training, and the rest of the faces are used for identification. We compare our method with original CMSM[2].

Table 1 shows the identification rate. The identification rate by our method is better in both illumination conditions.

Figure 9 shows examples of identification which failed by the original CMSM and succeeded by our method. Normalized face images used for identification by each method are shown in Figure 10, where (b) and (e) are generated by the original CMSM from input image in the upmost image and the middle image in figure 9 respectively. In (b), because shadows are cast around the right pupil and the shadow of the left nostril get smaller, the position of the right pupil is misaligned toward right, and that of the left nostril is misaligned upward. Therefore, a part of the mouth is out of the normalized face image. In (e), because the shadows of two nostrils get smaller, the position of nostrils is misaligned upward. Therefore, a part of the mouth is out of the normalized face image.

Normalized face (c) and (f) in the figure 10 are generated by our method. The mouth is contained in (c) and (f), and the position of the right eye is adjusted in (c).

5 Conclusion

In this paper, we have proposed a method of face identification under illumination from front or one side. The proposed method applies registration using the average face and the results of face detection in all frame of input. In addition, the proposed method classifies faces into three categories according to the ratio of the distance of two eyes to the vertical distance between the center of two eyes and the mouth, and generates the subspace in each category. We have shown that the proposed method is robust against various illumination conditions.

The future work includes identification under backlighting, evaluation with a larger number of persons, and normalization of illumination considering the shape of the face.

(a)One of input images (b)Identification result by the original CMSM (c)Identification result by our method

Figure 9: Examples of identification result

(a)Training face (b)Normalized face in the original CMSM (c)Normalized face in our method

(d)Training face (e)Normalized face in the original CMSM (f)Normalized face in our method

Figure 10: Examples of the normalized face image in each method

Acknowledgment

The author would like to thank TOSHIBA for the identification software rental.

References

[1] M.A. Turk and A. P. Pentland, "Face recognition using eigenfaces", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 586–591, June 1991

[2] Osamu Yamaguchi, Kazuhiro Fukui and Ken-ichi Maeda, "Face recognition using temporal image sequence", *Proceedings of the Third International Conference on Automatic Face and Gesture Recognition (FG98)*, pp.318-323, 1998

[3] Y. Araki, N. Shimada and Y. Shirai, "Detection of Faces of Various Directions in Complex Backgrounds", *Proceedings of International Conference on Pattern Recognition*, pp.409-412, 2002.