JOURNAL OF KOREA MULTIMEDIA SOCIETY, VOL. 6, NO. 4, JULY 2003

DETECTION OF FACIAL FEATURES IN COLOR IMAGES WITH VARIOUS BACKGROUNDS AND FACE POSES

Jae-Young Choi and Nak-Bin Kim

ABSTRACT

In this paper, we propose a detection method for facial features in color images with various backgrounds and face poses. To begin with, the proposed method extracts face candidacy region from images with various backgrounds, which have skin-tone color and complex objects, via the color and edge information of face. And then, by using the elliptical shape property of face, we correct a rotation, scale, and tilt of face region caused by various poses of head. Finally, we verify the face using features of face and detect facial features. In our experimental results, it is shown that accuracy of detection is high and the proposed method can be used in pose-invariant face recognition system effectively

Key words: *color image, face, detection, facial feature*

1. INTRODUCTION

Recently, the detection and recognition of human face from digital images has been researched. Especially, because face contains many information, the extraction of facial region and detection of facial features are important stages in face recognition study[1,2]. But the existing approaches have many restricted condition in rotation of face, tilt of face, and change of scale which were caused by the distance and position between camera and person. Moreover, in order to get the face region accurately, they need to set the background color of images the opposite of skin-color in advance[3].

So we propose the extraction method of face regions from color images that have skin-tone color backgrounds. Also a robust detection technique of facial features under various poses of face is presented.

First, our method extracts face candidacy region from the image by using color statistics which are made with color information of human face. At this time, to separate face color from skin-tone color which contained in non-face region, the edge information of YCbCr components is applied together.

On the extracted candidacy region, we make the principal axis of face the vertical axis of image by exploiting the elliptic nature of face, that is, we correct the angle of face in order to detect the facial features easily. And finally faces are detected and verified using facial geometric features.

To evaluate the proposed method, test images that have various poses of face and complex backgrounds are used. Through results, we confirmed that the presented technique can be applied to color images which contain multi situations robustly.

The remainder of this paper is organized as follows: section 2 describes extraction of face candidacy region using color information, correction of various face poses using angle distribution, and detection of facial features in order. Section 3 presents experimental results of the proposed method, and conclusions are drawn at last.

The authors are with the College of Software, Kyungwon Univ., Korea. E-mail : jy1001@web.kyungwon.ac.kr and nbkim@kyungwon.ac.kr

2. DETECTION OF FACIAL FEARURES

Our goal in this paper is to detect facial features. For achieving this goal, the first stage of proposed scheme extracts the face candidacy regions using information of the facial color and edge. And after correcting the angle of face axis in the second step, we detect the facial features using the horizontal edges and geometric features of face finally.

2.1 Extraction of Face Candidates

2.1.1 Face color segmentation

The color image is described one of many kinds of color models to express the relationship among the color components. The color models most often used for image processing are the RGB, the HSI, and the YCbCr model.

YCbCr color model is able to separate intensity value and chrominance components. Added to this, conversion of RGB into YCbCr is simple and fast. So the YCbCr model is used in this study, but only CbCr components are applied to proposed algorithm. It is for this reason that the variations of lighting condition can be reduced by eliminating the intensity value Y. The RGB to YCbCr conversion is defined as Eq. (1) [4].

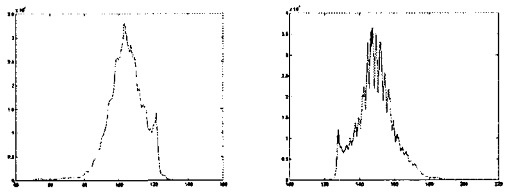
$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.29900 & 0.58700 & 0.11400 \\ -0.16874 & -0.33126 & 0.50000 \\ 0.50000 & -0.41869 & -0.08131 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

The process of extracting face candidacy region uses the sampling data which made with the YCbCr color information. To make this reference data, we collect 300 photos of the Orientals, and the data contain various photographing conditions as Table 1.

Table 1. Photographing conditions of sampling

Number of images		Classification		
		Places	Input devices	Time
300	150	Indoors	Digital camera	Day
	150	Outdoors	CCD camera	Night

The collected data applied to the Eq. (1), and the distributed graphs of CbCr components are got as shown Fig. 1. Its horizontal axis and vertical axis indicate the pixel value and the accumulative value of pixels respectively. In order to set the scope of distribution function, this paper applies the confidence interval 99% to distributed graph. As the results of this, the obtained limits of Cb and Cr are [63,125] and [128,182] respectively.



(a) Cb component (b) Cr component

Fig. 1. Distribution of CbCr components.

