A Study of the Effect of Illumination Conditions and Color Spaces on Skin Segmentation

Diogo Kuiaski, Hugo Vieira Neto, Gustavo Borba and Humberto Gamba
Post-graduation Program in Electrical Engineering and Industrial Informatics
Federal University of Technology – Paraná
Avenida Sete de Setembro, 3165
Curitiba-PR, CEP 80.230-910, Brazil
diogo.kuiaski@gmail.com, hvieir@utfpr.edu.br, gustavobborba@gmail.com, humberto@utfpr.edu.br

Abstract—**This work aims at investigating the influence of luminance information and environment illumination on skin classification. We explore Bayesian approaches to perform automatic classification of human skin pixels on digital images, using color features as input. Two probabilistic skin color models were built on different color spaces (RGB, normalized RG, HSI, HS, YCbCr and CbCr) and tested in a task of automatic pixel classification into skin and non-skin. Analyses of classification performance were done by presenting an illumination controlled image database containing images acquired in four different illumination conditions (shadow, sun, incandescent and fluorescent lights) to these classifiers. Our experiments show that building probabilistic skin color models using the CbCr color space generally improves performance of the classifiers and that best performance is achieved in shadow illumination.**

Keywords: skin segmentation, Bayes theory, image processing, color spaces.

I. INTRODUCTION

Automatic human skin segmentation plays an important role on many image processing applications such as face recognition [1], people tracking [2], biometric pattern recognition [3], offensive content detection [4] and gesture recognition [5]. The literature shows that, in these image processing methods, skin detection is often part of their preliminary steps.

Two pieces of information are needed in order to perform pixel classification, which are: (a) *a priori* information about the classes corresponding to each pixel; (b) appropriate visual features, which are intended to provide enough information to separate pixels in one class or another. Visual features are usually quantified in feature vectors, which constitute a point in a multidimensional space U — subsets of U constitute pixel classes.

In this paper each pixel of the input image is considered to belong to either of two classes, a skin class or a non-skin class. The visual feature used for classification is simply the color of each pixel, since skin color is easily recognized by people [6]. Moreover, color-based skin classifiers are robust to geometrical transformations and partial occlusions, and rely on fast algorithms [7].

Most of the difficulties in automatic skin pixel classification are due to environment illumination and camera sensor characteristics. Changes in environment illumination usually alter the resulting images acquired by the camera, causing colors to appear different. This is known as the color constancy problem [7]. Another matter that increases difficulties in automatic skin segmentation is the variation of skin color from person to person — ethnicity, age and gender are some of the factors that have influence. For all this reasons, creating a model that is able to express skin color based only on pixel information is a challenging task. Here we assume that it is possible to extract enough information from probabilistic models in order to perform pixel classification and investigate which color space minimizes the influence of environment illumination and intrinsic skin color variations in the performance of skin pixel classifiers.

This work uses models for skin color in six color spaces (RGB, normalized RG, HSI, HS, YCbCr and CbCr), which are described in section II. The two probabilistic models used are presented in section III. Section IV describes the process of joining probabilistic models and decision rules in order to build Bayesian classifiers. In section V we present an illumination-controlled skin database, which we used to conduct classification experiments. Classifier performances are assessed using ROC curves [8] and discussed in section VI.

II. COLOR SPACES FOR BUILDING SKIN MODELS

The sensor of a digital camera captures the light reflected from objects in the scene in a specific color space. There are basically two proprieties of the light that are represented by color spaces — luminance and chrominance [9]. Luminance is the measure of how many photons reach the sensor of the camera — the energy of the light source — and chrominance is the information about the spectral composition of the light source. Transformations from one color space to another may highlight specific characteristics of a particular color feature and, as a result, improve pixel classification.

The color of a pixel is a point within a three-dimensional color space, defined by a vector $c = [c_1, c_2, c_3]$. The com-

ponents of vector c represent luminance and chrominance information from a pixel. Usually luminance is represented by one of the dimensions and chrominance is represented by the other two dimensions of c. In this work we conduct experiments with reduced two-dimensional color spaces (chrominance information only) or three-dimensional color spaces (chrominance and luminance information).

A. Additive Color Spaces (RGB and normalized RG)

The RGB color space stems from the three types of color photoreceptors (cones) that are present in the normal human retina. Color representation in this model consists in the additive combination of the three primary colors red, green and blue. The RGB model was originally designed for use in cathode ray tube monitors and is still used in many image processing algorithms because of historical reasons related to the codification of image file formats.

