Indian Journal of Science and Technology, *Vol 9(32), DOI: 10.17485/ijst/2016/v9i32/92420, August 2016*

ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

ILUT based Skin Colour Modelling for Human Detection

Dushyant Kumar Singh* **and Dharmender Singh Kushwaha**

Department of Computer Science and Engineering, MNNIT Allahabad - 211004, Uttar Pradesh, India; dushaynt@mnnit.ac.in, dsk@mnnit.ac.in

Abstract

Objectives: Numerous techniques have been proposed in past for skin colour modelling and some out of these are utilized in human detection problem. This paper proposes a novel technique which uses Indexed Look-up-table (ILUT) for skin colour modelling. **Methods:** In the proposed technique, skin colour cluster on 2-dimensional Cb-Cr plane in YCbCr colour space is modelled with an ILUT. ILUT contains the lower and upper bounds of Cb values corresponding to each value of Cr in the skin colour cluster. Outliers in the cluster are removed by applying median filter, because they contribute to the wrong classification of skin colour and hence wrong detection. **Findings:** The proposed technique delivers reasonably good performance on True Positive Rate (TPR), False Positive Rate (FPR) and Accuracy parameters. Classification accuracy of 89.89% for proposed technique is almost comparable to that of other techniques in literature. Comparative results are shown in Table 2 of the paper. Classification complexity is one parameter on which the proposed technique outperforms the rest of the other skin colour modelling techniques and it is of O(MxN). MxN is the image resolution. Skin colour modelling using ILUT is also efficient in terms of space requirements as compared to other non-parametric methods of skin colour modelling. Applications: Least classification complexity makes this technique most appropriate for real time systems/ applications for detecting presence of human being.

Keywords: Classification Accuracy, Classification Complexity, Human Detection, Indexed Look-up-table, Skin Colour Modelling

1. Introduction

Automated detection of human objects is now becoming an important and integral part of a number of applications such as video surveillance, robotics, biometrics and human machine interface etc. Skin colour of human being which is mostly visible on face is a feature that fairly discriminates a human being from the other objects in an image. Skin colour occupies a specific range in each plane of the colour space[1]. The image pixels which lie in this range are recognized as skin pixels.

The range of skin colour varies based on age, ethnicity, and illumination of environment in which image is captured. Effects due to Illumination can be reduced by choosing an appropriate colour space. The colour spaces that have mostly been used in previous researches are HSV, YCbCr and YUV. RGB and L*a*b has also been used at some places. Comparative analysis in a research shows that YCbCr colour space gives better result as compared with others on maximum image databases[2].

Y of YCbCr is the luminance component and Cb,Cr are the chrominance components. The chrominance component has negligible impact of external illumination effect[3], that's why it is preferred to be used for skin colour modelling. The ranges of Cb and Cr for skin colour intensities are more compact as compared to ranges of components in other colour spaces[4]. This can well be tested on 'Thresholder app' of MATLAB for a given sample image. Compact range reduces the probability of overlapping of two distinct colours, in turn support accurate modelling. YCbCr colour space is hence used for skin colour modelling on our case.

Skin colour modelling is task of designing a classifier that can classify/distinguish skin colour intensities from non-skin intensities. Number of techniques has already been proposed in the literature for skin colour

** Author for correspondence*

modelling[5,6]. On the ground of classification methodology, these techniques can be categorized into three broad categories. Classification by thresholding is the first one, second is category of parametric methods and the third is non parametric methods category.

The idea of classification by thresholding is to explicitly define skin regions by deriving some threshold ranges of colour space components. The classification function could be the rules that define skin colour cluster based on thresholding done[7]. A set of rules on R, G and B components of RGB colour space are provided in[8,9]. The pixel is identified as skin if the conditions in the rules hold good. Another set of decision rules on RGB are proposed in[10,11]. The approach in[4] gives the linear and non-linear decision rules for Cb, Cr components of YCbCr space and U, V components of YUV space. Linear decision rules for H, S components of HSV space are provided in[12]. Challenge with this category of methods lies in finding the adequate decision rules.

