Pixel-Based Skin Color Classifier: A Review

Amit Kumar[1] and Shivani Malhotra[2]

Chitkara University, Punjab, India
amit.pandey@chitkara.edu.in , shivani.malhotra@chitkara.edu.in

Abstract

Vision based computer interface systems provides an easy way for human to communicate with computer or machine. At an early stage of development of these system i.e. detection and tracking of objects, some parameter or feature of human body are utilized for this purpose. Skin color has proven to be the most useful and robust parameter for detection and tracking of objects in real time. Skin color detection is basically divided into two class pixel based and region based. This paper presents an overview of pixel based skin color classifier proposed by various authors with their advantages and disadvantages. The performance assessment of these methods given by various authors is also discussed here.

1. Introduction

In numerous vision based application especially face recognition and hand gesture recognition, detection and tracking is an important stride in development. A number of approaches based on pattern recognition and image processing have been proposed for realizing potent and effective solution. Among all visible features of human body skin is of most importance and is widely used by researchers for formulating solution of above mentioned step in design vision based applications. Hence skin color detection and tracking have gained immense attraction these days [1, 5]. Due to fast processing and robustness of color cue, it's fast growing choice and widely adopted by many researchers. While designing system, that make use of skin color as a feature for detection and tracking, the researcher mainly faces two main problems. Firstly the selection of color space and secondly modeling of skin color in that color space to get better results of segmentation. Basically there are two types of approaches used for skin color detection namely pixel-based and region based. In pixel-based skin color detection, every pixel is classified as skin or non-skin pixels independently whereas in region based approach spatial arrangement of skin color pixels are taken into account while detecting skin pixels from non-skin pixels [7].

This paper discusses pixel-based skin color detection methods. Further these methods or classifiers are divided into three categories: explicitly defined skin regions, non-parametric methods and parametric methods. The Organisation of this paper is as follows, section II presents different skin color classifiers used for detection of skin pixel from non-skin pixel, section III gives brief discussion about all these methods including their advantages and dis-advantages, section IV deals with comparative assessment of these methods and section V concludes this paper.

2. Skin Color Classifiers

The main riddle in various vision based application like hand gesture recognition, face recognition, video indexing etc. is to distinguish skin pixels from non-skin pixels in

complex background. The use of skin pixels remained a challenge as tone and texture of skin pixel changes with the lightning conditions. Different imaging device gives different values of skin pixels for the same outlook resulting in intricate problem. Hence skin color detection is exigent problem for vision based applications [1-4].

Researchers have proposed various methods of skin color classifiers namely explicitly defined skin region, parametric methods and non-parametric methods as shown in Figure 1. These methods have their own advantages and disadvantages.
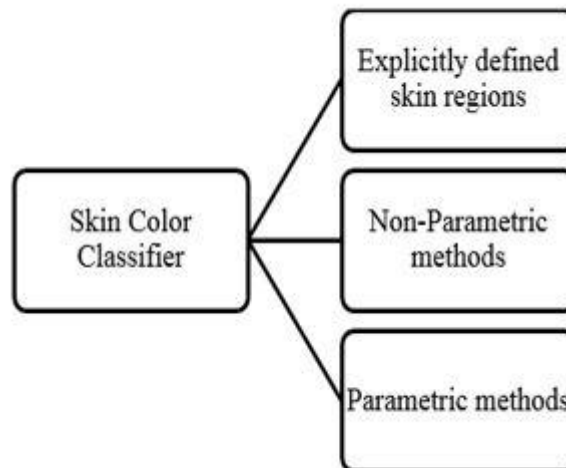

Figure 1. Various Skin Classifiers

2.1 Explicitly Defined Skin Region

Different individuals belonging to different ethnic groups, race and region noticeably have different skin color tone. But under the influence of same lightning conditions, skin color of different individuals falls in small area of color space [5]. So one of the straightforward approach is to explicitly define the decision boundary of skin color cluster for different color spaces. For example decision boundary of RGB color space is defined as:

$$R > 95 \text{ and } G > 40 \text{ and } B > 20 \text{ and}$$
$$\text{Max}\{R,G,B\} - \text{min}\{R,G,B\} > 15 \text{ and}$$
$$|R-G| > 15 \text{ and } R > G \text{ and } R > B$$

So the value of skin pixels falling in this range can be easily distinguished from non-skin pixels [a]. Similarly range used for HSV color space is H = [0, 50], S = [0.20, 0.68] and V = [0.35, 1.0] and for YCbCr color space is Cb = [77 127] and Cr = [133 173]. Due to simplicity of the implementation process this method is widely used by many researches. [6] [7].

