Face Detection in Color Images

Rein-Lien Hsu, *Student Member*, *IEEE*,
Mohamed Abdel-Mottaleb, *Member*, *IEEE*, and
Anil K. Jain, *Fellow*, *IEEE*

Abstract—Human face detection plays an important role in applications such as video surveillance, human computer interface, face recognition, and face image database management. We propose a face detection algorithm for color images in the presence of varying lighting conditions as well as complex backgrounds. Based on a novel lighting compensation technique and a nonlinear color transformation, 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 face detection over a wide range of facial variations in color, position, scale, orientation, 3D pose, and expression in images from several photo collections (both indoors and outdoors).

Index Terms—Face detection, face recognition, lighting compensation, color transformation, skin tone, facial feature map, Hough transform.

1 Introduction

HUMAN activity is a major concern in a wide variety of applications such as video surveillance, human computer interface, face recognition [7], [42], [47], and face image database management [36]. Detecting faces is a crucial step in these identification applications. Most face recognition algorithms assume that the face location is known. Similarly, face tracking algorithms (e.g., [10]) often assume the initial face location is known. Note that face detection can be viewed as a two-class (face versus nonface) classification problem. Therefore, some techniques developed for face recognition (e.g., holistic/template approaches [12], [30], [41], [27], feature-based approaches [6], and their combination [20]) have also been used to detect faces, but they are computationally very demanding and cannot handle large variations in face images.

Various approaches to face detection are discussed in [11], [24], [26], [45], [14]. The major approaches are listed chronologically in Table 1. For recent surveys on face detection, see [45] and [14]. These approaches utilize techniques such as principal component analysis, neural networks, machine learning, information theory, geometrical modeling, (deformable) template matching, Hough transform, motion extraction, and color analysis. The neural network-based [32], [33] and view-based [39] approaches require a large number of face and nonface training examples and are designed primarily to locate frontal faces in gray-scale images. Schneiderman and Kanade [35] extend their learning-based approach for the detection of frontal faces to profile views. A feature-based approach that uses geometrical facial features with belief networks [46] provides face detection for nonfrontal views. Geometrical facial templates and the Hough transform were incorporated to detect gray-scale frontal faces in real time applications [24]. Face detectors based on Markov random fields [23] and Markov chains [8], make use of the spatial arrangement of pixel gray values.

Categorizing face detection methods based on the representation used reveals that detection algorithms using holistic representations have the advantage of finding small faces or faces in poor-quality images, while those using geometrical facial features provide a good solution for detecting faces in different poses. A combination of holistic and feature-based approaches [13], [21] is a promising approach to face detection as well as face recognition. Motion [9], [11] and skin-tone color [11], [4], [38], [43], [12] are useful cues for face detection. However, the color-based approaches face difficulties in robustly detecting skin colors in the presence of complex background and different lighting conditions [12]. We propose a face detection algorithm that is able to handle a wide range of variations in static color images, based on a lighting compensation technique and a nonlinear color transformation. Our approach models skin color using a parametric ellipse in a two-dimensional transformed color space and extracts facial features by constructing feature maps for the eyes, mouth, and face boundary. Section 2 describes the face detection algorithm. Section 3 presents the detection results of our algorithm on several face databases. Conclusions and future work are described in Section 4.

2 Face Detection Algorithm

An overview of our face detection algorithm is depicted in Fig. 1, which contains two major modules: 1) face localization for finding face candidates and 2) facial feature detection for verifying detected face candidates. The algorithm first estimates and corrects the color bias based on a lighting compensation technique. The corrected red, green, and blue color components are then nonlinearly transformed in the YC_bC_r color space. The skin-tone pixels are detected using an elliptical skin model in the transformed space. The parametric ellipse corresponds to contours of constant Mahalanobis distance under the assumption of Gaussian distribution of skin tone color. The detected skin-tone pixels are iteratively segmented using local color variance into connected components which are then grouped into face candidates based on both the spatial arrangement of these components and the similarity of their color [4]. The size of a face candidate can range from 13×13 pixels to about three fourths of the input image size. The facial feature detection module rejects face candidate regions that do not contain any facial features such as eyes, mouth, and face boundary. A detected face enclosed by an ellipse with the associated eyes-mouth triangle is shown in Fig. 1.

2.1 Lighting Compensation and Skin Tone Detection

The appearance of the skin-tone color depends on the lighting conditions. We introduce a lighting compensation technique that uses "reference white" to normalize the color appearance. We regard pixels with the top 5 percent of the luma (nonlinear gamma-corrected luminance) values in the image as the reference white only if the number of these pixels is sufficiently large (> 100). The R, G, and B components of a color image are adjusted so that the average gray value of these reference-white pixels is linearly scaled

- *R.-L. Hsu and A.K. Jain are with the Department of Computer Science & Engineering, 3115 Engineering Bldg., Michigan State University, East Lansing, MI 48824. E-mail: {hsureinl, jain}@cse.msu.edu.*
- *M. Abdel-Mottaleb is with the Electrical and Computer Engineering Department, University of Miami, McAuthur Engineering Bldg., Room 518, Coral Gables, FL 33124. E-mail: mottaleb@miami.edu.*

Manuscript received 12 Mar. 2001; revised 4 Dec. 2001; accepted 22 Jan. 2002.
Recommended for acceptance by P. Belhumeur.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 113783.

