ROBUST FACE DETECTION FOR DIFFERENT CHROMATIC ILLUMINATIONS

Soo-Chang Pei, and Ching-Long Tseng
Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, R. O. C.
E-mail: pei@cc.ee.ntu.edu.tw

ABSTRACT

In this paper, we propose a robust face detection method in color images, and it is suitable for various races and under different illuminations. The pre-processing is to convert the 3-D RGB color histogram into 2-D R-Y/B-Y color histogram and to introduce the concept of flesh tone. The second procedure is the kernel processing, which is 2-D color histogram normalization, to eliminate the affine effect of the color histograms due to the illuminations changing. We propose the new standard flesh region by using statistics in the normalized domain in addition, then to extract the candidates of the flesh regions. And the post-processing is to use connected operators, some criterions and thresholds to obtain accurate face regions. Experimental results show that the face detection is successful and satisfactory for the synthetic images and the real images.

1. INTRODUCTION

Automatic human face detection is the major issue of facial image processing useful in various fields, such as model-based video coding, face identification, video indexing and retrieval, intelligent vision system, and human computer interface system [1]. Recently, works have begun in face detection and location in general images which have little or no constraint on the number, location, size, and orientation of human faces in the scenes. One approach detects face features at first, and then uses a ranking technique to get possible face locations. Another approach involves model-based algorithms that treat human faces as a whole unit consisting of inter-connected arcs. Other methods are based on ellipse fitting, neural networks, and multiresolution domain knowledge.

At the first step of many face detection algorithms, based on a model for the distribution of skin color, it detects skin-colored regions and creates a binary classification image which indicates the likelihood of each pixel to represent skin. Unfortunately, color of an object varies with changes in illumination source, illumination geometry, viewing angle, and miscellaneous sensor parameter, so color of the face in the image will deviate when the image is taken under different illuminations. Therefore, for the detection of faces in color images taken under different illuminations, the desired algorithm requires not only a technique to remove the effect of color deviation caused by varying illumination conditions first but also some criterions to detect the face position.

In this paper, we will propose a novel algorithm to detect the face under different illuminations. At first, to convert the color histogram of test images into an appropriate coordinate system. Then using 2-D color histogram normalization to remove the effect due to varying illuminations and proposing the new binary standard flesh region in the normalized domain. At last to adopt some connected operators and criterions to obtain actual facial regions in the final step.

2. MODELING COLOR IMAGE REPRESENTATION

We can observe that the n sensor measurements in a color imaging system at each image location (x,y) is [2]

$$\rho_i(x,y) = \int_\lambda I(\lambda) s(x,y,\lambda) f_i(\lambda) d\lambda, \quad 1 \le i \le n \tag{1}$$

where λ indicates wavelength, $I(\lambda)$ is the spectral power distribution of the scene illumination, $s(x,y,\lambda)$ is the spectral reflectance of the surface, $f_i(\lambda)$ is the spectral sensitivity of the ith sensor class. However, the surface reflectance $s(x,y,\lambda)$ can be approximated by

$$s(x,y,\lambda) = \sum_{j=1}^{m} \sigma_j(x,y) S_j(\lambda) \tag{2}$$

where $S_j(\lambda)$ are a set of m fixed basis functions and $\sigma_j(x,y)$ are the surface weighting coefficients corresponding to different locations. Combining Eqs. (1) and (2), we can find

$$\rho(x,y) = \mathbf{A}\sigma(x,y) \tag{3}$$

therefore,

$$\mathbf{A}_{ij} = \int I(\lambda) S_j(\lambda) f_i(\lambda) d\lambda \tag{4}$$

Obviously, **A** depends on the light spectrum wavelength λ but is independent of the image location (x,y).

