# TRANSFORMATION OF THE YCgCr COLOR SPACE FOR FACE DETECTION


*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

In this paper, a transformation based on the skin color distribution of the YCgCr color space presented in [10] is proposed for face detection improvement. Mainly, it is based on the rotation of the skin color distribution in the CgCr plane, so that a boundary box can be used for face segmentation. It has been tested in a set of portrait-like images with different lighting conditions and objects occluding the faces. Segmentation results have been studied using two types of decision thresholds in order to detect the whole head or just the skin region. Besides, the face detection results achieved in the segmentation process based on the transformation of the YCgCr color space has also been compared to those using the YCbCr and YCgCr color spaces. The detection results have been improved with the new decision thresholds presented in this paper.


## 1. INTRODUCTION

In the last years, different techniques have been proposed for segmenting faces in color images going from the morphological methods to neural networks. Recently, face detection schemes based on color information have been studied and compared to the results achieved by other previous techniques [1-3]. Human skin color has been considered as a distinguishing and effective feature in face detection, taking into account that the major differences between people having different skin color lie mostly in their intensity rather than their chrominance [4,5]. The main advantage in segmentation through the color characteristics is that facial detection can be performed independently on the size and the position of the face in the image [6].

However, an objection to its use is that the classification in regions of skin color pixels will fail if the face is partially detected or there are objects of a skin-like color in the background of the image [6-8]. Therefore, color information can be used in combination with filtering or morphological techniques in order to obtain a better performance in the segmentation process, so that the occluded or undetected regions may be added to the face detected area [1,4,8].

For the classification of the pixels in an image into skin and non-skin regions, several color spaces have been proposed, such as RGB, XYZ, CIE-Lab, HSV or YCbCr. In [10], a novel color space, YCgCr, based on YCbCr, was proposed. Color spaces which work with separated luminance and chrominance components, like HSV or YCbCr, seem to be more appropriate for face detection.

A transformation of the YCgCr color space is presented in this paper in order to improve the face segmentation results.

## 2. DEFINITION OF THE YCgCr COLOR SPACE

This new color space, YCgCr, proposed in [10] for face detection, is based on the application of colorimetry to television systems. Several color spaces (YIQ, YUV and YCbCr) have been defined for processing luminance and chrominance signals separately. As these color spaces are transmission oriented, the chrominance components were chosen trying to minimize the encoding-decoding errors, and the biggest differences were selected: (R-Y) and (B-Y). The YCgCr color space uses the smallest color difference (G-Y) instead of (B-Y) and is defined exclusively for analysis applications, mainly for face segmentation.

A human skin color model can be considered practically independent on the luminance and concentrated in a small region of the Cg-Cr plane, as can be seen in Figure 1. The simplest model used for the classification between skin and non-skin pixels is based in the definition of a chrominance bounding box [4]. Considering only color information, a pixel will be classified as skin if both of its color components are within each of these ranges. So, the technique used for face segmentation consists of defining maximum and minimum thresholds for each of the two chrominance components (Cg and Cr).

The YCgCr color space is based on YCbCr, but it differs on the use of the Cg color component instead of the Cb one. As the color space defined by CbCr is un-orthogonal, it was decided to use the smallest color difference, G-Y, (Cg component) instead of Cb to conduct the face segmentation.

Following ITU Rec. BT.601, YCgCr components can be derived from the YCbCr equations and expressed in

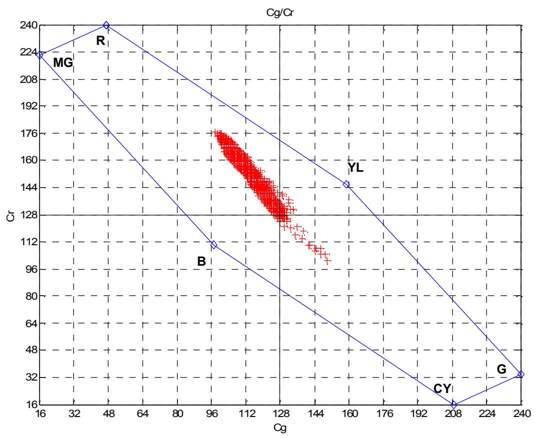


Fig. 1. Skin color pixels in the Cg-Cr plane.

