FAST FACE SEGMENTATION IN COMPONENT COLOR SPACE

Juan José de Dios

JuanJose.deDios@uclm.es
E. U. Politécnica
Universidad de Castilla-La Mancha
Cuenca, Spain

Narciso García

narciso@gti.ssr.upm.es
E.T.S. Ing. Telecomunicación
Universidad Politécnica de Madrid
Madrid, Spain

ABSTRACT

A fast face segmentation color-based method is presented in this paper. The characterization of the decision thresholds along the skin color distribution, constituting a non-rectangular decision region for skin detection in the YCgCr color space is analyzed. Besides, a transformation based on the skin representation in the chrominance plane is proposed for skin detection improvement. Basically, it is based on the rotation of the Cg and Cr axis in the chrominance plane, so that a boundary box can be used for face segmentation. Segmentation tests have been performed in a set of portrait-like images with different lighting conditions. Two types of decision thresholds have been obtained for detecting the whole head or just the skin region. The face detection results achieved in the segmentation process have also been compared to those using the YCgCr decision thresholds.

1. INTRODUCTION

Recently, investigation has been increased on face segmentation and recognition in color images, focused in new emerging applications where people location is required, e.g. videoconference or video coding. The difficulty of face detection depends on the complexity of the image and its application [1]. Several face segmentation techniques which work with a normalized face have been proposed in the last years, but some of them require that its location is previously known [2].

The main advantage of using face detection schemes based on color information is that it can be segmented independently on its size and position within the image [3]. Considering that people with different skin color have major differences in their intensity than in its chrominance [3-6], and that human skin color is concentrated in a small region of the color space, color can been considered as a distinguishing and effective parameter for face detection.

Nevertheless, the classification into skin and non-skin regions will fail if faces are partially occluded or there are skin-like objects in the background of the image, so faces are not correctly detected [3,5]. Color-based segmentation can be combined with other techniques in order to improve the segmentation results, so that the undetected areas may be added to the detected face [3,5].

Different color spaces, such as RGB, HSV or YCbCr, and pixel-based skin detection methods have been proposed for the individual classification of the pixels in an image as skin or non-skin, independently on its neighbors [5-8]. Color spaces with separated luminance and chrominance components, like the last two, seem to be more appropriate for face detection [7].

YCgCr, a novel color space based on YCbCr, was proposed in [10] for face detection. The decision thresholds used for skin segmentation will be characterized in this paper. A transformation of this space is also presented in order to obtain a better performance of the face segmentation process.

2. YCgCr COLOR SPACE

The YCgCr color space was proposed specifically for analysis applications, mainly for face segmentation. It is based on YCbCr, but it differs on the use of the Cg color component instead of Cb The color spaces used in television systems (YUV, YCbCr) are transmission oriented, so, in order to minimize the encoding-decoding errors, they use the biggest color differences: (R-Y) and (B-Y). The YCgCr color space uses the smallest color difference (G-Y) instead of (B-Y).

Besides, skin detection can be performed ignoring the luminance coordinate, as skin tone is more controlled by the chrominance than luminance components [6]. Therefore, the color analysis is simplified by reducing the space dimensionality, defining a color model concentrated in a small region of the Cg-Cr plane.

The simplest model used for classifying each pixel as skin or non-skin individually is based on the definition of a chrominance bounding box [5]. A pixel will be considered as skin if both of its color components are within each of the ranges defined by the maximum and minimum thresholds of the chrominance plane coordinates (Cg and Cr).

The YCgCr components can be obtained in a similar way than the YCbCr equations described in the ITU Rec. BT.601 and expressed in terms of Y', G'-Y' and R'-Y' components defined in the [0,1] range using the following matrix expression:

0-7803-8554-3/04/$20.00 ©2004 IEEE.

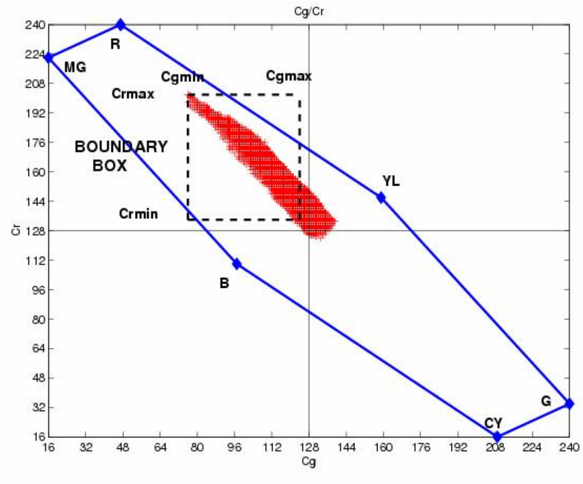

Fig. 1. Skin color pixels and boundary box in the Cg-Cr plane.

$$\begin{bmatrix} Y \\ Cg \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -81.085 & 112 & -30.915 \\ 112 & -93.768 & -18.214 \end{bmatrix} \cdot \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix}$$

The chrominance components are defined in the range [16,240], with excursions of ±112 and an offset of 128. Each component is coded in 8 bits.