However, the use of the RGB color space for skin color classification seems not to be such a very good choice [10]–[12], since this model presents a high correlation between its three channels and there is not a clear separation between luminance and chrominance components, i.e. luminance is diluted in all three color channels. Furthermore, the RGB color model is notably vulnerable to environment illumination variations [9]. However, in the work by Min C. Shin and others [10], promising results for skin detection using RGB color features were reported.

In order to reduce the effects of illumination, it is necessary to remove luminance information [6] — this is experimented with all color spaces used in this paper. Particularly for the RGB color space, which does not have a clear luminance component, a simple color space transformation is made, as shown in equations (1) and (2), in order to decouple luminance from chrominance.

$$r = \frac{R}{R + G + B}, \qquad (1)$$

$$g = \frac{G}{R + G + B}, \qquad (2)$$

where, R, G and B are the values of the RGB color channels red, green and blue, respectively, and r and g are the values of the so-called normalized R and G channels.

B. Perceptual Color Spaces (HSI and HS)

The HSI (Hue, Saturation and Intensity) color space tries to reproduce colors numerically as a human being would do perceptually. Hue represents the primary color of an object (green, yellow, blue, orange), saturation represents how pure this color is (deep blue, light blue), and intensity represents how bright it is. Hue and saturation represent chrominance and intensity represents luminance information. Therefore, it is possible to removed I component in an effort to reduce illumination effects in skin detection applications.

All images used in this work are coded in RGB and therefore it is necessary to transform them to HSI, according to equations (3), (4) and (5).

$$I = \frac{R + G + B}{3}, \qquad (3)$$

$$S = 1 - \frac{\min(R, G, B)}{I}, \qquad (4)$$

$$H = \arctan\left[\frac{\sqrt{3}(G - B)}{2R - G - B}\right], \qquad (5)$$

where, R, G and B are the RGB color channels normalized in range $[0, 1]$, and H, S and I are components of the HSI color space.

C. Orthogonal Color Spaces (YCbCr and CbCr)

The YCbCr color space is largely used in European television systems and consists of a linear transformation of the RGB color space [6]. Y represents luma or luminance, Cb represents chroma blue and Cr represents chroma red. Since this color space presents a decorrelation between luminance and chrominance information, it is called an orthogonal color space [7]. Having this in mind, we believe that removing the luminance component Y from this color space would yield good skin color classification results.

The transformation from RGB to YCbCr can be done using equations (6), (7) and (8).

$$Y = 0.299R + 0.587G + 0.114B, \qquad (6)$$

$$Cb = -0.168736R - 0.331264G + 0.5B, \qquad (7)$$

$$Cr = 0.5R - 0.418688G - 0.081312B, \qquad (8)$$

where, R, G and B are the RGB color channels normalized in range $[0, 1]$, and Y, Cb and Cr are the components of YCbCr color space.

III. PROBABILISTIC MODELS FOR SKIN COLOR

In order to perform skin color classification, we decided to use probabilistic models as previous knowledge. Such probabilistic models can be parametric or non-parametric. For skin pixel classification, parametric models are usually single Gaussian models or mixture Gaussian models [6]. On other hand, non-parametric models are usually color histograms built from many skin pixel observations.

Here we present two well known and established methods of modeling skin color, which we call "color histogram" and "naïve Bayes". Both methods are non-parametric and use only information about the color of pixels. The models represent the probability density functions of a color given that it is known to be from skin, $p(c|\text{skin})$, where c is the

pixel color feature vector. A probabilistic model for non-skin colors, $p(c|\text{non-skin})$, is also built. These models were built for every existing point in the color spaces used in this work.

A. Color Histograms

This type of probabilistic model is generated by counting the number of observations for each color in the database and dividing the result by the total count of pixels, thus obtaining the probability density function $p(c|\text{skin})$. The process is done with an auxiliary set of masks, each corresponding to an image in the database. In these masks, white pixels represent skin and black pixels represent non-skin in the corresponding location of the original input images. This set of auxiliary masks containing previous knowledge about the class of each pixel is called ground-truth.