The non-parametric methods estimate skin colour distribution from the training data. Methods in this category are distance-based segmentation, Look-up-table (LUT), Bayes classifier, ANN[13,14] and SVM[15]. Euclidean distance between test image pixels and mean of sample skin colour pixels is used as the measure of similarity in distance-based segmentation[16]. A threshold on distance is the classification rule for the method. Without using arithmetic operations, skin colour segmentation is achieved by LUT's. A 2D histogram is referred to as LUT, with each cell containing the frequency of particular colour value when processed on sample training images[2]. LUT is normalized to get cell values in the range (0 - 1). This Cell value represents the likelihood of colour to be skin[17]. The likelihood in previous method can be viewed as conditional probability of a colour to be skin i.e. . The technique proposed in [17–19] use Bayes rule for conditional probability to classify skin pixels.

As non-parametric methods are histogram based and therefore have much space complexity, the performance of these methods depends on size of training image set[20]. Parametric methods as single Gaussian, mixture of Gaussian's and boundary models resolve the performance issues of non-parametric methods. The cluster of skin pixels or skin colour distribution is modelled by a Gaussian joint probability distribution function (pdf). Posterior probability is used to classify pixel as skin or non-skin, which is the function of mean and covariance of gaussian[21,22]. Technique proposed in [23] has used mixture of Gaussian's for modelling the skin cluster. Beside these, some geometric boundary models has also been used to model skin cluster. In [24] have modelled the cluster by elliptical boundary function.

Though parametric methods reduce the complexity of training phase but still the testing phase involves complex arithmetic's on each pixel to classify as skin/non-skin. This testing phase complexity is the actual classification complexity of any method. The analysis of classification complexity of the various methods is a genuine contribution to the research proposed in this paper. This is discussed in the experimental result section i.e. section 4 of the paper. Indexed Look-up-table (ILUT) based classification proposed here is one promising solution, which has least classification complexity with the legitimate classification accuracy. Accuracy of the classifier is maintained by filtering the outlier's as outlier's increases the False Positive Rate (FPR) of classifier. This is explained in detail in the section 2 of the paper. Beside classification, pre-processing done in the proposed method is countable portion of research.

Rest of the paper is organised as follows. Section 2 details the method proposed here, with pre-processing, filtering and modelling as its subsections. Experimental results of skin colour classification using ILUT are presented in section 3. The discussion on classification complexities of various methods and as well the comparative analysis of accuracy results are also discussed in section 3. Finally, the paper is concluded in section 4 with directions to the future work.

2. Proposed Method

The method proposed here is novel and simple. The skin colour modelling is done using ILUT i.e. classification model used is ILUT. The kind of LUT's used earlier were the two dimensional array where each cell represents one Cb-Cr point. The value each cell contains is the frequency of Cb-Cr pixels occurring in images, when processed on the training dataset. The same is discussed in introduction section and references are provided for various researches that have made use of such LUT's.

The size of these LUT's is recognized to be 256x256, i.e. 65536 cells, each of which corresponds to a unique Cb-Cr pixel intensity. These 65536 cells correspond to both skin and non-skin regions. While ILUT is created for skin region points only, where number of cells limits to around 2000 only. Reduced size of ILUT reduces the space requirement and as well reduces the classification complexity to the order of size of image. This enables the classifier designed in this method to work in real time, for skin/non-skin region classification. The skin region recognized in the image denotes the presence of human being in the scene.

In details of the modelling ILUT as a part of proposed method is discussed in this section. The first step of the method is pre-processing, where the image dataset is pre-processed for acquiring Cb-Cr data against skin colour region in YCbCr color space. In next step, skin colour model is derived. The words model and classifier are used alternatively.

