Automatic Face Detection in Frontal Color Images

Ognian Boumbarov[1] and Strahil Sokolov[2]

Abstract - **In this work we propose an approach to automatic detection of faces in color images. The proposed approach consist these parts: color space selection, human skin-color modeling, and human skin-color segmentation to identify probable regions corresponding to human faces and facial feature extraction.**

Keywords - **Skin color model, Human skin segmentation, Facial features extraction.**

I. Introduction

Face tracking and face recognition are basic needs for many fields including surveillance and security, human-machine interfaces, object-based video coding, virtual reality, and automatic 3-D modeling. Although face tracking and recognition is a fundamental part in a fully automated facial analysis system, the first important step is detecting faces in static image and video sequences. There are many variations of image appearance such as pose variation (front and profile), occlusion, image orientation, lighting condition and facial expression [1-2].

In this work we propose an approach to automatic detection of faces in color images. The proposed approach consist of three parts: *color space selection* and *human skin modeling*, *human skin segmentation* to identify probable regions corresponding to human faces and *facial feature extraction*.

The paper is organized as follows. In the following Section the choice of suitable skin color space is described. In Section 3 the use of lighting compensation and single Gaussian modeling for the modeling human of skin is discussed. The face skin color segmentation is given in Section 4. The human skin segmentation employs a model-based approach to represent and differentiate the background colors and skin colors. The final facial feature detection is presented in Section 5. Experimental results evaluating the performance of the algorithm are given in Section 6.

II. Mathematical description

Color space selection

Color information is an important feature of human faces. Using skin color as a feature for detecting face skin regions has several advantages. In particular, processing color is faster than processing other facial features. The more so as color information is invariant to face geometrical transformations. However, even under a fixed ambient lighting, people have different skin color appearance.

[1]Ognian Boumbarov is with Faculty of Communication Technics and Technology, Technical University of Sofia, Kl. Ohridski 8, 1000 Sofia, Bulgaria, E-mail: olb@tu-sofia.bg

[2]Strahil Sokolov is with the Faculty of German Engineering Education and Industrial Management (FDIBA), Technical University of Sofia, Kl.Ohridski 8, 1000 Sofia, Bulgaria, E-mail: strahil@gmx.net

Furthermore, to successfully use skin color for face detection, we need to choose color space, in which human skin colors cluster tightly together and cluster in different color spaces. Different color spaces tend to enhance characteristics of images at the expense of reside remotely to background color[1-2]. Color space is a method of color information obtained from a static image and video sequences. Human skin color tends to others[1],[5],[6]. Some color spaces are discussed in this paper.

The RGB color space (or the normalized color space RGB) is one of the most widely used color spaces for processing and storing of digital image data [2],[5],[6].

However, high correlation between the color components, mixing of the chrominance and the luminance data make RGB color space an unfavorable choice for color analysis and recognition.

In the YCbCr color space, color is represented by luma, constructed as a weighting sum of RGB values, and two color difference values Cb and Cr that are formed by subtracting luma from RGB red and blue components [5 - 6]. The transformation simplicity and explicit separation of luminance and chrominance components make this color space very attractive for skin color modeling.

The family of color spaces HSI (V, L) – Hue, Saturation, Intensity (Value, Lightness) was introduced when there was a need for the user to specify color properties numerically[2]. *Hue* defines the dominant color of an area; *saturation* measures the colorfulness of an area in proportion to its brightness. The “intensity”, “lightness” or “value” is related to the color luminance. The explicit discrimination between luminance and chrominance components made this color spaces popular for skin color segmentation. Besides YCbCr color space, several other linear transform of the RGB color space was employed for human skin detection – YES, YUV and YIQ [2], [8], [9].

Many works on face skin detection discard the luminance component of the color space. This decision seems logical, because this is a dimensionality reduction. The goal of any color-based approach is diminishing the influence of the lighting conditions. The chrominance-only color analysis will make the approach partially independent from the lighting conditions.

