2010 Fourth International Conference on Genetic and Evolutionary Computing

Face detection based on feature analysis and edge detection against skin color-like backgrounds

Deng-Yuan Huang, Ta-Wei Lin, Chun-Ying Ho
Department of Electrical Engineering
Dayeh University
Changhua, Taiwan
kevin@mail.dyu.edu.tw

Wu-Chih Hu
Dep. of Computer Science and Information Engineering
National Penghu University of Science and Technology
Penghu, Taiwan
wchu@npu.edu.tw

***Abstract*—In this paper, we propose a method of face detection based on feature analysis and edge detection against skin color-like backgrounds. The proposed method consists of three phases including image preprocessing, skin color segmentation, and determination of face candidates. First, the objects in foreground are separated by image preprocessing. Second, the non-skin color objects are rejected by the method of skin color segmentation. Finally, both the elimination of skin color-like blobs in backgrounds and face detection are performed by the method of determination of face candidates. Experimental results show that the proposed method can effectively detect most of faces in skin color-like backgrounds. In addition, detection of face images with pose and expression variations or against dark backgrounds are also carried out.**

Keywords-face detection; skin-color segmentation; edge detection

I. INTRODUCTION

Recently, face detection has greatly attracted much attention of numerous researchers due to its wide variety of applications such as face recognition, human-computer interface (HCI), video surveillance, access control, and content-based image retrieval (CBIR). In real-world applications of face detection, speed and detection rate are two critical issues about which people most concern. During the past decade there has been a proliferation of literature on effective methodologies to improve the performance of detection rate. However, the speed of face detection is still a problem that seriously hinders it from being widely used in real-world applications. This problem is not completely solved until the algorithm of face detection based on AdaBoost and Haar-like features was proposed by Viola and Jones [1]. Moreover, face detection is often preceded by the extraction of skin-tone colors [2], since it is one of the most important cues of face features due to its invariance of face scales, poses, and facial expressions. However, color-based approaches are quite difficult to detect skin-tone color in the presence of skin color-like backgrounds. Therefore, we propose a novel method to robustly segment face clusters against skin color-like backgrounds.

In general, face detection based on skin-tone colors is significantly affected by illumination variations and complex backgrounds. To overcome such difficulties, Huang et al. [3] proposed an adaptive switching method to choose an optimal one from nine combinations of skin color models by a well-defined quality measure. The separated skin-tone clusters were further validated by the AdaBoost method with Haar-like features to determine whether human faces exist in those clusters or not. They used this method to track a sequence of images with the scenarios of dim-light, side-light, and back-light settings. The performance of their method achieves an average tracking time of about 145.4 ms/frame and a detection rate of 94.4%.

At the same time, Hu et al. [4] proposed a three-stage scheme to detect faces of pedestrians for intersection monitoring. Their method was mainly based on skin color as well as face features. Several testing results show the superiority of their proposed method especially in skin color-like backgrounds or varying illumination environments. Real-time performance is well achieved by their system since detection time is only 15.9 ms for each frame. However, many parameters used in their model are needed to be tuned for robustly detecting possible face candidates. Moreover, their method cannot operate well when the clusters of human skin colors are overlapped with those in skin color-like backgrounds.

In this paper, a robust method is proposed to solve the difficulties of segmenting face clusters from skin color-like backgrounds, especially in which face clusters are connected with those in skin color-like backgrounds. The remainder of this paper is organized as follows. In Section 2, the proposed method including three subsystems of image preprocessing, skin color segmentation, and face detection is described. Section 3 presents the result of face detection in skin color-like backgrounds. Conclusion and future work are given in Section 4.