This distributed graph seems a normal distribution but strictly speaking, is not a normal distribution function. In other words, suppose it is normal distribution, it is possible that the non-face color, which is not contained sample image, is belong to the limits.

$$f(x,y) = \begin{cases} 1 & if (63 < Cb < 125) \cap (128 < Cr < 182) \cap \\ & (R-G)_{\min} < R-G < (R-G)_{\max} \cap \\ & (G-B)_{\min} < G-B < (G-B)_{\max} \cap \\ & (R-B)_{\min} < R-B < (R-B)_{\max} \\ 0 & otherwise \end{cases} \tag{2}$$

In order to overcome this defect, the values of CbCr components with difference value of RGB are stored into the lookup table. The restricted conditions through the scope of CbCr and difference of RGB are shown as Eq. (2). By this way, each pixel in the image is labeled according to whether it is close to skin-tone or not. For example, if Cb and Cr values of pixel belong to the scope of face color then the pixel value of result image is 1 and if not so, the relevant pixel value of result image is 0.

The comparison between existing and proposed method is shown as Fig. 2. While the proposed

Fig. 2. Results of face color segmentation.
(a) Original image, (b) Extraction using existing method, (c) Extraction using proposed method

method extracts only human skin, the existing method extracts clothes which are similar to skin-tone color in addition to human skins.

2.1.2 Labeling

In case the images contain many people, the multiple face candidacy regions and noise regions could be detected. So the labeling is processed for segmenting multiple face candidacy regions and eliminating noises. In this time, if background has the skin-tone color, the face region and background region will be labeled the same number[5].

To prevent this matter, after composing face candidacy region and Laplacian edge, we execute labeling process. In this work Laplacian operator uses the Y, Cb, and Cr components. As shown Fig. 3, if the Laplacian edge is not used, the whole image is labeled to one because the pixel values

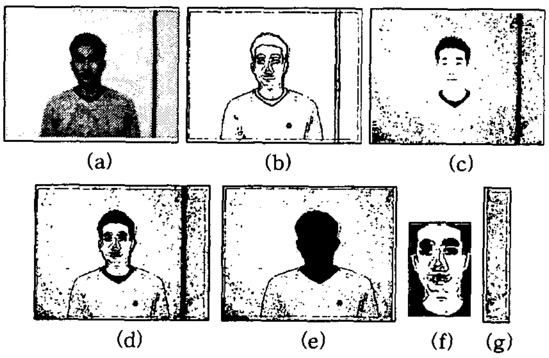


Fig. 3. Labeling process.
(a) Original image, (b) Edge image of YCbCr, (c) Image of face candidate (white is candidacy pixel), (d) Composite image of (b) and (c), (e)-(g) Result images of labeling (labeled regions are three).

of background and cloth are belong to the scope of skin-color.

If the labeled region is below a certain size then it is judged on noise and is eliminated. For setting threshold value, we analyze the distribution of labeling size and determine the value of threshold automatically as shown Eq. (3).

$$Th_{label} = \frac{1}{N}\sum_{i=1}^{N-1}(s_{i+1} - s_i) \tag{3}$$

- N : number of labels
- S_i : number of accumulative pixels of i-th label

2.2 Correction of Various Face Poses

One of the general features of face is that eyes and mouth elements contain a horizontal property. So the purpose of this stage is the correction of tilt face caused by various face poses in order to detect elements of eyes and mouth easily.

$$acc(\theta) = \sum f(x,y)$$
$$s.t\{x, y \mid f(x,y) \in T_{cr}, T_{cb} \quad and$$
$$\theta = \arctan[(fx - CP)/(fy - CP)]/\pi \times 180\} \tag{4}$$

- $f(x,y)$: Coordinate of face candidacy region by Eq. (2)
- CP : Center of gravity of all points in T_{cb}, T_{cr}
- T_{cb}, T_{cr} : Ranges of face color

Proposed algorithm is based on elliptical shape which is similar with facial shape. First, it measures the angles between center coordinate of face candidacy region and the other pixels which are belonged to the distribution scope of face color. And secondly, each angle is accumulated at one of the interval [0,359].

Then we obtained the values of cumulative pixels according to the degree of angle and drawn the accumulative graph as shown Fig. 4.

