Fusion Technique for Human Skin Detection

Mr. Shoeb I. Shaikh
PG student, Department of
Electronics & Telecommunication,
Saraswati College of engineering,
Kharghar, Navi Mumbai, India

Dr. Manjusha Deshmukh
Principal of Saraswati College of
engineering,
Kharghar, Navi Mumbai, India

***Abstract*—Human skin color is an important cue to infer variety of aspects including beauty, race, culture and age etc. Detection of skin color of human is most importance in numerous applications including gesture analysis, recognizing human by human and/or machine and face tracking. Detection of skin color pixels and non skin color pixels and its classification is quite challenging task. The human visual systems consolidate color opponency. Also, in an image the skin color is sensitive to various factors such as camera characteristics, ethnicity, hairstyle, makeup, shadows, illumination, motion background colors, also influence skin color appearance. A reliable human skin detection method that is adaptable to different human skin colors and illumination conditions is essential for better human skin segmentation. Although various human skin color detection solutions have been effectively applied, they prostrate with false skin detection and are not able to cope with the variety of human skin colors across different ethnic. Also, existing methods need high computational cost. This work aimed at providing a technique which provides more robust and accurate results with minimum computational cost. The suggested mixture technique combines a smoothed histogram and Gaussian model for automatic human skin detection in color images. This mixture technique reduces computational costs as no training is required and it also reduces the false positive rates and improves the accuracy of skin detection despite wide variation in illumination, ethnicity and background.**

Keywords— Nose localization, color extraction, Fusion technique, Skin detection.

I. INTRODUCTION

With the progress of information society today, images have become more and more important. Among them, skin detection plays an important role in a wide range of image processing applications from face tracking, gesture analysis, content based image retrieval systems to various human-computer interaction domains. In these applications, the search space for objects of interests, such as hands, can be reduced through the detection of skin regions. One of the simplest and commonly used human skin detection methods 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 pre defined ranges are selected as skin pixels. In this approach, 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, Bayesian classifiers and random forest. These aforementioned solutions, although, successfully applied to human skin detection; they still suffer from low accuracy, luminance invariant space and they required large training samples.

The proposed novel approach fusion framework uses product rules on two features, the smoothed histogram and Gaussian model to perform automatic skin detection. First of all, we employ Viola Jones object detection algorithm to detect nose region in an image. From the detected nose region we will extract skin tone values in RGB color space. Secondly, we are going to use smoothed histogram and Gaussian model to distribute skin and non-skin region, respectively. Finally, we will employ a fusion strategy framework using the product of 2 features to perform automatic skin detection.

The image pixels representation in a suitable color space is the primary step in skin segmentation in color images. In our approach, we choose the log RGB color space. The RGB color space consists of the three additive primaries: red, green and blue. Spectral components of these colors combine additively to produce a resultant color. The RGB model can be represented by a 3-dimensional cube with red, green and blue at the corners on each axis. Black is at the origin, white is at the opposite end of the cube. The gray scale follows the line from black to white. The RGB model simplifies the design of computer graphics systems . The red, green and blue color components are highly correlated.

For classification of skin and non-skin region we are using histogram and Gaussian model. Histogram acts as a graphical representation of the tonal distribution in a digital image. It plots the number of pixels for each tonal value. By looking at the histogram for a specific image a viewer will be able to judge the entire tonal distribution at a glance. The horizontal axis of the graph represents the tonal variations, while the vertical axis represents the number of pixels in that particular tone.

The other classifier which we are using is Gaussian model. Gaussian distribution is a very common continuous probability distribution. Gaussian distributions are important in statistics and are often used in the natural and social sciences to represent real valued random variables whose distributions are not known. Here we are using Gaussian model to classify skin and non-skin region.

The other sections of this work are structured as follows. Section 2 gives a brief description of related work. Section 3 introduces the problem definition and Section 4 describes our proposed human skin detection technique followed experimental results and conclusion.