Assume **A** and $\tilde{\mathbf{A}}$ corresponding to $I(\lambda)$ and $\tilde{I}(\lambda)$ are nonsingular matrices, then one can show the relationship of the two images

$$\tilde{\rho}(x,y) = \mathbf{M}\rho(x,y) \tag{5}$$

where $\mathbf{M} = \tilde{\mathbf{A}}\mathbf{A}^{-1}$. Let $H(\rho)$ and $H(\tilde{\rho})$ be n-dimensional color histograms corresponding to the same surface imaged under different illuminations $I(\lambda)$ and $\tilde{I}(\lambda)$. From Eq. (5), we can find

$$H(\tilde{\rho}) = H(\mathbf{M}\rho) \tag{6}$$

Thus, two n-dimensional color histograms are related by an affine transformation, when changing the illumination.

3. FLESH COLOR ANALYSIS

Based on Refs. [3]-[4], the fact of the color of human skin is quite different from many other natural objects is unique and well known. Based on Ref. [3], we can know that the chrominance components of skin pixels are clustered in small area of the chrominance plane, but the major difference between skin tones is intensity.

In a television receiver, the video signal is represented as a luminance (Y) component and two color difference components (R-Y and B-Y), and flesh tone is the most critical and common color reference in TV system [4]. Since the relation between the RGB and color difference values is a linear mapping, we will use the 2-D color difference components (R-Y and B-Y) to define the flesh-tone region which is shown in Fig. 1(a) [5]. The weights and threshold for each decision boundary are shown in Table 1. Note that we amend the values and sign of the table slightly in Ref. [5] and list the corrected values and sign in Table 1. Also, the decision boundary can be described by the following [5]

$$w_1(B-Y)+w_2(R-Y)+\theta=0 \tag{7}$$

However, the relation between the color difference components R-Y/B-Y and the 2-D chrominance components CbCr is

$$\begin{bmatrix} B-Y \\ R-Y \end{bmatrix}=\begin{bmatrix} 1.7721 & -0.0002 \\ 0 & 1.4019 \end{bmatrix}\begin{bmatrix} Cb \\ Cr \end{bmatrix}=\mathbf{S}\begin{bmatrix} Cb \\ Cr \end{bmatrix} \tag{8}$$

where $\mathbf{S}=\begin{bmatrix} 1.7721 & -0.0002 \\ 0 & 1.4019 \end{bmatrix}$ indicates the effect of skew of a general 2-D affine transformation; note that -0.0002 can be regarded as 0 corresponding to truncated error. Therefore, we can find that 2-D color histograms are related by an affine transformation in color difference components (R-Y and B-Y) coordinate due to the illuminations changing.

4. ALGORITHM FOR FACIAL REGION DETECTION

4.1 Pre-Processing: Convert into R-Y/B-Y Coordinate

First we convert the 3-D color histogram of RGB coordinate into 2-D color histogram of R-Y/B-Y coordinate for all images. Note that both R-Y and B-Y component values are centered around the origin and ranging from -2.5 to +2.5.

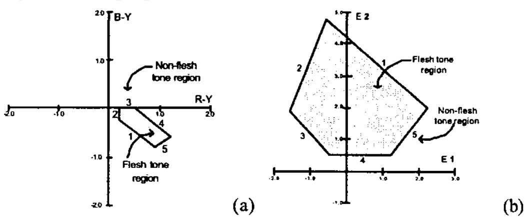

Fig. 1: (a) Classification regions for flesh tone in R-Y/B-Y [5], and (b) is the proposed version in normalized domain.

Category	Boundary	w_2	w_1	θ
Flesh tone (R-Y/B-Y)	1	0.1639	0.1408	0.0153
	2	-0.0515	0.0	0.0054
	3	0.0	0.0515	0.0
	4	-0.0877	-0.0735	0.0282
	5	-0.2703	0.2222	0.4098

Table 1: Weights and threshold for each decision boundary of Fig. 1(a)

Category	Boundary	w_2	w_1	θ
Flesh tone (Normalized)	1	-0.9943	-1	4.2210
	2	2.8277	-1	6.2733
	3	-1.0456	-1	0.0293
	4	-0.0247	-1	0.5412
	5	1.3034	-1	-1.1202

Table 2: Weights and threshold for each decision boundary of Fig. 1(b)

4.2 Kernel-Processing: 2-D Histogram Normalization

Refer to Ref. [6], we will apply the normalization algorithm to 2-D color histograms in color difference components R-Y/B-Y coordinate. In order to simplify and unify the notation, we will use the symbols (RY, BY) to replace the symbols $(r-y, b-y)$ in color difference components R-Y/B-Y coordinate in the following discussion. The procedure of normalizing the color histogram contains three major steps as described as below.

