



# Human Skin Detection :Using Fusion Approach with IBy Color Space

<sup>1</sup>Pratibha B. Khadse, <sup>2</sup>Shailaja Gaikwad

<sup>1,2</sup>Department of Electronics, Terna College of Engineering, India

**Abstract :** Skin detection is a process of finding skin colored pixels and regions in an image or a video. An effective skin detection process is the one which differentiates skin region under any illumination conditions and works well with persons of any ethnicity. In doing so choice of a suitable color space is of utmost importance. Different skin detection methods with different color space have been successfully applied but they are prone to false skin detection and also not able to cope up with the variety of human skin colors . Additionally requires high computational costs. In this paper a dynamic threshold approach with log opponent color space is implemented which reduces computational costs as no training is required, and it further improves the accuracy of skin detection. Additional feature is implemented i.e. fusion of a histogram model and Gaussian model for automatic human skin detection in color image(s) that gives good results when either of this method is employed individually.

**Index Terms**—Color space, dynamic threshold, fusion strategy, Gaussian Model, skin detection

## I. INTRODUCTION

Skin color and textures are important cues that people use consciously or unconsciously to infer variety of culture-related aspects about each other. Skin color and texture can be an indication of race, health, age, wealth, beauty, etc. [1]. However, such interpretations vary across cultures and across the history. Skin detection means detecting image pixels and regions that contain skin-tone color. Skin detection has a wide range of applications both in human computer interaction and content based analysis. Some applications such as: detecting and tracking of human body parts [2], face detection [3], naked people detection, people retrieval in multimedia databases [4] and blocking objectionable content [5], all involves skin detection as the primary step. The skin detection process has to be robust to large variations in appearance that can occur. Variation in skin color and shape is often affected by occlusion

(such as clothing, hair, eye glasses, etc.). Moreover, changes in intensity, color, and location of light sources also affect appearance of the skin. Other objects within the scene may even cast shadows or reflect additional light and thereby changing the skin color. Finally, there are many other objects which are easily confused with skin: certain types of wood, copper, sand as well as clothes often have skin-like colors. All this factors makes the skin detection process quite challenging for researchers.

One of the easiest and commonly used human skin detection technique is to define a fixed decision boundary for different color space components . Single or multiple ranges of threshold values for each color space components are defined and the image pixel values that fall within these predefined range(s) are selected as skin pixels. All this approaches are based on skin classification based on color information. As for any given color space, skin color occupies a part of such a space, which might be a compact or large region in the space. Other approaches are multilayer perceptron [6]–[8], Bayesian classifiers [9]–[11], and random forest [12]. These approaches uses single features, and are successfully applied to human skin detection but still suffer from high false detection that is low accuracy and requires large training stage for finding threshold value(s) for detecting human skin.

In this paper, we propose a fusion approach that is fusion of two features : the histogram model and Gaussian model to perform automatic skin detection. Firstly, we employ an online dynamic approach as in [13] with IBy color space to calculate the skin threshold value(s). thereby eliminating any training stage beforehand. Secondly, a histogram model and a Gaussian model are used to model the skin and non-skin distributions, respectively. Finally, a fusion of two features is employed to perform automatic skin detection.

