iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** I **Month of publication:** January 2022

DOI: https://doi.org/10.22214/ijraset.2022.39751

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue I Jan 2022- Available at www.ijraset.com

Color-Based Feature Extraction Method for Skin Segmentation Based on SVM

Harsha B. K.[1], G. Indumathi[2]
[1]*Department of ECE, CMR Institute of Technology (VTU RC), Bengaluru, Karnataka, India*
[2]*Department of ECE, Cambridge Institute of Technology, Bengaluru, Karnataka, India*

Abstract: ***Different colored digital images can be represented in a variety of color spaces. Red-Green-Blue is the most commonly used color space. That can be transformed into Luminance, Blue difference, Red difference. These color pixels' defined features provide strong information about whether they belong to human skin or not. A novel color-based feature extraction method is proposed in this paper, which makes use of both red, green, blue, luminance, hue, and saturation information. The proposed method is used on an image database that contains people of various ages, races, and genders. The obtained features are used to segment the human skin using the Support-Vector- Machine algorithm, and the promising performance results of 89.86% accuracy are then compared to the most commonly used methods in the literature.***
Keywords: Skin segmentation, SVM, feature extraction, digital images.

I. INTRODUCTION

This Skin detection is the path towards finding skin-toned pixels and areas in an image or a video. This is commonly used as a pre-processing step to find areas that possibly have human skin in the areas like faces, neck, hands and limbs in pictures. Many Computer Vision approaches have been proposed for skin detection. A skin detector transforms a given pixel into a reasonable concealing space on a consistent basis. Then, using a skin classifier, determine whether the pixel is skin or non-skin. A skin classification system describes a decision breaking point in the skin color class for training the model.

Recently, face detection [1-3] has acquired lots interest and has been an intensive studies topic because the major area exposed in most of the images containing human pictures will be the face. It is the crucial first step of many programs consisting of the face recognition [4], facial features analysis, surveillance systems [5], the paintings on video-conferencing, shrewd human-computer interaction, content-primarily based totally photograph retrieval systems, etc. Therefore, the effectiveness of face detection affects the overall performance of those systems. There had been numerous procedures proposed for face detection, which may be typically categorized as: Template matching methods, Feature-based methods, Knowledge-based methods, and Machine learning methods. Template matching method means the final decision comes from the similarity between input image and template. It is scale-dependent, rotation-dependent and computationally expensive. The presence or absence of a face in each skin region by using an eye detector based on a template matching scheme.

In this study, a comprehensive image data set has been used, which will be explained in Section 2.1 in detail. Red-Green-Blue-Luminance-Hue-Saturation (RGB-YHS) features have been used as predictor for making decision if the corresponding pixel is skin pixel or non-skin pixel. The detailed information for the proposed features will be as given in Section 2.2, 2.3 and 2.4. The chosen features have been generated from the chosen data set and the obtained features have been applied on Support-Vector-Machine with fine gaussian kernel, which is explained in Section 3 . The comparative skin segmentation performances were presented in Section 4. Finally, the presented article was concluded by Section 5 with analysis of obtained results.

II. MATERIALS AND METHODS

A. Schmugge Skin Image Data Set

In this study, for train and test of the chosen Support-Vector-Machine classifier with fine gaussian kernel, “Schmugge skin image data set with ground truth” is used [6]. Schmugge, Jayaram, Shin and Tsap presented comparative research for available studies in the literature for skin detection by alisting the performances on the data set they created [7]. Schmugge skin image data set includes 846 various images with their ground truth for skin segmentation. In the original ground truth images, ‘0’ was used to point the locations of exact skin pixels and ‘255’ was used to point exact nonskin pixels. In the original ground truth images, there are also exist some intermediate values between 0 and 255 for the pixels that cannot be sure whether or not they belong to the skin. In this study, these intermediate values were mapped to whichever of the values between 0 and 255 and binarized ground truth images were obtained. The skin pixels labeled as 0 is assumed as positive class and nonskin pixels labeled as 255 is assumed as negative class.