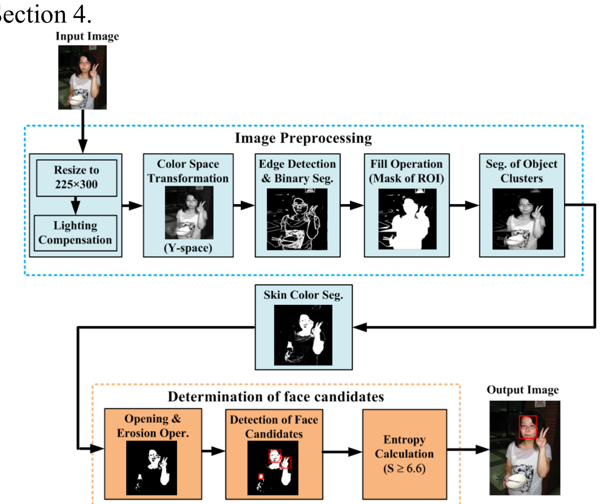

Figure 1. Flow diagram of the proposed method for face detection.

978-0-7695-4281-2/10 $26.00 © 2010 IEEE
DOI 10.1109/ICGEC.2010.175

II. THE PROPOSED METHOD OF FACE DETECTION

Figure 1 shows the flow diagram of the proposed method for face detection including three subsystems of image preprocessing, skin color segmentation, and determination of face candidates. The details of the three subsystems for face detection are described as follows.

A. Image preprocessing

The flow diagram of image preprocessing is shown in Fig. 1. To speed up the subsequent image processing, input images were first resized to 300×225 pixels by bilinear interpolation with a fixed aspect ratio. To keep face silhouettes in images not to be changed in the proposed method is quite critical. In this work, we resized the larger side of input images to be fixedly 300 pixels; therefore, the length of the other side could be obtained according to the aspect ratio of input images. Moreover, reference white was then performed because it is a popular method for lighting compensation in which the concept was first presented by Hsu et al. [2]. In their method, the top 5% of the luma values in images is regarded as the so-called reference white if the number of these pixels is sufficiently large (>192 pixels used in our work).

In contrast to RGB, YCbCr color space is luma-independent, leading to a better performance in varying lighting scenarios; therefore, YCbCr color space transformation [5] is carried out, and the luma, i.e., Y-space, of this color space is used to obtain a gray image. The formula of YCbCr color space transformation is expressed as follows:

$$\begin{cases} Y = 0.299R + 0.586G + 0.114B \\ Cb = -0.169R - 0.331G + 0.499B \\ Cr = 0.499R - 0.418G - 0.0813B \end{cases} \tag{1}$$

To detect object clusters in foreground as compact as possible by the proposed fill operation method described later, Sobel edge detection and following binary segmentation by TSMO method [6] were performed such that the edge pixels detected become more salient. Then, the object clusters enclosed by the edge points were filled with white pixels using the fill operation to create a mask of region of interest (ROI) that is operated with the preceding gray image by a logic AND operation to completely segment object clusters from backgrounds. The proposed fill operation method is described as follows.

In this work, we used a 3×3 mask to perform the fill operation. If the center point in the mask is a black pixel and there exist more than 4 white points (included) in the mask except the center point, the gray level of the center point is then set to a white pixel. We used this method to scan a binary image from top to bottom and left to right to obtain a resulting image A. At the same time, we also scanned the same binary image from top to bottom but right to left to obtain a resulting image B. Then, an AND logic operation is carried out on both images A and B, the fill operation is completed to fill the regions with white pixels that are enclosed by edge points.

B. Skin color segmentation

The object clusters were successfully segmented from backgrounds using the method of image preprocessing. To eliminate the non-skin color object clusters, the constraints of skin-tone color proposed by Garcia and Tziritas [5] were adopted in our work as follows.

$$\begin{aligned} & if\,(Y > 128) \\ & \quad \theta_1 = -2 + \frac{256 - Y}{16},\ \theta_2 = 20 - \frac{256 - Y}{16},\ \theta_3 = 6,\ \theta_4 = 8 \\ & if\,(Y \le 128) \\ & \quad \theta_1 = 6,\ \theta_2 = 12,\ \theta_3 = 2 + \frac{Y}{32},\ \theta_4 = -16 + \frac{Y}{16} \end{aligned} \tag{2}$$

and

$$\begin{aligned} & Cr \ge -2(Cb + 24),\ Cr \ge -(Cb + 17), \\ & Cr \ge -4(Cb + 32),\ Cr \ge 2.5(Cb + \theta_1), \\ & Cr \ge \theta_3,\ Cr \ge 0.5(\theta_4 - Cb), \\ & Cr \le \frac{220 - Cb}{6},\ Cr \le \frac{4}{3}(\theta_2 - Cb) \end{aligned} \tag{3}$$

