FACE DETECTION IN COLOR IMAGES

Rein-Lien Hsu[§], *Mohamed Abdel-Mottaleb**, *and Anil K. Jain* [§]

[§] Dept. of Computer Science & Engineering, Michigan State University, MI 48824
* Philips Research, 345 Scarborough Rd., Briarcliff Manor, NY 10510
* Electrical and Computer Engineering Dept., University of Miami, FL 33124
Email: {hsureinl, jain}@cse.msu.edu, mohamed.abdel-mottaleb@philips.com, mottaleb@miami.edu

ABSTRACT

Human face detection is often the first step in applications such as video surveillance, human computer interface, face recognition, and image database management. We propose a face detection algorithm for color images in the presence of varying lighting conditions as well as complex backgrounds. Our method detects skin regions over the entire image, and then generates face candidates based on the spatial arrangement of these skin patches. The algorithm constructs eye, mouth, and boundary maps for verifying each face candidate. Experimental results demonstrate successful detection over a wide variety of facial variations in color, position, scale, rotation, pose, and expression from several photo collections.

1. INTRODUCTION

Various approaches to face detection are discussed in [10]. These approaches utilize techniques such as neural networks, machine learning, (deformable) template matching, Hough transform, motion extraction, and color analysis. The neural network-based [11] and view-based [14] approaches require a large number of face and non-face training examples, and are designed to find frontal faces in grayscale images. A recent statistical approach [12] extends the detection of frontal faces to profile views by training two separate classifiers. Model-based approaches are widely used in tracking faces and often assume that the initial locations of faces are known. Skin color provides an important cue for face detection. However, the color-based approaches face difficulties in robust detection of skin colors in the presence of complex background and variations in lighting conditions. We propose a face detection algorithm which is able to handle a wide variety of variations in color images.

2. FACE DETECTION ALGORITHM

The use of color information can simplify face localization in complex environments [3,10]. An overview of our face detection algorithm is depicted in Fig. 1, which contains two major modules: (i) *face localization* for finding face candidates; and (ii) *facial feature detection* for verifying detected face candidates. Major modules of the algorithm are briefly described below.

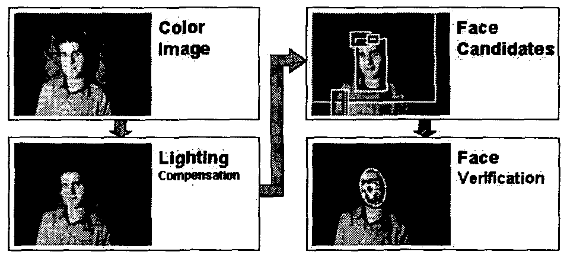

Figure 1: Face detection algorithm.

2.1. Lighting compensation and skin tone detection

The appearance of the skin-tone color can change due to different lighting conditions. We introduce a lighting compensation technique that estimates "reference white" to normalize the color appearance. We regard pixels with top *5* percent of the luma (nonlinear gamma-corrected luminance) values as the reference white if the number of these reference-white pixels is larger than 100. The red, green, and blue components of a color image are adjusted so that these reference-white pixels are scaled to the gray level of 255.

Modeling skin color requires choosing an appropriate color space and identifying a cluster associated with skin color in this space. Based on Terrillon et al.'s [15] comparison of the nine color spaces for face detection, we use the YCbCr space since it is widely used in video compression standards. Since the skin-tone color depends on luminance, we nonlinearly transform the YCbCr color space to make the skin cluster luma-independent. This also enables robust detection of dark and light skin tone colors. A parametric ellipse in the nonlinearly

0-7803-6725-1/01/$10.00 ©2001 IEEE

transformed Cb-Cr color subspace is used as a model of skin color. Figure 2 shows an example of skin detection.

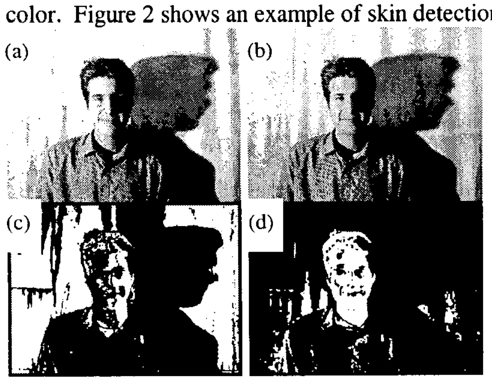

Figure 2: Skin detection: (a) a yellow-biased face image; (b) a lighting compensated image; (c) skin regions of (a) shown as pseudo-color; (d) skin regions of (b).

2.2. Localization of facial features

Among the various facial features, eyes and mouth are the most suitable features for recognition and estimation of 3D head pose [5]. Most approaches to eye and face localization [7, 13] are template based. However, our approach is able to directly locate eyes, mouth, and face boundary based on *measurements* derived from the color-space components of an image.