2.1 Pre-processing

An image dataset containing the human individual images, as well as the family images is taken for pre-processing. This is Pratheepan's dataset[4], which also has ground truth images with skin region marked. The image set of 32 images containing human individual in the image is used for training the classifier. The Cb-Cr data of skin region in the images of this image set is acquired by placing images and their ground truth adjacent. This Cb-Cr data is the Cb and Cr values of each skin pixel in an image. A pixel could be the skin pixel only if the pixel has some specific values of Cb for some specific values of Cr. Cb and Cr are therefore dependent variables and can be plotted as point (Cr, Cb) on the coordinate plane with Cb and Cr as two axes.

In practical scenario of our experiment, the number of acquired Cb-Cr data points is so large that it reaches something 5 lack data points[4]. With such large quantity of data, training a model/classifier is too much time taking. The situation becomes more adverse when complex classifiers as Support Vector Machine (SVM) or Artificial Neural Network (ANN) are used. Against this, histogram based approaches for training reduces the number of Cb-Cr data points to 256x256 i.e. 65536. This is because Cb and Cr in YCbCr colour space ranges from 0 to 256 for all possible colours in the nature. As discussed earlier, that all other methods using normal LUT's make use of all these 65536 data points, which results in increased space requirement. But, out of 65536 different colour intensities, the skin colour intensities only contribute to a smaller percentage. Around 2051 colour intensities were found to be lying in skin colour region when experimented on Pratheepan's dataset, where dataset covers images from various ethnic groups and remote geographical regions. The Figure 1 shows the scatter of Cb-Cr data points under skin colour region.

Figure 2(a) is the zoomed view of scatter in Figure 1. According to the scatter of Cb-Cr data points for skin region, the Cr value for skin pixel lies in range from 113 to 200 and Cb values lies in range 76 to 140. Scatter obtained forms an irregular shape and can be thought as a cluster of skin pixels. All remaining pixels can be treated as non-skin pixels. Few of the researchers[21-24] have identified this problem as a clustering problem and accordingly solutions are proposed. However, here in this paper a simple classification method is derived to model the cluster of skin pixels as seen in Figure 2.

The accuracy of classifier depends on how correctly the skin cluster is modelled. The pixels at boundary are sparsely scattered so are difficult to model. Many of these are so sparse that they become outlier to the model. These factors increase the False Positive and also decrease the True Positive of classification which in turn affecting the classifier accuracy. These outliers are like the noise to the distribution. For smoothing the blurred boundary of skin cluster and removing the outliers, the filtering is applied on cluster in Figure 2(a).

For filtering, a 3x3 filter with all weights 1 is taken to convolute the Cb-Cr coordinate plane. Filter function is median i.e. the centre pixel of sample in the image is replaced by median of convolution vector. And result of median filter obtained is shown in Figure 2(b).

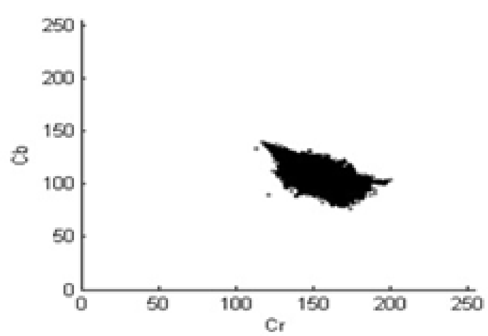

Figure 1. Scatter of skin region points.

