



Human Face Detection Using Skin Color Segmentation and Watershed Algorithm

Abdulganiyu Abdu Yusuf^{1,*}, Fatma Susilawati Mohamad², Zahraddeen Sufyanu³

¹National Biotechnology Development Agency (NABDA), Abuja, Nigeria

²Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin (Uni SZA), Terengganu, Malaysia

³Department of Computer Science, Federal University Dutse, Jigawa State, Nigeria

Email address:

abdulg720@gmail.com (A. A. Yusuf)

*Corresponding author

To cite this article:

Abdulganiyu Abdu Yusuf, Fatma Susilawati Mohamad, Zahraddeen Sufyanu. Human Face Detection Using Skin Color Segmentation and Watershed Algorithm. *American Journal of Artificial Intelligence*. Vol. 1, No. 1, 2017, pp. 29-35. doi: 10.11648/j.ajai.20170101.14

Received: May 15, 2017; **Accepted:** June 6, 2017; **Published:** July 24, 2017

Abstract: Face detection receives immense interest in computer vision to improve security and authenticity of a particular system. Color provides useful information at the early stage of face detection in a complex view. Such detection involves many complexities such as background, illumination, and poses. This study provides depth analysis on most prominent color models. The use of those color models can handle well-defined problems in face detection such as occlusions, poses, and illumination conditions. The application areas, techniques used, remarks as well as statistical conversion of the color models from Red Green Blue (RGB) color model are demonstrated. Moreover, a new framework for efficient face detection using skin color segmentation is proposed. The process involves transforming the face images from RGB to the selected color models; then segmentation is carried out by selecting a threshold value for each of the color models. Watershed algorithm is applied to isolate the facial feature from the background. Finally, lips area is localized as it may be missing during the detection process. Detection rate of up to 97.22% was obtained using standard database. The proposed framework targets a range of applications such as PC login security, passport authentication, and pornography filtering.

Keywords: Face Detection, Color Model, Watershed Algorithm, RGB

1. Introduction

Face detection is the first step in facial feature extraction and recognition process. The main purpose of face detection is to determine whether or not faces are present in an image and if present, return the location and content of each face. Therefore, both efficiency and accuracy of the process are improved by searching for the facial region instead of the whole image. Many researchers work in this area because it is a crucial step in many applications such as content-based image retrieval, Human Computer Interaction (HCI), security systems, face recognition and human crowd surveillance systems [1]. Humans can identify thousands of faces despite considerable changes in terms of skin color. This sort of information plays an important role during visual communication among humans. The use of facial information during interaction is achieved by the ability of humans to accurately identify and interpret faces in real-time. The

characteristics used to distinguish one color from another are luminance (the brightness) and chrominance (hue and saturation). Skin color is a valuable cue for human face detection in images since (i) it is fast, simple and has low computational cost (ii) it is invariant to image size and orientation (iii) it gives extra dimension compared to grayscale methods and (iv) it can classify entire image into face and non-face region [2].

According to Kumar in [2], the methods used for detecting faces in colored images include Chrominance based, Skin color based, AdaBoost based, Segmentation based and Neural Network based method. To utilize color as in visual cue in, image processing, multimedia, computer vision applications and computer graphics, an appropriate method for representing the color signal is needed. Different color spaces possess different characteristics as applied to various visual tasks. The term color space is sometimes referred to as color model. Examples of color spaces are Red Green Blue

(RGB), Hue Saturation Value (HSV), Chroma: Blue; Chroma: Red (YCbCr), Cyan Yellow Magenta and black (CYMK). There is no consensus on which color model is the most appropriate for skin color detection.

This study provides deep analysis on some prominent color models that handle well-defined problems in face detection such as noise, background complexity and illumination conditions. This consists of statistical conversion of the color models from RGB, with their corresponding advantages and disadvantages as well as their suitable application areas. Furthermore, a new framework for efficient face and facial feature detection using skin color segmentation is proposed.

The rest of the paper is organized as follows; Section 2 discusses current works on the color models. Section 3 demonstrates color models application areas, pros and cons, and their statistical conversions. An easy framework for face and facial feature detection is presented in Section 4. Result and discussion is given in Section 5. Finally, a conclusion is drawn in Section 6.