Eyes usually contain both dark and bright pixels in the luma component. Grayscale morphological operators (e.g., dilation and erosion) [8] can be designed to emphasize brighter and darker pixels in the luma component around eye regions. These operations have been used to construct feature vectors for a complete face at *multiple scales* for frontal face authentication [9]. In our eye detection algorithm, the grayscale dilation and erosion using a hemispheric structuring element at *an* estimated scale are applied independently to construct EyeMap in the luma. In addition, an analysis of the chrominance components indicated that high *Cb* and low *Cr* values are found around the eyes. The EyeMap in the chroma is constructed from *Cb*, the inverse of *Cr*, and the ratio *Cb/Cr*. The two resulting eye maps are combined by a multiplication operation. The resultant eye map brightens both the eyes and suppresses other facial areas, as can be seen in Fig. 3. Eye candidates are selected by using (i) a pyramid decomposition of the enhanced eye maps for coarse localizations and (ii) binary morphological closing and an iterative thresholding for fine localizations.

The mouth region contains more red component compared to the blue component than other facial regions. Hence, the chrominance component C_r is greater than C_b near the mouth areas. We further notice that the mouth has a relatively lower response in the C_r/C_b feature but a high response in C_r^2. Therefore, the difference between C_r^2 and C_r/C_b can emphasize the mouth regions. Figure 4 shows the mouth maps of the subjects in Fig. 3.

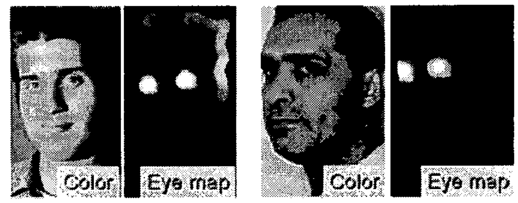

Figure 3: Construction of the eye maps for two subjects.

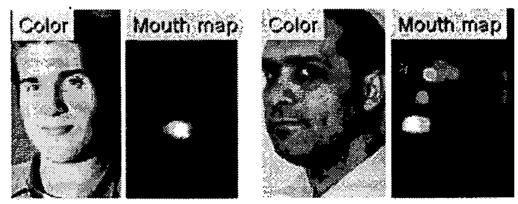

Figure 4: Construction of the mouth maps.

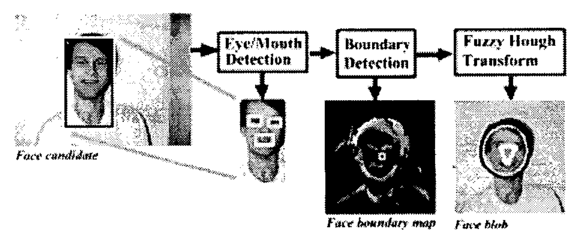

Figure 5: Face boundary and the eyes-mouth triangle.

The eyes and mouth candidates are verified by checking (i) luma variations of eye and mouth blobs; (ii) geometry and orientation constraints of eyes-mouth triangles; and (iii) the presence of a face boundary around eyes-mouth triangles. Based on the locations of eyes/mouth candidates, our algorithm first constructs a *face boundary* map from the luma, and then utilizes a Hough transform to extract the best-fitting ellipse. The fitted ellipse is associated with a quality measurement for computing eyes-and-mouth triangle weights. Figure 5 shows the boundary map which is constructed from both the magnitude and the orientation components of the luma gradient within the regions having positive orientations of the gradient orientations. The Hough transform is useful for the detection of parametric shapes; its efficiency depends on the dimensionality of the accumulator. An ellipse in a plane has five parameters: an orientation angle, two coordinates of the center, and lengths of major and minor axes. Since we know the locations of eyes and mouth, the orientation of the ellipse can be estimated from the direction of a vector that starts from midpoint between eyes to the mouth. The location of the ellipse center is estimated from the face boundary. Hence, we only require a two-dimensional accumulator for an ellipse in a plane. The ellipse with the highest vote is selected.

Each eye-mouth triangle candidate is associated with a weight that is computed from its eyes/mouth maps, ellipse vote and face orientation that favors upright faces and symmetric facial geometry (see [6] for details).

3. EXPERIMENTAL RESULTS

Personal photo collections usually contain color images that are taken under varying lighting conditions as well as with complex backgrounds. Further, these images may have quality variations and contain multiple faces with variations in color, position, scale, rotation, pose, and facial expression. We present detection results in Tables 1 and 2 on the HHI [1] and the Champion [2] databases, respectively. Figure 6 shows that our algorithm can detect multiple faces of different sizes with a wide variety of facial variations. Further, the algorithm can detect both dark skin-tone and bright skin-tone because of the nonlinear transform of the *Cb-Cr* color space.