From here, θ_{ik} means the temples of face, and generally a temple locates at 3/4 degree of distance between center of face and the crown of the head. And max($acc(\theta_i)$) means the crown of the head in Fig. 4, but not points at the top of the head

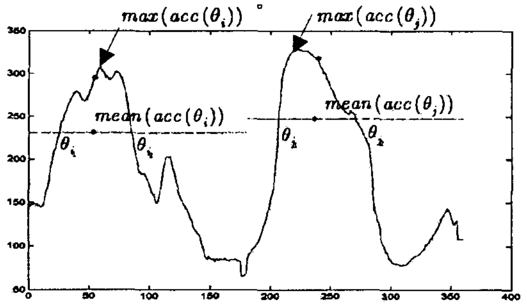


Fig. 4. Accumulative graph of face color according to the degree of angle.

exactly. It is just an assumption that the shape of face is modeled by an ellipse. θ_{jk} also means the chin joints like head part as shown Eq. (5).

$$\theta_{th} = \begin{cases} \theta_{i_k} \quad s.t \quad acc(\theta_{i_k}) = \max(acc(\theta_i)) \times 0.75 \\ \qquad\qquad if \quad 0 < \theta_{i_k}, \theta_i \le 180 \\ \theta_{j_k} \quad s.t \quad acc(\theta_{j_k}) = \max(acc(\theta_j)) \times 0.75 \\ \qquad\qquad if \quad 180 < \theta_{j_k}, \theta_j \le 360 \end{cases} \tag{5}$$

We will be able to calculate the angle of principal axis of the face using the temples and chin joints. The angle is calculated as following Eq. (6).

$$angle = \begin{cases} \theta_{i_{mean}} + (\theta_{j_{mean}} - \theta_{i_{mean}} - 180)/2 - 90 \\ \qquad if(\theta_{j_{mean}} - \theta_{i_{mean}}) \ge 180 \\ \theta_{i_{mean}} - (180 - \theta_{j_{mean}} + \theta_{i_{mean}})/2 - 90 \\ \qquad if(\theta_{j_{mean}} - \theta_{i_{mean}}) < 180 \end{cases}$$
$$, where \quad \theta_{i_{mean}} = (\theta_{i_1} + \theta_{i_2})/2$$
$$\theta_{j_{mean}} = (\theta_{i_1} + \theta_{i_2})/2 \tag{6}$$

The result of revising process of the image contained a tilt face pose is presented Fig. 5. As shown this Figure, the angle of facial principle axis is corrected -43° automatically by Eq. (4)~Eq. (6).

Fig. 5. Slanting face and result of automatic correction.
(a) Original image, (b) Region of face candidate, (c) Image of the pose correction.

2.3 Detection of Facial Features using Geometric Features of Face

2.3.1 Geometric features of face

The face, with elliptical shape, composed of eye, nose, mouth, and ear. The each ratio of all facial elements will be able to apply to the general person. Especially two eyes and a mouth are horizontal when the face puts straight.

Also, if we draw line from the center of eyes to the chin then we comes across the middle location of the mouth. Other relationships of features among the face structural elements are shown Fig. 6.

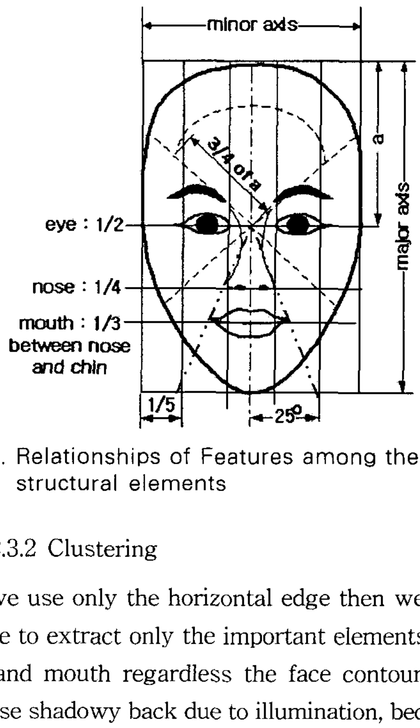
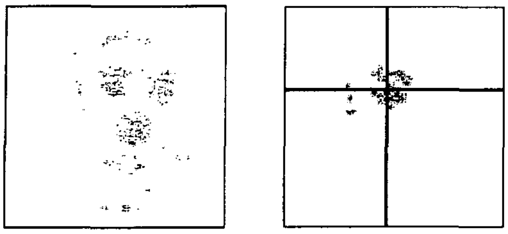


Fig. 6. Relationships of Features among the face structural elements

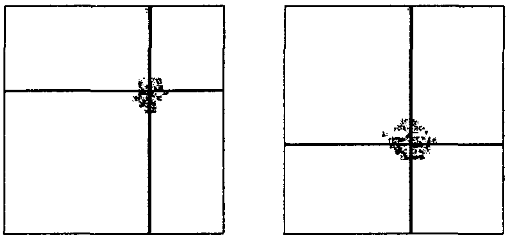
2.3.2 Clustering

If we use only the horizontal edge then we will be able to extract only the important elements like eyes and mouth regardless the face contour and the nose shadowy back due to illumination, because the principle elements such as eyes, eyebrows, and a mouth are formed of horizontality.