B. Naïve Bayes

Treating each dimension of a the color feature vector c as an independent variable, it is possible to calculate separate histograms for each color vector component c_i. The histograms obtained represent the probability density function of each color component given that it is known to be from skin, $p(c_i|\text{skin})$. In order to find $p(c|\text{skin})$ it is possible to use the equation 9.

$$p(c|\text{skin}) = \prod_{i=1}^{d} p(c_i|\text{skin}), \qquad (9)$$

where d is the dimensionality of the color space being used and c is the color feature vector.

We used a variant of the CompaqDB skin database — the CompaqDBeasy500 — to build up our probabilistic models for skin color. This image database was created by Jones & Rehg [13]. In the CompaqDBeasy500 there are 500 images containing skin pixel samples and another 500 images containing no skin pixel samples.

IV. BAYESIAN CLASSIFIERS

Bayesian classifiers are a combination of probability models and decision rules. In this work we used the decision rule proposed by Jones & Rehg [13], which works by applying a threshold value over $p(\text{skin}|c)$, as shown in equation 10. The transformation from the probabilistic model $p(c|\text{skin})$ to $p(\text{skin}|c)$ is computed by equation 11.

$$class(c) = \begin{cases} \text{skin, if } p(\text{skin}|c) > \Theta \\ \text{non-skin, otherwise} \end{cases} \qquad (10)$$

$$p(\text{skin}|c) = \frac{p(c|\text{skin})p(\text{skin})}{p(c)} \qquad (11)$$

The decision rule applied to $p(\text{skin}|c)$ is invariant to *a priori* probabilities $p(\text{skin})$ and $p(\text{non-skin})$, which only affect the choice of threshold Θ. Equation 11 can be rearranged to find the relationship between $p(c|\text{skin})$ and $p(c|\text{non-skin})$ as

shown in equation 12. The value of Θ can be then computed by means of equation 13.

$$\frac{p(c|\text{skin})}{p(c|\text{non-skin})} < \Theta \qquad (12)$$

$$\Theta = K \frac{1 - p(\text{skin})}{p(\text{skin})}, \qquad (13)$$

where K is a parameter in the range $[0, 1]$, which is used to normalize Θ.

Using the rule above in each probabilistic model described in section III, it is possible to build two classifiers: the Jones & Rehg classifier and the naïve Bayes classifier.

V. ILLUMINATION-CONTROLLED IMAGE DATABASE

The previous knowledge probabilistic models used in this work for the classifiers were built using the CompaqD-Beasy500, since CompaqDB is a large and diversified image database for skin color. However, this database was built disregarding any information about illumination. Therefore, in order to have additional information about illumination conditions, we decided to build our own structured skin image database, which besides images also contains meta information about illumination conditions, camera used, regions of interest, among other details related to the subject, such as skin color, age and gender. Since we intend to study the influence of illumination in skin color, our database was only used at test stages, using previous knowledge obtained from the compaqDBeasy500 in a training stage.

Our image database is called skinDB and currently contains 120 equally divided pictures in four environment illumination conditions: "sun" (direct sun light), "shadow" (no direct source of light), "incandescent" (indoor incandescent lamp light) and "fluorescent" (indoor fluorescent lamp light). Samples of images in each environment illumination condition can be seen in figure 1, where it is possible to observe significant changes in skin color due to variations in illumination.

Our database includes a set of ground-truth masks, but it is also possible to specify rectangular regions of interest which are filled with skin pixels. The database and correspondent metadata for each image can be downloaded from http://www.labiem.cpgei.cefetpr.br/Members/diogo.

One important point about the skinDB images used as input for the classifiers in this work is that for none of them the automatic white balance of the camera was disabled. This fact can result in serious variations in classification rates and will be better discussed in section VI.