2.2 Non-Parametric Methods

The main goal of non-parametric skin color modelling is to make an estimation about the skin color distribution through the training dataset which is devoid of using explicit model for skin color segmentation. Thus results obtained through these methods sometimes attributed in creation of Skin Probability Map (SPM). This skin probability map allocates a probability value to every point of isolated color space. The information carried by these probability values is extracted in order to distinguish between skin and non-skin color pixels in complex images [4, 8].

Figure 2 shows the three basic and foremost methods which comes under the non-parametric skin color taxonomy namely: standardized look up table (LUT) method, Bayes classifier and self-organizing map (SOM).
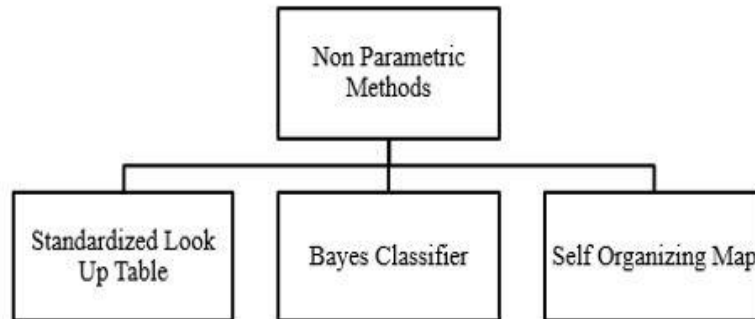

Figure 2. Showing Various Non-parametric Methods

2.2.1 Standardized Look up Table: Segmentation is an important step in various tracking algorithms. Approaches such as histogram based skin pixel segmentation is quite oftenly used in these algorithms. Normally the color space chosen for this purpose is chrominance color space. Further this color space is quantized into number of bins and these bins exhibits peculiar range of color component value duos for 2D or trios for 3D histograms. Thus these resulted bins representing 2D or 3D histograms are regarded as Look up table. The function of these bins is to store the number of times a specific color appears in the training images. At the end of training phase, histogram count obtained from these bins gives the normalized result. Now these normalized histogram values are converted to discrete probability distribution using equation (1):

$$P_{skin}(c) = \frac{skin\ (c)}{Norm} \tag{1}$$

In the above equation *skin* (c) provides the value of the histogram bin, correlative to color vector c and Norm gives the normalization coefficient *i.e.* addition of histogram values of all bins [9] or maximum value of the bin present [10]. These standardized or normalized values of look up table bins represent the likelihood that the correlative colors will only correlate with skin color [1, 10, 2, 9].

2.2.2 Bayes Classifier: In the context of statistical pattern classification, Bayes classifier is mostly used as it works on the principle of Bayesian decision rule for minimum cost in a system. This method has been used by many researchers for distinguishing skin color pixels from non-skin color pixels in YCbCr and RGB color space. The method can be explained as follows. Suppose *p(c/skin)* and *p(c/non-skin)* are the class conditional probability density functions (pdfs) of the skin and non-skin color pixels respectively. The class conditional probability density functions are estimated using histograms technique to distinguish between skin color and non-skin color pixels. The color c belongs to skin color if:

$$\frac{p\left(\frac{c}{skin}\right)}{p\left(\frac{c}{non-skin}\right)} \geq \theta \tag{2}$$

In equation 2 left hand side represents likelihood ratio and θ represents threshold. The theoretical value of θ that minimizes the classification cost is determined by *a priori* probabilities P (*skin*) and P (*non-skin*) of the two classes [12]:

$$\theta = \frac{\lambda_{fd}}{\lambda_{fr}}\frac{p(non-skin)}{p(skin)} \tag{3}$$

Where λ_{fd} and λ_{fr} gives the cost of false detection and false rejection respectively. Correct detections cost are assumed to be zero. The value of θ is determined experimentally. This means that choice of priori probability does not influence the general detector behavior as for any prior probability *p (skin)* it is possible to choose suitable value of $K=\frac{\lambda_{fd}}{\lambda_{fr}}$ that produces same detection threshold θ [1, 13].