TABLE 1
Major Face Detection Approaches

Authors	Year	Approach	Featurees Used	Head Pose	Test Databases	Minimal Face Size
Féraud et al. [11]	2001	Neural Networks	Motion; Color; Texture	Frontal and profile	Sussex; CMU; Web images	15×20
Maio et al. [24]	2000	Facial templates; Hough Transform	Texture; Directional images	Frontal	Static images	20×27
Garcia et al. [12]	1999	Statistical wavelet analysis	Color; wavelet coefficients	Frontal to near frontal	MPEG videos	80×48
Wu et al. [43]	1999	Fuzzy color models; Template matching	Color	Frontal to profile	Still color images	20×24
Rowley et al. [32], [33]	1998	Neural Networks	Texture	(Upright) frontal	CMU; FERET; Web images	20×20
Sung et al. [39]	1998	Learning	Texture	Frontal	Mug shots; CCD pictures; newspaper scans	19×19
Yang et al. [44]	1998	Multiscale segmentation; color model	Skin Color; intensity	Frontal	color pictures	NA
Colmenarez et al. [8]	1997	Learning	Markov processes	Frontal	FERET	11×11
Yow et al. [46]	1997	Feature; Belief networks	Geometrical facial features	Frontal to profile	CMU	60×60
Lew et al. [23]	1996	Markov Random Field; DFFS [27]	Most informative pixel	Frontal	MIT; CMU; Leiden	23×32

to 255. The image is not changed if a sufficient number of reference-white pixels is not detected or the average color is similar to skin tone. This assumption is reasonable not only because an image usually contains "real white" [31] pixels in some regions of interest (such as eye regions), but also because the dominant bias color always appears as "real white." Fig. 2 demonstrates an example of our lighting compensation method. Note that the yellow bias color in Fig. 2a has been removed, as shown in Fig. 2b. With lighting compensation, our algorithm detects fewer nonface pixels and more skin-tone facial pixels (see Figs. 2c and 2d).

Modeling skin color requires choosing an appropriate color space and identifying a cluster associated with skin color in this space. It has been observed that the normalized red-green (rg) space [5] is not the best choice for face detection [40], [34]. Based on Terrillon et al.'s [40] comparison of nine different color spaces for face detection, the tint-saturation-luma (TSL) space provides the best results for two kinds of Gaussian density models (unimodal and a mixture of Gaussians). We adopt the YC_bC_r space since it is perceptually uniform [31], is widely used in video compression standards (e.g., MPEG and JPEG) [12], and it is similar to the

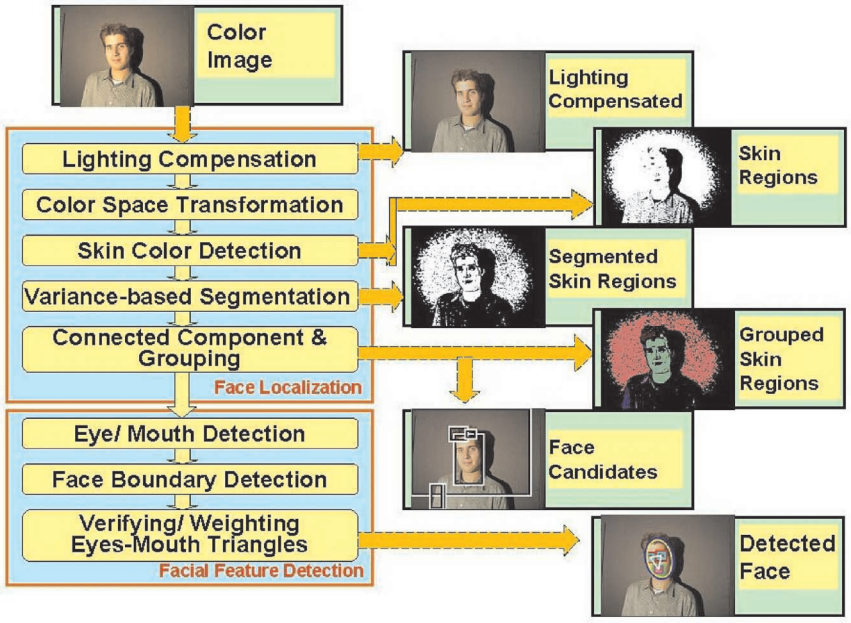

Fig. 1. Face detection algorithm.

Fig. 2. Skin detection: (a) a yellow-biased face image, (b) a lighting compensated image, (c) skin regions of (a) shown in white, and (d) skin regions of (b).

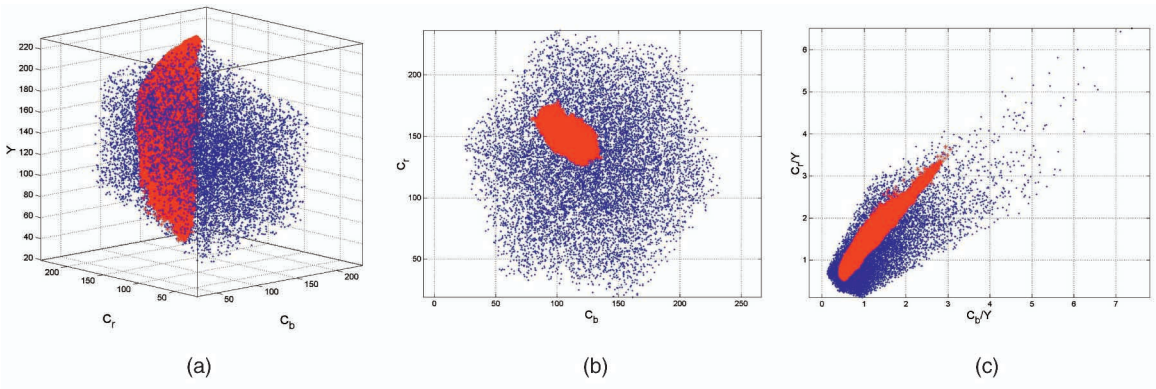

Fig. 3. The YC_bC_r color space (blue dots represent the reproducible color on a monitor) and the skin tone model (red dots represent skin color samples): (a) the YC_bC_r space, (b) a 2D projection in the C_bC_r subspace, and (c) a 2D projection in the (C_b/Y)-(C_r/Y) subspace.