As shown in Eqs. (2) and (3), Y-space obtained in Eq. (1) was used to determine the parameters of θ_1 to θ_4 by which values with Cb-space are further to restrict the ranges of Cr-space of being skin-tone color pixels. By this method, we successfully eliminated non-skin color object clusters but skin color-like blobs in backgrounds still existed in images that will be further removed by the following step of determination of face candidate.

C. Determination of face candidates

Determination of face candidates aims to eliminate both (1) skin color-like blobs in backgrounds and (2) skin color clusters without faces such as arm, hand, or leg and so forth. Figure 2 shows the detailed flow diagram of determination of face candidates. The image after skin color segmentation was performed by the operations of morphological opening and erosion such that the space between skin color-like blob in backgrounds and skin color cluster can be separated more largely, indicating that the two connected regions can be completely split. This way is beneficial to the application of 8-connected component labeling method that is often used in computer vision to detect connected regions in binary digital images to form a more compact region of face candidate. Subsequently, morphological dilation is conducted to keep the silhouettes of skin color clusters more close to their original shapes, which are slightly destroyed by the preceding operations of morphological opening and erosion. In the proposed method, detection of face candidates (see Fig. 2) includes 8-connected component labeling and the operation of morphological dilation.

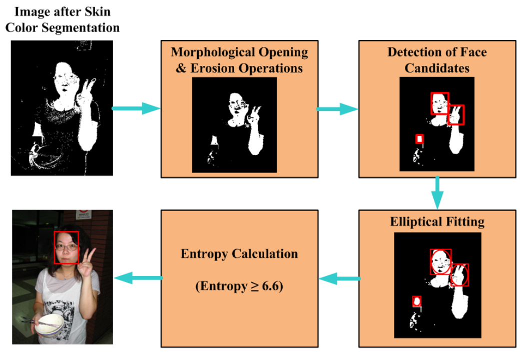

Figure 2. Detailed flow diagram of determination of face candidates.

The skin color clusters may include face parts with neck that is redundant in our case, or consists of other skin color regions without faces such as arm, hand, or leg that should be removed. As for the former, we proposed the following method to eliminate the neck part. As shown in Fig. 3(a), the skin color region drawn with a red rectangle box is determined using the 8-conneced component labeling method. The histogram of skin color pixels for the red rectangle box shown in Fig. 3(a) was then statistically completed (see Fig. 3(c)). Since the profile of human faces can be assumed to be approximately elliptical, the histogram appears to be a single mode of Gaussian distribution. Hence, the maximum value of histogram should exist in the centerline of faces, and the minimum values of histogram can be determined as the last valleys from both sides to the centerline whose values are less than $0.375V_{max}$ (see Fig. 3(c)). These three values are indicated by red dash lines shown in Fig. 3(c). Based on the two locations of x_1 and x_2 corresponding to the minimum values of V_1 and V_2, respectively, where $x_2>x_1$, the width of face region, say W_f ($=x_2-x_1$), can be determined. Since the aspect ratio of faces (=height/width) is assumed to be 4/3, the height of face region, say H_f, can be calculated as $4/3*W_f$. However, if the calculated H_f is larger than the height of original rectangle box, this rectangle box is not changed. By this method, the neck part of faces can be successfully eliminated, and the resulting face region is shown in Fig. 3(b).