2. Discussion of Previous Literature

Skin color represents the actual color tone of each image. A skin color intervals need to be determined to find out whether a pixel is in a color space is skin-colored or not. For example, the study by [3] define a range for pixels in the HSV color space by simply setting up two boundaries on the hue scale to detect swimmers in a swimming pool. In the same way, [4] defined intervals on the Cb and Cr value of the YCbCr color space to find skin pixels. The interval borders are chosen empirically by analyzing the color frequencies of 500 images featuring skin pixels. Duan *et al.* [5] also intersected two intervals which are defined on the I-component (i.e. red-orange tones) of the Luminance In-phase Quadrature (YIQ) color space and the angle of the vector on the UV plane of the Luminance Blue-luminance Red-luminance (YUV) color space. The main drawbacks of these approaches are the manual choice of color space and the

associated definition of the constraints are somewhat arbitrary.

Another way of defining a skin color model is to compute a skin color histogram and a non-skin color histogram from training images, where pixels are manually labeled as ‘skin’ or ‘non-skin’. Jones and Rehg [6] used a set of 13,640 images, where manually labeled skin pixels (in a subset of 4,675 images containing skin) to construct a 3D histogram with 256 bins for each color channel. They suggested that histograms with 32 bins for each color are sufficient and even slightly outperformed the more detailed histograms at classification experiments since the latter tends to cause over-fitting. The research conducted by Zeng *et al.* [7] computed RGB histograms for skin color detection for three different brightness clusters: bright, normal and dark images, based on the average image brightness. The results obtained were also compared to the color model used in [6] on skin pixels that have been manually annotated in 829 adult images.

Features that describe the shape of skin-colored image segments were created in [8], through the use of HSV histogram to find initial skin segments and grow them to include less likely skin pixels in the vicinity using empirically determined thresholds. Color histograms can also be evaluated fast, and it is usually updated easily with new training data. And they sometimes lack the ability to generalize beyond training data which is limited to a finite number of depicted people and lighting conditions.

Zhengming *et al.* [9] proposed a new technique to overcome the time consumption that can be applied in a real-time system. In this technique set of pixels were skipped and the remaining pixels were labeled as skin and non-skin pixels by using RGB color space. The performance of the histogram technique is degraded due to the presence of overlap between the skin and non-skin classes. The detection rate of this technique is slightly higher than the detection rate of Gaussian Mixture Model (GMM). However, a standard dataset is needed for good classification rate. Large number of images for training and testing is also required. Table 1 depicts the performance of some color spaces.

Table 1. Performance of some color spaces for skin detection.

Color space	Database	Evaluation method	Factors considered in the dataset	Best color space
HSV [3]	From web	Recognition rate	Skin tone, Background	---
YIQ and YUV [5]	From web	Detection rate	Skin background	---
RGB [10]	---	False acceptance rate/ False rejection rate	Region of interests (ROIs)	---
RGB [7]	From web	Precision and recall	Skin texture	---
YCbCr and RGB [11]	Edith Cowan University (ECU)	True positive/False positive	Skin tone	RGB
HS and YCbCr [12]	From web	Not revealed	---	N/A
YCgCr + YCgCb [13]	From web	Hit/False rate	Skin with complex background	N/A
YCbCr [1]	Not revealed	False positive/Detection rate	Background complexity	---
RGB, YUV, HSV, CMYK, and YCbCr [14]	Psychological Image Collection at Stirling (PICS)	Detection rate	Energy of the histogram of each component of the color space, the limit of skin range in each color space and the maximum energy of the color spaces.	YCbCr

From the survey, it can be concluded that

- Only few studies considered the problem of color space selection and provided justifications for optimality of

their choice.

- There is no fixed color model for skin color detection.
- Most of the techniques find it difficult to easily segment

- the facial features from the background and
- d. There are no standard databases for the benchmark comparison.

3. Application of Color Models

Color provides a useful source of information for a wide range of research areas such as segmentation, video analysis, and object recognition. Some of the original colors in an image might not be suitable for analysis. Hence, the colors need to be adjusted. Adjusting colors can be done by

transforming the colors from one space to another, while maintaining the image's original details and natural looks simultaneously [15]. Face detection using skin color involves two major steps: (i) representing the image using proper color space, and (ii) modeling the skin and non-skin pixels using inference methodology to obtain information from available skin samples and to deduce the results to given samples. To locate the candidate's facial regions, there is need to model the skin color which requires choosing an appropriate color space at first. Table 2 depicts comparative study of different color spaces used for face detection.