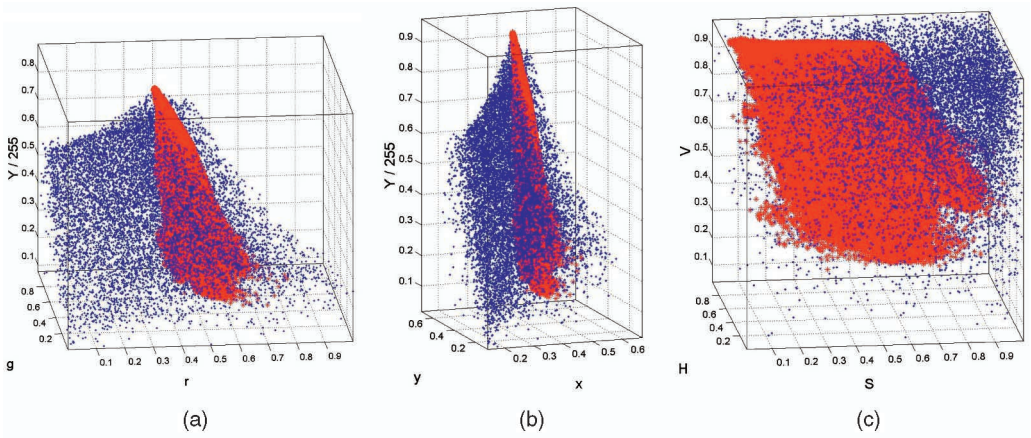

Fig. 4. The dependency of skin tone color: the skin tone cluster (red dots) is shown in (a) the rgY, (b) the $CIE\ xyY$, and (c) the HSV color spaces, where blue dots represent the reproducible color on a monitor. For a better presentation of cluster shape, we normalize the luma Y in the rgY and the $CIE\ xyY$ by 255, and swap the hue and saturation coordinates in the HSV space. The skin tone cluster is less compact at low luma values in (a) and (b) and low saturation values in (c).

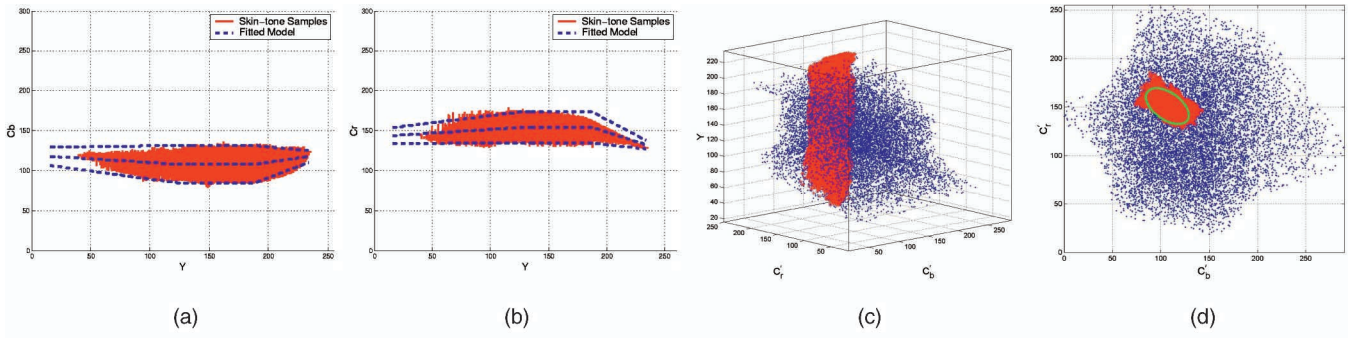

Fig. 5. Nonlinear transformation of the YC_bC_r color space: the skin tone cluster shown in (a) the YC_b subspace, (b) the YC_r subspace (c) the transformed YC_bC_r color space, (d) a 2D projection of (c) in the transformed C_bC_r subspace, in which the elliptical skin model is overlaid on the skin cluster. Red dots indicate the skin cluster. Three blue dashed curves in (a) and (b), one for cluster center and two for boundaries, indicate the fitted models.

TSL space in terms of the separation of luminance and chrominance as well as the compactness of the skin cluster. Many research studies assume that the chrominance components of the skin-tone color are independent of the luminance component [19], [25], [34], [38]. However, in practice, the skin-tone color is nonlinearly dependent on luminance. We demonstrate the luma

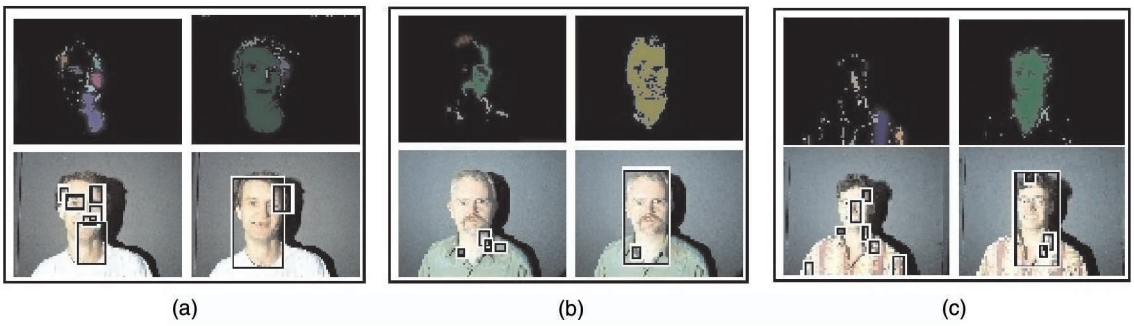

Fig. 6. Detection examples, with and without the transformation. For each of these examples, the images shown in (a) are skin regions and detections without the transformation, while those in (b) are results with the transformation.

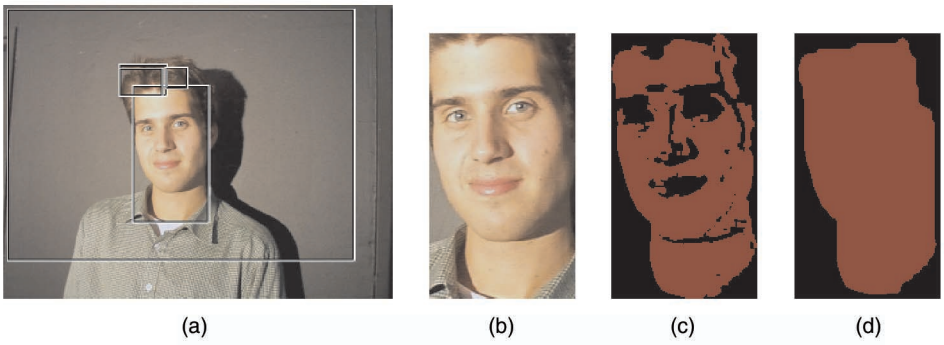

Fig. 7. Face mask: (a) four face candidates, (b) one of the face candidates, (c) grouped skin area, and (d) face mask.

