Extraction of Facial Features from Color Images

Miloš ORAVEC, Branislav KRIŠTOF, Michal KOLÁRIK, Jarmila PAVLOVIČOVÁ

Dept. of Telecommunications, Faculty of Electrical Engineering and Information Technology,
Slovak University of Technology, Ilkovičova 3, 812 19 Bratislava, Slovakia

milos.oravec@stuba.sk, jarmila.pavlovicova@stuba.sk, branislav.kristof@gmail.com

Abstract. *In this paper, a method for localization and extraction of faces and characteristic facial features such as eyes, mouth and face boundaries from color image data is proposed. This approach exploits color properties of human skin to localize image regions – face candidates. The facial features extraction is performed only on preselected face-candidate regions. Likewise, for eyes and mouth localization color information and local contrast around eyes are used. The ellipse of face boundary is determined using gradient image and Hough transform. Algorithm was tested on image database Feret.*

Keywords

Skin color segmentation, face detection, facial features extraction, morphological operations, ellipse detection, Hough transform, Feret database.

1. Introduction

The detection and localization of human faces in images [1-5] plays an important role in applications such as video surveillance, human – computer interface, automated tracking of the object (human) position, face recognition, face image database management and many others.

Most of existing skin segmentation approaches is based on skin color. Color information is fundamental for localization of human faces in still images with complex background. Skin regions are detected by searching for pixels they have skin color. The presence of skin color in the input image is detected by applying a skin color model based on different color space models. In the HSV model, it is possible to separate chrominance components from intensity component. The intensity component has no significant influence to final color appearance in the image. A lot of skin color could be found in regions, they belong to background or to other non-face objects, but they have the color similar to skin. This fact results in certain level of false localization of faces. To reduce the false face localization we analyze face aspect ratio and position. Thus non-face regions of abnormal shape like arms or background objects of preselected color are not taken into account. Remaining skin color regions are in the next processing considered as human face candidates. Gray scale morphological operations and information about selected colors appearance are applied to localize face features – eyes and mouth [8]. For face ellipse localization the Hough transform [10], [11] is used. Success of localization depends on various conditions such illumination, skin color tone, mouth color tone, presence of glasses, moustache, beard and others. We applied our localization system to face images of standard database FERET [9] and evaluated the achieved results.

Algorithm was implemented via MS Visual C++ 6.0 using CImg library, with features for image processing.

2. Skin Color Segmentation

As input images, we use color face images of standard database FERET [9]. First step of processing contains localization of face-candidate regions and creation of face-mask. Segmented face areas are marked and visualized by the minimum circumscribing rectangle. Then localization of eyes, mouth and face boundaries works only through each face-candidate region and corresponding face-mask.

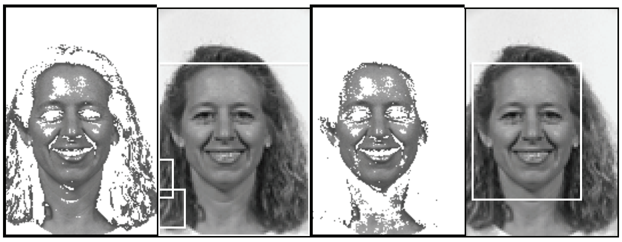
2.1 Skin Color Detection

The color image is divided to three essential planes: achromatic, transitional and chromatic one depending on hue, saturation and intensity values.

$$\begin{array}{lr} \text{Achromatic} & S \le 10\%,\ V \le 20\% \\ \text{Transitional} & 10\% < S < 20\%,\ V > 20\% \\ \text{Chromatic} & S \ge 20\%,\ V > 20\% \end{array} \qquad (1)$$

For testing of skin color presence in chromatic plane we used a modified criterion based on [6]. Our modification eliminated saturated red color from the skin model. The application of the criterion leads to better results of skin color segmentation. Skin color criterion is used as follows:

$$\begin{aligned} &T_{\text{hue1}} = 340° \le H \le T_{\text{hue2}} = 350° \\ &T_{\text{hue3}} = 0° \quad \le H \le T_{\text{hue4}} = 30° \\ &T_{\text{sat1}} = 20\% \quad \le S \le T_{\text{sat2}} = 90\% \\ &V \ge T_{\text{val}} = 35\% \end{aligned} \qquad (2)$$



a) b) c) d)