We choose YCbCr as the processing color space since it is perceptually uniform and separate the luminance and the chrominance components. In our implementation only the CbCr components are used to model the distribution of skin colors. The YCbCr space is of particular interest because it is widely used in still-images and video coding standards such as JPEG, MPEG, and H.263.

Modeling human skin

The purpose of skin color modeling is to build a decision rule that will discriminate between skin and not-skin pixels. The techniques for skin color modeling in YCbCr space can be classified into following categories: parametric, non-parametric, and semi-parametric. A parametric skin color model has a specific functional form (single or mixture Gaussian) [1].[2],[10] with adjustable parameters chosen to fit the model to the input data. A non-parametric model does assume any particular form (e.g. histogram thresholding). A semi- parametric approach applies a very general form with adaptive parameters systematically varied in number as wellas in value in order to create flexible models [11].

Our approach employs a two-dimensional normalized Gaussian probability distribution function for modeling skin color in Cb-Cr space.

Let P_{skin} (C) be the probability of a pixel with chrominace vector C = $[Cb,Cr]^t$ belonging to skin color class. We can shape the distribution of the chrominance components in a Cb-Cr skin space as follows [14]:

$$\text{Pskin (C)} = \frac{1}{2\pi\sqrt{Det\{K\}}}\exp\left[-\frac{(\vec{C}-\vec{\mu})^t K^{-1}(\vec{C}-\vec{\mu})}{2}\right]$$

where $\mu=[\mu_b, \mu_r]^t$ and $[K]=Cov(C, \mu)$ respectively are the mean and covariance matrix for the Cb-Cr skin space. The mean and covariance matrix are estimated from a training set. The chrominace vector C= $[Cb,Cr]^t$ is calculated for each color pixel (Cb(i,j), Cr(i,j)) at spatial position (i,j) from the two color components Cb and Cr of the color space YCbCr. The result obtained by skin color modeling with the normalized Gaussian probability distribution function is displayed on **Fig.3** as normalized grayscale map of expected skin regions.

Skin color segmentation

In order to detect a skin color region we must map the skin pixels into a region. Generating of binary skin/non-skin maps strongly depends on lighting conditions and the tuning of the camera. This is supported by the histograms of the normalized grayscale maps of expected skin regions, shown on **Fig.4.**

The abovementioned two facts impose that the skin-color segmentation is implemented via adaptive thresholding. In addition, there are 5 skin-types: white, black, yellow, brown, and red.

In the course of our work were researched various methods of automatic adaptive thresholding for the purpose of intensity segmentation of the grayscale maps of the expected face regions: 1D Fuzzy, 2D Fuzzy, NIR, QIR, and Otsu. Best results were generated by the Otsu-method, as shown on **Fig.5**.

Facial features extraction

In our approach a geometrical model of face representation is chosen, described in [11]. The purpose of this algorithm is automation of the process of feature extraction from full-face images and their representation as multidimensional vector.

There are various methods used for feature extraction. One of them is the method of adjustment of a flexible template [11, 12, 13]. According to this method, the form of the facial elements is approximated with a set of geometrical shapes, whose equations are known. In this way the eyes and the mouth are easily described. Their templates consist of parabolas representing their outlines, and the iris is approximated by a circumference. On the other hand, the eyebrows, the outline of the face, the hair and the nose, cannot be approximated by simple templates, due to their complex shapes. In the described face model are used extraction and analysis only of theeyes and mouth, using the algorithm of face region segmentation using skin color, described in the previous section. The resulting binary map of the expected skin regions of the face (**Fig.5**) considerably narrows the region of search (region of making a decision) and increases the accuracy of extraction of the facial features. Limiting the zones of interest in an input image speeds up the feature-extraction process. The approach for facial features extraction can work separately in recognition of full-face frontal images.