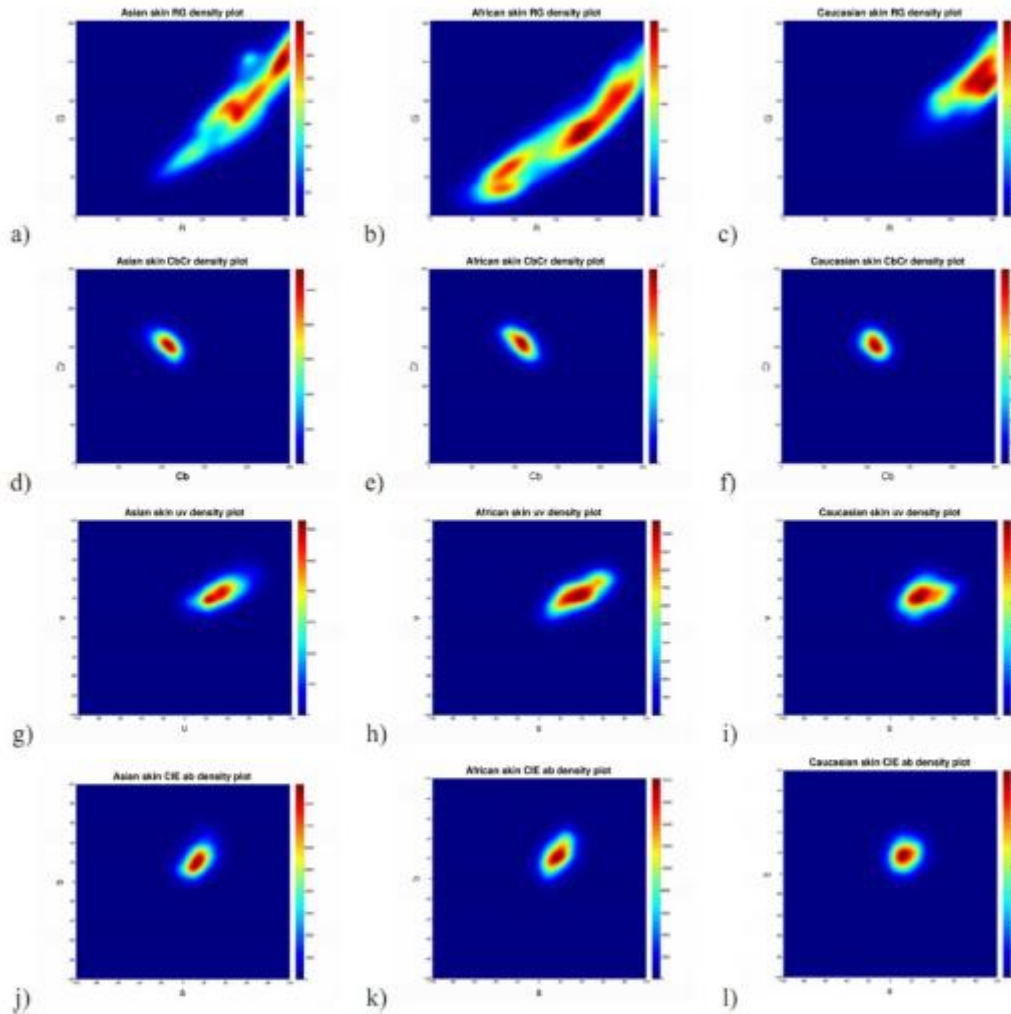


Figure. 1. Density plots of Asian, African, and Caucasian skin in different color spaces. Each row represents different color space and columns from left to right represent Asian, African, and Caucasian, respectively.

## II. RELATED WORK

Skin detection is the process of finding skin-color pixels and regions in an image or video. A skin detector typically converts a given pixel into a selected color space and then uses a skin classifier to label the pixel whether it is skin or non-skin. A skin classifier then defines a decision boundary of the skin-color class in the color space based on a training database of skin-color pixels.

For example, Sobottka and Pitas[14] used fixed range values on the HS color space where the pixel values belong to skin pixels in the range of  $R_H=[0,50]$  and  $R_S=[0.23, 0.68]$ . Wang and Yuan [15] used threshold values in RG space and HSV space where threshold values are set to be within the range  $R_r=[0.36, 0.465]$   $R_g=[0.28, 0.363]$ ,  $R_h=[0,50]$ ,  $R_s=[0.20, 0.68]$  and  $R_v=[0.35, 1.0]$  to differentiate skin and non-skin pixels. Dai and Nakano [16] used a fixed range on I component in YIQ space for detecting skin pixels from images containing mostly people with yellow skin. All the pixel values in the range,  $RI = [0; 50]$  are described as skin pixels in this approach. Chai and Ngan [17] proposed a face

segmentation algorithm in which they used a fixed range skin-color map in the CbCr plane. The pixel values in the range  $RCb = [77; 127]$ , and  $RCr = [133; 173]$  are defined as skin pixels. In these approaches, high false skin detection is a common problem when there are a wide variation in skin colors across different ethnicity, complex backgrounds, and high illumination. The skin color of people belonging to Asian, African, and Caucasian groups is different from one another as in Fig 1. hence some robustness has been achieved via the use of a suitable color space.