II. LITERATURE REVIEW

Skin detection is the process of finding skin color pixels and regions in an image or video. In images and videos, skin color is an indication of the existence of humans in media. In one of the early applications, detecting skin color regions was used to identify nude pictures on the Internet for content filtering. In another early application, skin detection was used to detect anchors in TV news videos for the sake of video automatic annotation, archival, and retrieval [1] [2]. Detecting skin pixels are rather computationally easy task and can be done very efficiently, a feature that encourages the use of skin detection in many video analysis applications. For example, in one of the early applications, detecting skin colored regions was used to identify nude pictures on the internet for the sake of content filtering [2]. The choice of the color space affects greatly the performance of any skin detector and its sensitivity to change in illumination conditions. As was highlighted by Forsyth and Fleck [3] the human skin color has a restricted range of hues and is not deeply saturated, since the appearance of skin is formed by a combination of blood (red) and melanin (brown, yellow). Therefore, the human skin color does not fall randomly in a given color space, but clustered at a small area in the color space. But it is not the same for all the color spaces. Variety of color spaces has been used in skin detection literature with the aim of finding a color space where the skin color is invariant to illumination conditions. RGB color space is the most commonly used color space in digital images. However, it is not perceptually uniform, which means distances in the RGB space do not linearly correspond to human perception. Despite these fundamental limitations, RGB is extensively used in skin detection literature because of its simplicity. For example, RGB is used by Rehg and Jones [4] and yield quite satisfying performance. Different classes of color spaces are the orthogonal color spaces used in TV transmission. This includes YUV, YIQ, and YCbCr. YIQ is used in NTSC TV broadcasting while YCbCr is used in JPEG image compression and MPEG video compression. One advantage of using these color spaces is that most video media are already encoded using these color spaces. Transforming from RGB into any of these spaces is a straight forward linear transformation [5]. All these color spaces separate the illumination channel (Y) from two orthogonal chrominance channels (UV, IQ, CbCr).Therefore, the location of the skin color in the chrominance channels will not be affected by changing the intensity of the illumination. Therefore, using such color spaces results in skin detectors which are invariant to human race. The simplicity of the transformation and the invariant properties made such spaces widely used in skin detection applications [6] [7] [8] [9] [10].

A variety of classification techniques have been used in the literature for the task of skin classification. A skin classifier is a one class classifier that defines a decision boundary of the skin color class in a feature space. The simplest way to decide whether a pixel is skin color or not is to explicitly define a boundary. Brand and Mason [11] constructed a simple one dimensional skin classifier; a pixel is labeled as a skin if the ratio between its R and G channels is between a lower and an upper bound.

In summary a skin detector typically transforms a given pixel into an appropriate color space and then uses a skin classifier to label the pixel whether it is skin or non skin [1]. A skin classifier defines a decision boundary of the skin color class in the color space based on a training database of skin color pixels. For example, sir Sobottka and Pitas [12] used fixed range values on the HSI (hue saturation intensity) color space and Wang and Yuan [13] used threshold values in RG space and HS Value space to differentiate skin and non skin pixels. In these approaches, high false skin detection is a common problem when there are a wide variety of skin colors across different ethnicity, complex backgrounds, and high illumination. As we know the skin color of people belongings to Asian, African, and Caucasian groups is different from one another and ranges from white, yellow to dark. On the other hand sir U. Yang, B. Kim, and K. Sohn [14] showed us that some robustness may be achieved via the use of luminance invariant color spaces [15]. However, such an approach can only cope if the change in skin color distribution is within a narrow set of conditions [16]. Other methods through which human skin can be detected are multilayer perceptron [17] [18], Bayesian classifiers [19] [20] [21] [24], and random forest [22]. In multilayer perceptron based skin classification, a neural network is trained to learn the complex class conditional distributions of the skin and non skin pixels [17]. Sebe et al. [19] used a Bayesian network with training data of 60000 samples for skin modeling and classification. Although these solutions had been very successful, they suffer from a tradeoff between precision and computational complexity. Sir Ran Tan and Sir Chee Sang Chan proposed a fusion approach for skin detection [23]. In summary, the fusion proposed method has two advantages in comparison to the state of the art solutions. First of all, proposed skin detection method calculates skin pixels in real time. With this, a training stage can be eliminated. Second, fusion strategy also reduces computational costs.