Figure 1 represents the primary and complementary colors in the YCgCr color space. Between the red and yellow colors, the skin color region is also represented. Most of the pixels are concentrated in a small area of the 2nd quarter of the Cg-Cr plane, and distributed in the direction that connects the Red and Cyan colors.

3. CHARACTERIZATION OF THE DECISION THRESHOLDS

To determine the maximum and minimum thresholds for each of the two chrominance components (Cg and Cr) that define the boundary box, facial regions need to be manually extracted from a set of training images. Fifty-five images acquired by different sources: digital cameras, scanned photographs, software edited images, etc. and obtained under different lighting conditions were used as a training set for modeling the human skin color in the YCgCr space.

The maximum and minimum values of Cg (Cgmin, Cgmax) and Cr (Crmin, Crmax) were computed for each face region within every training image. They were characterized as Gaussian distributions, and their statistics were obtained: minimum, maximum, mean, and standard deviation values of each color component.

The best segmentation results were achieved with the boundary box defined by the minimum values of (Cgmin, Cgmax) and maximum of (Crmin, Crmax) thresholds of the color components obtained from the training images: [76,125] for Cg and [136,202] for Cr [10] (see Figure 1).

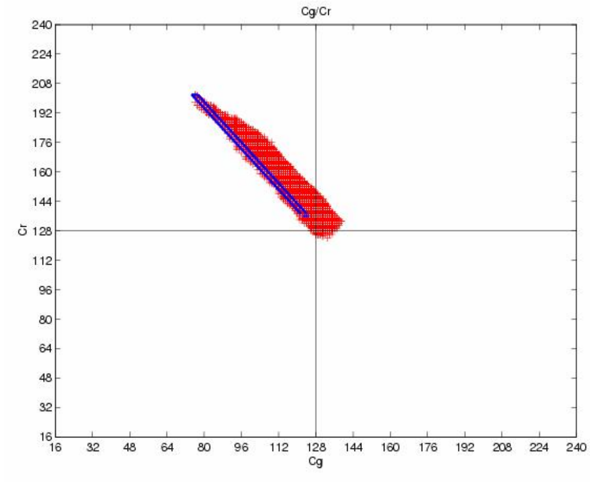

Fig. 2. Non-rectangular decision region in the Cg-Cr plane.

Similar results were achieved using a set of portrait-like images (AR face database [9]) with homogeneous background, containing human faces with different facial expressions, lighting conditions, objects occluding faces (like sunglasses and scarf) and several combinations of them.

Considering the skin color distribution depicted in Figure 1, it is possible to achieve a better performance of the face segmentation in the YCgCr color space by defining a boundary box in the direction of the line that connects the Red and Cyan colors, where most of the skin values are concentrated.

New decision thresholds constituting a non-rectangular region can be defined, as it is represented in Figure 2, so that the Cr thresholds are the same vertical limits of the boundary box, while the Cg thresholds, Cgmin and Cgmax lines, are parallel to the Red-Cyan line. Hence, an individual pixel will be classified as skin if it is within the Crmin and Crmax thresholds and the Cgmin and Cgmax lines, according to the following equations:

$$Cr\min \leq Cr \leq Cr\max$$

$$\frac{(305 - Cr)}{1,38} \leq Cg \leq \frac{Cr\min + 1,38 * Cr\max - Cr}{1,38}$$

Instead of using two binary masks for face segmentation based on the Cg and Cr thresholds of the boundary box, in the case of a non-rectangular decision region, for every Cr value, a binary mask is used for each Cgmin-Cgmax pair of Cg thresholds. The final mask for face segmentation is obtained as the intersection of all the binary masks.