After getting the horizontal edge (see Fig. 7.a) from face candidacy region, in order to detect the feature location (see Fig. 7.b-d) of the face, this paper uses K-means algorithm and Iterative Self-Organizing Data Analysis Techniques A (ISODATA) algorithm[6].



(a) Horizontal edge image (b) Clustering right eye



(c) Clustering left eye (d) Clustering mouth region

Fig. 7. Processing results of k-means algorithm and ISODATA algorithm;

2.3.3 Detection of facial features

After clustering, it readjusts the location of the eyes by Eq. (7).

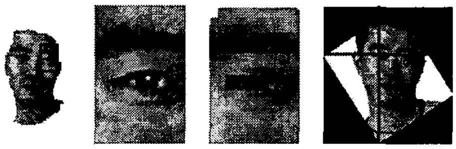
$$New_{eye}(x,y) = \arg\min f_B(x,y)$$
$$,where \quad Old_{eye}(x) - T_s \le x \le Old_{eye}(x) + T_s$$
$$Old_{eye}(y) - T_c \le y \le Old_{eye}(y) + T_c \qquad (7)$$

- $New_{eye}(x,y)$: New coordinates of readjusted eye
- $f_B(x,y)$: Result image of 3×3 block mean processing in gray scale image of candidacy region
- $Old_{eye}(x)$: The value of vertical axis of center coordinate clustered
- $Old_{eye}(y)$: The value of horizontal axis of center coordinate clustered
- T_c : The distance of Partitioning while ISODATA algorithm is working
- T_s : The distance of merging while ISODATA algorithm is working

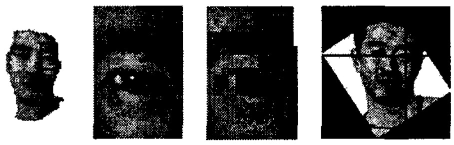
Because the location is get from the horizontal edges of eyebrow and eye, the location of detected eye, which was selected the center of cluster, is not actual position. Also, in the case of person in glasses, the errors of estimation may occur by the rim of spectacles. So the readjusting the location of the eyes is needed.

T_c and T_s are set up automatically during the clustering, and reused in the process of readjusting to extract eye part as shown Fig. 8.

After readjusting the eye position, if it draws down perpendicularly from center coordinate of inter-eye then it gets location of mouth. At the same time, even if feature point of mouth is not located exactly, we judge the mouth region if the feature point of mouth is located in scope which is the range from the perpendicular line to left and right 25° respectively as Fig. 6.



(a) Readjusting right eye



(b) Readjusting left eye

Fig. 8. Images of readjusted eyes.

3. EXPERIMENTAL RESULTS AND DISCUSSION

It is difficult to evaluate the performance of different methods because different benchmark data sets and criteria are applied. In our method, an evaluation for detecting the facial feature has been conducted for 90 images with 126 faces from different sources which have various backgrounds and face poses.

Some examples are shown in Fig. 9. In each image, (a)-(f) images have various face poses, but the positions of eyes and mouth were detected, also in (h) image though the hand and face segmented in one region we detected the features through readjusting the eyes. In the case of (i) image,