VI. RESULTS AND DISCUSSION

The primary results of our classification experiments are obtained in the form of skin maps, which are binary images with white pixels indicating what the classifier considered as skin-pixels and black pixels indicating non-skin. In figure 2
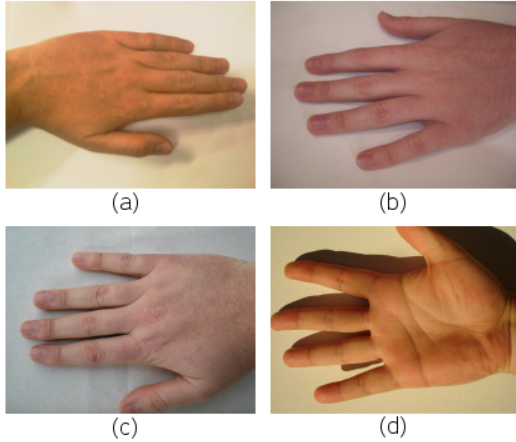
Figure 1. Image samples from skinDB: (a) Image under fluorescent light ; (b) Image under incandescent light; (c) Image under shadow; (d) Image under sun light
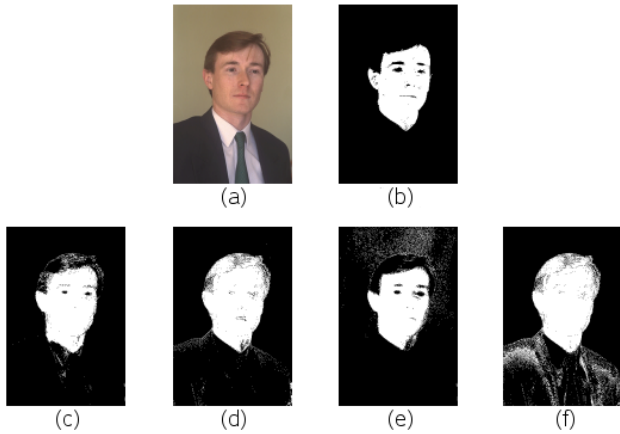

Figure 2. Classification example: (a) Original image; (b) Ground-truth; (c) Result using the Jones & Rehg Classifier and the CbCr color space; (d) Result using the Jones & Rehg Classifier and the normalized RG color space; (e) Result using the naïve Bayes Classifier and the CbCr color space; (f) Result using the naïve Bayes Classifier and the normalized RG color space.
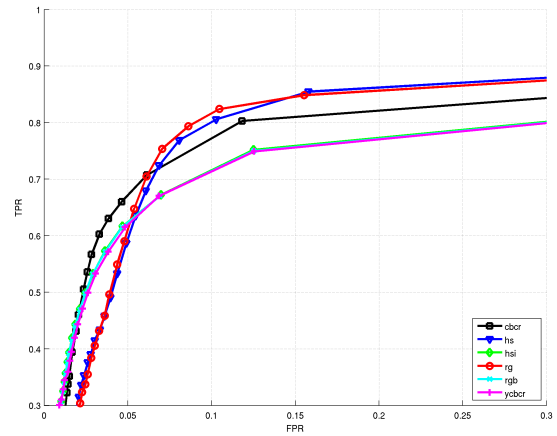

Figure 3. ROC curves for the Jones & Rehg classifier in all used color spaces.
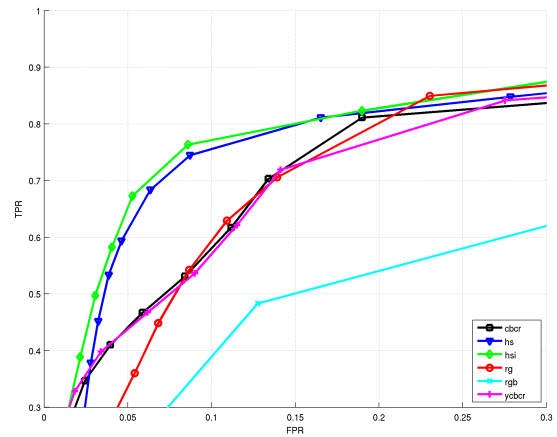
Figure 4. ROC curves for the naïve Bayes classifier in all used color spaces.

there is an example of an image from the CompaqD-Beasy500 database, with its ground-truth and some skin maps for different models, color spaces and threshold values.

The evaluation of classifier performance is made trough ROC curves, which holds information about the rate of incorrect classification (FPR - false positive rate) versus the rate of correct classification (TPR - true positive rate). Each point of the curve refers to a different threshold value. The desired value of threshold is the one which minimizes the value of FPR while maximizing the value of TPR. Having said so, curves that show sharper knees close to the point (0,1) indicate better classification results. The values of TPR and FPR are expressed in percentage of pixels correctly and incorrectly classified, respectively.

Before entering into the merit of evaluating the effects of environment illumination conditions in skin classification, we performed an overall skin classification evaluation. This evaluation was done using the entire skinDB database for both classifiers (Jones & Rehg and naïve Bayes) and all six color spaces (RGB, normalized RG, HSI, HS, YCbCr and CbCr) that were presented earlier. Figure 3 shows the ROC curves obtained for the Jones & Rehg classifier and figure 4 shows the ROC curves obtained for the naïve Bayes classifier.

