STUDIA UNIV. BABEŞ–BOLYAI, INFORMATICA, Volume **LVII**, Number 3, 2012

AN EVALUATION OF COLOR SPACES USED IN SKIN COLOR DETECTION

IOAN-GABRIEL MIRCEA

Abstract. The purpose of the present paper is to determine the best choices for color segmentation techniques as basis for skin detection and to review the main approaches in subsequent texture segmentation. The main goal is to determine which color space offers the best clustering for the skin color space and therefore a compared analysis of the most popular color spaces used in the literature is conducted. The results are statistically processed in order to offer the possibility of significant conclusions to be drawn. However, the color segmentation alone does not provide efficiently accurate results when it comes to skin detection so it needs to be followed by efficient texture segmentation that would reduce false positives.

1. INTRODUCTION

An efficient skin detection software is a very useful computer vision tool, either as a stand-alone product or as a starting point for several other more complex software. Such a tool needs to be efficient with respect to both time and resources especially if it is part of a lightweight software intended for the average computer user. Skin detection can be viewed as the core process of more elaborate actions such as face and figure recognition, motion detection, explicit content filtering etc. and its efficiency directly impacts the overall process. This paper is intended to be a sketch for the devise of an efficient skin detection tool that is efficient enough to be easily run on today's usual hardware.

Before the actual development of the skin detection software, many preliminary choices need to be made. After thoroughly browsing the current bibliography on the matter one could conclude that, in fact, skin detection is usually conducted as a sequence of filtering actions upon the pixels of an image: first, only the pixels in the image that have skin-like color are selected,

Received by the editors: September 17, 2012.
2010 *Mathematics Subject Classification.* 68U10.
1998 *CR Categories and Descriptors.* I.4.6 [**Image Processing**]: Segmentation — *Pixel classification*.
Key words and phrases. computer vision, skin detection, color space, texture.

but usually this segmentation gives a high false positive rate and this rate is improved by a subsequent texture filtering stage which selects only the portions of the image which have skin-like texture. If the result obtained after the fore mentioned filters have been applied is still unsatisfactory, additional filters may be employed to reduce the false positive rate - FAR. With this in mind, the current article tries to find which is the best way to conduct color segmentation.

First, a brief overview of different skin color detection techniques is presented. The outcome of each technique is however influenced by the choice of the underlying color space used for color modeling and thus the following section in the article presents the most popular five color spaces used in the literature, highlighting their utility in skin color detection.

This introduces the main objective of the paper: establishing which of the proposed color spaces gives the best results in skin detection. A statistical evaluation is conducted and the results are commented, in order to assess whether statistically significant conclusions may be drawn from the evaluation outcomes.

Finally, as a prelude to an eventual follow-up article certain techniques for texture segmentation encountered in the literature are presented since they would represent the object of the future compared analysis and conclusions are drawn.

2. SKIN COLOR DETECTION

The literature presents several directions for skin color detection which are highlighted thereafter. The first direction consists in explicitly defining the boundaries of the skin color cluster inside the color space. It is a fast method but it requires a good choice for a color space and an efficient set of decision rules that would help achieve high enough detection rates. This approach was used in papers such as [13] and [5] which emphasize how critical the choice for a color model and a set of decision rules is for this kind of approaches.

Another direction is represented by the estimation of the skin color probability distribution based on information obtained from training data. One of the most thorough papers in this respect is [8], referencing a research conducted at The Cambridge Research Laboratory of Compaq Computer Corporation. Having access to a wide database of internet images, a statistical skin color model was created by use of a simple histogram technique. From a total of more than a billion pixels a statistical skin model was constructed by use of two histograms: for skin and non-skin pixels. The likelihood of a pixel being a skin pixel or not was computed via a standard likelihood ratio:

$\frac{P(rgb|skin)}{P(rgb|\neg skin)} > \emptyset, 0 < \emptyset < 1, P(rgb \mid skin) = \frac{s[rgb]}{Ts}, P(rgb \mid \neg skin) = \frac{n[rgb]}{Tn}$

where s[rgb],n[rgb] represent the skin,respectively non-skin pixels in the histogram bin for a certain color rgb, Ts,Tn are the total number of pixels in the skin ,respectively non-skin histograms and Øis the acceptance threshold.

The likelihood of a pixel being classified either as a skin or non-skin pixel can also be expressed using the Bayes rule, that aids the computation of "a probability of observing skin, given a concrete color value" [16]:

$$P(skin \mid c) = \frac{P(c|skin)P(skin)}{P(c|skin)P(skin)+P(c|\neg skin)P(\neg skin)} \geq \emptyset$$

These approaches are usually employed when a huge amount of training data can be obtained.