©IJRASET: All Rights are Reserved

300 of 846 images in Schmugge data set are landscape images, which do not contain any skin pixels. Therefore, their ground truth images consist entirely of values of '255'. 21 of 846 images are community images, which contain multiple people with diverse age, racial and gender characteristics and resolutions of these community images are generally higher than other images. 171 of 354 images belong to female individuals with diverse age and racial characteristics. 167 of these 171 images contain only faces and other 4 contain additionally arm, hand, leg etc. 354 of 846 images belong to male individuals with diverse age and racial characteristics. 342 of these 354 male images contain only faces and other 12 contain additionally arm, hand, leg etc.
The detailed information about these images are listed in Table I.

TABLE I: Detailed content explanation of Schmugge skin image data set

Category	Number of images	Total no. of pixels	Total no. of skin pixels in binarized ground truth images	Total no. of nonskin pixels in binarized ground truth images
Landscape	300	3,862,410	0	3,862,410
Community	21	9,177,544	1,816,387	7,361,157
Female	171	7,792,182	2,001,988	5,790,194
Male	354	12,861,765	4,463,134	8,398,631
Total	846	33,693,901	8,281,509	25,412,392

B. Color Representations

Colors are defined in different color spaces and converted into multi-dimensional digital or analog signals suitable for the purpose. The most common color spaces can be listed as Red-Green-Blue (RGB), Cyan-Magenta-Yellow-Black (CMYK), Luminance-Blue Projection-Red Projection (YUV), Luminance-Orange Blue Range- Purple Green Range (YIQ), Luminance-Blue Difference- Red Difference (YCbCr), Hue-Saturation-Value (HSV) and Hue-Saturation-Lightness (HSL). In this work, RGB, HSV and YCbCr representations are introduced and the proposed RGB-YHS features obtained from these representations are explained.

C. RGB Representation

Title RGB is the abbreviation of "Red-Green-Blue". RGB refers to creating various colors by mixing three light tones. Unlike the subtractive model "Cyan-Magenta-Yellow-Black (CMYK)", RGB is an additive model. This representation was firstly proposed in 1855 [8] and used for color photography in 1861 [9] by James Clerk Maxwell. RGB representation has evolved according to the technology used in photography, television, and the latest personal computers. The RGB notation used in this study is an 8-bit digital RGB notation, which uses 8 bits for each color band, so each color band has 28 different possible shades. Since there are 3 different bands, 224 different colors can be defined.

D. Proposed RGB-YHS Representation

The RGB representation is open to improve for skin segmentation, because it is affected by shadow and light changes, causing skin pixels to be confused with nonskin pixels. In this study, all the features of RGB representation are inherited, and Saturation and Value features of HSV and Luminance feature of YCbCr are included. The reason for choosing the luminance(Y) component of YCbCr instead of the Value (Brightness) component of the HSV representation is that luminance is an objective measurement and the brightness is the opposite subjective measurement, which is dependent to human perception. In fact, luminance and brightness are strongly correlated concepts, but luminance is predicted to be a healthier measurement than brightness. Because of this objective feature of Luminance, it is also preferred instead of Lightness in HSL representation, because lightness is also a subjective measurement unit based on human perception like brightness. The reason why the blue difference (Cb) and red difference (Cr) features are not preferred is that the information carried by these features can be calculated from the linear combination of R, G and B. Therefore, it is predicted that the classifier will also extract the information that it can obtain from Cb and Cr from Red-Green-Blue directly.

III.SUPPORT-VECTOR-MACHINES

Support-Vector-Machine (SVM) is a classification method, which was originally proposed for two class problems [10]. SVM aims to find the best hyperplanes that distinguish data samples of one class from the data samples of the other class. In the original SVM, this separating hyperplane is assumed as linear. The best separating hyperplane for SVM obtains the largest margin between the two classes. Support vectors are samples of the classes closest to the separating hyperplane and these samples are on the border of the respective classes. The distance between the planes passing through the support vectors and parallel to the separating hyperplane determines the length of the margin.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