The overall skin pixel classification ROC curves shown in figures 3 and 4 indicate that, in most of the cases, two-dimensional color spaces containing only chrominance information provide better results, which in this case are RG/HS for the Jones & Rehg classifier and HSI/HS for the naive Bayes . Comparing our results using the Jones

& Rehg classifier with the work of others [13], [14], two-dimensional color spaces yielded poorer results than tree-dimensional color spaces. This is possibly due to the fact that skinDB is a small database which was especially designed and whose images were acquired in a controlled way, having skin classification studies in mind.

Another important result is that ROC curves for classifications using the Jones & Rehg classifier and any of the tree-dimensional color spaces yield very similar results (figure 3). This shows that there is no loss of information due to the transformation from RGB to other tree-dimensional color spaces, as one should expect.

When the naïve Bayes classifier ROC curves are analyzed, the only two color spaces that maintain similar results are the HS and HSI color spaces (figure 4). These similar result for HS and HSI shows that these color spaces have a higher independence between color components, i.e. the probability of a pixel corresponding to skin does not depend on the complete color probability but on independent color component probabilities. Something similar occurs with CbCr and YCbCr color spaces. The results may be poorer than the ones obtained for the Jones & Rehg classifier, but the corresponding ROC curves remain close to each other, showing that the Y component does not affect the result and the independence between YCbCr channels can be supposed. Finally, the results in figure 4 show that RGB color space has a high correlation between its color channels, opposed to what is observed for HSI and YCbCr.

The comparison of ROC curves for both classifiers using images acquired in four different illumination conditions results in an interesting discussion. First of all, we analyzed the influence of illumination for the Jones & Rehg classifier, whose results are shown in figures 5, 6, 7 and 8. We also analyzed the influence of illumination for the naïve Bayes classifier, whose results are shown in figures 9, 10, 11 and 12.

A. Jones & Rehg Classifier

The results for the Jones & Rehg classifier show that three-dimensional color spaces yield better classification rates than two-dimensional color spaces from the origin to a specific false detection rate point on the ROC chart. When the FPR is lower than 5% under sun and incandescent light, three-dimensional color spaces yield better classification. However, when higher false detection rates are acceptable, two-dimensional color spaces constitute a better choice. Under shadow illumination, the same happens, but for a lower FPR (approximately 2,5%). Another interesting result is that for all illumination conditions, three-dimensional color spaces yield extremely similar ROC curves.

Using images taken under sunlight (figure 5), the CbCr color space provides the best results, 86% TPR at 10% FPR. Sunlight illumination yielded the most divergent results, especially for normalized RG and HS color spaces.
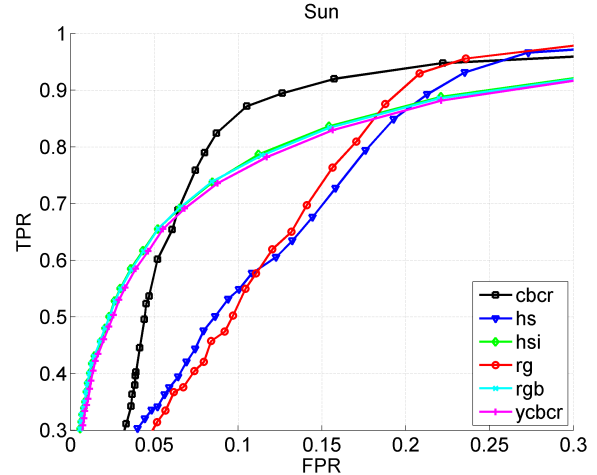

Figure 5. ROC curves for the Jones & Rehg classifier using images taken under sunlight.
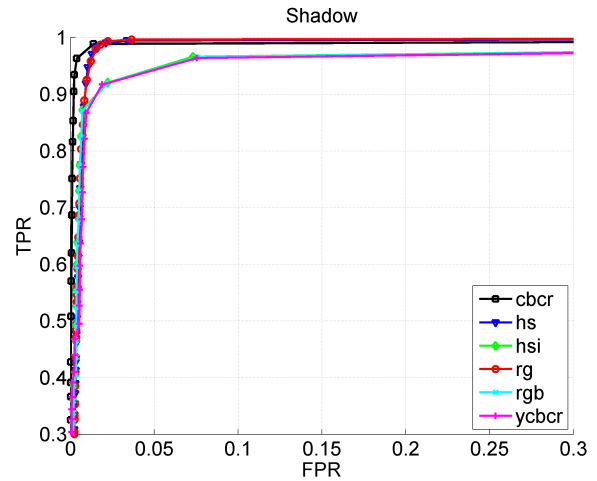

Figure 6. ROC curves for the Jones & Rehg classifier using images taken under shadow.