Other than the color space, it is noticeable that most of the fixed threshold based skin classifiers were successfully applied in controlled imaging conditions, e.g. segment white and yellow skins. The fixed threshold values vary from one color space to another and also under different illumination conditions. It is thus very difficult to find a range of threshold values that covers all the subjects of different skin color. This technique is less accurate in case of shadows and situations where the skin color is not distinguishable from background. Thus in order to handle uncontrolled imaging conditions fuelled a need to have a dynamic threshold classification

technique and so dynamic threshold approach [19] is used.

### III. METHOD

Figure 2. shows the block diagram of the method used for skin detection

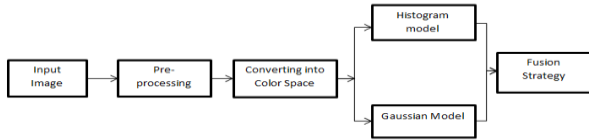


Figure 2. Block diagram of the Proposed System

#### A) Pre-processing

Initially face region is cropped from the image so as to get the optimal threshold values for skin detection, as we are using the assumption that the face and body of a person always share the same colors. The cropped region now has smooth (i.e., skin) and non-smooth regions (i.e., eyes, eye brown, mouth, etc.). As we are only interested in smooth regions, canny edge detection is employed to remove non-smooth regions. Then, the detected edge pixels are further dilated using a dilation operation to get the optimal non-smooth regions. The calculated non-smooth region is then subtracted from the face region and smooth region is obtained. Finally we get smooth regions in the cropped face, generating a new image  $S'$ .

#### B) Color Space

An image can be represented in a number of different color space models. It is important to select the right color space for modeling human skin color. Here we propose the use of the LO color space because of two reasons: first, color opponency is perceptually relevant as it has been proved that the human visual system uses an opponent color encoding; and second, in this color space, the use of logarithms renders illumination change to a simple translation of co-ordinates. The region obtained from preprocessing step is then converted into LO color space.

The conversion formula is given by:

$$I = L(G)$$

$$R_g = L(R) - L(G)$$

$$B_y = L(B) - (L(R) + L(G)) / 2$$

$$\text{Where } L(x) = 105 + \log_{10}(x+1)$$

#### C) Histogram Model

An image Histogram is a type of histogram that acts as a graphical representation of the lightness/color distribution in a digital image. It plots the number of pixels for each value. Image histograms can be useful tools for thresholding.

The lower and upper boundaries of the dynamic threshold values are found for both the color channels,  $I$  and  $B_y$ . Based on these calculated dynamic threshold

values each and every pixel in the color image is classified as skin and non-skin.

#### D) Gaussian Model

The Gaussian model is a sophisticated model that is capable of describing complex-shaped distributions and is popular for modeling skin-color distributions.

The threshold skin-color distribution in the histogram would be modeled through normal probability distribution functions defined as

$$y = f(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Where  $\mu$  is mean of the function and  $\sigma$  is the standard deviation.

The graph of a Gaussian is a characteristic symmetric "bell curve" shape that quickly falls off towards zero. Gaussian functions are widely used in statistics where they describe the normal distributions, in signal processing where they serve to define Gaussian filters, in image processing where two-dimensional Gaussians are used for Gaussian blurs.

The mean and standard deviation is calculated for each color channel i.e.  $I$  and  $B_y$ . Finally compared to the values of the main image to get the final detected skin regions.

#### E) Fusion Strategy

Finally in order to increase the effectiveness and robustness of the skin detection algorithm, a fusion strategy is proposed by integrating the two incoming single features into a combined single representation. This can be done by applying AND rule to both models.

## IV. RESULTS AND ANALYSIS

The performance of the proposed approach under different conditions, such as fusion strategy, color spaces, and a comparison with the state-of-the-art methods. Experiments are conducted using The Pratheepan's dataset [13] which consists of a set of images downloaded randomly from Google. These random images are captured with a range of different cameras using different color enhancements and under different illuminations. Some images are taken from the ETHZ pascal dataset and some are taken randomly from the internet. And some are from Stottinger dataset [18].

The Proposed system is the fusion of two individual skin detection methods that is Histogram Method and the Gaussian method. The fusion result is compared with the individual methods as shown in the Figure 3. Black being non-skin region while white being skin region. Algorithm assumes that the face and body shares the same color.