Step 1: Computing the mean vector $\mathbf{c}=[C_{RY}\ C_{BY}]^T$ *and the covariance matrix* $\mathbf{M}$ *of the color histogram*

Step 2: Finding out the eigenvalues and aligning the coordinates with eigenvectors of $\mathbf{M}$

Assume $\mathbf{e}_1=[e_{1RY}\ e_{1BY}]^T$ is the eigenvector associated with the eigenvalue λ_1 of $\mathbf{M}$, and $\mathbf{e}_2=[e_{2RY}\ e_{2BY}]^T$ is the eigenvector associated with the eigenvalue λ_2 of $\mathbf{M}$. Consequently, we can construct a rotational matrix $\mathbf{E}$

$$\mathbf{E}=\begin{bmatrix} e_{1RY} & e_{1BY} \\ e_{2RY} & e_{2BY} \end{bmatrix} \tag{9}$$

In order to transform the coordinate system lying on the eigenvectors of $\mathbf{M}$, as shown as Eq. (10)

$$\begin{bmatrix} \hat{R}Y \\ \hat{B}Y \end{bmatrix}=\mathbf{E}\begin{bmatrix} RY-C_{RY} \\ BY-C_{BY} \end{bmatrix} \tag{10}$$

where $[\hat{R}Y\ \hat{B}Y]^T$ denote the new coordinates. To avoid the reflection of the transformed pattern to the original color histogram, so $\mathbf{E}$ must be rewritten in the form

$$\mathbf{E}=\begin{bmatrix} e_{1RY} & e_{1BY} \\ -e_{1BY} & e_{1RY} \end{bmatrix} \tag{11}$$

Step 3: Rescaling the new coordinates using the corresponding eigenvalues of $\mathbf{M}$

We can scale each component independently in order to make covariance matrix equal to a scaled identical matrix

$$\begin{bmatrix} \hat{R}'Y \\ \hat{B}'Y \end{bmatrix}=\mathbf{W}\begin{bmatrix} \hat{R}Y \\ \hat{B}Y \end{bmatrix}=\mathbf{W}\mathbf{E}\begin{bmatrix} RY-C_{RY} \\ BY-C_{BY} \end{bmatrix} \tag{12}$$

where $\mathbf{W}=\begin{bmatrix} k/\sqrt{\lambda_1} & 0 \\ 0 & k/\sqrt{\lambda_2} \end{bmatrix}$ is called a scaling matrix, which preserves the overall size of the normalized shape; and k is constant. Then we obtain the most compact color histogram after the transformation in (12).

Afterwards, we just find an angle by which we rotate the compact color histogram into a normalized form that is invariant under rotation. Refer to Ref. [6] and make use of some tensor theories to solve the rotational problem. The detailed description is in Ref. [6]. Here, we only conclude concisely. Assume $\bar{t}^1$ and $\bar{t}^2$ are the first-order tensors of the normalized color histogram, and t^1 and t^2 are the first-order tensors of the compact color histogram. The rotational problem can be defined as follows

$$\begin{bmatrix} \bar{t}^1 \\ \bar{t}^2 \end{bmatrix}=\begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}\begin{bmatrix} t^1 \\ t^2 \end{bmatrix} \tag{13}$$