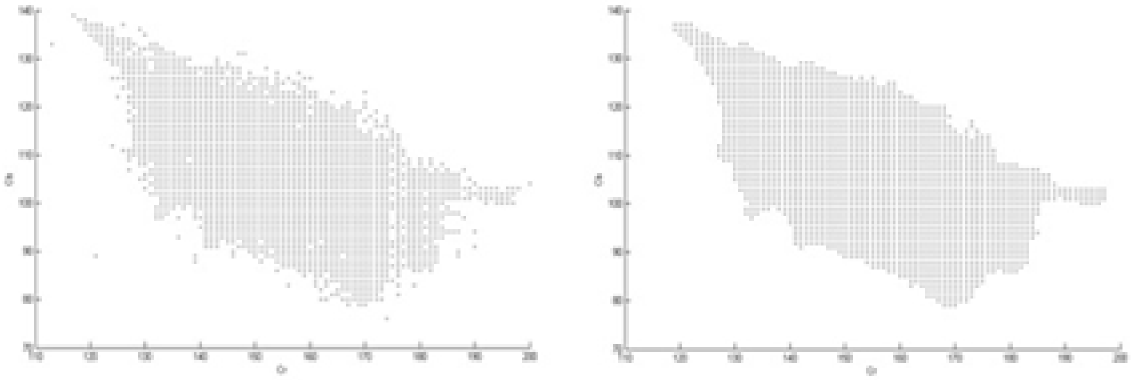

Figure 2. Skin region scatter (a) before filtering, (b) after filtering.

2.2 Proposed ILUT

ILUT proposed here is a 256x2 matrix, where 256 rows designate 256 levels of Cr values and two columns corresponds to the Cb values. Having only 2 columns for Cb against 256 different levels is the only main idea of reduction of LUT size. Two columns contain minimum Cb and maximum Cb for each of 256 levels of Cr. The complexity of classification with this model reduces to 'O(MxN)' i.e. the size of image. Data points used to train this ILUT are Cb-Cr points in skin cluster obtained after filtering i.e. cluster of Figure 2(b). With just two comparisons the skin pixel is classified while it takes just 1 comparison for non-skin pixel. This results in faster rate of classification and enables it to be feasible for real-time applications. The classification accuracy and time of classification results are shown in the next section.

3. Experimental Results

Experiment is done on images from Pratheepan's dataset[4]. The dataset contains images of individuals and family images from various ethnic groups from all around the world. The training of ILUT is done with 32 image set of individuals. Later, testing is performed on both images with individuals and family images as well. The testing results of classification are shown in Figures 3 and 4. Figure 3 contains the image results of skin classification in individual's images while Figure 4 contains the results of family images.

Figure 3. Images with individuals from Pratheepan's Image dataset, original image in 1st column, classification results without filtering in 2nd column is, and results with filtering in 3rd column.

Figure 4. Classification results on family images, row 1 are original images and row 2 are results with filtering.

Pixels of the test image are classified either as skin or non-skin. Output of the classifier hence generated is a binary image. How correctly the skin and non-skin pixels are classified as skin and non-skin decides the accuracy of classifier. The quantitative measurement parameters to this are True Positive Rate (TPR), False Positive Rate (FPR) and Accuracy.

$$TPR = \frac{TP}{TP + FN} \quad (1)$$

$$FPR = \frac{FP}{FP + TN} \quad (2)$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (3)$$

where,
TP is the number of skin pixels classified as skin
FN is the number of skin pixels classified as non-skin
FP is the number of non-skin pixels classified as skin, and
TN is the number of non-skin pixels classified as non-skin

TPR and FPR calculation is done for each of the test results of 32 images in both cases, one without using filter before training the ILUT and other when median filter is applied before training. The plot of this is shown in Figure 5. Figure 6 shows the plot of average TPR and FPR in two cases.

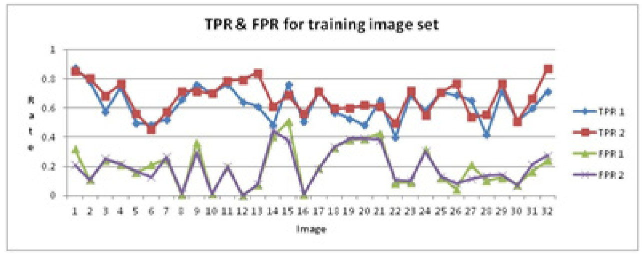

Figure 5. TPR and FPR for training image set; TPR1 and FPR1 in case of without filtering and TPR2 and FPR2 in case of with filtering.