2.2.3 Self-Organizing Map: Self-Organizing Map abbreviated as SOM, is a mastermind of Kohonen in 80's and has now become as one of the most well-liked approaches of unsupervised artificial neural network. Various researchers have proposed the SOM detector which is used for differentiating between skin and non-skin pixels. SOM detector is tested under different color spaces for different datasets by researchers and their result shows that SOM detector do not show the lucid performance changes while using distinct color-spaces. Researchers also point up that SOM method requires substantially less resource than histogram and Gaussian mixture models, hence it can be competently implemented for run time usages with the help of SOM hardware.

2.3 Parametric Methods

Due to some shortcomings in non-parametric methods used in skin color classification, there is a need of more covenant skin models. These models must be able to simplify and incorporate with less training data hence fosters the evolution of parametric skin color distribution methods.

Figure 3 represents various parametric methods used for skin color taxonomy. These methods are namely single Gaussian, mixture of Gaussian and elliptical boundary model [1-2]. A brief discussion about all these is mentioned in this section.
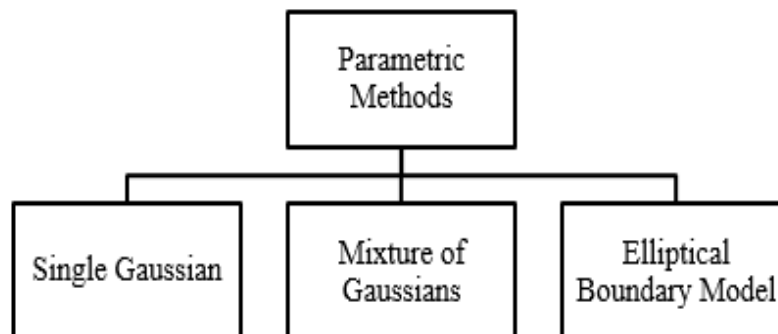

Figure 3. Showing Various Parametric Methods

2.3.1 Single Gaussian: With the influence of constant lightning conditions, a small cluster of skin color is formed for distinct entity within the same color space. That is why in the presence these lightning conditions, skin color taxonomy of distinct entities can be modeled using Gaussian distribution in standardized color space. Thus skin color distribution is formed using elliptical Gaussian joint probability density function (pdf), expressed as:

$$p(c) = \frac{1}{\sqrt{2\pi|\varepsilon|}}exp\left[-\frac{1}{2}(c-\mu)^T\varepsilon^{-1}(c-\mu)\right] \tag{4}$$

In the above equation (4), c is the color space, ε and μ are mean vector and covariance matrix respectively. These parameters are estimated for the entire samples of the training data in a color space using equations (5).

$$\mu = \frac{1}{n}\sum_{j=1}^{n} c_j \qquad \text{and}$$

$$\varepsilon = \frac{1}{n-1}\sum_{j=1}^{n}(c_j - \mu)(c_j - \mu)^T \qquad (5)$$

Maximum likelihood estimation approach is used here for the estimation of these parameters. Thus probability p(c) gives the primary measure of skin color likeliness and pixels are grouped normally through comparison of these skin color pixels with certain threshold values obtained experimentally against training data [1, 11, 2].