III. PROBLEM DEFINATION

Human skin color detection plays an important role in infer variety of cultural aspects, race, health, age, wealth, beauty, etc. Detecting human skin color is of utmost importance in numerous applications such as, steganography, recognizing human by human and/or machine, and to various human computer interaction domains. The existing skin detection methods using HS, SV, HV, YCb, YCr, CrCb and I R_g B_y color spaces, are prone to false skin detection and are not able to cope with the variety of human skin colors across different ethnic, illumination, camera condition, background condition & individual characteristics. Most of the research efforts on skin detection have focused on visible spectrum imaging. Skin color detection in visible spectrum can be a very challenging task as the skin color in an image is sensitive to various factors.

To overcome the above mentioned drawbacks of existing methods to detect human skin, the proposed fusion framework approach can help greatly. The proposed novel approach fusion framework uses product rules on two features, the smoothed histogram and Gaussian model to perform automatic skin detection. First of all, we employ Viola Jones algorithm to detect nose region in an image and from this nose region we will extract skin tone in RGB color space. Therefore, the proposed method does not require any training stage beforehand. Second, a histogram with smoothed densities and a Gaussian model are used to model the skin and non-skin distributions, respectively.

IV. METHODOLOGY

The simple procedure adopted for skin color pixel detection begins with obtaining RGB color images with face regions. These input images are pre-processed in 4 stages as shown in Fig. 4.1. For classification purpose fusion technique is adopted.

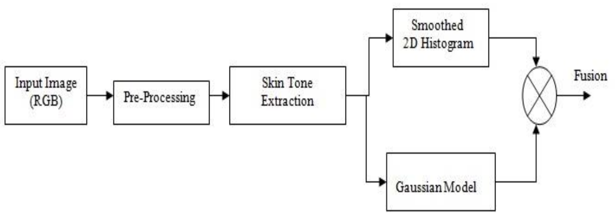


Fig 4.1. Skin color pixel detection technique using fusion approach.

A. Pre-processing

Firstly RGB color images obtained under varied conditions including face regions are taken as database images. Secondly Voila Jones algorithm is applied to detect nose region.

The Viola–Jones object detection framework is the first object detection framework to provide competitive object detection rates in real time. The algorithm has mainly 4 stages.

1. Haar Features Selection
2. Creating Integral Image
3. Adaboost Training algorithm
4. Cascaded Classifiers

B. Skin Tone Extraction

Skin color pixel detection is based on identification of skin color. In skin tone extraction, color values of image pixels in nose region are viewed as an ensemble of skin color samples. It is well established that the distribution of colors in an image is often a useful cue. An image can be represented in a number of different color space models. Therefore, it is important to choose the appropriate color space for modeling human skin color. We propose the use of RGB color space. From detected nose region we have extracted skin tone values in RGB. For this we have first considered nose region detected image and we have taken centre point of this nose image. Then we have extracted RGB color intensity values from the centre point of the nose image.

C. Skin Detection

Skin detection can be performed by skin segmentation and classification by using classifiers. In this work we have used smoothed histogram model through Gaussian model and by employing fusion strategy we have combined results of both the models. The smoothed histogram based skin segmentation at pixel is given as

$$I_{Hist} = \begin{cases} 1 & if \ \ I_{skin} = 1 \\ 0 & if \ \ I_{skin} < 1 \end{cases} \quad (2)$$

Where

$$I_{skin} = \frac{(I_{Rskin} + I_{Gskin} + I_{Bskin})}{3} \quad (3)$$

$$I_{Rskin} = \begin{cases} 1 & if \ \ I_{Rvalue} < 5 \\ 0 & if \ \ I_{Rvalue} \geq 5 \end{cases} \quad (4)$$

$$I_{Rvalue} = I_{Rhist} - R_{mean} \quad (5)$$

Where

I_{Rhist} =Histogram of red color intensity of image frame.