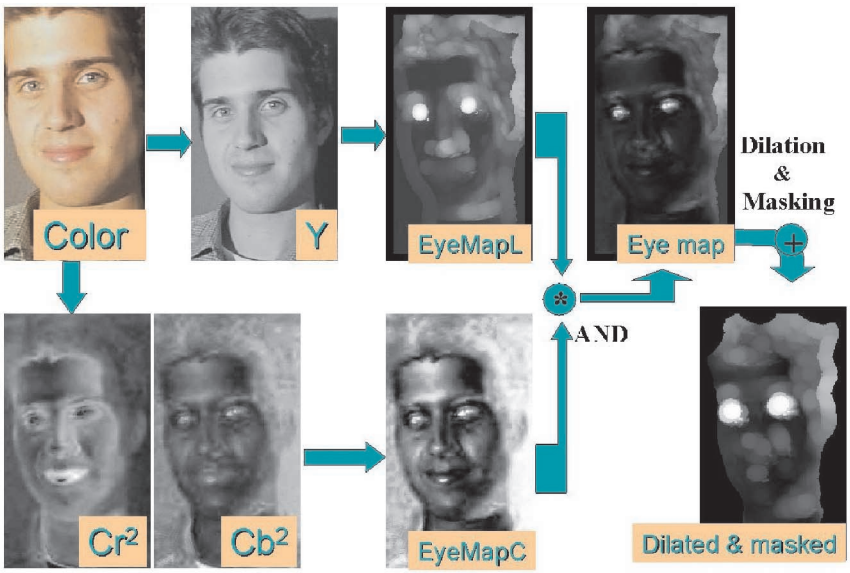

Fig. 8. Construction of eye maps.

dependency of skin-tone color in different color spaces in Figs. 3 and 4, based on skin patches ($853,571$ pixels) collected from nine subjects (137 images) in the Heinrich-Hertz-Institute (HHI) image database [1]. Detecting skin tone based on the cluster of training samples in the C_bC_r subspace, shown in Fig. 3b, results in many false positives. Face detection based on the cluster in the (C_b/Y)-(C_r/Y) subspace, shown in Fig. 3c, results in many false negatives. Therefore, we nonlinearly transform the YC_bC_r color space to make the skin cluster luma-independent. This is done by fitting piecewise linear boundaries to the skin cluster (see Figs. 5a and 5b). The details of the model and the transformation are described in Appendix A. The transformed space, shown in Figs. 5c and 5d, enables a robust detection of dark and light skin tone colors. Fig. 6 shows that more skin-tone pixels with low and high luma are detected in the transformed subspace than in the C_bC_r subspace.

2.2 Localization of Facial Features

Among the various facial features, eyes and mouth are the most prominent features for recognition and estimation of 3D head pose [15], [29]. Most approaches for eye localization [17], [22], [37] are template-based. However, we directly locate eyes, mouth, and face boundary based on their feature maps derived from both the luma and chroma of an image. We consider only the area covered by a *face*

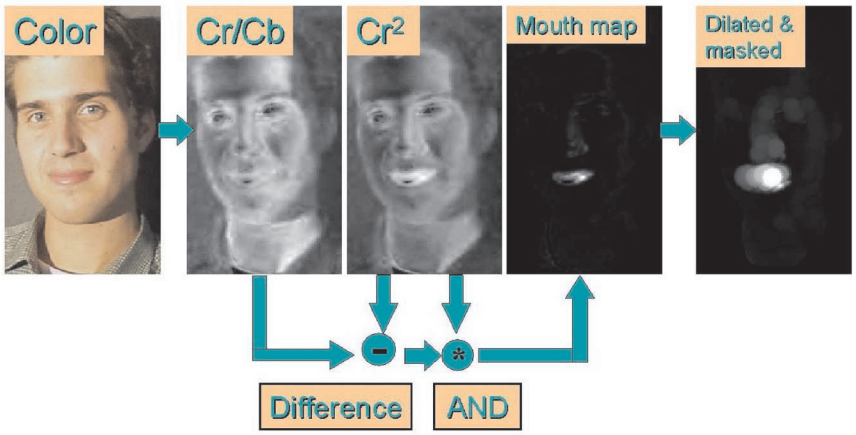

Fig. 9. Construction of the mouth map.

mask that is built by enclosing the grouped skin-tone regions with a pseudo convex hull. Fig. 7 shows an example of the face mask.

2.2.1 Eye Map

We first build two separate eye maps, one from the chrominance components and the other from the luminance component. These two maps are then combined into a single eye map. The eye map from the chroma is based on the observation that high C_b and low C_r values are found around the eyes. It is constructed by

$$EyeMapC = \frac{1}{3}\Big\{(C_b^2) + (\tilde{C}_r)^2 + (C_b/C_r)\Big\}, \quad (1)$$

where C_b^2, $(\tilde{C}_r)^2$, C_b/C_r all are normalized to the range $[0, 255]$ and $\tilde{C}r$ is the negative of C_r (i.e., $255 - C_r$). Since the eyes usually contain both dark and bright pixels in the luma component, gray-scale morphological operators (e.g., dilation and erosion) [18] 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 faces at multiple scales for frontal face authentication [20]. We use gray-scale dilation and erosion with a hemispheric structuring element to construct the eye map from the luma as follows:

$$EyeMapL = \frac{Y(x,y) \oplus g_\sigma(x,y)}{Y(x,y) \ominus g_\sigma(x,y) + 1}, \quad (2)$$

where the gray-scale dilation $\oplus$ and erosion $\ominus$ operations on a function $f : \mathcal{F} \subset R^2 \longrightarrow R$ using a structuring function $g : \mathcal{G} \subset R^2 \longrightarrow R$ are defined in [18]. The eye map from the chroma is enhanced by histogram equalization and then combined with the eye map from the luma by an AND (multiplication) operation, i.e., $EyeMap = (EyeMapC)\mathbf{AND}(EyeMapL)$. The resulting eye map is then dilated, masked, and normalized to brighten both the eyes and suppress other facial areas, as shown in Fig. 8. The locations of the eye candidates are initially estimated from the pyramid decomposition of the eye map and then refined using iterative thresholding and binary morphological closing on this eye map.