Table 1. Comparison between our proposed method and used individual methods

Classifier	Accuracy	Precision	F-score	True Positive Rate	False Positive Rate
Fusion	0.8975	0.8790	0.8263	0.7796	0.1210
Histogram	0.8072	0.6268	0.7626	0.9735	0.3732
GMM	0.8902	0.8145	0.8075	0.8006	0.1855



Figure 3. Column from left to right represents original images , Gaussian model’s result , Histogram result and Fusion approach’s result.

From the table 1 it can be seen that the Fusion approach has the highest accuracy , precision and F-score then the individual approaches. Moreover, it can also be observed that the fusion strategy has lower false positive rate compared to the single feature approach. Figure 3 visualises the difference as well. The histogram approach is not able to detect most of the skin regions while Gaussian model performs well compared to histogram method .

Apart from comparing with the individual methods that are involved, the proposed method is also compared with the other state-of-art methods .Figure 4 shows the visual comparison while table 2 shows the associated improved values.

Figure 4 shows the visual comparison between Proposed Approach and the Random Forest[12] approach. For random forest [12], 1990 image frames are randomly chosen for training and remaining images are used for testing. From those 1990 images, around 3 million pixels were randomly chosen and 15 trees were trained. Each tree extracts 70% of the pixels, respectively, for training , which consumed time in minutes.

As seen from Figure4 ,in all the images Random forest method failed to detect the skin regions while the proposed approach detects it but with some discrepancy as well. In order to increase the accuracy of random forest on this dataset, huge training samples and/or more

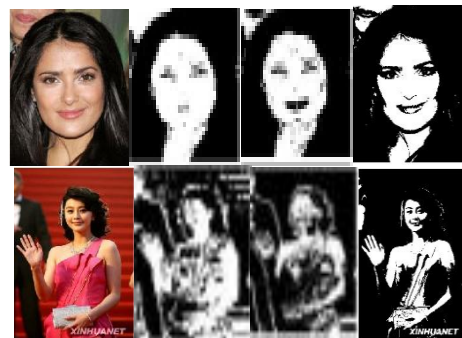
trees will be needed. This will cause higher computational power as number trees increases and time consuming during training. Quantitatively from the table 2, which compares the fusion method with other state-of-art methods, proves that the proposed method has improved accuracy , precision and recall.

Table 2. Comparison between Different Classifier performance

Classifier	Accuracy	Precision	Recall
Proposed Method	0.892	0.971	0.633
Random Forest [12]	0.877	0.738	0.740
Static Threshold [19]	0.833	0.474	0.557



Figure 4. Comparison between Random Forest [12] and Proposed Method. (a)Original Image. (b) Random Forest Result and (c) Proposed Method Result



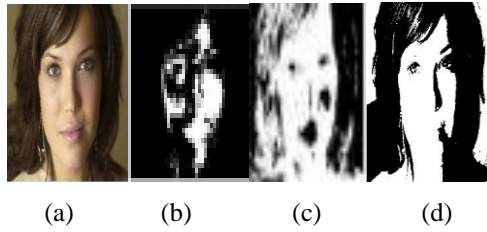


Figure 5. (a) Original Image (b) Method[19] (c) Method [13] and (d) Proposed Method

Figure 5 shows the comparison between two methods [13] and [19] with our method. It is observed that our method is robust in terms of illumination, background image, different camera characteristic, and different ethnicity. It is able to detect skin region appropriately then the other two method and with least noise.

For example in image 2 (read from top to bottom) the method [13] and [19] are not able to detect even the face region. While there is noise in mostly the skin detected images. One advantage of our method is it doesn't require training phase hence reducing the computational cost when compared to [13] and [19].

Here we have tested the proposed system using different color space i.e. CbCr color space. It can be observed that IBy color space shows better true positive rate than CbCr space from Fig and its associated parameter values. However, we selected IBy because it has been proven that the human visual system uses an opponent color coding.