Comparing the performance of the non-rectangular decision region and the boundary box in the YCgCr color space, as it is represented in Figure 3 for a man and a woman faces, the face detected area was improved in the

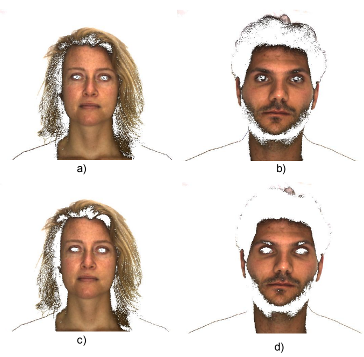

Fig. 3. Comparison of the segmentation results using a: a) , b) non-rectangular decision region, c), d) boundary box.

former case, as the eyes, the inside and outside parts of the face (including the eyebrows area in the man image) are included. In both cases, the best results were achieved using the minimum Cg and the maximum Cr decision thresholds.

The main disadvantage of a non-rectangular decision region is that the processing time is increased, being more than twice the time in the boundary box case.

A transformation of the YCgCr color space is also proposed in this paper in order to improve the face segmentation results by refining the decision thresholds in the skin color distribution and reducing the overall processing time.

4. TRANSFORMATION OF THE YGCR COLOR SPACE

The transformation of the YCgCr color space basically consists of defining a new chrominance plane, based on the skin color distribution in the Cg-Cr plane represented in Figure 1. One of the axis of the new chrominance plane Cg'-Cr', defined in the direction where most of the skin pixels are distributed, will constitute the vertical axis, Cr'. The other axis, Cg', will be perpendicular to Cr'.

So, it will be necessary to rotate 30º clockwise the Cg and Cr axes to define the Cg'-Cr' plane (Cgy-Cry in Figure 4). Besides, the center of the hexagon formed by the primary and complementary colors should be shifted to match the center of coordinates, recalculated to (128,128), and the axes range will be [0, 255].

Therefore, the new color space YCg'Cr' will be defined by the following equations:

$$Cg' = Cg * \cos 30^{\circ} + Cr * \sin 30^{\circ} - 48$$
$$Cr' = -Cg * \sin 30^{\circ} + Cr * \cos 30^{\circ} + 80$$

The first idea was considering the boundary box represented in Figure 4, based on the directly transformed non-rectangular decision region in Cg-Cr (see Figure 2), but the detected region was not as good as expected.

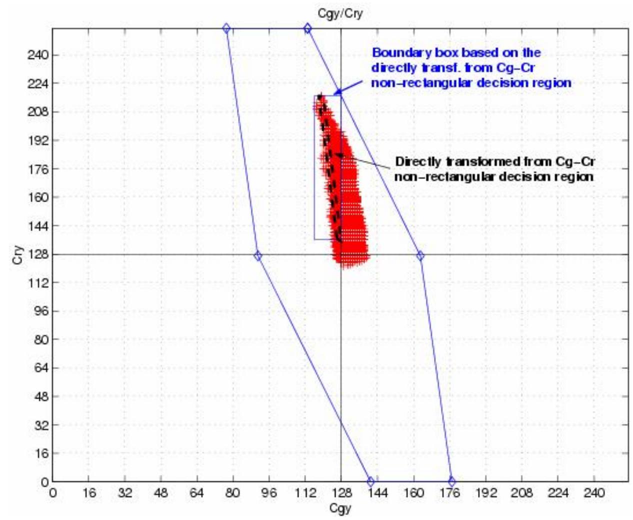

Fig. 4. Non-rectangular decision region and boundary box based on it transformed to the Cg'-Cr' plane.

The training process was repeated for the transformed color space in order to improve the segmentation results, so the minimum and maximum Cg' and Cr' values for any individual image were obtained. The thresholds were also characterized as Gaussian distributions, obtaining their statistics: mean, minimum, maximum, etc. for Cg' and Cr'.

5. EXPERIMENTAL RESULTS

The AR face database [9] was used for this study in order to compare the achieved results with the same test sets used in YCbCr and YCgCr. Tests using different thresholds for Cg' and Cr': mean, minimum, maximum, centroids, modified-mean (Cg'-½(Cg'mean-Cg'min) , Cr' + ½(Cr'max -Cr'mean)), minimum for Cg' and maximum for Cr', etc. were performed.

Analyzing the skin detection areas achieved by using different thresholds, it was shown that the Cr' thresholds are critic for the segmentation process. When Cr'min is reduced, the decision region is near the center and includes more pixels, so the whole image is detected as skin. The Cg' thresholds are within a small range, so that either the whole image or nothing can be segmented using this color component mask.

In the transformed color space, the best results are achieved using two types of decision regions: maximum ([125,140] for Cg' and [136,217] for Cr') and modified-mean thresholds ([119,131] and [128,199] for Cg' and Cr', respectively). The segmentation masks used for determining the skin pixels for each of the color components and both of them are represented in Figure 5 using these two thresholds.