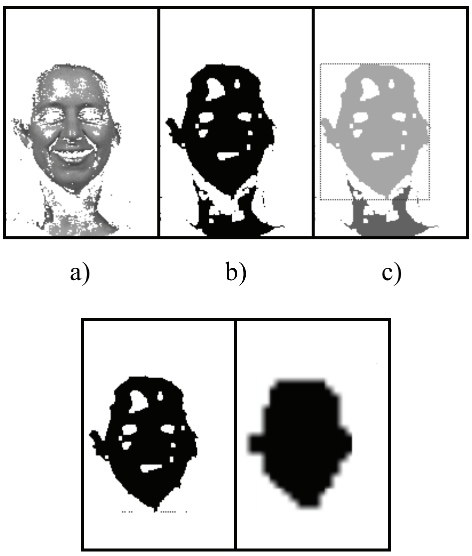
Fig. 1. Example of application of the original and modified criterion for skin color segmentation, a) skin color by original criterion, b) output image with marked face candidate regions, c) skin color by modified criterion (2), d) output image with marked face candidate regions after application of modified criterion.

2.2 Face Region Candidates

Now we have one or more rectangles in the image, they contain regions of skin color. After the skin color segmentation step the binary image of skin color and non-skin color areas is created. Then, different color is assigned to each closed region of skin. The regions which aspect ratio does not correspond to the face dimensions are eliminated [8]. Remaining regions are marked as face candidates. They are then processed further.

2.3 Face-mask

Found skin regions in the image contain many discontinuities and small gaps. It is useful to smooth them out. The smoothed face-mask is created using binary morphological adjustment of skin color segment in each face candidate region. The procedure is illustrated in Fig. 2.



a) b) c)

d) e)

Fig. 2. Face-mask creation: a) segmented skin color, b) binary skin color map, c) skin color segments distinguished by color, d) binary image of the face candidate, e) face mask after binary morphological processing (opening and closing).

3. Facial Features Extraction

In the image areas denoted by the face masks, algorithms for facial features extraction are applied. We focused our attention to two of them, eyes and mouth. Both types of features are represented by typical colors. Because of that, we utilize color information in the YCbCr color-space.

3.1 Eyes Localization

Human eyes contain white color that contrasts with the surrounding face. To extract eye regions, both components – luminance and chrominance can be used. We create the final eyemap as a combination of two fundamental eye maps. One map is designed using the luminance component by use of greyscale morphology and the second map is designed using the chrominance components based on color properties [4].

Since eyes contain both dark and light pixels, greyscale morphological operations (dilation, erosion) [7] can be used to highlight these pixels around eyes in luminance component. Greyscale dilation and erosion is applied using hemispherical structural function (Fig. 3).

Dilation of function $f(x)$ by structural function $g(x)$ is denoted $(f \oplus g)(x)$ and defined as

$$(f \oplus g)(x) = \sup_{t \in G \cap \check{D}_{-x}} \{f(x-t)+g(t)\} \tag{4}$$

Erosion of function $f(x)$ by structural function $g(x)$ is denoted $(f \Theta g)(x)$ and defined as

$$(f \Theta g)(x) = \inf_{t \in G \cap D_{-x}} \{f(x+t)-g(t)\} \tag{5}$$

where $f\!:\boldsymbol{D} \subset \boldsymbol{R}^n \rightarrow \boldsymbol{R}$

$g\!:\boldsymbol{G} \subset \boldsymbol{R}^n \rightarrow \boldsymbol{R}$

$D_x = \{x+t : t \in \mathrm{D}\}$

$\check{D} = \{x : -x \in \mathrm{D}\}$

sup(f) is supremum (the highest value) of function f and inf(f) is infimum (the lowest value) of function f. For discrete functions, when function f is a finite set of points, there are max{} and min{} used instead of sup(f) and inf(f).

Hemispherical structural function is constructed as follows

$$g_\sigma(x) = |\sigma| \left((1-||x/\sigma||^2)^{1/2}-1\right) \quad ||x|| \le \sigma \tag{6}$$

where σ is the measure parameter.

Construction of eye map from luminance component:

$$EyeMapL = \frac{Y(x,y) \oplus g_\sigma(x,y)}{Y(x,y) - g_\sigma(x,y) + 1}\,. \tag{7}$$

Hemispherical structural function is implemented

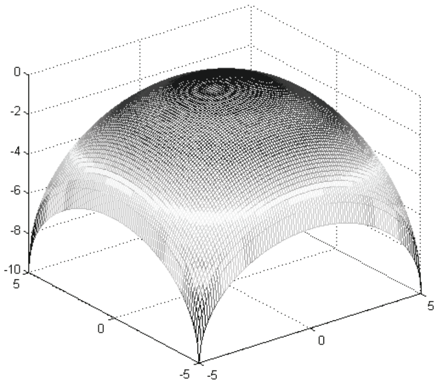


Fig. 3. Hemispherical structural element for morphological opening and closing.

according to (6), where the radius of hemisphere $\sigma = 3$ in our experiments. The radius was chosen experimentally with regard to eye proportion. Face candidate region is resized proportionally to the size of 100 pixels width. For the accurate eye segments a structural function diameter should not be greater than an eye diameter.