Table 2. A comparative study of different color space used in face detection.

Color Space Categories	Advantages	Disadvantages	Remarks	Application Area
Basic color spaces (RGB and Normalized RGB)	The basic color spaces nullify brightness discrepancies among pixels.	The basic color spaces give high false positive and negative rates.	They are not the best color space for skin detection because some other non-skin objects may have their RGB value	Computer Graphics, Image processing, Analysis, Storage
	The basic color spaces contain no luminance information	Detecting skin region is discontinuous in basic color spaces due to lightning condition	Loss of data since luminance information is eliminated.	
Perceptual Color Spaces (HSV, Hue Saturation Intensity (HSI) and Hue Saturation Luminance (HSL))	Using perceptual color spaces, skin area is easily detected when it falls within the threshold	Perceptual color spaces are mostly unable to detect people with different skin region	Perceptual color space give good performance under simple image background	Human visual perception, Computer graphics, Computer Vision, Image Analysis, Human vision, Image editing software, Video editor
	High detection performance	Sometimes false positive occurrence		
Orthogonal Color spaces (YCbCr, YUV, and YIQ)	The intensity component of orthogonal color space can easily be accessed	The detection rates of orthogonal color space decrease when face merges with ground color.	It could be a good choice for face detection under simple background	TV broadcasting, Video System, Digital video
	Perceptually uniform and widely used in the video compression method. Having separation of luminance.	Face difficulties in case of low -luma and high-luma skin tone.		

4. Conversion of RGB to Other Color Models

Colors are represented in terms of three primary colors: Red (R), Green (G) and Blue (B). It is an additive model because all colors are created from these primary colors. The red, green and blue color components are highly related; for example, red (255, 0, 0) and green (0, 255, 0) combined in the same amounts to create yellow (255, 255, 0). The amount of each primary color gives its intensity but if all components are of highest intensity, then white color is resulted. The RGB model simplifies the design of computer graphics systems but is not suitable for all applications. And it is not the most used in skin detection nowadays because it is not very robust when lighting condition changes [16]. The face images in this study are converted from RGB to YUV, HSV, YIQ and YCbCr color spaces using Equations 1, 2, 3 and 4, respectively.

- a. From RGB to YUV

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

Where $\begin{bmatrix} Y \\ U \\ V \end{bmatrix}$ are the components of YUV color space, $\begin{bmatrix} R \\ G \\ B \end{bmatrix}$ are the components of RGB color space and $\begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix}$ is the multiplication factor.

b. RGB to HSI

$$I = \frac{1}{3}(R + G + B)$$

$$S = 1 - \frac{3}{(R+G+B)}[\min(R, G, B)] \quad (2)$$

$$H = \cos^{-1} \left\{ \frac{0.5[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right\}$$

If $B < G$, then $H = 360 - H$

Where R, G, B are the components of RGB color space and H, S, I are the components of HSV color space.

c. RGB to YIQ

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.586 & -0.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3)$$

Where $\begin{bmatrix} Y \\ I \\ Q \end{bmatrix}$ are the component of YIQ color space, $\begin{bmatrix} R \\ G \\ B \end{bmatrix}$ are the component of RGB color space and $\begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix}$ is the multiplication factor.

d. RGB to YCbCr

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4)$$

Where YCbCr are the components of YCbCr color space, RGB are the components of RGB color space and $\begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix}$ is the multiplication factor.

5. Proposed Framework

The proposed framework for face detection is based on color information, so the input image must be color image rather than grayscale images. Figure 1 describes the process of the proposed framework.