Because the compact color histogram is rotated around the center point of the color histogram, the tensor $\bar{t}^1$ becomes zero. Therefore, by setting $\bar{t}^1=0$ for the normalization algorithm

$$\bar{t}^1=0=t^1\cos\alpha+t^2\sin\alpha \tag{14}$$

We get the relation which is represented by third-order central moments of the compact color histogram

$$\tan\alpha=-\frac{t^1}{t^2}=-\frac{u_{12}+u_{30}}{u_{03}+u_{21}} \tag{15}$$

We choose ϕ in $\{\alpha, \alpha+\pi\}$ such that $\bar{t}^2>0$ in order to uniquely determine the rotational angle in the normalization algorithm. Combining with Eq. (12), we can obtain the whole normalization transformation from the original to the normalized color histogram

$$\begin{bmatrix} \hat{R}'Y \\ \hat{B}'Y \end{bmatrix}=\begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}\begin{bmatrix} \frac{k}{\sqrt{\lambda_1}} & 0 \\ 0 & \frac{k}{\sqrt{\lambda_2}} \end{bmatrix}\begin{bmatrix} e_{1RY} & e_{1BY} \\ -e_{1BY} & e_{1RY} \end{bmatrix}\begin{bmatrix} RY-C_{RY} \\ BY-C_{BY} \end{bmatrix} \tag{16}$$

According to the statistics of hundreds of test images which are taken under nearly normal (standard) illumination, by adopting the normalization algorithm, we find out the pixels settling down the flesh-tone region in R-Y/B-Y coordinate gather at a particular cluster in the normalized domain. Therefore, we can recommend another new binary standard flesh region in the normalized domain as shown as Fig. 1(b) to extract the candidates of the flesh regions after these illumination-changed images doing the 2-D color histogram normalization algorithm. The weights and thres hold for each proposed decision boundary are shown in Table 2. Similarly, the decision boundary can be described by

$$w_1 \bullet E2 + w_2 \bullet E1 + \theta = 0 \quad (17)$$

where $E1$ is the direction of the first eigenvector of M (base 1), and $E2$ is the direction of the second eigenvector of M (base 2); the values and sign of Table 2 are only suitable for the normalized domain which is obtained from the R-Y/B-Y coordinate centered around the origin and ranging from −2.5 to +2.5.

When the new standard flesh region is obtained and the test images are transformed by the above transformation in Eq. (16), it can be extracted actual flesh regions of the test images by according to the overlaps between new standard flesh region and normalized forms to obtain a binary image.

4.3 Post-Processing: Connected Operators and Criterions

Connected operators are nonlinear filters that eliminate some parts of the image while preserving the contour of the remaining parts. Fig. 2 shows the concept of binary connected operators. Therefore, here, we take turns to use closing, opening and closing in order to connect and reduce the number of facial candidate regions. The last work is to use the face shape property to eliminate the non-face regions significantly to derive the facial candidates.

We first compute for each connected region C the best-fit ellipse E on the basis of moments. An ellipse is exactly defined by the center $(\bar{x}, \bar{y})$, its orientation angle θ, the minor axis a, and the major axis b. Fig. 3 shows the minimal bounding rectangular (MBR), the ellipse model, and the facial candidate regions and its best-fit ellipse.

The procedure of eliminating the non-face regions consists of two steps described below.

1. $\dfrac{\text{black pixels within the MBR}}{\text{total pixels within the MBR}} > TH1$
2. $TH2 < b/a < TH3$

Note that these criterions can be used to simplify and analyze the binary image at different threshold levels to avoid the inexact facial region detection; generally by experiments, *TH1* is 0.52 at least, *TH2* is about 0.48, and *TH3* doesn't exceed in 0.83. In summary, the whole proposed face detection method could be depicted as shown as Fig. 4.

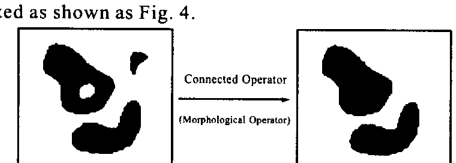

Fig. 2: The concept of the binary connected operator.

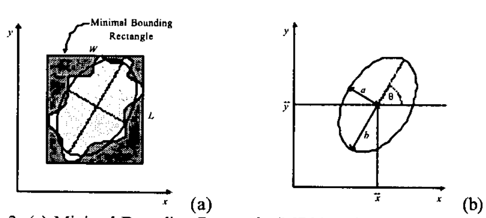

Fig. 3: (a) Minimal Bounding Rectangle (MBR) and approximation of a region by an ellipse and (b) the ellipse model.