2.3.2 Mixture of Gaussians (MoG): A more refined form of single Gaussian model is mixture of Gaussian (MoG). This model has the ability to portray the complex shaped distributions easily. The probability density function (pdf) for this case is mentioned below:

$$p(c) = \sum_{i=1}^{N} w_i \frac{1}{\sqrt{2\pi|\varepsilon|}} exp\left[-\frac{1}{2}(c - \mu)^T \varepsilon^{-1}(c - \mu)\right] \qquad (6)$$

In equation (6), N is the number of Gaussian and w_i is the weight factor, which is the contribution of ith Gaussian. This model is trained using expectation maximization (EM) algorithm which is iterative in nature. Same methodology is adopted here for the classification of skin and non-skin color pixels as in the case of single Gaussian model (SGM) [1, 2].

2.3.3 Elliptical Boundary Model: Gaussian mixture models requires much of computational task. So in order to reduce this [14] proposed a new model called elliptical boundary model. The performance of this model is similar to that of MoG and computationally is as simple as training a SGM. Equation (7) describes the elliptical boundary model as follows:

$$\phi(c) = (c - \psi)^T \Lambda^{-1}(c - \psi) \qquad (7)$$

Here color vector is represented by c and model parameters are represented using ψ and Λ. These are defined as below:

$$\psi = \frac{1}{n}\sum_{i=1}^{n} c_i$$

$$\Lambda = \frac{1}{N}\sum_{i=1}^{n} f_i (c_i - \mu)(c_i - \mu)^T \qquad (8)$$

In the above equation N denotes total samples in the training data set, f_i is the number of samples in the chrominance space c_i and μ is the mean of the chrominance vector in training data set. Pixel having chrominance c belongs to skin pixel if $\phi(c) < \Theta$, where Θ is the some threshold value calculated experimentally using Otsu thresholding algorithm [17].

3. Classifier Discussion

The skin color detection classifiers are divided into three broad categories namely explicitly defined skin region, parametric methods and non-parametric methods. These classifiers have some advantages and disadvantages too on one another. Table 1 shows the pros and cons of above mentioned methods. Among all methods discussed above explicitly defined skin region is intuitive and robust in terms of implementation and complexity.

Table 1. Advantages and Disadvantages of Skin Color Classifiers

Type of classifier	Advantages	Disadvantages
Explicitly defined skin regions	Easy Implementation Simple approach. Less storage memory is required.	Difficulty in selection of color space. Require sufficient decision rules experimentally.
Parametric Methods	Able to interpolate training data. Less storage space is required.	Slow in training Allied on shape of skin distribution.
Non-Parametric Methods	Fast in training Theoretically unallied from shape of skin distribution.	More storage space is required Incapacity to interpolate the training data.

4. Comparative Assessment

In order to assess the performance of different skin color detection classifiers these have to be tested in identical environment. But various authors have taken their own environmental conditions and database. Most widely used database is Compaq database.

Table 2. Skin Color Classifiers Performance Reported by Various Authors

Method	True Positive	False Positive
Bayes SPM in RGB [8]	93.4%	19.8%
Bayes SPM in RGB [9]	90%	14.2%
Maximum Entropy Model in RGB [15]	80%	8%
Gaussian Mixture Model in RGB [9]	80%	90.5%
SOM in TS [17]	78%	32%
Elliptical boundary model in CIE-XY [14]	90%	20.9%
Single Gaussian in YCbCr ([14]	90%	33.3%
Gaussian Mixture in IQ [14]	90%	30%
Thresholding of I axis in YIQ [8]	94.7%	30.2%

Table 2, shows the comparative assessment of the performance of above mention methods for their true positive (TP) and false positive (FP) rates. The most promising results are shown by Bayes SPM in RGB and maximum entropy model.

5. Conclusion

The main aim of this paper is to discuss the various methods for pixel based skin color detection proposed by various authors. This paper also presents the advantages and disadvantages along-with their performance in terms of true positive and false

Positive rates. A classifier is good if it has lower false positive rate for a given true positive rate. So among all, Bayes SPM in RGB has shown good performance.

References

[1]   V. Vezhnevets, V. Sazonov and A. Andreeva, "A survey on pixel-based skin color detection techniques", GRAPHICON03, (**2003**), pp. 85–92.

[2]   P. Kakumanu, S. Makrogiannis, and N. Bourbakis, "A survey of skin-color modeling and detection methods", Pattern recognition, vol. 40, no. 3, (**2007**), pp. 1106-1122.