Fig. 4. *EyeMapL* – Eye map computed from the luminance component.

3.2 Eye Map from Chrominance Component

The eye map creation from the chrominance component is based on the fact, that there are usually high values of blue chrominance component *Cb* and relatively low values of red chrominance component *Cr* around eyes.

The map is constructed as follows:

$$EyeMapC = \frac{1}{3}\left\{(Cb)^2 + \left(\hat{C}r\right)^2 + \left(\frac{Cb}{Cr}\right)\right\} \qquad (8)$$

where $\hat{C}r$ is negative *Cr*, i.e. $\hat{C}r = 255 - Cr$ Cb^2, $(\hat{C}r)^2$ and *Cb/Cr* are normalized to the range [0,255].

Fig. 5. Map of eyes selected from chrominance components *EyeMapC*.

3.3 Mouth Map

The procedure of mouth map construction is based on the fact that the color in mouth region consists of stronger red component and weaker blue component than other parts of the face. It means there are higher values of chrominance component *Cr* than chrominance component *Cb* in the mouth region. In the mouth region, we achieve also a relatively low response of *Cr/Cb* characteristic, but a high value of Cr^2.

$$MouthMap = Cr^2\left(Cr^2 - \eta\frac{Cr}{Cb}\right)^2, \qquad (9)$$

$$\eta = 0.95\left(\frac{\frac{1}{n}\sum Cr(x,y)^2}{\frac{1}{n}\sum\frac{Cr(x,y)}{Cb(x,y)}}\right)$$

where Cr^2 and *Cr/Cb* are normalized to the range [0,255] and *n* is the number of pixels of face region (in our case the number of pixels of face-mask). Parameter η is a ratio of Cr^2 average to *Cr/Cb* average.

a) b)

c) d)

Fig. 6. Mouth map creation using face candidate region from(Fig. 1.(d), a) component Cr^2, b) *Cr/Cb*, c) mouth map, d) resulting thresholded mouth map.

3.4 Eyes-Mouth Triangle Assignation

The face mask is applied both to the eye map obtained from luminance and the eye map obtained from chrominance and the maps are graphically adjusted in order to achieve a better localization result. Subsequently, eye maps are multiplicatively combined:

$$EyeMap = (EyeMapC)\text{AND}(EyeMapL) \qquad (10)$$

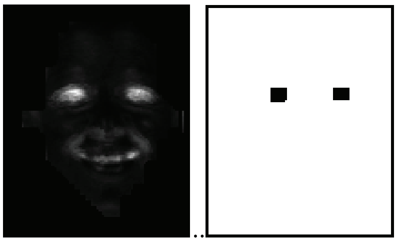


Fig. 7. Final *EyeMap and* Final binary eye map.

Finally, eye map is then normalized and thresholded by ISODATA (Iterative Self-Organizing Data Analysis Techniques). For elimination of small ineligible segments, binary morphological operations (erosion, dilation, opening) are applied. In this way, the final binary eyemap is created, which is shown in Fig. 7.

Analogous to eye map construction, mouth map is thresholded by ISODATA algorithm and adjusted by binary morphological erosion and dilation.

Centers of segments on both binary maps are linked in order to create the eye-mouth triangle shown in Fig. 12.

4. Face Boundary Localization

Human face has approximately an elliptical shape, hence we can use Hough transform for ellipse detection to find face boundaries. Gradient image as an output of certain edge detector is used as input image for Hough transform.

Fig. 8. Original picture and gradient image using Sobel edge detector.

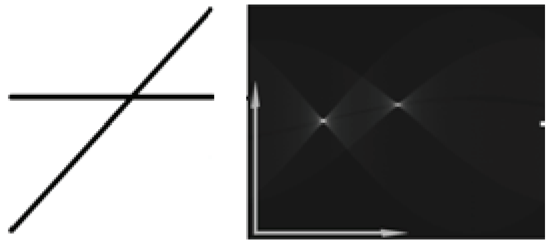
4.1 Hough Transform

Hough transform enables to determine position and proportion of any pattern of a known shape from its contours. The method is based on the transformation between a picture space (x,y) and a parameter space. For the detection of straight lines in parametric Hough space, we use the following equation

$$r = x \cdot \cos\theta + y \cdot \sin\theta \qquad (10)$$

where $\theta \in [0, 2\pi]$ and $r \geq 0$ are two parameters of a line in a 2D parametric space.