The algorithm for automated extraction of the facial features “eyes” and “mouth” comprises the following basic steps:

- Finding the eyes, or more precisely drawing their surrounding rectangles;
- Estimation of the mouth location by its surrounding rectangle;
- Precise adjustment of the form and location of both eyes and the mouth by setting up flexible templates;
- Extraction of the vector with the features, saving it to a file, designed for neural-network recognition

III. EXPERIMENTAL RESULTS

Photos are usually taken under various lighting conditions. This is the reason why they have variations in quality, color, position, pose and facial expression. Our algorithm has been evaluated on various images from the World Wide Web. We present the evaluation on a set of images from our local database.The experiment took place in the following steps. First, a training set of ten skin samples is loaded, shown on **Fig.1**. Next, the sequence of original images is shown on **Fig.2.** A grayscale map of the expected skin region, shown on **Fig.3,** is generated with the normalized Gaussian probability distribution function. After finding the optimal threshold value with the algorithm of Otsu for every image we obtained a binary mask of the expected skin region, **Fig.5.** Using the binary mask we extract the expected skin region, ingnoring the non-skin regions of the image, as can be seen on **Fig.6.** Finally we use the limited region of the skin to search for facial features, with the algorithm described the previous section. The templates have been precisely adjusted to the eyes and mouth regions of every full-face frontal image, **Fig.7.**

The experiment was implemented on a system, running WindowsXP SP1with a 2.6 GHz Pentium 4. Both algorithms have shown good robustness and reasonable accuracy for the photos from our test sets. The whole facial feature set is detected automatically. The overall feature extraction time with the described algorithms is about 1-2 seconds per one photo on our system.

IV. CONCLUSION

We have presented a method for face detection in frontal images and feature extraction from a detected face region. First of all, our method generates a normalized Gaussian probability distribution of face region in the image.After generating an optimal mask, the face region is being extracted from the original image and is subjected to the algorithm for detection of facial features.The detected features are saved in a file, used for recognition by a neural network. Results for some photos of our test set ave been presented. Our approach is particularly suitable for embedding in surveillance and security systems. Our goal is to implement a system that detects faces and facial features and uses them as patterns for recognition in images and video streams.

ACKNOWLEDGEMENT

This work has been sponsored by the Project 311-7/2003 R & D Department, TU-Sofia.

REFERENCES

[1]. M.Yang, D. Kriegman, N. Ahuja, "Detecting Face in Image: A Survey, IEEE Trans. on PAMI, vol.1, pp. (34-58), January, 2002.

[2]. V.Vezhnevets, V. Sazonov, A. Andreeva, "A survey on pixel-based skin color detection techniques", Proc. Graphicon-2003, September 2003.

[3]. D. Mayo, D. Miltonic, Real-time face location o gray-scale static images, Pattern Recognition, vol.33, no.9, pp.1525-1539, September 2000.

[4]. H. Rowley, S. Beluga, T. Canada, Neural Networks based face detection, IEEE Trans. PAMI, vol.20, pp. 23-38, January 1998.

[5]. K. Sung, T. Poggie, Example-based learning for view-based human face detection, IEEE Trans. PAMI, vol.20, no.1, January 1998.

[6]. A., Jain, Fundamentals of digital image processing, Prentice – Hall, 1989.

[7]. W., Pratt, Digital I mage processing. AWiley-Interscience Publication. John Wiley&Sons, Inc. 2001.

[8]. E. Saber, A. Tekalp, Frontal-view face detection and facial features extraction using color, shape and symmetry based cost functions, Pattern Recognition Letters, vol.9, pp669-680, 1998.

[9]. F. Marques, V. Vilaplana, A morphological approach for segmentation and tracking of human faces, Int. Conf. on Pattern Recognition (ICPR'00), vol1, pp.5064-5068, 2000.

[10]. Rein-Lien Hsu, M. Abdel-Mottaleb, A. Jain, Face detection in color space, IEEE Trans. PAMI, vol.24, no.5, pp.696-706, 2002.S.Phung, D.Chai, A. Bouzerdoum, A universal and robust human skin color model using neural networks, IEEE, vol. pp.2844-2899, 2001.

[11]. O.Boumbarov, I.Ankov, Feature extraction for face recognition on frontal images, Proc. of International Scientific Conference on Energy and Information Systems and Technologies'2001, vol.2, pp.593-599, Bitola, Macedonia.