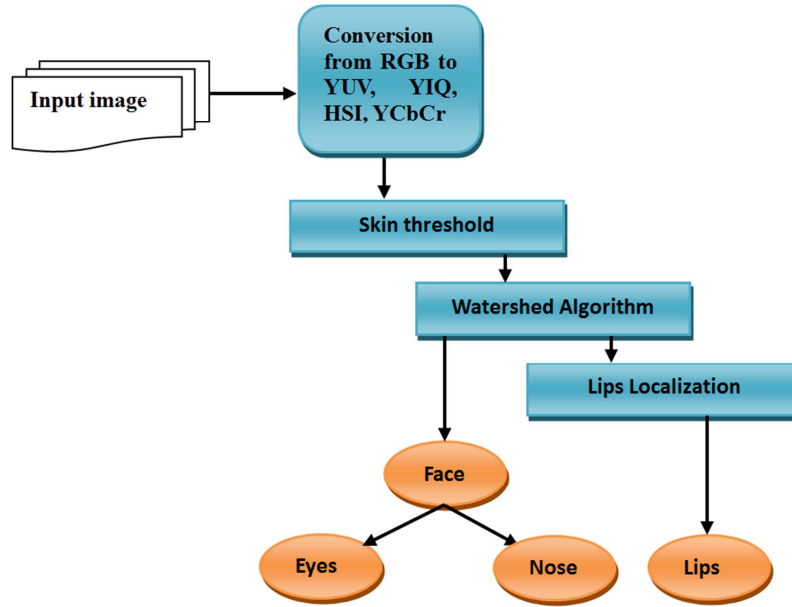


Figure 1. Framework of the proposed detection system.

a. Input Image

Input images with different backgrounds are selected. The selection should be from standard databases to enable comparison with benchmark results and to be deployed in the real-world applications. The level of contrast in an image may vary due to poor illumination or improper setting in the acquisition sensor device. Therefore, there is need to manipulate the contrast of an image in order to compensate for difficulties in image acquisition. Also, the images become much clearer and noiseless. The pixel values of a low-contrast image or high-contrast image are stretched by extending the dynamic range across the whole image spectrum. The function “stretchlim” in Matlab simulator is

used to find limit to contrast stretching of image. Then, the size of the image is scaled down to reduce the resolution of face image. This reduction in size is done to reduce complexity during training process.

b. Skin Threshold

By default, images are represented in RGB space color. The problem with RGB is that it is easily affected by illumination and pose changes. So, there is need to transform the images from RGB to another color model in order to separate the luminance and chrominance information. The segmentation is done by choosing threshold values for each channel of the color model. The candidate's face region can be found using Equation 5.

$$Face_{candidate} = \begin{cases} 1, & a \ll Face_a \ll 255, b \leq Face_b \ll 255, \text{ and } c \leq Face_c \ll 255 \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

Where $a \ll Face_a \ll 255$ is the threshold value for the component of the color space $Face_a$, ranging from 0 to 255, $b \leq Face_b \ll 255$ for $Face_b$ and $c \leq Face_c \ll 255$ for

$Face_c$.

For color model conversion, different equations are present for different color models. In this study, Equations 1, 2, 3 and

4 are used to convert face images from RGB to YUV, HSI, YIQ and YCbCr respectively.

c. Watershed Algorithm

After skin color segmentation there remains some small noise. Those are reduced by using watershed algorithm. The watershed algorithm is the method of choice for image segmentation in the field of mathematical morphology [17]. It is the process of isolating objects in the image from the background by partitioning the image into disjoint regions, such that each region is homogeneous with respect to some property, such as grey value or texture. Watershed algorithm has been employed in two stages. The first stage is to isolate the biggest connected white areas and merges small isolated white areas to the background. The second stage is to separate the facial features (in black) from the background (in grayscale). At this stage, the facial features such as a pair of eyes and nose can be seen clearly. But lips color commonly interferes with skin color and may cause missing lips during detection.

d. Lips Localization

There are several methods for mouth region localization. Among them only 'skin color segmentation' depends on color properties of the mouth region. From the literature, there are lack of mouth detection results specifically dealing with thick beard, mustache and open mouth. Lips color region contains a stronger red component and weaker blue

and green component than the other facial regions. The potential lips area can be localized using [18] formula in Equation 6.

$$LipMap = \left\{ \left(\frac{r}{r+b} \right) * \left(1 - \frac{r}{r+b} \right) \right\}^2 \quad (6)$$

Where b and r are the blue and red chrominance information of the selected color space. The skin segmented image may contain any part of the face like a pair of eyes and the mouth.

6. Results and Discussion

The experiment is conducted on Psychological Image Collection at Stirling (PICS) database. This database consists of 369 images with subjects between 1 to 18 images. The images resolution varied from 336 x 480 to 624 x 544 pixels. A total number of 72 images were used for training and 369 images testing. Equal number of males and females were selected for the training. The images are renamed in the order Pm1 to Pm18 for males and Pf1 to Pf18 for females. Then images are preprocessed by contrast stretching and then reduced to the size 270 × 200 pixels. Figure 2 shows the results of detection process of a face image for each of the selected color models.