A. SVM with fine Gaussian Kernel

The original maximum Support-Vector-Machine algorithm proposed by Vapnik in 1963 was designed as a linear classifier, as mentioned earlier. However, it was observed that it was not sufficient to classify real life data. In real life data, the hyperplane between two classes could be indented or curved, which can be defined only by a nonlinear function. In order to solve this problem, Aiserman, Braverman, Emmanuel and Rozonoer first proposed kernel trick in 1964 [11-13]. Bernhard Boser, Isabelle Guyon and Vladimir Vapnik improved this proposed kernel trick and applied it to SVM in 1992 [14-16]. The resulting nonlinear classifier was replaced by a nonlinear kernel function of each data sample. Thus, a maximum-margin hyperplane fit into a transformed feature space is achieved. Although the classifier has a linear hyper plane in the transformed feature space, it shows nonlinear property in the original sample space. The mostly used nonlinear kernel types can be listed as quadratic, cubic and gaussian. The other name of the gaussian kernel is radial basis function (RBF). The general equation of the gaussian kernel is given in (1).

$$k(\vec{x_i},\vec{x_j}) = e^{-\frac{\left\|\vec{x_i}-\vec{x_j}\right\|^2}{2\sigma^2}} \quad (1)$$

In the fine tuning of the Gaussian kernel, Matlab offers 3 different options: coarse, medium and fine. The fine gaussian kernel used in this study allows for a detailed separation of classes from each other by defining kernel scale as $\sqrt{P}/4$, where P is the number of predictors.

IV. COMPARATIVE RESULTS

As shown in Table 1, Schmugge skin image data set contains 33,693,901 number of pixels, where 8,281,509 of them are skin pixels and 25,412,392 of them are nonskin pixels. Due to these numbers, it can be said that 24.58% of pixels belong to skin and 75.42% of pixels belong to nonskin regions, which means that the skin class (positive class) is minority class and nonskin class (negative class) is majority class. These ratios cause unbalanced classification problem. In this study, the unbalanced classification problem is solved by training the fine gaussian SVM with equal number of randomly selected skin and nonskin pixels. 16564 random pixels whose 8282 of them belong to skin and 8282 of them nonskin pixels are chosen to train the SVM. When the pixels are randomly selected, it is guaranteed that it is absolutely unknown which images they belong to, and the chosen pixels are also shuffled among themselves so that the classifier does not memorize. The size of train set is 0.05% of the size of entire set, which means shallow learning is applied to the SVM with fine gaussian kernel. On the other hand, the trained classifier is tested by whole set of landscape images, whole set of community images, whole set of female images, whole set of male images and finally entire Schmugge data set, respectively. An example for a skin segmentation on a landscape image is illustrated in Figure 1.



(a) (b) (c) (d) (e) (f)

Figure 1: Fine Gaussian SVM predictions for a landscape image (a) Original image (b) Ground truth (c) Prediction by RGB features (d) Prediction by HSV features (e) Prediction by YCbCr features (f) Prediction by RGB-YHS features

All classifiers based on RGB, HSV, YCbCr and RGB-YHS features mistakenly detected some skin pixels due to their similarity to the skin pixels in the training set. Especially, YCbCr based Fine Gaussian SVM confused for dark beach pixels in the given example. Overall performances of the fine Gaussian SVM for 300 landscape images are given in Table 2.

TABLE III: Performances of SVM with fine gaussian kernel for landscape images

Features	Average Accuracy	Average Sensitivity	Average Specificity	Average F1 score
RGB	0.8986	1	0.8986	0.0535
HSV	0.8920	1	0.8920	0.0535
YCbCr	0.8973	1	0.8973	0.0502
RGB-YHS	0.8946	1	0.8946	0.0500

RGB-YHS features give a performance between HSV and RGB for landscape images via accuracy and specificity. The average sensitivities of all classifiers are 1 for landscape images, because there is no positive class sample to sense in landscape images. The existence of false positives causes F1 scores to converge 0. An example for a skin segmentation on a community image is illustrated in **Error! Reference source not found.**.