[12]. N. Roeder, X.Li, Accuracy Analysis for Facial Feature Detection, Pattern Recognition, vol.29, pp.143-157, 1999.

[13]. J.Deng, F.Lai, Region-Based Template Deformation and Masking for Eye-Feature Extraction and Description, Pattern Recognition, vol.30, pp.403-419, 1997.

[14]. Ed. A. D. Poularikas, The Handbook of Formulas and Tables for Signal Processing, CRC Press LLC, 1999.

Fig.1 Experimental training set of ten color images of faces

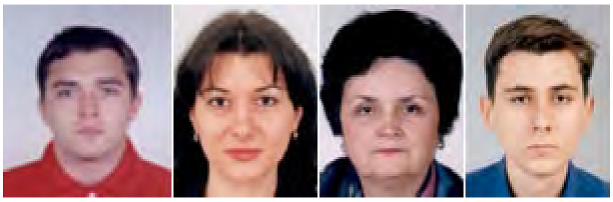

Fig.2 Original images of faces

Fig.3 Normalized grayscale maps of expected skin regions

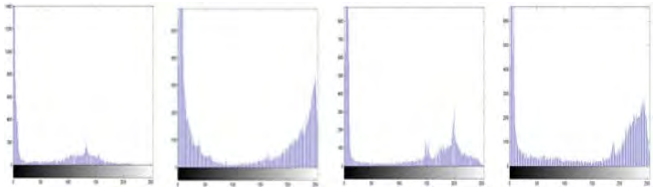

Fig.4.Histograms of normalized grayscale maps of expected skin regions

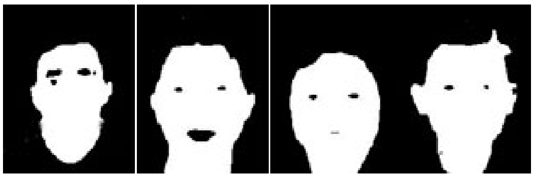

Fig.5 Binary map of expected skin regions

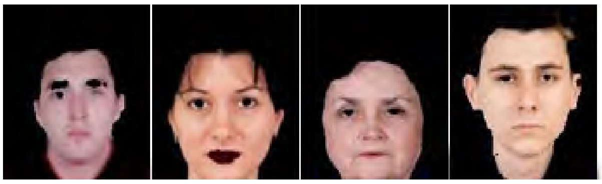

Fig.6 Colormap of expected skin regions

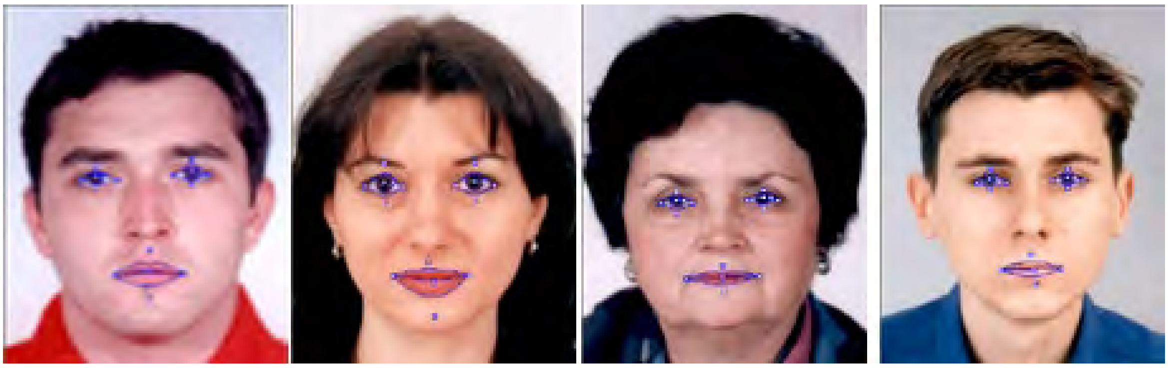

Fig.7 Templates adjusted to original image