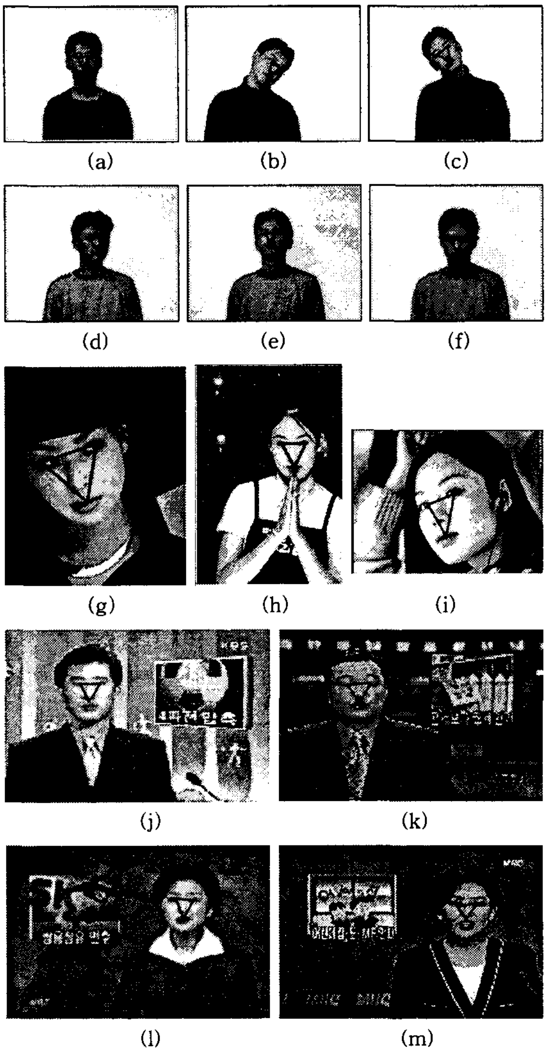


Fig. 9. Results of facial features detection. (a)-(f) Various face poses, (g)-(i) TV/drama photos, (j)-(m) News frame shots.

although background contains many skin-tone colors, the proposed algorithm separated the face part from the arm part using process of labeling with the Laplacian edge, and detected eyes and mouth accurately. Like above, the facial features were detected in (j)-(m) images in spite of noisy frame and various backgrounds.

As shown the test results we could find the features using correction of tilt angle in images contained various face poses, and also separate face and background using edge information of YCbCr in case that the background is adjacent to face with same skin-tone color. Moreover facial features are detected regardless of scale.

Total rate of detection is 91.1% as shown Table 2, and the rate came out high when the device is digital camera or the faces are the front like mug-shot. The false images are the case as follows. Fist, the one part of face is damaged during a segmentation of face candidacy region and secondary, the mistake occurs in clustering by noise.

Table 2. Total experimental results

Classification	Number of total images	Correctly detection	False	Detection rate (%)
Snap shots / News	42	39	3	92.8
Movie / TV drama	48	43	5	89.6
Total	90	82	8	91.1

4. CONCLUSION AND FUTURE WORK

This study overcomes the defects of existing methods by setting the candidacy regions using difference value of RGB in order to eliminate non-face component in the scope of normal distribution function. Moreover, through using the information of color and the Laplacian edge of YCbCr simultaneously, the proposed method can more efficiently separate the face regions from the backgrounds with similar face color than the existing methods. In the detection stage of facial features, the detection of facial features accurately by correcting a tilt face. This idea comes from the principle that usually the eyes and mouth contained many horizontal edges. Further more, after finding positions of features by K-means and ISODATA algorithm, it can search the position of eyes accurately through readjusting positions of features.

Proposed method complements the weak points of existing works that must fix the position of face at the frame of camera. So it is more effectively used in application field such as observation and identity-confirm system that the person does not know whether he get taken.

As an extension of this work, we will study the recognition of individual using fuzzy or neural network.

5. REFERENCES

[1] M-H. Yang and N. Ahuja, "Detecting Human Faces in Color Images," IEEE Proc. Image Processing, pp.127-139, 1998.

[2] K. Sobottka and I. Pitas, "Extraction of Facial Regions and Features using Color and Shape Information," IEEE Proc. Pattern Recognition, vol, pp.C421-C425, 1996.

[3] D. Chai K.N. Ngan, "Locating facial region of a head-and-shoulders color image," IEEE Proc. Automatic Face and Gesture Recognition, pp127-129, 1998.

[4] R. Crane, *A Simplified approach to Image Processing*, Prentice Hall, pp.19-20, 1997.

[5] E. Saber. A. M. Tekalp. and G. Bozdagi, "Fusion of color and edge information for improved segmentation adn edge linking," in Int. Conf. Acoustics, Speech, and Signal Processing, 1996.

[6] J. Tou and R. Gonzalez, *Pattern Recognition Principles*, Addition Wesley, N.Y., 1974.

Jae-Young Choi

received the B.S. degree in mathematics from Kyungwon University, Korea, in 1995 and the M.S. degrees in Computer Science from Kyungwon University, in 1999. He is now working toward to Ph.D. degree in Computer Science at Kyungwon University. His current research interests include image processing, multimedia, pattern recognition, and computer vision.

Nak-Bin Kim

received the B.S. degree, M.S. degree, and the Ph.D. degree in Computer Engineering from Korea University, Korea, in 1976, 1982, and 1985, respectively. He was a research professor at the Carnegie Mellon University, USA in 1996-1997. He is currently a Professor of Software College, Kyungwon University, Korea. His primary research interests include computer architecture, image processing, and multimedia processing.

For information of this article, please send e-mail to: nbkim@kyungwon.ac.kr(Nak-Bin Kim)