157-159

Recognition of Human Front Face in Grayscale Images

Mohamed Rizon[1], Phaklen Ehkan[2], A. Nasir Che Rosli[2].
[1]School of Mechatronics Engineering, [2]School of Computer and Communication Engineering
Northern Malaysia University College of Engineering
01000 Kangar, Perlis, Malaysia

***Abstract*-We propose a new technique to recognize the unknown person in grayscale face images for which the position, scale and image-plane rotation of the face are unknown. The proposed system detects the iris of both eyes and normalizes the position, scale and image-plane rotation of the face using iris positions. The algorithm measures the degree of matching between the image and the face template of each person using normalized cross-correlation value. Recognition of a human face is done using the highest degree of matching.**

Keywords: face recognition, iris detection, template matching.

1. INTRODUCTION

Recognition of human faces has many applications such as security systems, mug shot matching and model-based video coding. There are two approaches for face recognition: template based approach [1]～[3] and feature based [1], [4].

In this paper, we propose a template-based face recognition system. The proposed system can be applied to a face image in which the position, scale and image-plane rotation of the face are unknown. The proposed system first detects the irises of both eyes and then normalizes the position, scale and image-plane rotation of the face using the positions of the irises of both eyes.

The iris is a darker region than its surrounding and the shape of the iris is a circle. Using these properties of the iris, the proposed system measures the iris-likeness of each blob. In addition, the proposed system measures the iris-likeness of each blob-pair using a template of both eyes obtained by cutting off the region of both eyes from a face image.

2. THE PROPOSED FACE RECOGNITION SYSTEM

We assume the followings about the input image.

1) The image is an intensity image.

2) The image is a head-shoulder image with plain background and the face in the image is a frontal face.

3) The size and the image-plane rotation of the face in the image are unknown although their lower bounds and upper bounds are given in advance.

4) The irises of both eyes appear in the image.

Below, we show the face recognition system proposed in this paper.

2.1. Modeling stage

For each person, we prepare an intensity image of his frontal face and cut off a region consisting of both eyes and eyebrows from the image. The region is called the face template of the person and is denoted by $T(u,v)$, $0 \leqq u \leqq m-1$ and $0 \leqq v \leqq n-1$ We prepare another template $T_E(u,v)$, $-p \leqq u \leqq m-1+p$ and $-p \leqq v \leqq n-1+p$, for each person. (p was set to 10). The template T_E is called the extended template.

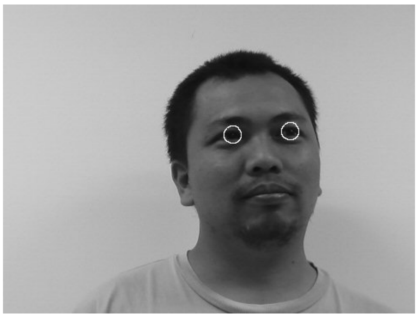


Figure 1. The positions of the irises of both eyes detected by the iris detection algorithm from input image

2.2. Face recognition stage

(1) Extract the irises of both eyes of the unknown person in the input image using the algorithm shown in the next section. Let (x_1, y_1) and (x_2, y_2) denote the positions of the irises of the left and right eyes (see Fig.1).

(2) For each person in the database, do the following substeps (2-1) to (2-4).

(2-1) Let (u_1, v_1) and (u_2, v_2) denote the positions of the irises of the left and right eyes in the face template of the current person. Then, apply an affine transform to the input image so that the positions of the irises (x_1, y_1) and (x_2, y_2) in the input image have the coordinates (u_1, v_1) and (u_2, v_2) in the new image.

(2-2) Let $T_E(u,v)$, $-p \leqq u \leqq m-1+p$ and $-p \leqq v \leqq n-1+p$, denote the extended template of the current person. Then, cut off a window $W(u, v)$, $-p \leqq u \leqq m-1+p$ and $-p \leqq v \leqq n-1+p$, from the input image after transformation.
Moving the face template T over the window W, compute the normalized cross-correlation value C by

$$C = \frac{E(TW_T) - E(T)E(W_T)}{\sigma(T)\sigma(W_T)} \quad (1)$$

W_T denotes a patch of W matched to T. E(T) and σ (T) are the average and the standard deviation of the intensities of pixels inside T. E(W_T) and σ (W_T) are similarly defined. E(TW_T) denotes the average of the pixels-by-pixel products of T and W_T.

(2-4) Find the maximum value of C over all pixels inside W and let the maximum value of C denote the degree of match between the current person and the input image.

(1) Find a person that has the best match to the input image. We determine this person to be the unknown person in the image.