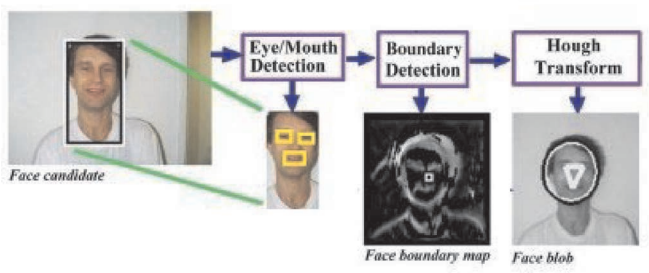

Fig. 10. Computation of face boundary.

2.2.2 Mouth Map

The color of mouth region contains stronger red component and weaker blue component than other facial regions. Hence, the chrominance component C_r is greater than C_b in the mouth region. We further notice that the mouth has a relatively low response in the C_r/C_b feature, but it has a high response in C_r^2. We construct the mouth map as follows:

$$MouthMap = C_r^2 \cdot \left(C_r^2 - \eta \cdot C_r/C_b\right)^2, \quad (3)$$

$$\eta = 0.95 \cdot \frac{\frac{1}{n}\sum_{(x,y)\in\mathcal{FG}} C_r(x,y)^2}{\frac{1}{n}\sum_{(x,y)\in\mathcal{FG}} C_r(x,y)/C_b(x,y)}, \quad (4)$$

where both C_r^2 and C_r/C_b are normalized to the range $[0, 255]$, and n is the number of pixels within the face mask, $\mathcal{FG}$. The parameter η is estimated as a ratio of the average C_r^2 to the average C_r/C_b. Fig. 9 shows the construction of the mouth map for the subject in Fig. 8.

2.2.3 Face Boundary Map and Face Score

We form an *eye-mouth triangle* for all possible combinations of the two eye candidates and one mouth candidate. Each eye-mouth triangle is verified by checking 1) luma variations and average gradient orientations of eye and mouth blobs, 2) geometry and orientation of the triangle, and 3) the presence of a face boundary around the triangle. A face score is computed for each verified eye-mouth triangle based on its eyes/mouth maps, ellipse vote, and face orientation that favors upright faces and symmetric facial geometry [16]. The triangle with the highest score that exceeds a threshold is retained. Fig. 10 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 (i.e., have counterclockwise gradient

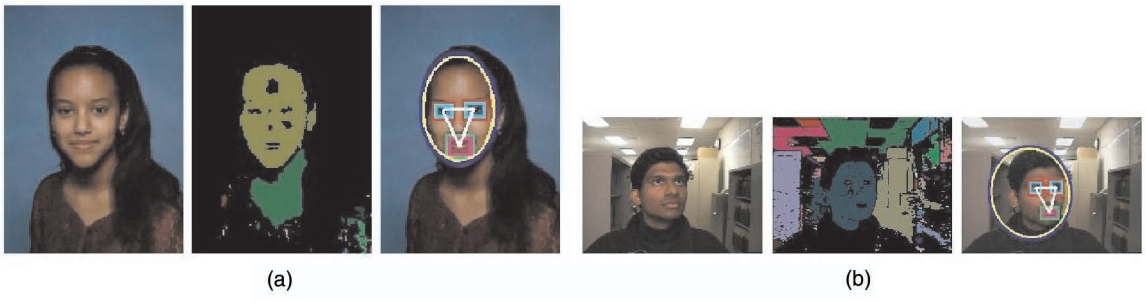

Fig. 11. Face detection examples containing dark skin-tone faces. Each example contains an original image, grouped skin regions, and a lighting-compensated image overlaid with detected face and facial features.

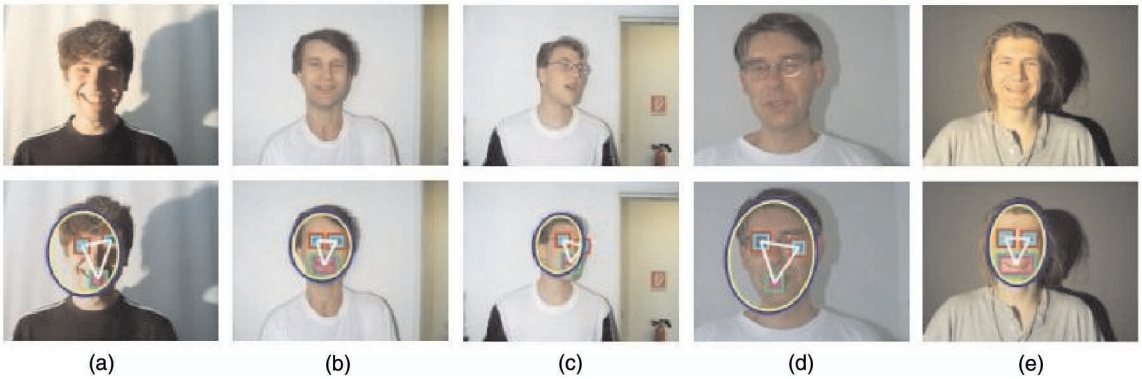

Fig. 12. Face detection results in the presence of facial variations and eye glasses. Each example contains an original image (top) and a lighting-compensated image (bottom) overlaid with detected face.