A different way of establishing skin color distribution is by constructing statistical models for skin color likelihood by use of 2-dimensional Gaussian models or by the mixture-of-Gaussians model. This approach is used when the training data is scarcer but still enough to offer viable information to construct the likelihood function. This gives a statistical measure of "how skin-like a certain color is" [10] using the following formula:

$$P(c \mid skin) = \frac{1}{2|\Sigma S|^{1/2}} e^{-\frac{1}{2}(c-\mu_S)^T \Sigma S^{-1}(c-\mu_S)}$$

where

$$\mu_S = \frac{1}{2}\sum_{j=1}^{n} c_j \text{ and } \Sigma S = \frac{1}{n-1}\sum_{j=1}^{n}(c_j-\mu_S)^T(c_j-\mu_S)^T$$

are the mean and co-variance measures obtained from the training data which contains n color samples - c_j

The mixture-of-Gaussians method employs a number of k mixture components for better approximation of the underlying data:

$$P(c \mid skin) = \sum_{i=1}^{k} \pi_i P_i(c \mid skin)$$

The performance boost achieved by using this latter approach in comparison with the single Gaussian approach is perfectly illustrated in [3]. Another important reference on the matter is [6] which uses Expectation Maximization as a technique for parameter adaptation.

According to [16] , the disadvantage with the first approach is that a good set of rules is difficult to find empirically, while the non-parametric approach requires a huge amount of training data which needs to be thoroughly selected and the parametric approach ignores the non-skin statistics from their model which makes them prone to high false positive rates.

3. MOTIVATION FOR COLOR SPACE EVALUATION

Deciding which colorspace gives the best results with respect to skin detection is of utmost importance to the entire skin detection purpose. An efficient chromatic segmentation is not only a good start towards an efficient detection rate but since the color analysis is both the first and the most important part of the detection process and therefore the impact of the choice of color space for representation can deeply influence the overall efficiency of the process. Therefore, a statistical evaluation of the efficiency of the main color spaces used in skin detection is not only justified but also relevant and important to the field. The results obtained from the presented comparison may offer valuable statistically significant information to researchers and developers regarding the viable options for color spaces in the field of skin detection.

It is optimal that the skin cluster formed in the color space is contiguous and somehow compact and although [8] proposed the classic RGB as color space, there are different other color spaces that may express skin clusters in a more compact way, the most popular of which are presented hereafter.

4. COLOR SPACES FOR SKIN DETECTION

After a review of the bibliography, a few color models stood out as being the most popular with respect to skin detection. An important aspect when constructing a color space for skin detection is luminance isolation from the chrominance. Luminance is not considered important when it comes to skin; what we are interested in is chroma since light incidence may alter the luminance of a skin pixel but its chrominance may still lay within skin color boundaries. This aspect is very important when it comes to skin detection since the elimination of the light component would also drastically reduce the size of skin cluster in the color space.

4.1. **Normalized RGB.** The normalized RGB space is computed from the RGB space by obtaining pure colors in two dimensions r, b - from the RGB components:

$$r = \frac{R}{R+G+B} \text{ and } b = \frac{B}{R+G+B}$$

In this way we reduce the color space by expressing the blue and red components as a function of all the RGB components, thus eliminating a dimension in representation. It is useful for a color space to represent the skin space as compact as possible and normalized RGB offers a good clustering. As a result, normalized RGB is successfully used for skin detection - see [2].

4.2. **HSV.** Another popular color space, used in many papers is HSV (Hue-Saturation-Value), a cylindrical color space. Hue defines the dominant color of an area; saturation measures the colorfulness of an area in proportion to its brightness [14]. The Value represents the luminance of the color and is not of interest to our purpose. This separation of luminance gives a more clear representation of the chrominance in the Hue-Saturation space than in the previous normalized RGB method. HSV can be obtained from RGB by application of the following rules:

$$H = \arccos\left(\frac{\frac{1}{2}((R-G)+(R-B))}{\sqrt{((R-G)^2+(R-B)(G-B)}}\right) \; S = 1 - \left(3\frac{\min(R,G,B)}{(R+G+B)}\right) \; V = \frac{1}{3(R+G+B)}$$

4.3. **YCbCr.** "YCrCb is an encoded nonlinear RGB signal, commonly used by European television studios and for image compression work" [16]. This is a clear choice for skin detection since it separates luminance efficiently and it is easy to be obtained from RGB and vice versa:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The difference between this color space and HSV is that, due to the easy transformation to and from the RGB space it spreads the values of its Cr and Cb components in a range between 0 and 255 giving a clear choice for the number of bins needed in an eventual histogram approach, while the cylindrical shape of the HSV space requires a linear transformation in order to construct the histogram bins.