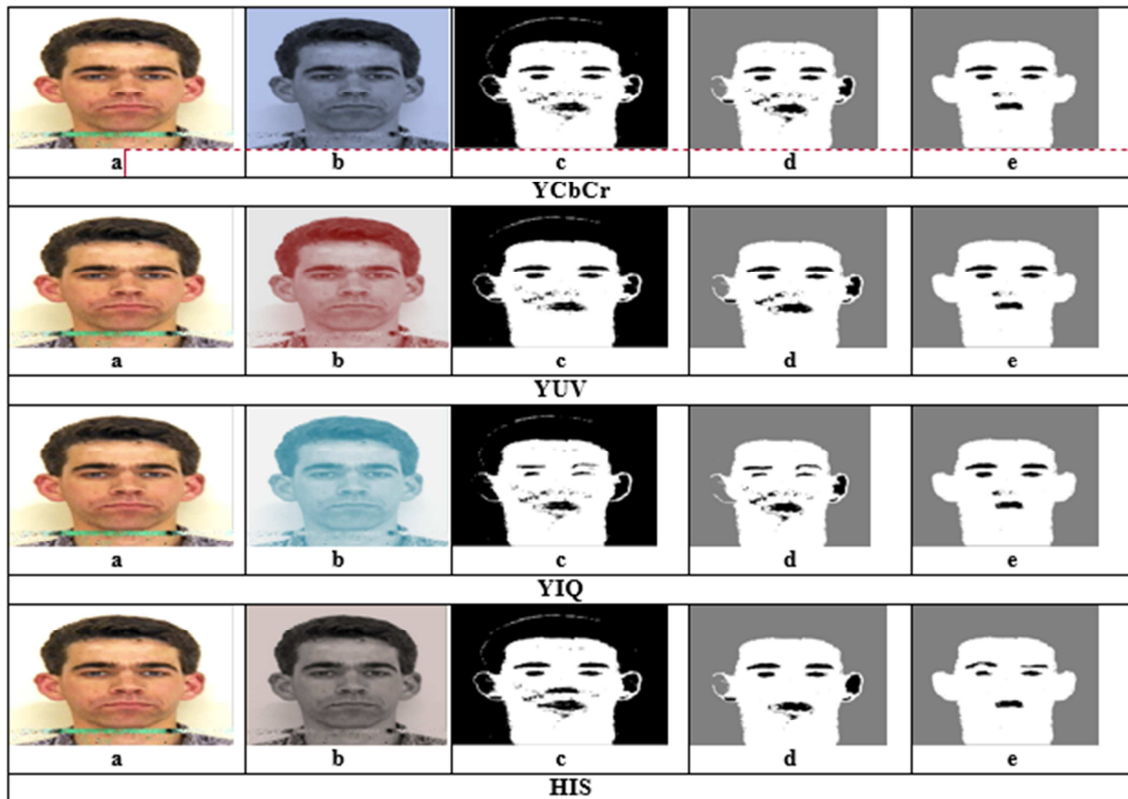


Figure 2. Sample of face image undergoing the stages of detection.

In Figure 2, the input images (a) are preprocessed and converted to the color models to produce images (b). Then, skin threshold is applied on the converted images (b) to

produce images (c), where the white color represents the skin area. Watershed algorithm was used at this point, and the biggest connected white area was isolated and small isolated

white areas were merged to the background. Then the background and facial features background were isolated as shown in gray in images (d) while the facial feature areas were shown in black. The watershed reduces the background noise and fills the missing pixels in the face region resulting in one connected face region. Finally, the lips and the pair of the eyes localized in images (e). Table 3 shows result of the face detection.

Table 3. Face detection result.

Color model	No. of Images	Perfect Detection	False Detection	Detection Rate %
RGB	360	292	68	81.11
YCbCr	360	350	10	97.22
YUV	360	346	14	96.11
YIQ	360	335	25	93.06
HSI	360	333	27	92.50

The table compares face detection result of the aforementioned color models. Different parameters are considered such as the number of images for testing, the number of images that are correctly detected, the number of images that are not accurately detected and their detection rates. Highest detection rate was obtained using YCbCr color model. Table 4 gives a comparison with some current related work.

Table 4. Comparison with some current studies.