Table 1: Detection results on the HHI image database.
FP: False Positives, DR: Detection Rate.

Head Pose	Frontal	Near-Frontal	Half-Profile	Profile	All
No. of images	66	54	75	11	206
Image size	640 x 480 (pixel)				
Stage 1: Grouped Skin-region					
No. of FP	3145	2203	3781	277	9406
DR (%)	95.45	98.15	96.00	100	96.60
Time (sec)	2.99 (average) ± 0.87 (s. d.) on a 860MHz CPU				
Stage 2: Facial Feature Location					
No. of FP	4	6	14	3	27
DR (%)	89.40	90.74	74.67	18.18	80.58
Time (sec)	39.93 (average) ± 28.93 (s. d.)				

Table 2: Detection results on the Champion image database.

Head Pose	Frontal, Near-frontal, Half profile		
No. of images	227	Size (pixel)	150 x 220
Stage 1: Grouped Skin-region			
No. of FP	5582	DR (%)	99.12
Time (sec)	0.15 (average) ± 0.08 (s. d.) on a 860MHz CPU		
Stage 2: Facial Feature Location			
No. of FP	14	DR (%)	91.63
Time (sec)	9.29 (average) ± 7.77 (s. d.)		

4. CONCLUSIONS AND FUTURE WORK

We have presented a face detection algorithm for color images using a skin-tone color model and facial features. Our method first corrects the color bias by a novel lighting compensation technique that automatically estimates the reference white pixels. We overcome the difficulty of detecting the low-luma and high-luma skin tones by applying a nonlinear transform to the YCbCr color space. Our method detects skin regions over the entire image, and then generates face candidates based on the spatial arrangement of these skin patches. The algorithm constructs eye/mouth/boundary maps for verifying each face candidate. Detection results for several photo collections have been presented. Our goal is to design a system that detects faces and facial features, allows users to edit detected faces, and uses the facial features as indices for retrieval from image and video databases.

5. REFERENCES

[1] MPEG7 content set from Heinrich Hertz Institute, <http://www.darmstadt.gmd.de/mobile/MPEG7/Documents/N2466.html>.

[2] Champion database, <http://www.libfind.unl.edu/alumni/champions>.

[3] M. Abdel-Mottaleb and A. Elgammal, "Face detection in complex environments from color images," *IEEE ICIP*, pp. 622-626, Oct. 1999.

[4] J. Canny, "A computational approach to edge detection," *IEEE Trans. PAMI*, vol. 8, pp. 679-698, Nov. 1986.

[5] T. Horprasert, Y. Yacoob, and L.S. Davis, "Computing 3-D head orientation from a monocular image," *Proc. Int'l Conf. Automatic Face and Gesture Recognition*, pp. 242-247, Oct. 1996.

[6] R.-L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, "Face detection in color images," *Tech. Report MSU-CSE-01-7*, Michigan State Univ., March 2001.

[7] W. Huang, Q. Sun, C.-P. Lam, and J.-K. Wu, "A robust approach to face and eyes detection from images with cluttered background," *ICPR*, vol. 1, pp. 110-114, Aug. 1998.

[8] P.T. Jackway and M. Deriche, "Scale-space properties of the multiscale morphological dilation-erosion," *IEEE Trans. PAMI*, vol. 18, pp. 38-51, Jan. 1996.

[9] C. Kotropoulos, A. Tefas, and I. Pitas, "Frontal face authentication using morphological elastic graph matching," *IEEE Trans. Image Processing*, vol. 9, pp. 555-560, April 2000.

[10] D. Maio and D. Maltoni, "Real-time face location on gray-scale static images," *Pattern Recognition*, vol.33, no. 9, pp. 1525-1539, Sept. 2000.

[11] H.A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE Trans. PAMI*, vol. 20, pp. 23-38, Jan. 1998.

[12] H. Schneiderman and T. Kanade, "A statistical method for 3D object detection applied to faces and cars," *IEEE CVPR*, June 2000.

[13] F. Smeraldi, O. Carmona, and J. Bigün, "Saccadic search with gabor features applied to eye detection and real-time head

tracking," *Image and Vision Computing*, vol. 18, no. 4, pp. 323-329, 2000.

[14] K.K. Sung and T. Poggio, "Example-based learning for view-based human face detection," *IEEE Trans. PAMI*, vol. 20, pp. 39-51, Jan. 1998.

[15] J.C. Terrillon, M.N. Shirazi, H. Fukamachi, and S. Akamatsu, "Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition*, pp. 54-61, 2000.

Figure 6: Face detection results on the HHI and the Champion databases, and a collection of family photos. The detected faces, represented as ellipses, are overlaid on the color-compensated images. There are a few false positives and negatives in the family group photographs.