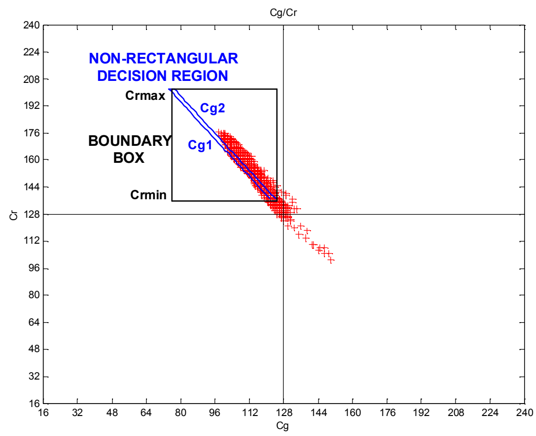


Fig. 2. Boundary box and non-rectangular decision region.

terms of Y', G'-Y' and R'-Y' components in the range [0,1] using the following equations:

$$Y = 16 + 219 \cdot Y'$$

$$Cg = 128 + 112 \cdot \left[ \frac{1}{1-0.587}(G'-Y') \right]$$

$$Cr = 128 + 112 \cdot \left[ \frac{1}{1-0.299}(R'-Y') \right]$$

Luminance and chrominance are coded in 8 bits. The chrominance components are defined in the range [16,240].

The primary and complementary colors in the YCgCr color space are also represented in Figure 1. The skin color region is concentrated in a small area corresponding to the 1st and 2nd quarters of the Cg-Cr plane (between the red and yellow colors).

## 3. MODIFICATION OF DECISION THRESHOLDS

To obtain the boundary box, facial regions need to be manually segmented in an RGB image to determine the above mentioned maximum and minimum thresholds for each of the two chrominance components (Cg and Cr). These decision thresholds must be precisely determined, so that all the face pixels are detected and the pixels in the background are excluded from the detection area. Not only human skin color must be taken into account, as the elements on the background could be considered as face pixels. Therefore, the selection of the segmentation decision values will achieve the detection of either nothing or the whole image.

Fifty-five images acquired by different sources: digital cameras, scanned photographs, software edited images, etc. and obtained under different lighting conditions were used in [10] as a training set for modeling the human skin color in the YCgCr space. Figure 1 represents the Cg and Cr values of an individual face.

Initially, maximum and minimum values of Cg (Cgmin, Cgmax) and Cr (Crmin, Crmax) were computed for the face region of 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 results were always achieved applying the segmentation with the minimum Cg and the maximum Cr thresholds to the training images and also using a set of portrait-like images (AR face database [9]).

The boundary box represented in Figure 2 corresponds to 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.

Considering the skin color distribution, 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. A non-rectangular decision region has been defined (see Figure 2).

The Cr thresholds are the same vertical limits of the boundary box, while the Cg thresholds are parallel to the Red-Cyan line. So, pixels will be classified as skin if they are within the Crmin and Crmax thresholds and the Cgmin and Cgmax lines, in a parallel direction to the line connecting the Red and Cyan colors, as it is described in the following equations:

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

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

In a previous work, the performance of the non-rectangular decision regions was compared to the boundary box in the YCgCr color space. The face detected area was improved in the 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 using non-rectangular decision regions is that the processing time is more than twice the time for face detection in the boundary box case.

A transformation of the YCgCr color space is presented 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. FUNDAMENTALS OF THE COLOR SPACE TRANSFORMATION

Based on the skin color distribution in the Cg-Cr plane represented in Figure 1, a new chrominance plane will be defined where the skin color region is distributed along the vertical axis and closer to the center.

So, it will be necessary to rotate the Cg and Cr axes 30º clockwise, as can be seen in Figure 3. Besides, the center of the hexagon formed by the primary and complementary colors should be moved to the origin of coordinates. This area should be shifted in both directions (48 pixels upwards and 80 pixels in the horizontal direction), using the following equations:

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

They can be directly computed from Y’, G’-Y’ and R’-Y’, so there is no need to conduct a two step calculation.

One of the consequences of this transformation is that the vertical and horizontal axes will be defined from 0 to 255 instead of the range [16,240] for the Cg-Cr plane. It is important to consider this condition for implementing the inverse transformation, as the pixel information of both vertical extremes could be lost. Figure 3 represents the primary and complementary colors in the Cg’-Cr’ plane (showed as Cgy-Cry) of the transformed color space, where the center of coordinates has been recalculated to (128,128).

The initial idea was considering the directly transformed decision region (see Figure 3) from the Cg-Cr boundary box represented in Figure 2. When using for the segmentation a boundary box based in the left and right vertices of this region in the Cg’ direction, the face was partially detected, as can be seen in Figure 4.