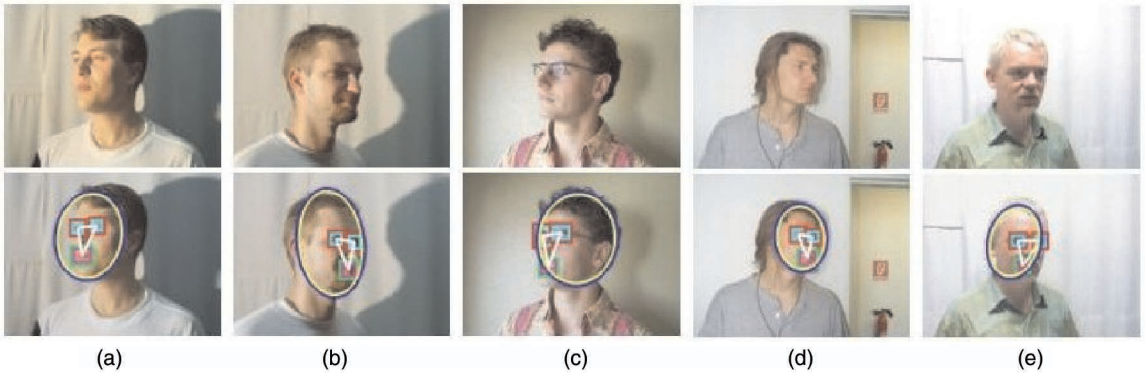

Fig. 13. Face detection results on half-profile faces (some with facial hair). Each example contains an original image (top) and a lighting-compensated image (bottom) overlaid with detected face.

orientations). Finally, we utilize the Hough transform to extract the best-fitting ellipse. Since we know the locations of eyes and mouth, and can estimate the orientation of the ellipse from eyes and mouth, we need only a two-dimensional accumulator for estimating the ellipse center. The accumulator is updated by perturbing the estimated center by a few pixels for a more accurate localization of the ellipse.

3 Experimental Results

We have evaluated our algorithm on several face image databases, including family and news photo collections. Face databases designed for face recognition, including the FERET face database [28], usually contain gray-scale mugshot-style images and, therefore, in our opinion, are not suitable for evaluating face detection algorithms. Most of the commonly used databases for face detection, including the Carnegie Mellon University (CMU) database, contain gray-scale images only. Therefore, we have constructed our database for face detection from MPEG7 videos, the World Wide Web, and personal photo collections. These color images have been taken under varying lighting conditions and with complex backgrounds. Further, these images contain multiple faces with variations in color, position, scale, orientation, 3D pose, and facial expression.

Our algorithm can detect multiple faces of different sizes with a wide range of facial variations in an image. Further, the algorithm can detect both dark skin-tone and bright skin-tone because of the nonlinear transformation of the C_bC_r color space. All the algorithmic parameters Fig. 11 demonstrates that our algorithm can successfully detect dark skin faces. Fig. 12 shows the results for subjects with some facial variations (e.g., closed eyes or open mouth) and for those

TABLE 2
Detection Results on the HHI Image Database (Image Size 640×480) on a 1.7GHz CPU

Head Pose	Frontal	Near-Frontal	Half-Profile	Profile	Total
No. of images	66	54	75	11	206
Stage 1: Grouped skin regions					
No. of FP	3145	2203	3781	277	9406
DR (%)	95.45	98.15	96.00	100	96.60
Time (sec): average $\pm$ s. d.	1.56 ± 0.45				
Stage 2: Rectangle merge					
No. of FP	468	287	582	39	1376
DR (%)	95.45	98.15	96.00	100	96.60
Time (sec): average $\pm$ s. d.	0.18 ± 0.23				
Stage 3: Facial feature detection					
No. of FP	4	6	14	3	27
DR (%)	89.40	90.74	74.67	18.18	80.58
Time (sec): average $\pm$ s. d.	22.97 ± 17.35				

FP: False Positives, DR: Detection Rate.

TABLE 3
Detection Results on the Champion Database (Image Size $\sim 150 \times 220$) on a 1.7GHz CPU

Stage	Grouped skin regions	Rectangle merge	Facial feature detection
No. of images	227		
No. of FP	5582	382	14
DR (%)	99.12	99.12	91.63
Time (sec): average $\pm$ s. d.	0.080 ± 0.036	0.012 ± 0.02	5.78 ± 4.98

FP: False Positives, DR: Detection Rate.

who are wearing glasses. Fig. 13 demonstrates that our algorithm can detect nonfrontal faces as long as the eyes and mouth are visible in half-profile views. Face can also be detected in the presence of facial hair. A summary of the detection results (including the number of false positives, detection rates, and average CPU time for processing an image) on the HHI MPEG7 image database [1] (see Figs. 12 and 13) and the Champion database [2] (see Fig. 11a) are presented in Tables 2 and 3, respectively. The HHI database contains 206 images, each of size 640×480 pixels. Subjects in the HHI database belong to several racial groups and the lighting conditions (including overhead lights and side lights) change from one image to another. Further, these images contain frontal, near-frontal, half-profile, and profile face views of different sizes. A detected face is a *correct* detection if the detected locations of the eyes, the mouth, and the ellipse bounding a human face are found with a small amount of tolerance, otherwise it is called a *false positive*. The detection rate is computed by the ratio of the number of correct detections in a grallery to that of all human faces in the gallery. The detection rate on the HHI database after the first two stages (before facial feature extraction) is ~ 97 percent for all the poses. After the third stage, the detection rate decreases to 89.40 percent for frontal faces, and to 90.74 percent for near-frontal faces, and to 74.67 percent for half-profile faces. The reason for this decrease in detection rate is the removal of those faces in which the eyes or mouth are not visible. However, we can see that the number of false positives is dramatically reduced from $9,406$ after the skin grouping stage to just 27 after the feature detection stage for the whole database containing 206 images.

The Champion database was collected from the Internet, and contains 227 *compressed* images which are approximately 150×220 pixels in size. Because most of the images in this database are captured in frontal and near-frontal views (portraits with large faces), we present the total detection rate in all poses in Table 3. Figs. 14 and 15 show the detection results on a subset of a collection of family photos (total of 55 images); Fig. 16 shows results on a subset of news photos (total of 327 images) downloaded from the Yahoo news site [3]. As expected, detecting faces in family group and news pictures is more challenging, but our algorithm is able to perform quite well on these images. Detection rate on the collection of 382 family and news photos (1.79 faces per image) is 80.35 percent and the false positive rate (the ratio of the number of false positives to the number of true faces) is 10.41 percent. Please see more results at http://www.cse.msu.edu/~hsureinl/facloc/.