The best classification performance for the Jones & Rehg classifier was obtained using images taken under shadow (figure 6). All color spaces that were used reached TPR values that are higher than 95% for FPR values lower than 2.5%. We believe that this occurs because the camera's automatic white balance has less influence under shadow, consequently preserving skin color.

For images taken under incandescent light (figure 7), the results were not as good as under shadow and not as poor as under sunlight. Once again, the hypothesis is that the automatic white balance of the camera is responsible for the variation in classification rates. For incandescent illumination, the CbCr color space yielded the best classification rate, 92.5% TPR at 10% FPR.

Table I summarizes the classification rates obtained for the Jones & Rehg classifier for all illumination conditions and
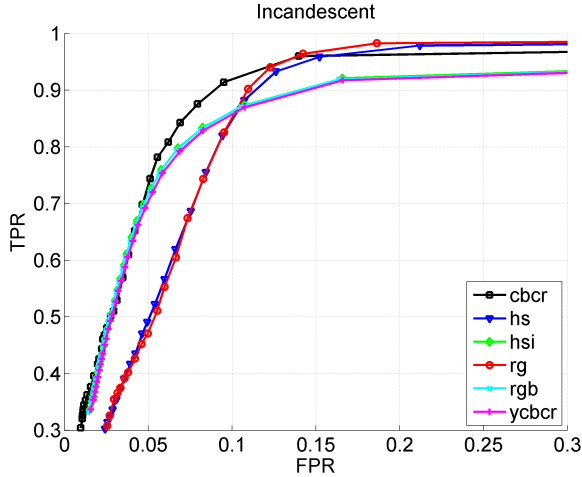
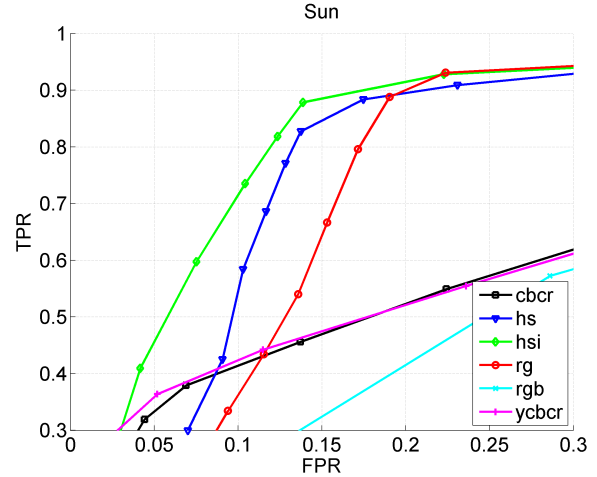Figure 7. ROC curves for the Jones & Rehg classifier using images taken under incandescent light.

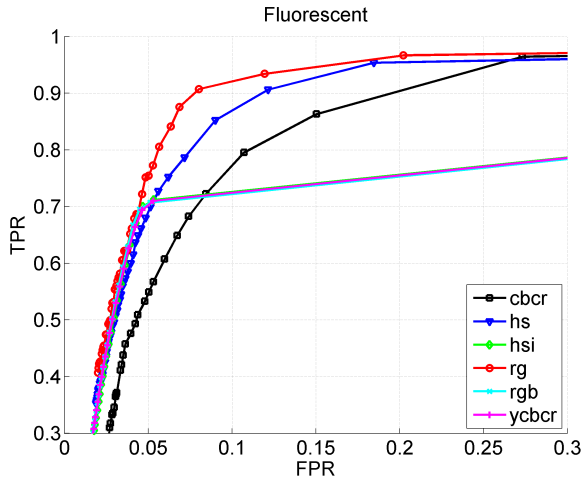Figure 9. ROC curves for the naïve Bayes classifier using images taken under sunlight.