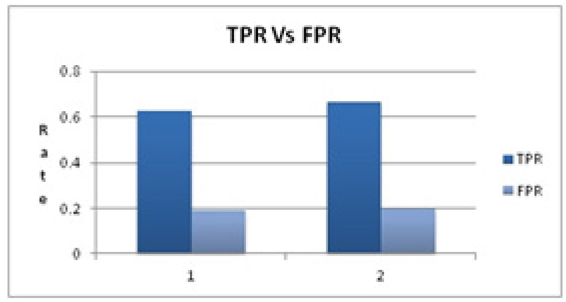

Figure 6. Average TPR and FPR in two cases.

Accuracy of the classifier in above two cases is tabulated in Table 1. It can be observed from the table that, applying filter on skin cluster improves the classifier accuracy. Table 2 gives a comparison of accuracy achieved by various previously discussed methods with that of proposed methods. It outperforms several of those methods but not all. Then another factor to taken care is Classification and computational complexities of the methods, where proposed method outperforms others. No time calculation is done for each method as time of execution of any algorithm depends on computer system performance. That's why complexity here is shown in asymptotic notations. Analyzing data of Table 2, we can also determine that the proposed method is best to work in real time.

Table 1. Comparison of TPR, FPR and accuracies in two cases

Classifier	TPR	FPR	Accuracy
ILUT (without filter)	0.624	0.191	0.837
ILUT (with filter)	0.668	0.199	0.8989

Table 2. Comparison of accuracies and classification complexities for various skin colour modelling methods

Classifier	Accuracy	Classification Complexity	Remark
Proposed Approach	0.8989	O(MxN)	Training simplest in terms of space & time both
Linear Fitting[8-10]	0.7958	O(MxN)	
LUT[17]	0.8641	O((MxN)^2)	Classification not achieved in real time
Non-linear fitting[3]	0.8832	O((MxN)^n), n is degree of fitting curve	
Distance-Based Segmentation[16]	0.8792	O((MxN)^2)	
ANN[13,14]	0.8931	O(MxNxO(activation function))	Training phase is too much complex i.e. time taking
SVM[15]	0.9027	O(MxNxO(kernal))	
Gaussian PDF[21,22]	0.9070	O(MxNxO(d)^3), d is order of covariance matrix	
Elliptical Boundary Model[24]	0.9013	O(MxNxO(Locus))	

4. Conclusion

The method proposed here for skin colour modelling is very efficient in terms of classification complexity, which makes it to be used for automatic human detection in a real-time application. The automatic detection of human being in any scene/image requires the approximation of Blob (binary large objects) found in processed binary image to the human posture. The classification accuracy achieved by ILUT is also reasonably high, which can give a higher rate of correct detection of human being. The use of YCbCr colour space makes the ILUT modelling and classification unaffected by ambient light illumination effects.

The constant space requirement and the constant classification complexity with good accuracy make the ILUT most appropriate solution to human detection in real-time systems. Moving human object can be detected and tracked even with moving/stormy background is one added advantage of the skin colour modelling based human detection. One minor drawback the skin colour modelling based techniques faces is high FPR and in turn low accuracy in case of skin colour background or background objects. The resolution to this problem is kept under directions to the future work.