4.4. **YCgCr and YCgCb.** Some useful variations of the YCbCr color space are presented in [17]. These two color spaces are similar to YCbCr but substitute R-Y and B-Y respectively with G-Y representing good alternatives for skin segmentation. They can be obtained from RGB values normalized to [0,1] through the following transformations:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -81.085 & 112 & -30.915 \\ -37.797 & -74.203 & 112 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -81.085 & 112 & -30.915 \\ 112 & -93.786 & -18.214 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

These two spaces promise to offer good results in skin detection since they include all the advantages of the YCbCr space enhanced by the additional filtering efficiency of a combined approach.

5. Related Approaches

Although the performance of different color spaces on skin detection has already been assessed in papers such as [1], [4] or [15], the present approach tackles the problem in a slightly different manner since the present goals are different. The research presented in [1] aimed to prove that, if an optimal detector was employed for any of the color spaces - RGB, YCbCr, HSV, CbCr - they would give the same detection rate, but, in contrast, the present approach analyzes only 2-dimensional spaces and evaluates their performance on the same benchmark. Paper [4] follows the same purpose as the present research but it does not include color spaces such as HSV, YCgCr and YCgCb which have obtained good scores in the present evaluation. However the evaluation methodology used in the present paper is, to a certain extent, similar to the one from [4]. A very thorough analysis on the matter is conducted in [15], including a wider rage of color spaces and also considering both the 2-dimensional and 3-dimensional cases, but the paper evaluates the color spaces with respect to several metrics computed on the skin color clusters and not with respect to their impact on the efficiency of the segmentation process, which is the sought purpose of the proposed approach.

6. Color space Evaluation

Given the importance of color space in the process of color skin classification,the purpose of the present paper is to determine which of the previously presented color spaces is the best choice for the construction of skin detection software.

6.1. **Skin Sample Database Construction.** Inspired by the methodology presented in [8], we constructed a database containing only pixels coming from human skin portions selected from various images. Around 1400 skin samples of 50*50 pixels were collected from different images on the internet, depicting skins belonging to different races, from different body areas and in different illumination conditions.

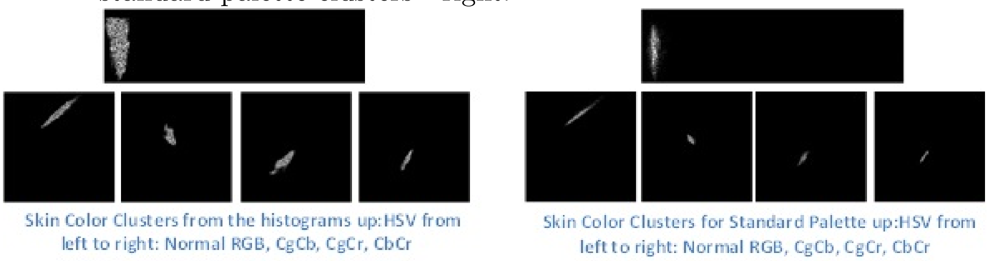
Although a more standardized benchmark might have been used for the research, the choice for a custom benchmark is influenced by the need of a source of skin pixels from various skin-tones and in various illuminations and the pictures provided by the most of the benchmarks -[9]- are usually shot in the same lighting conditions thus reducing the needed generality.

The selected pixel database is used to benchmark the performance of each of the fore mentioned color spaces. In order to do that, histograms are constructed in each color space in order to reveal the skin color clusters. At first the obtained results need to be compared against a chromatic reference for the human skin. A skin color palette used as reference in the domain of visual arts was chosen - see Figure 1.

Figure 1. Skin color palette used as reference.

The shape of the clusters formed in each of the color space depicting all the shades in the reference palette is compared to the corresponding previously constructed histograms.

Figure 2. Comparison between histogram clusters - left - and standard palette clusters - right.



The comparison is illustrated by Figure 2, that shows that while the shapes on the left tend to exceed the contour of the ones on the right, they follow the same orientation and the same overall shape, meaning that the constructed color histograms clearly superpose over the colors in the standard palette, giving a reasonably accurate representation of skin color. A total of 3497497 skin pixels were used to construct each histogram.