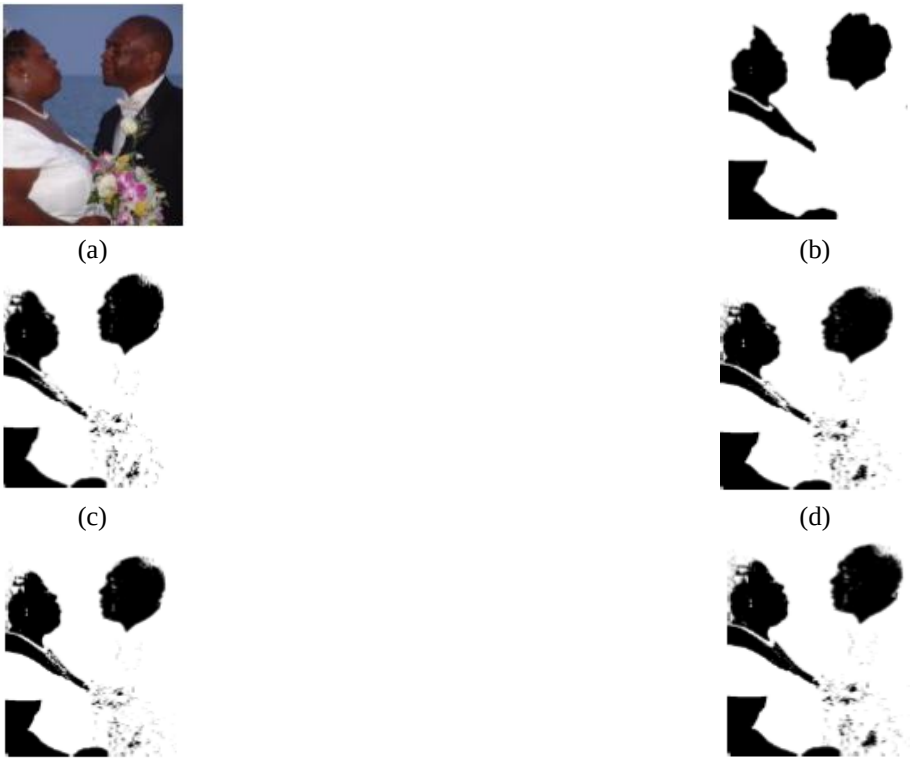


Figure 2: Fine Gaussian SVM predictions for a community image (a) Original image (b) Ground truth (c) Prediction by RGB features (d) Prediction by HSV features (e) Prediction by YCbCr features (f) Prediction by RGB-YHS features

The fine gaussian SVM classifier, trained with YCbCr features, also perceives hair as skin for male image as in the female image. The proposed RGB-YHS features decrease the false positives on lips and eyes as in the female image. The shadow under the chin and above the neck was mistakenly perceived as a nonskin region for all other features except YCbCr. The detailed performances for 354 male images are given in Table III.

TABLE IIIII: Performances of SVM with fine gaussian kernel for male images

Features	Average Accuracy	Average Sensitivity	Average Specificity	Average F1 score
RGB	0.8676	0.8962	0.8355	0.8392
HSV	0.8702	0.8950	0.8427	0.8411
YCbCr	0.8476	0.9249	0.7748	0.8297
RGB-YHS	0.8681	0.9036	0.8296	0.8414

Best F1 score is obtained also by proposed RGB-YHS features for male images, which means that RGB-YHS features optimize the balance between sensitivity and specificity as in female images. On the other hand, YCbCr gives also maximum sensitivity and minimum specificity in male images. The detailed performances for all 846 images in Schmugge skin image data set are given in Table IV.

TABLE IVV: Performances of SVM with fine gaussian kernel for entire Schmugge skin image data set

Features	Average Accuracy	Average Sensitivity	Average Specificity	Average F1 score
RGB	0.8790	0.9270	0.8641	0.5498
HSV	0.8782	0.9258	0.8660	0.5512
YCbCr	0.8645	0.9450	0.8263	0.5421
RGB-YHS	0.8781	0.9305	0.8601	0.5501

According to Table IV, the proposed RGB-YHS features show similar performance with HSV in terms of accuracy. The proposed RGB-YHS show better sensitivity than RGB and HSV. The sensitivity of RGB-YHS is less than YCbCr, but the specificity of RGB-YHS is much better than YCbCr. The F1 score of RGB-YHS is better than RGB and YCbCr.