Table II
TPR IN PERCENTAGE FOR THE NAÏVE BAYES CLASSIFIER AT 10% FPR

	Sun	Shadow	Incandesc.	Fluoresc.
RGB	23.3%	42.9%	45.5%	77.4%
RG	36.2%	95.2%	43.4%	<20%
HSI	71.5%	99.5%	85.2%	42.9%
HS	54.5%	99.8%	74.2%	44.4%
YCbCr	42.4%	99.6%	88.9%	26.0%
CbCr	41.4%	99.5%	89.0%	75.3%

Figure 8. ROC curves for the Jones & Rehg classifier using images taken under fluorescent light.

Table I
TPR IN PERCENTAGE FOR THE JONES & REHG CLASSIFIER AT 10% FPR

	Sun	Shadow	Incandesc.	Fluoresc.
RGB	76.3%	96.5%	86.3%	75.5%
RG	52.1%	99.6%	85.2%	92.1%
HSI	76.5%	96.5%	86.3%	75.5%
HS	54.7%	99.5%	84.7%	87.0%
YCbCr	75.5%	96.4%	85.6%	75.5%
CbCr	87.1%	99.0%	91.9%	77.3%

color spaces. In this table, the FPR value is fixed at 10% as is commonly done in the literature, for a better comparison of TPR values.

B. Naïve Bayes Classifier

The naïve Bayes classifier supposes that color components are statistically independent from each other. Although it is known that in most cases statistically independence does not hold, the naïve Bayes classifier is used to investigate which are color spaces with the most statistically independent components for skin classification, while preserving machine resources (memory and processing power).

As previously stated, color channels in RGB color space are highly correlated. The effect of this characteristic in the performance of the naïve Bayes classifier can be observed in figures 9, 10 and 11, since for all curves the poorest results are from the RGB color space. For yet unknown reasons, the ROC curve obtained for images taken under fluorescent light (figure 12) shows relatively good results for the RGB color space.

For images taken under sun light (figure 9), the best results were achieved using the HSI, HS and normalized RG color spaces, reaching 90% TPR at 20% FPR.

Under shadow (figure 10), the results were very similar to the ones obtained for the Jones & Rehg classifier, except for the RGB color space. Acquiring skin images under shadow results in more than 95% TPR at less than 5% FPR for the naïve Bayes classifier. This indicates that all used color spaces but RGB show little channel correlation, but illumination conditions distort the appearance of skin pixels.

Table II summarizes the TPR values obtained for all color spaces and illumination conditions experimented in this work.
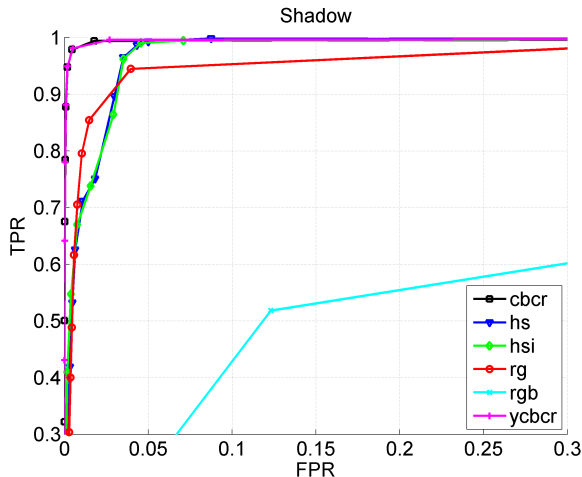
Figure 10. ROC curves for the naïve Bayes classifier using images taken under shadow.
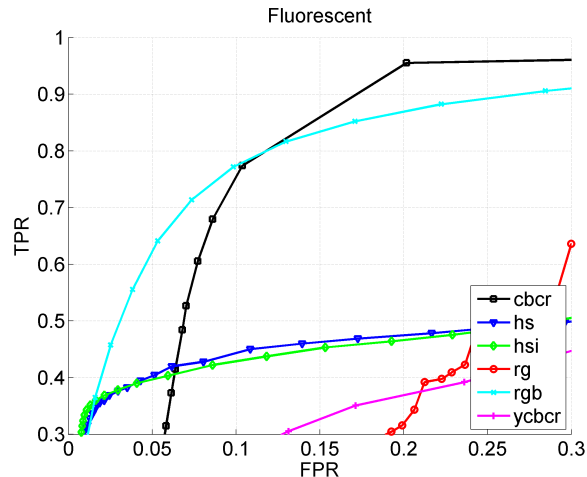