6.2. **Testing the model.** Once the histograms were constructed and validated against the standard palette, the next step was to test their performance on real images from the internet. A selection of 155 pictures of considerable size was made such that it contained important amounts of human skin as well as background that might contain skin-like colors. It is important to select a percentage of the pictures to contain materials that resemble skin because we want to test which color space achieves better results even in this situation which influences the false acceptance rate. In order to construct the ideal frame of reference,the human skin pixels in each picture were then manually segmented and a set of another 155 matching pictures containing only those pixels was created. The two paired sets of pictures were used to give the statistical measure of the performance of each color space by the computation of two statistics: the TAR (True Acceptance Rate) and FAR (False Acceptance Rate) of each color space histogram, which are computed by the following rules:

$$TAR = \frac{numberofrealskinpixelsdetectedbythemethod}{numberofpixelsinonly-skinimage}$$

$$FAR = \frac{numberoffalselydetectedpixels}{numberofnon-skinpixelsinonly-skinimage(defaultwhite)}$$

The actual test consisted in computing the True Acceptance Rate and False Acceptance Rate for each color space on all the pictures in the set. A measure of discrepancy between the TAR and FAR was used to assess the performance of each color space on each picture. For each picture, only the best performance was taken into account for measuring which of the color spaces performed best in the most cases. Also, the mean and standard deviation were computed for each color space in order to offer a more compelling picture of their performance. All the significant data obtained from the tests is presented in the table in Figure 3.

FIGURE 3. Table containing test performances for each color space.

average/stdev of TAR	average/stdev of FAR	# of max values
YCbCr		
0.787018561	0.188420976	27
0.2217745	0.171314158	
Normal-RGB		
0.901019324	0.268400145	25
0.158630908	0.199100371	
CgCb		
0.798440968	0.206448848	33
0.215034102	0.17783193	
CgCr		
0.847099832	0.215081078	41
0.193123592	0.188059618	
HSV		
0.846940685	0.277192945	29
0.174494866	0.186165776	

The analysis of the data shows that the two variations of the YCbCr color space, CgCb and CgCr have a slight advantage over the other spaces with respect to the test score, meaning that these spaces obtained better FARs even on the pictures containing substantial skin-like regions in their background. However, the differences are not significant enough to induce a clear conclusion.

The statistics show that, for the YCbCr space the lowest FAR has been achieved but also the lowest TAR suggesting that this color space is a little bit too drastic with the classification. At the other extreme we have the Normal RGB color space that achieves the highest TAR but also one of the highest FARs meaning it is too permissive. The best scores are achieved by the three spaces that maintain a high TAR and also keep the FAR at an acceptable level regardless of the nature of the picture evaluated.

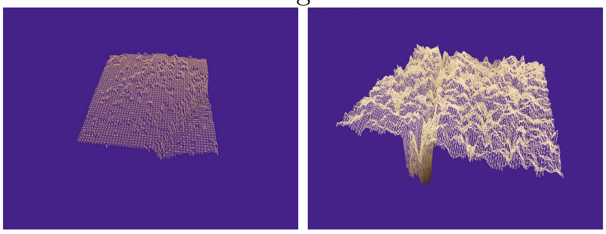
One important aspect that is considered an advantage on the CgCb , CgCr side is the fact that, as it was mentioned in the description of the HSV color space, its cylindrical shape makes it rather difficult to construct bins for the Hue component. All the conclusions drawn from the analysis of the data in the table are perfectly reflected in Figure 4 which illustrates the performance of color skin classification in all the evaluated color spaces.

FIGURE 4. Table containing test performances for each color space.

7. Texture Detection - a way of reducing the FAR

As seen in the results of the previous experiment, while the TAR rate was somewhat satisfactory, a FAR of around 20% is achieved in most of the color spaces. This is due to the fact that the usual environment of pictures representing human skin contains materials that have skin-like colors without being skin. That is why further filtering needs to be applied on the image in order to reduce the False Positives. Many natural materials that resemble human skin in color : sand, some animal's fur, etc. have entirely different textures compared to the texture of the skin. Texture classification can be therefore used to eliminate some of the false positives obtained from color classification. The task is simplified by the fact that usually skin has a smooth texture with slight degradation in luminosity while other materials of skin-like color have a more rugged aspect. A Skin sample is compared to a skin-colored sample with respect to texture in Figure 5.