V. CONCLUSIONS

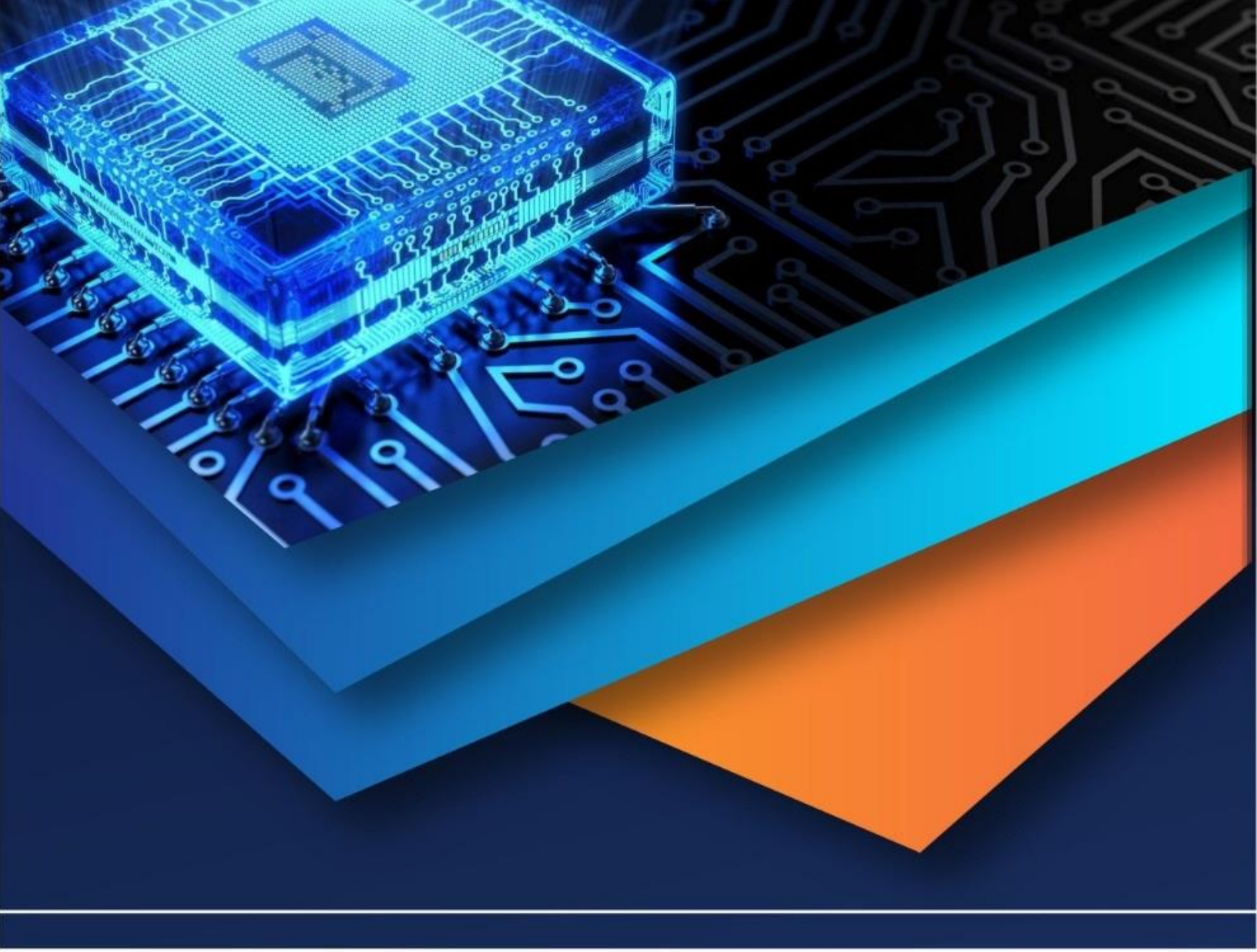
The advantage of the proposed RGB-YHS features is especially evident in male and female pictures containing only facial pixels. The success of the proposed RGB-YHS in community images was only able to surpass YCbCr, but this problem can be solved by training the fine gaussian SVM classifier with more samples of community images in future studies. In addition, the applied shallow learning should be deepened in order to prevent the wrong skin pixel detection in the landscape pictures in future studies. It is not prevented that intense shadows on the skin cause false nonskin pixel recognition in all color representations. In this case, it is thought that color information is not sufficient. This situation is thought to be solved by a region-based study, not pixel-based study. Although region-based studies are available in the literature, the success of this proposed research can also be increased in future studies by including region-based approaches. It is thought that high performance will be achieved with deep learning methods, but due to the nature of the deep learning, it will increase the computational complexity of the training stage to very high values. By conclusion, the study conducted with the proposed method showing promising performances on the selected dataset is considered to be open to development by expanding the proposed features, improving the classifier, and using a more comprehensive training set.