As mentioned earlier, the skin color clusters that may not contain any faces should be eliminated. Since the aspect ratio of faces of 4/3 is assumed, it indicates that if the width of skin color clusters is larger than its height, the skin color clusters that may be an arm, a hand or a leg should be rejected. However, the clusters of containing face candidate, say Wr, can be fitted to an ellipse, say We, shown in Fig. 4. Furthermore, if $(Se+Sp+Sa)/3<\rho$, the skin color clusters are rejected, where Se, Sp, and Sa represent the sensitivity (true positive rate=TP/(TP+FN)), the specificity (true negative rate= TN/(TN+FP)), and the spatial accuracy (=1-(FP+FN)/(TP+FN)), respectively. These values are estimated from the elliptical regions We, where TP, TN, FP and FN represent the total number of true positive, true negative, false positive and false negative pixels, respectively. The value of ρ is a threshold set to 0.7.

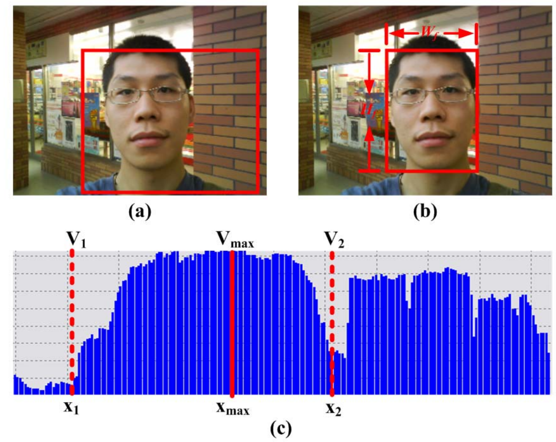

Figure 3. Results of determination of face candidates, (a) Original rectangle box of face candidates, (b) Modified rectangle box of face candidates, (c) Histogram of skin color pixels in rectangle box in (a).

The face candidates accepted by elliptical fitting may contain no faces, and these face candidates should be further removed. In general, face regions usually has higher entropy than those of other skin regions such as arms or hands. The entropy S is a histogram feature that is defined as $-\Sigma p(g)\ln[p(g)]$, $g=0,1,\ldots,255$, and $p(g)=N(g)/M$, where $N(g)$ is the number of pixels at gray level g, and M is the total number of pixels in the image. In this work, $S\geq 6.6$ is a criterion for accepting skin color clusters as face regions.

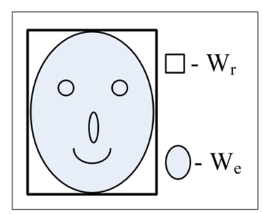

Figure 4. The face template used for elliptical fitting.

III. Experimental Results and Discussion

To extract object clusters in foregrounds and eliminate most of skin color-like pixels in backgrounds, the proposed image preprocessing method was carried out. As shown in Fig. 5, the results of skin color segmentation obtained by the proposed method (Figs. 5(b) and (e)) and Garcia et al. [5] (Figs. 5(c) and (f)) show that the proposed method has less fraction of skin color-like pixels in backgrounds due to the use of image preprocessing. Note that only the method of skin color segmentation was used in [5]. Furthermore, more results of face detection against skin color-like backgrounds are shown in Fig. 6. This result was also compared with those of Huang et al. [3] and Hu et al. [4]. As shown in this figure, many skin color-like clusters were detected by [3] since their method primarily focused on the issues of varying illumination. Moreover, when the overlapped area between faces and skin color-like clusters was larger, missing or false detection of faces (see column 3 in Fig. 6) appeared in [4]; however, when the overlapped area was relatively small, their method still operated well (see the test image with 5 people). The results of comparison show the superiority of the proposed method. Additionally, face images in relatively dark backgrounds (see Fig. 7) and with different poses as well as expressions (see Fig. 8) are also conducted by the proposed method. As shown in Figs. 7 and 8, higher performance of detection rate can be achieved.

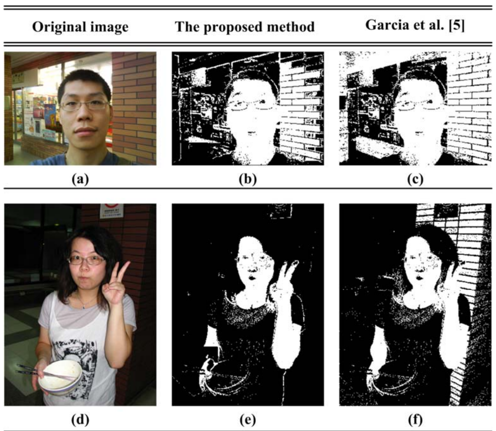

Figure 5. Results of skin color segmentation by the proposed method and Garcia et al. [5].