Figure 12. ROC curves for the naïve Bayes classifier using images taken under fluorescent light.
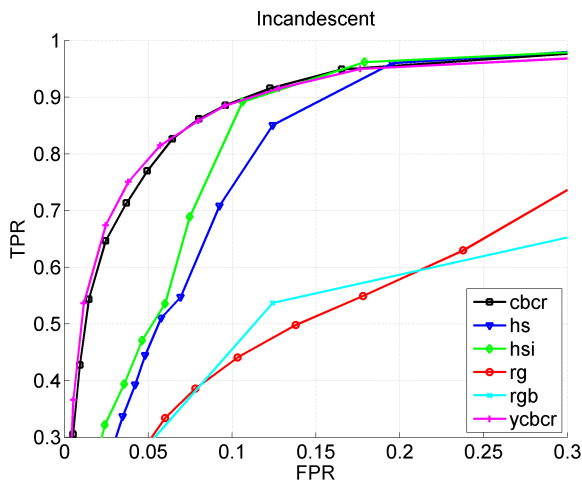

Figure 11. ROC curves for the naïve Bayes classifier using images taken under incandescent light.

VII. CONCLUSION

Skin color is a very attractive visual feature for applications involving the detection of skin pixels, since it provides fast and robust processing. However, illumination conditions and automatic white balance camera features have a large influence on skin color, leading us to investigate which are the existing color models that minimize this influence on classification performance.

Our experiments show that the use of two-dimensional color spaces yield better results by discarding luminance information and thus reducing the influence of illumination. Results obtained with two-dimensional color spaces show that using partial color information is sufficient for reasonably good skin-pixel classification.

Further investigation is needed to clear up apparently contradicting results between our experiments and experiments

reported by others [13], [14]. Future work also includes improving the skinDB database by using other control parameters, such as camera white balance, type of background, anonymous identification of subjects photographed and distance from the camera. This more controlled scenario would contribute to make skin color variations mostly due to the effects of illumination.

VIII. ACKNOWLEDGEMENTS

REFERENCES

[1] D. Chai and K. N. Ngan, "Face segmentation using skin-color map in videophone applications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 4, pp. 551–564, June 1999.

[2] M.-H. Yang and N. Ahuja, "Detecting human faces in color images," in *Proc. Int. Conf. on Image Processing*, vol. 1, 1998, pp. 127–130.

[3] P. Yan and K. W. Bowyer, "Biometric recognition using 3D ear shape," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 8, pp. 1297–1308, August 2007.

[4] M. M. Fleck, D. A. Forsyth, and C. Bregler, "Finding naked people," in *Proc. 4th European Conf. on Computer Vision*, vol. 2, 1996, pp. 592–602.

[5] J. Han, G. M. Award, A. Sutherland, and H. Wu, "Automatic skin segmentation for gesture recognition: Combining region and support vector machine active learning," in *Proc. 7th Int. Conf. on Automatic Face and Gesture Recognition.*, 2006, pp. 237–242.

[6] V. Vezhnevets, V. Sazonov, and A. Andreeva, "A survey on pixel-based skin color detection techniques," in *Proc. GraphiCon 2003*, 2003, pp. 85–92.

[7] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modeling and detection methods," *Pattern Recognition*, vol. 40, no. 3, pp. 1106–1122, March 2007.

[8] T. Fawcett, "ROC graphs: Notes and practical considerations for researchers," HP Laboratories, Tech. Rep., 2004.

[9] M. Störring, "Computer vision and human skin colour," Ph.D. dissertation, Aalborg University, Denmark, 2004.

[10] M. C. Shin, K. I. Chang, and L. V. Tsap, "Does colorspace transformation make any difference on skin detection?" in *Proc. 6th IEEE Workshop on Applications of Computer Vision*, 2002, pp. 275–279.

[11] B. D. Zarit, B. J. Super, and F. K. H. Quek, "Comparison of five color models in skin pixel classification," in *Proc. Int. Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems*, 1999, pp. 58–63.

[12] A. Albiol, L. Torres, and E. J. Delp, "Optimum color spaces for skin detection," in *Proc. Int. Conf. on Image Processing*, vol. 1, 2001, pp. 122–124.

[13] M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection," *International Journal of Computer Vision*, vol. 46, no. 1, pp. 81–96, January 2002.

[14] S. L. Phung, A. Bouzerdoum, and D. Chai, "Skin segmentation using color pixel classification: Analysis and comparison," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 1, pp. 148–154, January 2005.