5. References

1. Praveen K, Makrogiannis S, Bourbakis N. A survey of skin-color modeling and detection methods. Pattern Recognition. 2007; 40(3):1106–22.
2. Benjamin DZ, Super BJ, Quek FHQ. Comparison of five color models in skin pixel classification. Proceedings International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, Washington; 1999. p. 1–6.
3. Tan WR, Chan CS, Yogarajah P, Condell J. A fusion approach for efficient human skin detection. IEEE Transactions on Industrial Informatics. 2012; 8(1):138–47.
4. Vadakkepat P, Lim P, De Silva LC, Jing L, Ling LL. Multimodal approach to human-face detection and tracking. IEEE Transactions on Industrial Electronics. 2008; 5(3):1385–93.
5. Ganesan P, Rajini V, Sathish BS, Kalist V, Basha SKK. Satellite image segmentation based on YCbCr color space. Indian Journal of Science and Technology. 2015; 8(1):35–41. DOI: 10.17485/ijst/2015/v8i1/51281.
6. Mehdi N, Sepidname G, Eizi A, Amani A. A new skin color detection approach based on fuzzy expert system. Indian Journal of Science and Technology. 2015; 8(21):1–11. DOI: 10.17485/ijst/2015/v8i21/50606.
7. Wang CW, Ke, Ming L. Skin color modeling for face detection and segmentation: a review and a new approach. Multimedia Tools and Applications. 2014; 74(1):321–45.
8. Peter P, Solina F. An automatic human face detection method; 1999. p. 122–30.
9. Borut B, Solina F, Peer P. 15 seconds of fame: an interactive, computer-vision based art installation. Proceedings of the 12th annual ACM International Conference on Multimedia. 2004; 1:198–204.
10. Wei-Che C, Wang MS. Region-based and content adaptive skin detection in color images. International Journal of Pattern Recognition and Artificial Intelligence. 2007; 21(5):831–53.

11. Esmaeil K, Tabatabaie ZS. A hybrid face detection approach in color images with complex background. Indian Journal of Science and Technology. 2015; 8(1):49. DOI: 10.17485/ijst/2015/v8i1/51337.
12. Selin MB, Bulut M, Atalay V. Projection based method for segmentation of human face and its evaluation. Pattern Recognition Letters. 2002; 23(14):1623–9.
13. Lam PS, Chai D, Bouzerdoum A. A universal and robust human skin color model using neural networks. International Joint Conference on Neural Networks. IJCNN'01, Washington. 2001; 4:2844–9.
14. Alaa YT, Jalab HA. Increasing the reliability of skin detectors. Scientific Research and Essays. 2010; 5(17):2480–90.
15. Jinguang H, Awad G, Sutherland A. Automatic skin segmentation and tracking in sign language recognition. Computer Vision, IET. 2009; 3(1):24–35.
16. Rafael CG, Woods RE, Eddins SL. Digital image processing using MATLAB. Pearson Education India; 2004.
17. Michael JJ, Rehg JM. Statistical color models with application to skin detection. International Journal of Computer Vision. 2002; 46(1):81–96.
18. Douglas C, Phung SL, Bouzerdoum A. A Bayesian skin/non-skin color classifier using non-parametric density estimation. Proceedings of the 2003 International Symposium on Circuits and Systems. ISCAS'03. 2003; 2:464–7.
19. Zhanyu M, Leijon A. Human skin color detection in RGB space with Bayesian estimation of beta mixture models.18th European Signal Processing Conference (EUSIPCO-2010), Aalborg; 2010. p. 1204–8.
20. Vladimir V, Sazonov V, Andreeva A. A survey on pixel-based skin color detection techniques. Proceedings on Graphic. 2003; 3:1–7.
21. Rein-Lien H, Abdel-Mottaleb M, Jain AK. Face detection in color images. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2002; 24(5):696–706.
22. Bernd M, Wien M. Segmentation and tracking of facial regions in color image sequences. Visual Communications and Image Processing 2000; 2000.
23. Frank YS, Cheng S, Chuang C-F, Wang PSP. Extracting faces and facial features from color images. International Journal of Pattern Recognition and Artificial Intelligence. 2008; 22(3):515–34.
24. Jae YL, Yoo SI. An elliptical boundary model for skin color detection. Proceedings of the 2002 International Conference on Imaging Science, Systems, and Technology; 2002.