While an equation of an ellipse has five parameters, (two center coordinates, lengths of half-axes and angle) we should construct 5D parameter space to find the solution. It would be computationally very expensive.



a) b)

Fig. 9. Illustration of Hough transform. Input image (a) and Hough parametric space (b) Lightest positions in Hough space represent parameters of detected lines.

If we take advantage of symmetry of an ellipse, we can reduce the problem of ellipse detection to the problem of lines detection [11]. We look for two perpendicular lines of symmetry of an ellipse. In other words, we test each pair of points in a predefined distance in edge image for symmetry. Because the ellipse is symmetric around its half-axes, the most pairs of points should be symmetrical right around these axes.

In Hough space, we can define parameters s=r and θ, which represent the line of symmetry of a pair of points:

$$s = xs_{\mathrm{m}} \cos\theta + y_m \sin\theta \qquad (11)$$

where x_m, y_m are coordinates of the point in the center of a pair of points. In parametric space, θ and s are defined as follows:

$$\theta = \arctan\left(\frac{y_2 - y_1}{x_2 - x_1}\right), \qquad (12)$$

$$s = \frac{\cos\theta \cdot (x_1 + x_2) + \sin\theta \cdot (y_1 + y_2)}{2}.$$

Each pair of points, that is symmetrical around half-axis, has identical values of s and θ. The most frequently occurred parameters correspond to the main half-axis. The adjacent axis of ellipse has to be perpendicular to the main axis and at the same time it should be the line of symmetry of the investigated ellipse. The point of the intersection of these axes designates the ellipse center and the angle of main axis relates to the tilt of the ellipse with respect to x-axis in image space.

In our implemented algorithm, only pairs of points with predefined distance higher than D_{min} are calculated. It has two reasons:

- Computation of points lying too close is sensitive to noise.
- Lower computation requirements.

Localization of the face boundary ellipse is applied only over previously localized face candidates; hence the algorithm searches always for only one ellipse at a time. An example of the presented process is shown in Fig. 10.

Ellipse detection can be inaccurate on two or more overlapped faces. The result is also strongly affected by a picture size. The same picture with a different resolution can lead to a different result. An example is shown in Fig. 11.

Fig. 10. Localized face boundary ellipse.

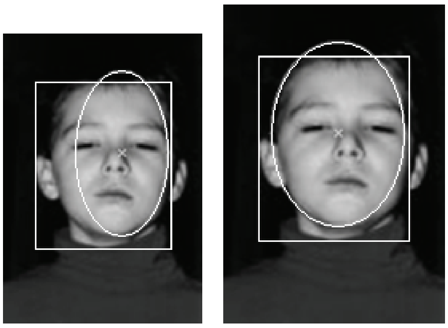


Fig. 11. Example of face boundary localization results dependent on image size.

5. Face Database

For our experiments, we used Color FERET (The Facial Recognition Technology) face database [9]. The so called FERET program runs from 1993 and it is sponsored by Department of Defense (DoD) Counterdrug Development Technology Program through the Defense Advanced Research Products Agency (DARPA). Color FERET database was released on October 31, 2003 and it contains 11338 facial images collected by photographing of 994 subjects at various angles.

We randomly chose 300 images of 165 different faces containing every type of face (European, African and Asian, women and men) at various head orientation, faces wearing glasses and having moustache or beard.

6. Results and Conclusion

The proposed method results from conjunction of modified methods for human face extraction and facial features localization. In Tab. 1, we present our results of face localization based on color. Algorithm localized in the first step (skin color detection) all faces correctly, but other 118 nonface objects marked also as faces. In the second step, localization of eyes and mouth were performed only in face candidate regions. Eye localization was successful for the most test images and occasionally failed (particularly for faces wearing glasses). False negative eyes were found on side faces with an exposed ear. It is due to similarity of eyemaps where an ear appears such as an eye. Faces with strong red pigment or faces with pale color of lips may be problematic for mouth detection algorithm because of likeness of surroundings. Results of our analysis of correct face localization are shown in Tab. 2. Results were evaluated for each kind of face separately. The relative number of each face type in tested set is also shown.