R_{mean} =Extracted red color intensity from nose region.

The above formula explains detection process of skin region with the help of smoothed histogram method, where I_{Hist} gives detected skin pixels values and I_{Rskin}, I_{Gskin} and I_{Bskin} gives detected skin pixels values in red, green and blue color space respectively.

And the Gaussian model is given as

$$I_{gm} = \begin{cases} 1 & if \ \ P > 0.95 \\ 0 & else \end{cases} \quad (6)$$

Where $P = \left(P_r + P_g + P_b\right)/3$ (7)

$$P_r = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(I_r - R_{mean})^2}{2\sigma^2}} \quad (8)$$

σ^2 = Variance of red color intensity of image frame.

R_{mean} = Mean intensity of red color of image frame.

I_r = Red color intensity of image frame.

The above formula explains detection process of skin region with the help of Gaussian method, where I_{gm} gives detected skin pixels values in the given image and P_R, P_G and P_B are the probability distribution function values of Gaussian in red, green and blue color space respectively.

D. Fusion or Combined Single Representation

Fusion strategy, involves integration of two incoming single features into a combined single representation by using product rule. The fusion rule is given in Equation (9)

$$I_{Fusion} = I_{Hist} * I_{gm} \quad (9)$$

Where in (i) I_{Fusion} represents result of fusion method, (ii) I_{Hist} represents result of smoothed 2D histogram (iii) I_{gm} represents result of Gaussian model.

In brief, our proposed algorithm means and includes pre-processing stage, skin tone extraction and classification of pixels by using fusion technology as detailed here under. (i) First RGB color face images are considered (ii) Secondly nose region of given image is detected and localized by using Viola Jones algorithm. (iii) RGB color of nose regions are extracted. (iv) Smoothed histogram model through Gaussian join probability distribution function is used for classification purpose. (v) The two incoming single features are integrated into a combine single representation by using the fusion rules.

V. RESULTS

The methodology discussed earlier is simulated using MATLAB R2014a Version. The results are obtained using programming in MATLAB. For experimental purpose and performance analysis dataset from databases and downloaded images randomly from google are used. For testing purpose we have considered images with face regions. Results and conclusions are drawn as follows.