FIGURE 5. Comparison between skin texture - left - and non-skin texture with skin color -right.



We can accept a slight decrease in the TAR as long as the FAR is reduced significantly after the application of the texture filter, but the choice of the texture measure is crucial. It has to be computationally efficient and also it has to enhance the differences between skin texture and any other texture. The most usual approaches found in current literature are the GLCM, used in [11] or Gabor wavelets [7], but they require a certain amount of computation and parametrization to work efficiently. A non-parametric solution is presented in [12] which uses feature distributions for texture segmentation. The texture classification technique needs to be both exact and fast since the texture segmentation stage is the one the needs to improve the False Acceptance Rate.

8. CONCLUSIONS

After a review of the most popular techniques and color spaces used in skin detection, taking into account the results of the experiment evaluating the performance of different color spaces as basis for skin detection, a choice can be made to use a combined color space CgCr, CgCb in skin color detection since it would bring a good color clustering necessary for imposing a high TAR. However, the obtained FAR is still unsatisfactory if only one form of classification is used and it is of utmost necessity to improve on the FAR by use of an efficient texture filter.

References

[1] Alberto Albiol, Luis Torres, and Edward J. Delp. Optimum color spaces for skin detection, 2001.

[2] Hani K. Almohair, Abd Rahman Ramli, Elsadig A. M, and Shaiful J. Hashim. Skin detection in luminance images using threshold technique skin detection in luminance images using threshold technique abstract. In *International Journal of The Computer, the Internet and Management*, volume 15, pages 25 –32, 2007.

[3] Tiberio S. Caetano, Silvia D. Olabarriaga, and Dante Augusto Couto Barone. Performance evaluation of single and multiple-gaussian models for skin color modeling. In *Proceedings of the 15th Brazilian Symposium on Computer Graphics and Image Processing*, SIBGRAPI '02, pages 275–282, Washington, DC, USA, 2002. IEEE Computer Society.

[4] Jamal Ahmed Dargham Chelsia Amy Doukim and Ali Chekima. Comparison of three colour spaces in skin detection. *Borneo Science*, 24(3):75–81, 2009.

[5] Giovani Gomez and Eduardo F. Morales. Automatic feature construction and a simple rule induction algorithm for skin detection. In *In Proc. of the ICML Workshop on Machine Learning in Computer Vision*, pages 31–38, 2002.

[6] Ming hsuan Yang and Narendra Ahuja. Gaussian mixture model for human skin color and its applications in image and video databases. In *booktitle*, pages 458–466, 1999.

[7] Zhiwei Jiang, Min Yao, and Wei Jiang. Skin detection using color, texture and space information, 2007.

[8] Michael J. Jones and James M. Rehg. Statistical color models with application to skin detection. *Int. J. Comput. Vision*, 46(1):81–96, January 2002.

[9] A. M. Martinez and R. Benavente. The AR Face Database. Technical report, CVC, June 1998.

[10] Bernd Menser and Mathias Wien. Segmentation and tracking of facial regions in color image sequences. In King N. Ngan, Thomas Sikora, and Ming-Ting Sun, editors, *VCIP*, volume 4067 of *Proceedings of SPIE*, pages 731–741. SPIE, 2000.

[11] Anal Kumar Mittra. Automated detection of skin disease using texture features. In *International Journal of Engineering Science and Technology*, pages 4801–4808, 2011.

[12] Timo Ojala and Matti Pietikäinen. Unsupervised texture segmentation using feature distributions. *Pattern Recognition*, 32(3):477–486, 1999.

[13] Peter Peer, Jure Kovac, and Franc Solina. Human skin colour clustering for face detection, 2003.

[14] Charles A. Poynton. *Frequently asked questions about colors.* publisher, 1997.

[15] Min C. Shin, Kyong I. Chang, and Leonid V. Tsap. Does colorspace transformation make any difference on skin detection? In *In IEEE Workshop on Applications of Computer Vision*, pages 275–279.

[16] Vladimir Vezhnevets, Vassili Sazonov, and Alla Andreeva. A survey on pixel-based skin color detection techniques. In *IN PROC. GRAPHICON-2003*, pages 85–92, 2003.

[17] Shi Yuexiang. Zhang Zhengzhen. Skin color detecting unite ycbcg color space with ycgcr color space. note, 2009.

Babeş-Bolyai University, Department of Computer Science, 1 M. Kogălniceanu St., 400084 Cluj-Napoca, Romania

E-mail address: mims0509@scs.ubbcluj.ro