The presented method provides good results in facial features localization of non frontal face images and also on faces with glasses. Face and facial features localization based on color is very fast with low computational requirements.

Face kind	Correct localization		False localization	
European	247	300	106	118
African	26		3	
Asian	27		9	

Tab. 1. Results of face localization based exclusively on color information. All 300 faces were successfully localized and 118 nonface regions were localized as faces.

Face kind		Quantity %	Successful localization %		
European	male	57	80,88	80,51	80,73
	female	25,30	79,68		
African	male	5	60	70,05	
	female	3,70	83,64		
Asian	male	5,70	95,88	96,22	
	female	0,70	99		
Indian	male	2,33	80,86	83,25	
	female	0,33	100		
wearing glasses		6,33	59,50		
faces with beard		9,33	35,40		

Tab. 2. Results of eye–mouth triangle localization. 80.73% eye–mouth triangles were localized correctly over our test set containing 300 randomly chosen faces. Partial results for different face type are also shown. No triangle was found in first step's false localized face regions.

Acknowledgements

The research described in the paper was financially supported by the Slovak Grant Agency VEGA under grant No. 1/3117/06.

References

[1] KAKUMANU, P., MAKROGIANNIS, S., BOURBAKIS, N. A survey of skin-color modeling and detection methods. *Pattern Recognition*, 2007, vol. 40, p. 1106 – 1122.

[2] YANG, M. H., KRIEGMAN, D. E., AHUJA, N. Detecting Faces in Images: A Survey. *IEEE Tranactions on Pattern Analysis and Machine Intelligence*, 2002, vol. 24, no. 1, p. 34-58.

Fig. 12. Example results of localization at some different kinds of faces.

[3] SANDEEP, K., RAJAGOPALAN, A. N. Human face detection in cluttered color images using skin color and edge information. In *Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing* (ICVGIP'02). Ahmedabad (Dec. India), 2002, p. 230-235.

[4] REIN-LIEN, HSU K., MOHAMED, ABDEL-MOTTALEB, ANIL K. JAIN. Face detection in color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2002, vol. 24, p. 696 to 706.

[5] PHUNG, S. L., BOUZERDOUM, A., CHAI, D. Skin segmentation using color pixel classification: analysis and comparison. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 2005, vol. 27, no. 1, p. 148-154.

[6] HERODOTOU, N., PLATATONIS, K. N., VENETSANOPOULOS, A. N. *Image Processing Techniques for Multimedia Processing*. CRC Press, Boca Raton, Florida, 2000, p. 97-127, ISBN 0-8493-3492-6.

[7] JACKWAY, P. T., DERICHE, M. Scale-space properties of the multiscale morphological dilation-erosion. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 1996, vol. 18, p. 38-51.

[8] WAQAR, M., NOMAN, A., CHUNG-TSE MAR Face Detection Project, 2003, http://www.stanford.edu/~wmohsin/ee368group10.pdf

[9] FERET Database, http://www.itl.nist.gov/iad/humanid/color-feret/home.html, NIST, 2003.

[10] FISCHER, R., PERKINS, S., WALKER, A., WOLFART, E. Hough Transform, University of Edinburgh UK 2003, http://homepages.inf.ed.ac.uk/rbf/HIPR2/hough.htm.

[11] LEI, Y., WONG, K. C. Ellipse detection based on symmetry. Nanyang Technological University, Singapore, *Pattern Recognition Letters* 20 1999, p. 41-47.

About Authors...

Miloš ORAVEC received the MSc., PhD. and Assoc. Prof. degrees in Telecommunication Engineering from the FEI SUT in Bratislava in 1990, 1996 and 2002, respectively. He is with the Department of Telecommunications FEI SUT Bratislava. He is a member of the IET. His research interests include image processing, neural networks and communication networks.

Branislav KRIŠTOF received the MSc. degree in Telecommunication Engineering from the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology (FEI SUT) in Bratislava in 2007. Now he is the PhD student at the same faculty.

Jarmila PAVLOVIČOVÁ received the MSc., PhD. and Assoc. Prof. degrees in Telecommunication Engineering from the FEI SUT in Bratislava in 1986, 2002 and 2006 respectively. She is with the Department of Telecommunications of FEI SUT Bratislava. Her research interests include image processing, especially image segmentation.

Michal KOLÁRIK received the MSc. degree in Telecommunication Engineering from the Faculty of Electrical Engineering and Information Technology, Slovak University of Technology (FEI SUT) in Bratislava in 2007.