3. DETECTION OF THE IRISES

For each pixel (x,y) inside the face region, the algorithm computes a cost C(x, y) by

$$C(x,y) = C_1(x,y) + C_2(x,y) \quad (2)$$

Let S(x,y) denote the square region with center (x,y) and side-length d. Then, $C_1(x,y)$ is given by

$$C_1(x,y) = \sum_{j=y-d/2}^{y+d/2} V_r(j) + \sum_{i=x-d/2}^{x+d/2} V_c(i) \quad (3)$$

$V_r(j)$ and $V_c(i)$ are the mean crossing numbers of row j and column i which are defined as follows.

Let I(i,j) denote the intensity values of pixels (i,j) in the image. Then, for each row j, $V_r(j)$ represents the number of pixels (i,j), $x-d/2 \leqq i \leqq x+d/2$, such that one of the I(i−1,j) and I(i,j) is greater than μ plus K and the other is smaller than μ minus K where μ is the average intensity of pixels (i,j), $x-d/2 \leqq i \leqq x+d/2$ and K is a constant. $V_c(i)$ denotes the number of pixels (i,j), $y-d/2 \leqq j \leqq y+d/2$, such that one of the I(i,j−1) and I(i,j) is greater than μ plus K and the other is smaller than μ minus K.

$C_2(x,y)$ is a function which evaluates the intensity difference between the central part and the boundary parts of S(x,y).

After extracting the valley region from the face region, the proposed algorithm computes costs C(x,y) for all pixels (x,y) inside the valley region using the feature template proposed by [6]. The costs are given by (2). Next, as feature points, the proposed algorithms selects pixels (x,y) that give the local maxima of C(x,y).

In the proposed iris detection algorithm, the irises are modeled by circles. Let (x,y) and r denotes the centers and the radii of the irises modeled by circles.

According to [7], the separability η between regions R_1 and R_2 of the eye template is given by

$$\eta = \frac{B}{A} \quad (4)$$

$$A = \sum_{i=1}^{N} (P_i - \overline{P}_m)^2 \ , \quad B = n_1(\overline{P}_1 - \overline{P}_m)^2 + n_2(\overline{P}_2 - \overline{P}_m)^2$$

where

n_k $(k = 1,2)$: the number of pixels inside R_k,

$N=n_1+n_2$,

$\overline{P}_k$ $(k = 1,2)$: the average intensity inside R_k,

$\overline{P}_m$:the average intensity inside the union of R_1 and R_2,

P_i :the intensity values of pixels i.

Let η (x, y, r) denote the separability between two regions R_1 and R_2 in the template with size r placed at pixel (x, y). As the candidates for the irises, we select triplets (x, y, r) which give the local maxima of η (x, y, r). The circles defined by these triplets are called blobs in this paper.

Let $B_i=(x_i, y_i, r_i)$ denote blobs obtained by the procedure shown in the last section. First, we apply the Canny edge detector [9] to the head region and then measures the fit of blobs to the edge image using the circular Hough transform [10]. We give the equation of a circle by

$$(x-a)^2 + (y-b)^2 = r^2 \quad (5)$$

where (a, b) is the circle center and r is the radius.

Let P denote an edge point lying on the boundary of the circle corresponding to the iris. If P has the position (x, y) and the orientation θ and the error of θ is at most Δθ, the center(a, b) of the circle lies on the arc given by R={(a, b)|a=x+r cos(t), b=y+r sin(t), $\theta + \pi - \Delta\theta \leqq t \leqq \theta + \pi + \Delta\theta$}. Below, we denote the arc defined above by (x, y, r, $\theta + \pi - \Delta\theta$, $\theta + \pi + \Delta\theta$). In the experiments, Δθ was set to 45 [degrees].

First, for each $r \in \{r_i - 1, r_i, r_i + 1\}$, and for each edge point P inside the square region with center (x_i, y_i) and side-length 2r, enumerating all integer points (a,b) on the arc (x, y, r, $\theta + \pi - \Delta\theta$, $\theta + \pi + \Delta\theta$) where (x,y) and θ are the position and the orientation of P, vote for (a, b, r). Next, select (a, b, r) with the largest vote. We denote the largest vote by V(i), which is called the vote for B_i.

Next, we first place the template of Fig.2(a) at the position (x_i, y_i) on I(x, y) and then compute the separabilities $\eta_{23}(i)$ and $\eta_{24}(i)$ by using Eq.(4), where $\eta_{kl}(i)$ denotes the separability between regions R_k and R_l. Similarly, we compute the separabilities $\eta_{25}(i)$ and $\eta_{26}(i)$ using the template of Fig.2(b).