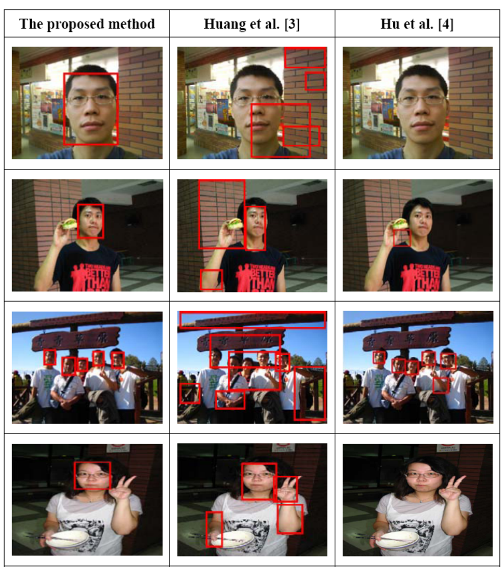

Figure 6. Results of face detection against skin color-like backgrounds by the proposed method, Huang et al. [4] and Hu et al. [4].

Figure 7. Testing samples in relatively dark illumination.

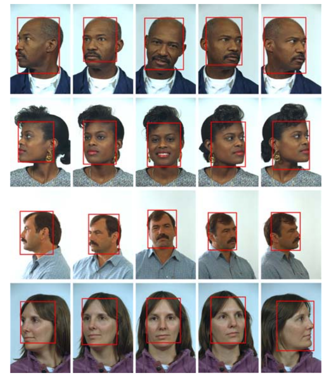

Figure 8. Testing samples in FERET face database [7] with different poses and expressions.

IV. CONCLUSIONS

In this paper, the face detection method against skin color-like backgrounds has been successfully presented. The proposed method also works well in relatively dark backgrounds. Moreover, higher performance of detection rate can be achieved even though faces are with different poses or expressions. The comparative results show the superiority of the proposed method.

ACKNOWLEDGEMENTS

This research was fully supported by a grant from National Science Council, Taiwan, under contract NSC99-2221-E-212-022.

REFERENCES

[1] P. Viola, and M. Jones, "Rapid object detection using a Boosted cascade of simple features," In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR2001), Hawaii, USA, 2001, pp. 511-518.

[2] R.L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, "Face detection in color image," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 5, 2002, pp. 696-706.

[3] D.Y. Huang, W.C. Hu, and M.H. Hsu, "Adaptive skin color model switching for face tracking under varying illumination," In: Proc. IEEE 4th International Conference on Innovative Computing, Information and Control (ICICIC2009), Kaohsiung, Taiwan, 2009, pp. 326-329.

[4] W.C. Hu, C.Y. Yang, D.Y. Huang, and C.H. Huang, "Real-time and reliable face detection in intersection monitoring by integration of skin color and facial features," In: Proc. IEEE 4th International Conference on Innovative Computing, Information and Control (ICICIC2009), Kaohsiung, Taiwan, 2009, pp. 1160-1163.

[5] C. Gracia, and G. Tziritas, "Face detection using quantized skin color regions merging and wavelet packet analysis," IEEE Transactions on Multimedia, Vol. 1, No. 3, 1999, pp. 264-277.

[6] D.Y. Huang, and C.H. Wang, "Optimal multi-level thresholding using a two-stage Otsu optimization approach," Pattern Recognition Letters, Vol. 30, No.3, 2009, pp. 275-284.

[7] P.J. Phillips, H. Moon, P.J. Rauss, and S. Rizvi, "The FERET evaluation methodology for face recognition algorithms," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, 2000, pp. 1090-1104.