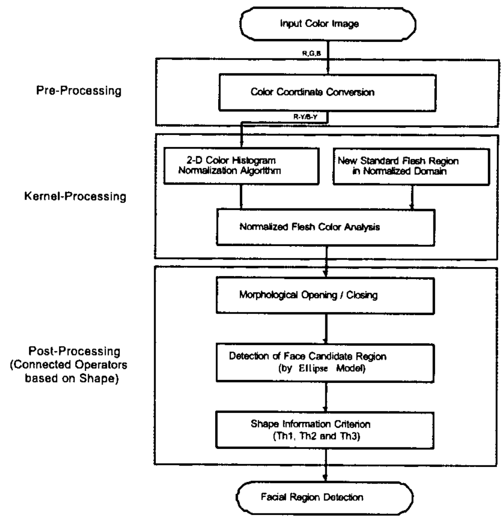

Fig. 4: Flow chart of the proposed face detection algorithm.

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

In these test images, from Website and some natural images, there are two synthetic images and the two other real images. Note that the size of all images is 128×128.

(1) The synthetic images, taken under different illuminations, are simulated by computer

. *Experimental Setup*

We choose the first images which are taken under near to standard illumination as the standard template image from Website and digitizing a natural image, then we will use some image processing software to imitate the illuminant-changed images. Therefore, additional seven images of each picture were fabricated under lighter, darker, red, yellow, green, blue and magenta by the current software called Paint Shop Pro 5.0 for several illumination adjustments.

. *Experimental Results*

In Fig. 5, there is a set of computer simulated experimental results of facial region detection for the Caucasian and Mongoloid, respectively. The resultant images of each group are aligned in a horizontal row, for example as label (a), for the single Caucasian (Man), the most left figure shows the test color image that is taken under magenta illumination. Next, the second figure shows the color histogram of test image in R-Y/B-Y coordinate system. Note that the range formed with the red line segments in the second figure is the standard flesh region. Afterwards, it can be extracted flesh regions of the test images by according to the overlaps between the two color regions; the result is shown as the third figure. In the fourth figure, it shows normalized color histograms of the test image and new standard flesh region in normalized domain; the range formed with the red line segments represents the proposed standard flesh region in normalized domain. Then, it can be also extracted flesh regions of the test images by according to the overlaps between the new standard flesh region and the normalized color histogram, as shown as the fifth figure. To continue adopting the third step of the whole-proposed algorithm, we can obtain the result of the facial detection as the most right figure. The whole experiments of the different image under different simulation conditions are shown in another group labeled (b). The image is fabricated under green illuminations by Paint Shop Pro 5.0 to regard as the tested images. There are seven other illumination' esults for each case not displayed here for space saving. Compare the third figure with the fifth figure of extracted flesh regions of each group experiments, we can easily find the successful rate of the fifth figure after normalization is better than the third figure without it for all illumination conditions. It's very effective with 2-D color histogram normalization to reduce the affine effect of the color histograms caused by the different illuminations.

(2) The near real images, taken under various and unknown illuminations, are obtained from Website

. *Experimental Setup*

Here, since the test images are obtained from Website; therefore, the color of skin and the illuminations are multiple, various and unknown. These images seem to be taken under different illuminations; the slanted color appearances are existed in the test images originally.

. *Experimental Results*

Similarly, Fig. 6 shows the experiment results of the near real images taken under various and unknown illuminations. These experimental results show that the facial detection is also satisfactory for such cases, no matter what the illuminations or the races are.

6. CONCLUSIONS

We propose a robust algorithm to detect the facial regions in color images which are taken under different illuminations. The algorithm requires not only a technique to remove the effect of color deviation caused by varying illumination conditions but also some criterions to detect the face position. In the first experiment, there are two synthetic images taken under different illuminations, which are simulated by computer. On the other hand, in order to be close to the realistic, the other two real images are taken under multiple, various and unknown illuminations, which are obtained from Website. From these experimental results, the segmented facial regions figures are quite successful and satisfactory for both cases. Worthy to be mentioned, the algorithm is suitable for use in the facial regions detection of various races (Mongolian, Caucasian, and Negroid).