Figure 2. Comparison between CbCr and IBy space (a) Original Image. (b) CbCr Result and (c) IBy Result

Table 3. Comparison between IBy and CbCr color space

Color Space	Accuracy	True Positive Rate	False Negative Rate
IBy	0.9011	0.8587	0.1413

CbCr	0.6658	0.8032	0.1968
------	--------	--------	--------

### V. CONCLUSION

Thus a dynamic approach, with a fusion framework based on histogram model and Gaussian model has been proposed for automatic human skin detection in images. The results when compared with other methods found to be better. However the proposed method does fails for improper eye detection and optimal cropping of the face. When a false face region is obtained, false dynamic thresholds will be generated resulting in false detection of skin regions.

### REFERENCES

- [1] A. M. Elgammal, C. Muang, and D. Hu, "Skin Detection," in Encyclopedia of Biometrics, Germany, Berlin, Springer, pp. 1218–1224, 2009.
- [2] Antonis A. Argyros and Manolis I.A. Lourakis, "Realtimetracking of multiple skin-colored objects with a possibly moving camera," in ECCV, 2004, pp. 368–379.
- [3] J. Cai and A. Goshtasby, "Detecting human faces in color images," Image and Vision Computing, vol. 18, pp. 63–75, 1999.
- [4] Liang-Liang Cao, Xue-Long Li, Neng-Hai Yu, and Zheng-Kai Liu, "Naked people retrieval based on adaboost learning," in MLC 2002, 2002, vol. 2, pp. 1133–1138.
- [5] Julian Stöttinger, Allan Hanbury, Christian Liensberger, and Rehanullah Khan, "Skin paths for contextual flagging adult videos," in ISVC (2), 2009, pp. 303–314.
- [6] D. Brown, I. Craw, and J. Lewthwaite, "A SOM based approach to skindetection with application in real time systems," in Proc. Brit. Mach. Vis. Conf., 2001, pp. 491–500.
- [7] M.-J. Seow, D. Valaparla, and V. K. Asari, "Neural network based skin color model for face detection," in Proc. Appl. Image Pattern Recognit. Workshop, 2003, pp. 141–145.
- [8] S. L. Phung, D. Chai, and A. Bouzerdoum, "A universal and robust human skin colour model using neural network," in Proc. Int. Joint Conf. Neural Netw., 2001, vol. 4, pp. 2844–2849.
- [9] N. Sebe, I. Cohen, T. S. Huang, and T. Gevers, "Skin detection: A Bayesian network approach," in Proc. Int. Conf. Pattern Recognit., 2004, pp. 903–906.
- [10] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," Mach. Learn., vol. 29, pp. 131–163, Nov. 1997.

- [11] M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection," *Int. J. Comput. Vision*, vol. 46, no. 1, pp. 81–96, 2002.
- [12] R. Khan, A. Hanbury, and J. Stottinger, "Skin detection: A random forest approach," in *Proc. Int. Conf. Image Process.*, Hong Kong, 2010, pp. 4613–4616.
- [13] P. Yogarajah, A. Cheddad, J. Condell, K. Curran, and P. McKeivitt, "A dynamic threshold approach for skin segmentation in color images," in *Proc. Int. Conf. Image Process.*, 2010, pp. 2225–2228.
- [14] K. Sobottka and I. Pitas, "A novel method for automatic face segmentation, facial feature extraction and tracking," *Signal Process.: Image Commun.*, vol. 12, no. 3, pp. 263–281, 1998.
- [15] Y. Wang and B. Yuan, "A novel approach for human face detection from color images under complex background," *Pattern Recognit.*, vol. 34, no. 10, pp. 1983–1992, 2001.
- [16] Y. Dai and Y. Nakano, "Face-texture model based on GSD and its application in face detection in a color scene," in *Pattern Recognition*, 1996, pp. 1007–1017.
- [17] D. Chai and K.N. Ngan, "Face segmentation using skin color map in videophone applications," in *IEEE Trans. Circuits Syst. Video Technol.*, 1999, pp. 551–564.
- [18] J. Stottinger, A. Hanbury, C. Liensberger, and R. Khan, "Skin paths for contextual flagging adult video," in *Proc. Int. Symp. Visual Comput.*, 2009, pp. 903–906.
- [19] A. Cheddad, J. Condell, K. Curran, and P. McKeivitt, "A skin tone detection algorithm for an adaptive approach to steganography," *J. Signal Process.*, vol. 89, pp. 2465–2478, 2009.