In order to improve the segmentation results, the same training process was repeated for the transformed color space for defining the best decision region, so the minimum and maximum Cg’ and Cr’ values were obtained. They were also characterized as gaussian distributions, and their statistics were obtained: minimum,

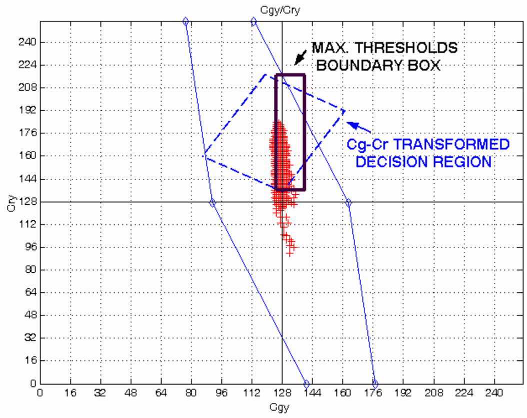


Fig. 3. Skin color pixels, directly transformed Cg-Cr boundary box and boundary box in the transformed color space.

Fig. 4. Mask and segmented face with transformed thresholds.

maximum, mean, and standard deviation values of each color component.

## 5. EXPERIMENTAL RESULTS

This study has been conducted 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.

The experimental results using different thresholds as the ones described in [10] are represented in Figure 5. The Cg’ thresholds are within a small range, so that either the whole image or nothing can be segmenting using this color component mask. The Cr’ thresholds are critic. The main difference with the previously used color spaces is that the best segmentation results are not achieved using the same type of decision thresholds (minimum for Cg and maximum for Cr).

In the transformed color space, the best segmentation faces are achieved using two types of decision regions: maximum thresholds (continuous boundary box in Figure 3) and modified-mean thresholds (Cg’- ½(Cg’mean - Cg’min), Cr’+ ½ (Cr’max -Cr’mean)). The skin-like region is better detected with the maximum thresholds, while the detected face is bigger in the second case, as the eyes and beard areas are also included, though the hair and part of the clothes are detected too. These two thresholds were used for the rest of the work.

Fig. 5. Segmented faces using different thresholds: mean, minimum, maximum, modified-mean, gaussian (mean+std), and min. for Cg' and max. for Cr' values.

The processing time of the segmentation process is very similar to the case of using a boundary box in YCgCr. It is also reduced to half the time in the case of the non-rectangular decision region.

The segmentation results were also compared to the ones achieved using different color spaces, like YCbCr, or YCgCr. The more complete face is achieved using modified-mean thresholds in YCg'Cr' as the whole face and hair (including the beard) is detected, though part of the clothes and some skin-like pixels are not detected, as can be seen in Figure 6. When using maximum thresholds the segmentation results are very similar to the other color spaces, as only the skin-like pixels are detected.

## 6. CONCLUSIONS

A transformation of the YCgCr color space 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, new decision thresholds using a set of training images have been computed. The segmentation process has been compared to YCgCr and YCbCr color spaces.

The best results are achieved using maximum or modified-mean thresholds in the new transformed color space. This is the main difference with the previous used color spaces, as the minimum for Cg' and maximum for Cr' possible thresholds do not yield to the best results.

More skin-like face pixels are detected using the maximum possible thresholds, while the hair and clothing regions are undetected. With the modified-mean thresholds most of the face is detected, including also the eyes and part of the hair. The processing time is also reduced compared to the case of non-rectangular decision regions and very similar to the previous boundary box case. Besides, it must be pointed that the new Cg' and Cr' thresholds provide better results.

Our future work will consist of improving the face segmentation results by refining the decision thresholds and trying to complete the segmenting mask with the use of other techniques (e.g. morphological).

Fig. 6. Segmented faces using different color spaces: YCbCr, YCgCr, and modified-mean and maximum Cg'Cr' thresholds.

## 7. REFERENCES


[1] R-L Hsu, M. Abdel-Mottaleb, A.K. Jain, "Face detection in color images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, n.5, pp. 696-706, May 2002.

[2] S. Srisuk, W. Kurutach, "A new robust face detection in color images", Proceedings of the 5th IEEE Int. Conference on Aut. Face and Gest. Recogn. (FGR'02), pp. 306-311, May 2002.

[3] C-W. Lin, Y-J Chang, Y-C. Chen, "Low complexity face-assisted video coding", IEEE Int. Conference on Image Processing ICIP'00, pp. 207-210, Vancouver, September 2000.

[4] 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.

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

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

[7] A. Albiol, L. Torres, E. J. Delp, "An Unsupervised Color Image Segmentation Algorithm for Face Detection Applications", IEEE Int. Conf. on Image Processing ICIP'01, Thessaloniki, Oct. 2001.

[8] D. Chai, K.N. Ngan, "Face Segmentation Using Skin-Color Map in Videophone Applications", IEEE Trans. on circuits and systems for video technology, v.9, n. 4, pp. 551-564. June 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", IEEE Int. Conf. on Image Processing ICIP'03, vol. 3, pp. 909-912, Barcelona, Sept. 2003.