Finally, the cost of each blob B_i is given as follows.

$$C_b(i) = C_1(i) + C_2(i) + C_3(i) + C_4(i) \quad (6)$$

$$C_1(i) = \frac{V_{max}}{V(i)}, \quad C_2(i) = \frac{|\eta_{23}(i) - \eta_{24}(i)|}{\eta_{23}(i) + \eta_{24}(i)}$$

$$C_3(i) = \frac{|\eta_{25}(i) - \eta_{26}(i)|}{\eta_{25}(i) + \eta_{26}(i)}, \quad C_4(i) = \frac{U(i)}{U_{av}}$$

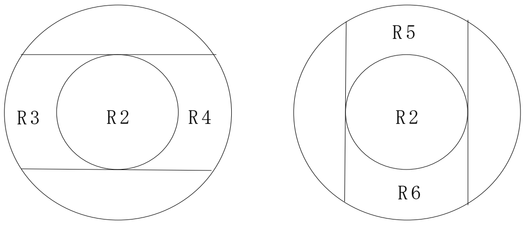
where

V(i) : the vote for B_i given by Hough transform,

V_{max} : the maximum of V(i) over all blobs,

U(i) : the average intensity inside B_i,

U_{av} : the average of U(i) over all blobs.



(a) (b)

Figure 2. The templates used to compute $C_2(i)$ and $C_3(i)$

For each pair of blobs B_i and B_j , let d_{ij} and θ_{ij} denote the length and the orientation of the line connecting the centers of B_i and B_j . The proposed algorithm computes a cost for each pair of blobs B_i and B_j such that $L/4 \leqq d_{ij} \leqq L$ and $-30^\circ \leqq \theta_{ij} \leqq 30^\circ$ where L is the difference between the x-positions of the left and right contours of the head. And, a pair of blobs with the smallest cost is determined to be the irises of both eyes.

The cost of a blob-pair B_i and B_j is given by

$$F(i, j) = C_b(i) + C_b(j) + 1/R(i, j) \qquad (7)$$

where $C_b(i)$ and $C_b(j)$ are the costs computed by (6) and R(i, j) is the normalized cross-correlation value computed by using an eye template.

For each pair of blobs B_i and B_j, the algorithm computes R(i, j) by using the following procedure.

1) Apply an affine transform to the input image so that the centers of B_i and B_j are placed at the centers of the irises in the template, respectively.

2) Compute the normalized cross-correlation value R(i,j) between the template and the image by using (1).

3) If R(i,j) is smaller than 0.1, then set R(i, j) to 0.1.

4. EXPERIMENTAL RESULTS

We made experiments using face images of 80 pictures of near-front faces to evaluate the performance of the proposed face recognition system.

The success rate of the proposed system was 97.5%. All failures occurred in face recognition were due to failures of the iris detection. We can say from the results of the experiments that the performance of the proposed system is very sensitive to the iris positions.

5. CONCLUSIONS

We proposed a new technique to recognize the unknown person in grayscale face images for which the position, scale and image-plane rotation of the face are unknown. The proposed system first detects the irises of both eyes and then normalizes the position, scale, image-plane rotation of the face in the image using the positions of the irises of both eyes. After that, the algorithm measures the degree of match between the image and the face template of each person using the normalized cross-correlation value and determines a person with the highest degree of match to be the unknown person in the image.

REFERENCES

[1] R. Brunelli and T. Poggio. 1993. Face recognition: features versus templates. IEEE Trans. PAMI, 15(10): 1042-1052.

[2] D.J. Beymer. 1994. Face recognition under varying pose. Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, Seattle, WA, USA, June 21-23, 1994 : 756-761.

[3] M. Turk and A. Pentland. 1991. Eigenfaces for recognition, J. Cognitive Neuroscience, 3(1) : 71-86.

[4] S.Y. Lee, Y.K. Ham and R.H. Park. 1996. Recognition of human front faces using knowledge-based feature extraction and neuro-fuzzy algorithm. Pattern Recognition, 29(11) : 1863-1876.

[5] A. Pentland, B. Moghaddam, T. Starner. 1994. View-based and modular eigenspaces for face recognition. Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, Seattle, WA, USA, June 21-23, 1994 : 84-91.

[6] C.H. Lin and J.L. Wu. 1999. Automatic facial feature extraction by genetic algorithms. IEEE Trans. Image Processing, 8(6) : 834-845.

[7] K. Fukui and O. Yamaguchi. 1997. Facial feature point extraction method based on combination of shape extraction and pattern matching. IEICE Trans., J80-D-II(8) : 2170-2177 (in Japanese).