[3]   Y. Dai and Y. Nakano, "Face-texture model based on SGLD and its application in face detection in a color scene", Pattern Recognition, vol. 29, no. 6, (**1996**), pp. 1007–1017.

[4]   G. Gomez and E. Morales, "Automatic feature construction and a simple rule induction algorithm for skin detection", Proc. of the ICML workshop on Machine Learning in Computer Vision, (**2002).**

[5]   J. Yang, W. Lu, and A. Waibel, "Skin-color modeling and adaptation", Springer Berlin Heidelberg, (**1997**).

[6]   D. Chai and K. N. Ngan, "Locating facial region of a head-and-shoulders color image, ICFGR98, (**1998).**

[7]   Y. Wang and B. Yuan, "A novel approach for human face detection from color images under complex background", Pattern Recognition, vol. 34, no. 10, (**2001**), pp. 1983-1992.

[8]   J. Brand and J. Mason, "A comparative assessment of three approaches to pixel level human skin-detection", In Proc. of the International Conference on Pattern Recognition, vol. 1, (**2000**), pp. 1056–1059.

[9]   M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection", In Proc. of the CVPR '99, vol. 1, (**1999**), pp. 274–280.

[10] Zarit, Benjamin D., Boaz J. Super, and F. K. H. Quek, "Comparison of five color models in skin pixel classification", Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, (**1999**), Proceedings, International Workshop on. IEEE, (**1999).**

[11] J.-C. Terrillon and S. Akamatsu, "Comparative performance of different chrominance spaces for color segmentation and detection of human faces in complex scene images", Vision Interface, vol. 99, (**1999).**

[12] R. O. Duda, P. E. Hart, and D. G. Stork, "Pattern Classification", New York: John Wiley & Sons, Inc., (**2001).**

[13] S. L. Phung, A. Bouzerdoum, and D. Chai, "Skin segmentation using color and edge information", Signal Processing and Its Applications, 2003, Proceedings, Seventh International Symposium on. vol. 1, IEEE, (**2003**).

[14] S. L. Phung, A. Bouzerdoum, and D. Chai, "Skin segmentation using color and edge information", Signal Processing and Its Applications, 2003, Proceedings, Seventh International Symposium on, vol. 1, IEEE, (**2003**).

[15] B. Jedynak, H. Zheng, M. Daoudi, and D. Barret, "Maximum entropy models for skin detection", Tech. Rep. XIII, Universite des Sciences et Technologies de Lille, France, (**2002**).

[16] D. Brown, I. Craw, and J. Lewthwaite, "A som based approach to skin detection with application in real time systems", In Proc. of the British Machine Vision Conference, (**2001**).

[17] S. Verma, R. Chandra, H. J. Kim and G. S. Tomar, "Lossless Image Compression Using Motion Image", VIE 06, (**2006**) September, pp. 1-4.

[18] S. S. Bedi, G. S. Tomar and S. Verma, "Robust Watermarking of Image in the Transform Domain using Edge Detection", IEEE International Conference on simulation UKSIM 2009, (**2009**), March 25-29, pp. 233-238.

[19] N. Ostu, "A threshold selection method from gray-level histogram", IEEE Trans. Systems Man Cybemet. SMC-8, (**1978**), pp. 62-66.

Authors

Amit Kumar, he received B. Tech degree in Electronics and Communication Engineering from Kurukshetra University, Haryana (India) in 2009. He is currently pursuing M.E. in Electronics and communication engineering from Chitkara University, Punjab (India). He is currently working in the area of Digital Image Processing and Pattern recognition. His work is mainly related to Gesture recognition and he had presented several papers regarding this in various conferences.

Shivani Malhotra, she received M.Tech degree in Electronics and Communication Engineering from Punjab University, Punjab (India) in 2008. She is currently pursuing PhD. in Electronics and communication engineering from Chitkara University, Punjab (India). She is currently working in the area of Synthesis and characterization of radar absorbing materials, microwave characterization of ferrites and Gesture recognition and she had presented several papers regarding this in various Conferences