Face Detection Based on Skin Color Segmentation and SVM Classification

Hwei-Jen Lin, Shwu-Huey Yen, Jih-Pin Yeh, and Meng-Ju Lin
Department of Computer Science and Information Engineering, Tamkang University
hjlin@cs.tku.edu.tw

Abstract

This paper proposes an improved version of our previously introduced face detection system based on skin color segmentation and neural networks. The new system uses a support vector machine (SVM) based method for verification.

1. Introduction

Recently, face detection has received much attention and has been an extensive research topic [1~4]. It is the important first step of many applications such as the face recognition system, facial expression analysis, surveillance systems. We previously proposed an approach for detecting human faces in color images under different illumination condition, scale, rotation, with/without glasses, based on classification by neural networks. In this paper, we modify the approach by using SVM for classification. Firstly, skin color segmentation is performed to find skin color regions. Secondly, possible face blocks are located by using some restrictions on these regions. Thirdly, eye detection and matching are carried out on each possible face blocks to obtain face candidates, which are then verified by an SVM classifier.

2. Face Candidate Searching

We obtain face candidates through three steps: skin color segmentation, locating possible face blocks using geometry property of faces, and eye detection for filtering out some non-face blocks

2.1. Skin color segmentation

The distribution of skin colors clusters in a small region of the chromatic color space and thus skin color detection is firstly performed to reduce the computational complexity. Among numerous color spaces, RGB color space is sensitive to the variation of intensity, and thus it alone is not sufficient to use only RGB color space to detect skin color.

Due to the fact that both the normalized RGB space and HSV color space can reduce the lighting effect, we combine them with the RGB color space to detect skin pixels. Since it suffices to represent color using only two values in the normalized RGB color space, we use the values of $r = R/(R+G+B)$ and $g = G/(R+G+B)$. A pixel is labeled as a skin pixel if its color values conform to the following constraints:

$$R > G, |R - G| \geq 11, 0.33 \leq r \leq 0.6, 0.25 \leq g \leq 0.37,$$
$$0.12 \leq S \leq 0.7, 0.3 \leq V \leq 1.0, \text{and} (340 \leq H \leq 359 \text{ or } 0 \leq H \leq 50)$$

As a result, we can generate a binary skin map where the white points represent the skin pixels and the black points represent the non-skin pixels. Then, we apply median filter and morphological opening operation to the skin map to eliminate small skin blocks. Afterward, utilizing connected component operation to find out all connected skin regions and each of the skin regions is labeled by a bounding box.

2.2. Locating possible face blocks

To check if a skin region contains a face, we use the following three constraints: (1). The area of the region is greater than 30×30, (2). The ratio of the height to the width of the bounding box, denoted by *Aspect Ratio,* is between 0.8 and 2.6, and (3). The percentage of the skin pixels in the bounding box is greater than 40%. Any skin region satisfying these constraints is regarded as a possible face block; otherwise, a non-face region.

2.3. Eye detection

In general, the intensity of eye is darker than that of other facial features in a face and it does not belong to skin region. Utilizing this property of eyes, we can find out some eye pixels and present them with white points in the possible face block. Median filter is applied to remove noises and then connected component operation is performed to find all so-called eye-like

978-0-7695-3266-0/08 $25.00 © 2008 IEEE
DOI 10.1109/SSIRI.2008.23

blocks circumscribed with a bounding box, and then examined by the three following conditions to verify if it contains an eye: (1). The first condition is that the *Aspect Ratio* of the block is between 0.2 and 1.67, (2). The percentage of the skin pixels in the box is greater or equal to 30%, (3). The ratio of the width of the eye-like block to the width of the possible face block is between 0.028 and 0.4. These parameters came from experimental results. Any eye-like block satisfying these conditions is called an eye block. Finally, each pair of eye blocks is matched by locating the centroids of the two eye blocks, calculating the horizontal distance, denoted by D, between two centroids, and then applying the following matching criteria: (1). The ratio of D to the width of the face is between T_1 and T_2, (2). The two eyes are located at the upper portion of a face, and (3). The sizes of the two eyes are nearly equal. In our experiments, we set $T_1 = 0.2$, $T_2 = 0.65$. For a matched eye pair, we clip a face candidate based on the face model as shown in Figure 1.

3. Face candidate verification using SVM

For a clipped face candidate, we utilize an SVM based classifier to finalize verification. The SVM used the radial basis function (RBF) kernel $K(x_i, x_j) = \exp(-\gamma (x_i - x_j)^2)$, where we set $\gamma = 20$ for our experiments. There are 300 face and 300 non-face grayscale training samples. The size of each training sample is 20×20. They are manually clipped from photographs, digital cameras, or internet. Before training, the gray value I_i of each pixel in the training samples is normalized to I_i', as shown in Equation (1). The testing samples are also performed the above normalization before verified by the SVM.

$$\bar{I} = \frac{1}{N}\sum_{i=1}^{N} I_i, \qquad I_i' = (I_i - \bar{I}) + 128 \qquad (1)$$

4. Experimental results

We tested our system on 1817 color images, taken from digital cameras, scanners, the World Wide Web, and the Champion dataset. These testing images include both indoor and outdoor scenes under different lighting conditions and backgrounds, and have sizes varying from 105×158 to 640×577. 2615 face candidates of sizes varying from 35×30 to 370×275 are clipped from these images. Our previous and present methods have detection rates 88.26% and 96.37%, and false acceptance rates 5.81%, and 3.63%, respectively. Figure 2 shows some detection results of our system.

5. Discussions and conclusions

This paper proposes a human face detection system based on skin color segmentation and SVM classification. Experimental results show that the proposed system results in better performance than our previously proposed method [4], which has been proved to outperform the method proposed by Fröba and Küblbeck [1], in terms of correct detection rate and capacity of coping with the problems of lighting, scaling, rotation, and multiple faces.

References

[1] Bernhard Fröba and Christian Küblbeck, "Face Detection and Tracking using Edge Orientation Information," *SPIE Visual Communications and Image Processing*, 2001, pp. 583-594.

[2] Ing-Sheen Hsieh, Kuo-Chin Fan, and Chiunhsiun Lin, "A Statistic Approach to the Detection of Human Faces in Color Nature Scene," *Pattern Recognition*, Vol. 35 , 2002, pp. 1583-1596.

[3] Y. Suzuki and T. Shibata, "Multiple-clue Face Detection Algorithm using Edge-based Feature Vectors," *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004 (ICASSP '04)*, Vol. 5, 17-21 May 2004, pp. V - 737-740.

[4] Hwei-Jen Lin, Shu-Yi Wang, Shwu-Huey Yen, and Yang-Ta Kao, "Face Detection Based on Skin Color Segmentation and Neural Network," in *Proceedings of the International Conference on Neural Networks and Brain (ICNN&B'05)*, Beijing, China, Oct.13-15, 2005, pp. 1144-1149.

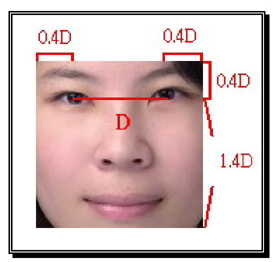

Figure 1 Face model

(a) (b)

Figure 2 Some face detection results