REFERENCE

[1] H. Wang and S. F. Chang, "A highly efficient system for automatic face region detection in MPEG video," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 7, no. 4, pp. 615-628, Aug. 1997.

[2] G. Healey and D. Slater, "Global color constancy: recognition of objects by use of illumination invariant properties of color distributions," *J. Opt. Soc. Am. A*, vol. 11, pp. 3003-3010, Nov. 1994.

[3] M. J. Jones and J. R. Rehg, "Statistical color models with application to skin detection," Tech. Rep. CRL 98/11, Compaq Cambridge Research Lab., 1998.

[4] E. J. Lee and Y. H. Ha, "Automatic flesh tone reappearance for color enhancement in TV," *IEEE Trans. on Consumer Electronics*, vol. 43, no. 4, pp. 1153-1159, Nov. 1997.

[5] S. Satyanarayana and S. Dalal, "Video color enhancement using neural networks," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 6, pp. 295-307, June 1996.

[6] S. C. Pei and C. N. Lin, "Image normalization for pattern recognition," *Image and Vision Computing*, vol. 13, pp. 711-723, Dec. 1995.

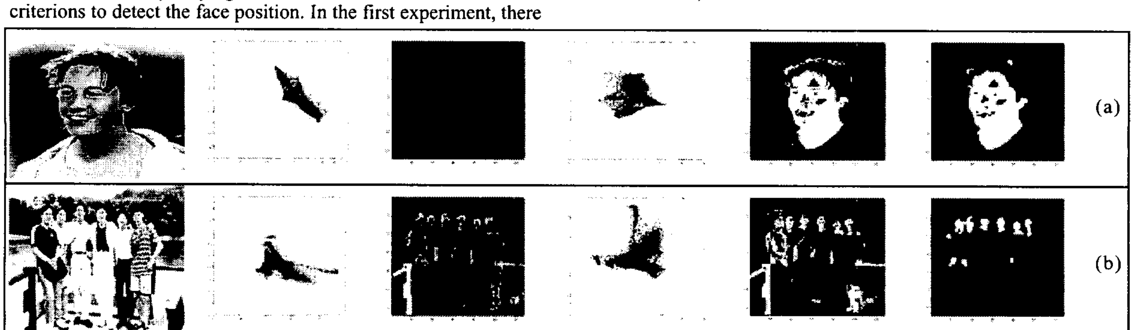

Fig. 5: Examples of facial region detection for the **Caucasian and Mongoloid**. Images under *(a) Magenta (b) Green* illumination, *by computer simulation*.

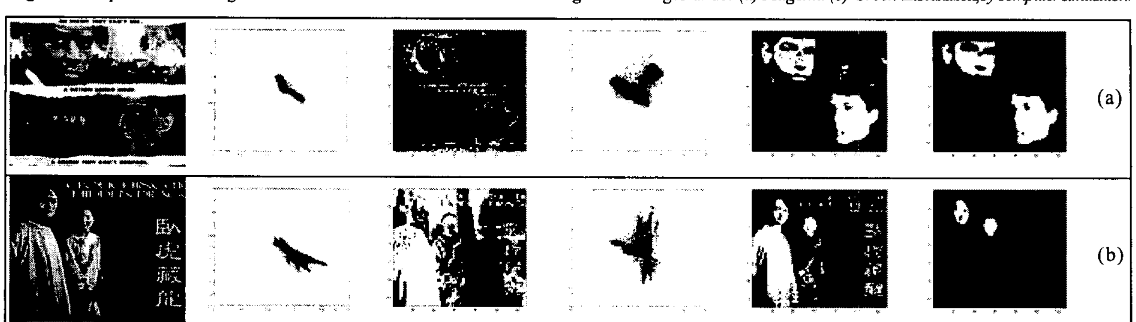

Fig. 6: Examples of facial region detection for the **multiple races** under **various unknown illuminations**. Illumination close *(a) Blue (b) Red* illumination.