REFERENCES

[1] L. Liu et al., “Deep Learning for Generic Object Detection: A Survey,” Int. J. Comput. Vis., vol. 128, no. 2, pp. 261–318, 2020, doi: 10.1007/s11263-019-01247-4.
[2] Z. Cai, Q. Fan, R. S. Feris, and N. Vasconcelos, “A unified multi-scale deep convolutional neural network for fast object detection,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2016, vol. 9908 LNCS, pp. 354–370, doi: 10.1007/978-3-319-46493-0_22.
[3] D. G. Ganakwar, “A Case Study of various Face Detection Methods,” Int. J. Res. Appl. Sci. Eng. Technol., vol. 7, no. 11, pp. 496–500, 2019, doi: 10.22214/ijraset.2019.11080.
[4] K. W. Wong, K. M. Lam, and W. C. Siu, “An efficient color compensation scheme for skin color segmentation,” Proc. - IEEE Int. Symp. Circuits Syst., vol. 2, pp. 676–679, 2003, doi: 10.1109/iscas.2003.1206064.
[5] H. J. Lin, S. Y. Wang, S. H. Yen, and Y. T. Kao, “Face detection based on skin color segmentation and neural network,” Proc. 2005 Int. Conf. Neural Networks Brain Proceedings, ICNNB’05, vol. 2, pp. 1144–1149, 2005, doi: 10.1109/icnnb.2005.1614818.
[6] skin image Data set with ground truth. [Online]. Available: https://www.researchgate.net/publication/257620282_skin_image_Data_set_with_ground_truth
[7] Schmugge SJ, Jayaram S, Shin MC, Tsap LV. Objective evaluation of approaches of skin detection using ROC analysis. Computer vision and image understanding 2007; 108(1-2): 41-51.

[8] Maxwell JC. On the Theory of Colours in Relation to Colour-blindness. 1855.
[9] Maxwell JC. On the theory of three primary colours. 1861: Royal Institution of Great Britain.
[10] Hunt RW. The specification of colour appearance. I. Concepts and terms. Color Research & Application 1977; 2(2):55-68.
[11] Smith AR. Color gamut transform pairs. ACM Siggraph Computer Graphics 1978; 12(3): 12-19.
[12] Wharton W, Howorth D. Principles of television reception. 1967: Pitman.
[13] Rec. ITU-R BT.601-5. Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-screen 16:9 Aspect Ratios. (1982-1986-1990-1992-1994-1995): Section 3.5.
[14] Vapnik V. Pattern recognition using generalized portrait method. Automation and remote control 1963; 24:774-780.
[15] Aiserman M, Braverman EM, Rozonoer L. Theoretical foundations of the potential function method in pattern recognition. Avtomat. i Telemeh, 1964; 25(6): 917-936.
[16] Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. In: Proceedings of the fifth annual workshop on Computational learning theory; 1992: 144-152.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)