Fig.5.1. Original Image

Fig.5.2. Nose detected region

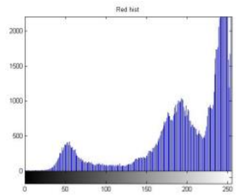


Fig.5.3. Histogram of Red color intensity

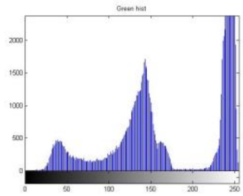


Fig.5.4. Histogram of Green color intensity

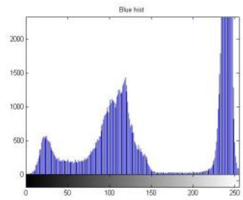


Fig.5.5. Histogram of Blue color intensity

Fig.5.6. Histogram model detected skin region

Fig.5.7. Gaussian model detected skin region

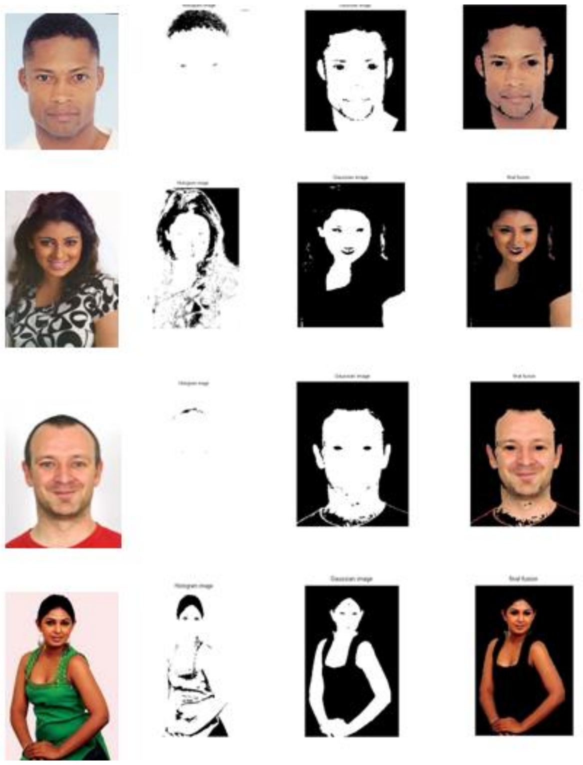
Fig.5.8. Fusion model detected skin region

From above figures we can apparently see that the proposed fusion approach gives better result compared to histogram model and Gaussian model, when we apply them individually on original image. Now we need to see, whether the proposed fusion technique can work on variety of human skin colors across different ethnicity or not. Figures 5.9 shows the results of fusion technique performed on variety of human skin colors across different ethnicity.

TABLE 5.1. Comparison between fusion & non fusion approach using computer vision dataset

Classifier	TPR %	TNR %	FPR %	FNR %	Accuracy%
Histogram Approach	85.43	90.31	14.56	9.68	88.46
Gaussian Approach	87.02	91.61	12.97	8.38	89.87
Fusion Approach	**90.76**	**94.73**	**9.23**	**5.26**	**93.22**

Table 5.1 shows the comparison results of histogram method, Gaussian method and our proposed fusion method. Table has been made by taking dataset from computer vision database. For testing purpose we have chosen total 70 frames from dataset with pixel size 641×350. The accuracy percentage is calculated based on TPR and FPR. Comparison shows that fusion approach has the highest accuracy and TPR. Moreover, it can also be visualized that the fusion strategy has low false positive rate (FPR) compared to the single feature approach.



(a) (b) (c) (d)

Fig.5.9. Comparative results, from left to right (a) Original Image, (b) Histogram result, (c) Gaussian result, (d) Fusion result

a. Discussion

In brief, our proposed algorithm means and includes pre-processing stage, skin tone extraction and classification of pixels by using fusion technology as detailed here under (i) First RGB color face images are considered (ii) Secondly nose region of given image is detected and localized by using Viola Jones algorithm. (iii) RGB color of nose regions are extracted. (iv) Smoothed histogram model through Gaussian join probability distribution function is used for classification purpose. (v) The two incoming single features are integrated into a combine single representation by using the fusion rules.

VI. CONCLUSION & FUTURE WORK

Fusion framework based on histogram and Gaussian model has been proposed to detect human skin automatically in images. As exhibited in experiments, the proposed method gives good accuracy in different conditions, background model, and ethnicity. With this it shows potential to be applied to a range of applications such as gesture analysis. One drawback of this proposed approach is that its success relies on nose detection algorithm. However, this is the general problem faced by all other researchers who work in this domain.

Our future work is focused on building a better pre-processing method and more reliable skin detection approach.

REFERENCES

[1] P. Kakumanua, S. Makrogiannisa, and N. Bourbakis, “A survey of skin color modeling and detection methods,” Pattern Recognition”, Volume 40,Issue no. 3, pages 1106–1122, 2007.

[2] S. Mitra and T. Acharya, “Gesture recognition: A survey,” IEEE Transactions on system, man and cybernetics ,part C: applications and reviews., Volume 37, Issue no. 3, pages 311–324, May 2007.

[3] Fleck, M.M., Forsyth, D.A., Bregler, C.: “Finding naked people”. In: Proceedings of the European Conference on Computer Vision (ECCV). (1996) 593– 602.

[4] Jones, M.J., Rehg, J.M. “Statistical color models with application to skin detection,” International Journal of Computer Vision (IJCV) **46**(1) (2002) 81–96.