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 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 YC_bC_r 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. Our algorithm constructs eye, mouth, and boundary maps to verify the face candidates. Detection results on several

Fig. 14. Face detection results on a subset of four family photos. Each image contains multiple human faces. The detected faces are overlaid on the color-compensated images. False negatives are due to extreme lighting conditions and shadows. Notice the difference between the input and color-compensated images in terms of color balance.

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 use these detected facial features as indices for identification and retrieval from image and video databases.

APPENDIX A

NONLINEAR TRANSFORMATION OF CHROMA AND THE SKIN MODEL

In the YC_bC_r color space, we can regard the chroma (C_b and C_r) as functions of the luma (Y): $C_b(Y)$ and $C_r(Y)$. Let the transformed chroma be $C_b'(Y)$ and $C_r'(Y)$. The skin color model is specified by the centers (denoted as $\overline{C_b}(Y)$ and $\overline{C_r}(Y)$) and spread of the cluster (denoted as $Wc_b(Y)$ and $Wc_r(Y)$) (See Fig. 5) and is used for computing the transformed chroma.

$$C_i'(Y) = \begin{cases} \left(C_i(Y) - \overline{C_i}(Y)\right) \cdot \frac{Wc_i}{Wc_i(Y)} + \overline{C_i}(K_h) & \text{if } Y < K_l \text{ or } K_h < Y, \\ C_i(Y) & \text{if } Y \in [K_l, K_h], \end{cases} \tag{5}$$

$$Wc_i(Y) = \begin{cases} WLc_i + \frac{(Y - Y_{min}) \cdot (Wc_i - WLc_i)}{K_l - Y_{min}} & \text{if } Y < K_l, \\ WHc_i + \frac{(Y_{max} - Y) \cdot (Wc_i - WHc_i)}{Y_{max} - K_h} & \text{if } K_h < Y, \end{cases} \tag{6}$$

Fig. 15. Face detection results on a subset of seven family photos. Each image contains cluttered background. The detected faces are overlaid on the color-compensated images. False negatives are due to extreme lighting conditions and shadows. Note that the bias color in the original images has been compensated in the resultant images.

$$\overline{C_b}(Y) = \begin{cases} 108 + \frac{(K_l - Y)\cdot(118-108)}{K_l - Y_{min}} & \text{if } Y < K_l, \\ 108 + \frac{(Y - K_h)\cdot(118-108)}{Y_{max} - K_h} & \text{if } K_h < Y, \end{cases} \tag{7}$$

$$\overline{C_r}(Y) = \begin{cases} 154 - \frac{(K_l - Y)\cdot(154-144)}{K_l - Y_{min}} & \text{if } Y < K_l, \\ 154 + \frac{(Y - K_h)\cdot(154-132)}{Y_{max} - K_h} & \text{if } K_h < Y, \end{cases} \tag{8}$$

where C_i in (5) and (6) is either C_b or C_r, $Wc_b = 46.97$, $WLc_b = 23$, $WHc_b = 14$, $Wc_r = 38.76$, $WLc_r = 20$, $WHc_r = 10$, $K_1 = 125$, and $K_h = 188$. These parameter values are estimated from training samples of skin patches from a subset of the HHI images. Y_{min} and Y_{max} values in the YC_bC_r color space are 16 and 235, respectively.

The elliptical model for the skin tones in the transformed $C_b'C_r'$ space is described in (9) and (10), and is depicted in Fig. 5.

$$\frac{(x - ec_x)^2}{a^2} + \frac{(y - ec_y)^2}{b^2} = 1, \tag{9}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} C_b' - c_x \\ C_r' - c_y \end{bmatrix}, \tag{10}$$

where $c_x = 109.38$, $c_y = 152.02$, $\theta = 2.53$ (in radian), $ec_x = 1.60$, $ec_y = 2.41$, $a = 25.39$, and $b = 14.03$ are computed from the skin cluster in the $C_b'C_r'$ space.

Fig. 16. Face detection results on a subset of 24 news photos. The detected faces are overlaid on the color-compensated images. False negatives are due to extreme lighting conditions, shadows, and low compression quality.

ACKNOWLEDGMENTS

This research was supported by US Office of Naval Research grant no. N 00014-01-1-0266.

REFERENCES

[1] MPEG7 Content Set from Heinrich Hertz Institute, http://www.darmstadt.gmd.de/mobile/hm/projects/MPEG7/Documents/N2466.html. Oct. 1998.
[2] The Champion Database, http://www.libfind.unl.edu/alumni/events/breakfast_for_champions.htm. Mar. 2001.
[3] Yahoo News Photos, http://dailynews.yahoo.com/h/g/ts/. Dec. 2001.
[4] M. Abdel-Mottaleb and A. Elgammal, "Face Detection in Complex Environments from Color Images," *IEEE Int'l Conf. Image Processing*, pp. 622-626, 1999.
[5] L.M. Bergasa, M. Mazo, A. Gardel, M.A. Sotelo, and L. Boquete, "Unsupervised and Adaptive Gaussian Skin-Color Model," *Image and Vision Computing*, vol. 18, pp. 987-1003, Sept. 2001.
[6] R. Brunelli and T. Poggio, "Face Recognition: Features vs. Templates," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1042-1052, Oct. 1993.
[7] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *Proc. IEEE*, vol. 83, pp. 705-740, May 1995.
[8] A.J. Colmenarez and T.S. Huang, "Face Detection with Information Based Maximum Discrimination," *IEEE Int'l Conf. Computer Vision and Pattern Recognition*, pp. 782-787, 1997.
[9] A. Colmenarez, B. Frey, and T. Huang, "Detection and Tracking of Faces and Facial Features," *IEEE Int'l Conf. Image Processing*, pp. 657-661, Oct. 1999.
[10] D. DeCarlo and D. Metaxas, "Optical Flow Constraints on Deformable Models with Applications to Face Tracking," *Int'l J. Computer Vision*, vol. 38, no. 2, pp. 99-127, July 2000.
[11] R. Féraud, O.J. Bernier, J.-E. Viallet, and M. Collobert, "A Fast and Accurate Face Detection Based on Neural Network," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 1, pp. 42-53, Jan. 2001.
[12] C. Garcia and G. Tziritas, "Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis," *IEEE Trans. Multimedia*, vol. 1, no. 3, pp. 264-277, Sept. 1999.