The skin-like region is better detected with the maximum thresholds while the face region is bigger in the

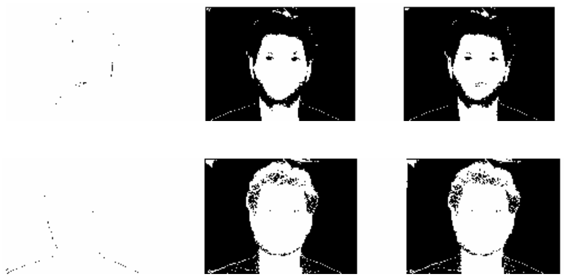

Fig. 5. Detection masks for Cg’ (left column), Cr’ (center) and both of them (right), with maximum (top) and modified-mean thresholds (bottom).

second case, as the eyes and beard areas are also included, though the hair and part of the clothes are detected too.

The segmentation results were also compared to the ones achieved using YCgCr with the boundary box and the non-rectangular decision regions (see Figure 3). The more complete face is obtained using modified-mean thresholds in YCg’Cr’, as can be seen in Figure 6.

The best segmentation results are achieved using maximum thresholds, as only the skin-like pixels are detected and the objects in the background can also be eliminated as it is represented in the right image of Figure 6, that is a frame of a TV daily news program.

The processing time of the segmentation process is very similar to the case of the boundary box in YCgCr. The main difference with this color space is that the best segmentation results are not achieved using the minimum for Cg and maximum for Cr possible thresholds.

6. CONCLUSIONS

Decision thresholds for face segmentation have been characterized in the YCgCr color space and a transformation of it has been proposed in this paper, basically consisting of the rotation of the Cg-Cr plane. As the segmentation results using the directly transformed thresholds are not good enough, a new training process has been carried out.

The segmentation process has been compared to the boundary box and non-rectangular decision regions cases in YCgCr. The main difference is that the same type of decision thresholds do not yield to the best results, as they are achieved using maximum or modified-mean thresholds in YCg’Cr’. The processing time is also reduced compared to the case of the non-rectangular decision region.

Using the maximum possible thresholds more skin-like pixels are detected, while with the modified-mean thresholds most of the face is detected.

Our future work will consist of improving the face segmentation results by refining the maximum decision thresholds masks using morphological techniques.

Fig. 6. Segmented faces using maximum Cg’Cr’ (top) and modified-mean (bottom) thresholds.

7. REFERENCES

[1] M. Abdel-Mottaleb, A. Elgammal, “Face detection in complex environments from color images", Proc. of the IEEE Int. Conf. on Image Proc. ICIP'99, pp 622-626, Kobe, Oct. 1999.

[2] L. Torres, J.Y. Reutter, L. Lorente, “The importance of the color information in face recognition", Proc. of the IEEE Int. Conf. on Image Proc. ICIP'99, pp 627-631, Kobe, Oct. 1999.

[3] M.H. Yang, N. Ahuja, “Detecting Human Faces in Color Images", Proceedings of the 1998 IEEE International Conf. on Image Processing ICIP'98, pp. 127-139, Chicago, Oct. 1998.

[4] B. Menser, M. Brünig, “Segmentation of human faces in color images using connected operators", Proc. of the IEEE Int. Conf. on Image Proc. ICIP'99, pp. 632-636, Kobe, Oct. 1999.

[5] M-H. Yang, D. J. Kriegman, N. Ahuja, “Detecting Faces in Images: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, 24, pp. 34-58, January 2002.

[6] V. Vezhnevets, V. Sazonov, A. Andreeva, “A Survey on Pixel-Based Skin Color Detection Techniques”, Graphicon-2003, Moscow, Russia, Sept. 2003.

[7] J-C. Terrillon, M.N. Shirazi, H. Fukamachi, S. Akamatsu, “Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images", Proc. of the 4th IEEE Int. Conf. on Aut. Face and Gest. Recogn. FG'00, pp. 54-63, March 2000.

[8] B. D. Zarit, B. J. Super, F.K.H. Quek, “Comparison of five color models in skin pixel classification", Int. Workshop on Recognition, Analysis and Tracking of Faces and Gestures on Real-Time Systems (ICCV'99), pp.58-63, Sept. 1999.

[9] A.M. Martinez and R. Benavente, "The AR face database", CVC Tech. Report #24, 1998.

[10] J.J. de Dios, N. García, "Face Detection Based on a New Color Space YCgCr", Proc. of the IEEE Int. Conf. on Image Processing ICIP’03, vol. 3, pp. 909-912, Barcelona, Sept. 2003.