[5] Burger, W., Burge, M. “Digital Image Processing, an Algorithmic Introduction Using Java.” 2008 Springer Verlag New York. ISBN 978-1-84628-379-6.

[6] Shin, M.C., Chang, K.I., Tsap, L.V. “Does color space transformation make any difference on skin detection?” In: WACV '02: Proceedings of the Sixth IEEE Workshop on Applications of Computer Vision, Washington, DC, USA, IEEE Computer Society (2002) 275.

[7] Zheng, Q.F., Zhang, M.J., Wang, W.Q. “A hybrid approach to detect adult web images.” In: PCM (2) Springer-Verlag Berlin Heidelberg. (2004) 609–616.

[8] Albiol, A., Torres, L., Delp, E. “Optimum color spaces for skin detection.” In Proceedings of the International Conference on Image Processing (ICIP). (2001) I: 122–124.

[9] Lee, Y., Yoo, S.I. “An elliptical boundary model for skin color detection.” In Proceedings of the International Conference on Imaging Science, Systems, and Technology. (2002).

[10] Senior, A., Hsu, R.L., Mottaleb, M.A., Jain, A.K. “Face detection in color images.” IEEE Transaction on Pattern Analysis and Machine Intelligence (PAMI). **24**(5) (2002) 696–706.

[11] Brand, J., Mason, J. “A comparative assessment of three approaches to pixel based skin detection method.” GML computer vision (2005).

[12] K. Sobottka and I. Pitas, “A novel method for automatic face segmentation, facial feature extraction and tracking,” Signal Process.: Image Communication”, Volume 12,Issue no. 3, pages. 263–281, 1998.

[13] Y. Wang and B. Yuan, “A novel approach for human face detection from color images under complex background,” Pattern Recognition”, Volume 34,Issue no. 10, pages. 1983–1992, October 2001.

[14] P. Vadakkepat, P. Lim, L. De Silva, L. Jing, and L. L. Ling, “Multi-modal approach to human-face detection and tracking,” IEEE Transactions Industrial Electronics, Volume 55,Issue no. 3, pages 1385–1393, March 2008.

[15] U. Yang, B. Kim, and K. Sohn, “Illumination invariant skin color segmentation,” “Industrial Electronics and application”, ICIEA and 4th IEEE conference on May 2009, pp. 636–641.

[16] S. Jayaram, S. Schmugge, M. C. Shin, and L. V. Tsap, “Effect of color space transformation, the illuminance component, and color modeling on skin detection,” in Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Volume 2, pages. 813–818.

[17] D. Brown, I. Craw, and J. Lewthwaite, “A SOM based approach to skin detection with application in real time systems,” in Proceedings of the British Machine Vision Conference, 2001, pages. 491–500.

[18] S. L. Phung, D. Chai, and A. Bouzerdoum, "A universal and robust human skin colour model using neural network," Neural Networks. Proceedings. IJCNN '01. International Joint Conference, 2001, Volume 4, pages. 2844–2849.

[19] N. Sebe, I. Cohen, T. S. Huang, and T. Gevers, "Skin detection: A Bayesian network approach," Pattern Recognition. ICPR. Proceedings of the 17th International Conference., August 2004, pages. 903–906.

[20] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," Machine Learning,Volume 29, pages 131–163, Nov. 1997.

[21] M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection," Computer Vision and Pattern Recognition, IEEE Computer Society Conference, Volume 46,Issue no 1, pages. 81–96, 2002.

[22] R. Khan, A. Hanbury, and J. Stoettinger, "Skin detection: A random forest approach,"in Image Processing (ICIP), 2010 17th IEEE International Conference, pages 4613–4616.

[23] Wei Ren Tan, and Chee Seng Chan, "A Fusion Approach for Efficient Human Skin Detection", IEEE Transactions Industrial Informatics, Volume 8, No. 1, pages 138–147, February 2012.

[24] P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modeling and detection methods", *Pattern Recognition*, Volume. 40, pages 1106 – 1122, 2007.