[13] M. Grudin, "On Internal Representation in Face Recognition Systems," *Pattern Recognition,* vol. 33, pp. 1161-1177, 2000.
[14] E. Hjelm and B.K. Low, "Face Detection: A Survey," *Computer Vision and Image Understanding,* vol. 83, no. 3, pp. 236-274, Sept. 2001.
[15] 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.
[16] R.-L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, "Face Detection in Color Images," Techical Report MSU-CSE-01-7, Michigan State Univ., Mar. 2001.
[17] W. Huang, Q. Sun, C.-P. Lam, and J.-K. Wu, "A Robust Approach to Face and Eyes Detection from Images with Cluttered Background," *Proc. Int'l Conf. Pattern Recognition,* vol. 1, pp. 110-114, Aug. 1998.
[18] P.T. Jackway and M. Deriche, "Scale-Space Properties of the Multiscale Morphological Dilation-Erosion," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 18, no. 1, pp. 38-51, Jan. 1996.
[19] M. Jones and J.M. Rehg, "Statistical Color Models with Application to Skin Detection," *Technical Report Series,* Cambridge Research Laboratory, Dec. 1998.
[20] C. Kotropoulos, A. Tefas, and I. Pitas, "Frontal Face Authentication Using Morphological Elastic Graph Matching," *IEEE Trans. Image Processing,* vol. 9, pp. 555-560, Apr. 2000.
[21] K.M. Lam and H. Yan, "An Analytic-to-Holistic Approach for Face Recognition Based on a Single Frontal View," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 20, no. 7, pp. 673-686, July 1998.
[22] K.M. Lam and H. Yan, "Locating and Extracting the Eye in Human Face Images" *Pattern Recognition,* vol. 29, no. 5, pp. 771-779, 1996.
[23] M.S. Lew and N. Huijsmans, "Information Theory and Face Detection," *Proc. Int'l Conf. Pattern Recognition,* pp. 601-605, Aug. 1996.
[24] 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.
[25] B. Menser and M. Brunig, "Locating Human Faces in Color Images with Complex Background," *Intelligent Signal Processing and Comm. Systems,* pp. 533-536, Dec. 1999.
[26] M. Pantic and L.J.M. Rothkrantz, "Automatic Analysis of Facial Expressions: The State of the Art," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 22, no. 12, pp. 1424-1445, Dec. 1996.
[27] A. Pentland, B. Moghaddam, and T. Starner, "View-Based and Modular Eigenspaces for Face Recognition," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 84-91, 1994.
[28] P.J. Phillips, H. Moon, S.A. Rizvi, and P. J. Rauss, "The FERET Evaluation Methodology for Face-Recognition Algorithms," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 22, no. 10, pp. 1090-1104, Oct. 2000.
[29] A. Nikolaidis and I. Pitas, "Facial Feature Extraction and Determination of Pose," *Pattern Recognition,* vol. 33, pp. 1783-1791, 2000.
[30] E. Osuna, R. Fréund, and F. Girosi, "Training Support Vector Machines: an Application to Face Detection," *Proc. IEEE Conf. Computer Vision and Pattern Recognition* pp. 130-136, June 1997.
[31] C.A. Poynton, *A Technical Introduction to Digital Video.* John Wiley & Sons, 1996.
[32] H.A. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 20, no. 1, pp. 23-38, Jan. 1998.
[33] H.A. Rowley, S. Baluja, and T. Kanade, "Rotation Invariant Neural Network-Based Face Detection," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 38-44, 1998.
[34] E. Saber and A.M. Tekalp, "Frontal-View Face Detection and Facial Feature Extraction Using Color, Shape and Symmetry Based Cost Functions," *Pattern Recognition Letters,* vol. 19, pp. 669-680, 1998.
[35] H. Schneiderman and T. Kanade, "A Statistical Method for 3D Object Detection Applied to Faces and Cars," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 746-751, June 2000.
[36] A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-Based Image Retrieval at the End of the Early Years," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 22, no. 12, pp. 1349-1380, Jan. 2000.
[37] 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.
[38] K. Sobottka and I. Pitas, "A Novel Method for Automatic Face Segmentation, Facial Feature Extraction and Tracking," *Signal Processing: Image Comm.,* vol. 12, pp. 263-281, 1998.
[39] K.K. Sung and T. Poggio, "Example-Based Learning for View-Based Human Face Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 20, no. 1, pp. 39-51, Jan. 1998.
[40] 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. Face and Gesture Recognition,* pp. 54-61, 2000.
[41] M.A. Turk and A.P. Pentland, "Face Recognition Using Eigenfaces," *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* pp. 586-591, June 1991.
[42] *Face Recognition: From Theory to Applications,* H. Wechsler, P. Phillips, V. Bruce, F. Soulie, and T. Huang, eds. Springer-Verlag, 1998.
[43] H. Wu, Q. Chen, and M. Yachida, "Face Detection from Color Images Using a Fuzzy Pattern Matching Method," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 21, no. 6, pp. 557-563, June 1999.
[44] M.-H. Yang and N. Ahuja, "Detecting Human Faces in Color Images," *Proc. IEEE Int'l Conf. Image Processing,* pp. 127-139, Oct. 1998.
[45] M.-H. Yang, D. Kriegman, and N. Ahuja, "Detecting Faces in Images: A Survey," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 24, no. 1, pp. 34-58, Jan. 2001.
[46] K.C. Yow and R. Cipolla, "Feature-Based Human Face Detection," *Image and Vision Computing,* vol. 15, no. 9, pp. 713-735, Sept. 1997.
[47] W. Zhao, R. Chellappa, A. Rosenfeld, and P.J. Phillips, "Face Recognition: A Literature Survey," CVL Technical Report, Center for Automation Research, Univ. of Maryland at College Park, ftp://ftp.cfar.umd.edu/TRs/CVL-Reports-2000/TR4167-zhao.ps.gz. Oct. 2000.