Authors	Color space	Detection Rate %
(Ekta & Saroj, 2013) [19]	RGB	65.00
	HIS	70.00
	YCbCr	90.00
(Kumar, 2014) [20]	RGB	78.00
	HSI	70.00
	YCbCr	80.00
(Chandrashekar, 2012) [21]	RGB	59.50
	HIS	61.50
	YCbCr	72.50
(Singh et al., 2015) [1]	YCbCr	93.33
	RGB	81.11
	HIS	92.50
PICS (Proposed Method)	YIQ	93.06
	YUV	96.11
	YCbCr	97.22

From the table above, it is clear that higher detection rate was obtained using the proposed method applying the same color models.

7. Conclusion and Future Study

The color models presented in this study are utilized in different application areas in daily life. Each of the color models have a specific representation space and components as well as the ability to be transformed from one color space to another using standard formula. Some prominent color models with the demonstration of each color system characteristics were reviewed. It was deduced that the selection of a proper color model for a specific application depends on the properties of the model and the nature of application. Statistical conversion of the color models from RGB was also presented. Hence, a framework for efficient

face detection using skin color segmentation is proposed. A detection rate of up to 97.22% was obtained using standard database. In future, the technique will be improved by providing a platform for determining the range and proper color model.

References

- [1] Singh, P., Singh, S. K. & Gaba, N. (2015). YCbCr Technique based Human Face Recognition. *International Journal of Advance Research and Innovation*, 3 (1): 171 – 174.
- [2] Kumar, A. (2014). An Empirical Study of Selection of the Appropriate Color Space for Skin Detection. *In International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*, 725–730.
- [3] Liao, W. H. & Liu, M. J. (2004). Robust Swimming Style Classification from Color Video. *In Proceedings of the International Computer Symposium*, 541–546.
- [4] Arentz, W. A. & Olstad, B. (1994). Classifying Offensive Sites based on Image Content. *Computer Vision and Image Understanding*, 295–310.
- [5] Duan, G. L., Cui, W. & Zhang, H. (2002). Adult Image Detection Method based on Skin Color Model and Support Vector Machine. *In Proceedings of the 5th Asian conference on computer vision*, 797–800.
- [6] Jones, M. J. & Rehg, J. M. (2002). Statistical Color Models with Application to Skin Detection. *In International Journal of Computer Vision*, 81–96.
- [7] Zeng, W., Gao, W., Zhang, T. & Liu, Y. (2004). Image Guarder: An Intelligent Detector for Adult Images. *In Proceedings of the Asian Conference on Computer Vision*, 198–203.
- [8] Bosson, G. Cawley, Y. C. & Harvey, (2002). Non-retrieval: Blocking Pornographic Images. *In Proceedings of the International Conference on Image and Video Retrieval*, 50–60.
- [9] Zhengming, L., Zhan, T. & Zhang, J. (2010). Skin Detection in Color Images. *In IEEE Computer Engineering and Technology (ICCET), 2010 2nd International Conference on*, 1.
- [10] Yang, J., Fu, Z., Tan, T. & Hu, W. (2004). A Novel Approach to Detecting Adult Images. *In Proceedings of the 17th International Conference on Pattern Recognition*, vol 4, pp 479–482.
- [11] Phung, S. L. Bouzerdoum, A. & Chai, D. (2005). Skin Segmentation using Color Pixel Classification: Analysis and Comparison. *In IEEE Transaction on Pattern Analysis and Machine Intelligence*, 27 (1).
- [12] Qiong, L. & Guang-Zheng, P. (2010). A Robust Skin Color Based Face Detection Algorithm. *In 2nd International Asia Conference on Informatics in Control, Automation and Robotics*.
- [13] Zhengzhen, Z. & Yuexiang, S. (2009). Skin Color Detecting unite YCgCb color space with YCgCr color space. *In Image Analysis and Signal Processing, 2009. International Conference*, 221–225.

- [14] Fatma S. M., Abdulganiyu A. Y. & Zahraddeen S. (2015). Evaluation of Suitable Color Model for Human Face Detection. *International Journal of Advanced and Applied Sciences*. 2 (12), 46-50.
- [15] Liu, C. C. (2012). A Global Color Transfer Scheme Between Images based On Multiple Regression Analysis. *International Journal of Innovative Computing, Information and Control*, 8 (1A): 167-186.
- [16] Chaves-Gonzalez, J. M., Vega-Rodriguez, M. A, Gomez-Pulido, J. A. & Sanchez-Pérez, J. M. (2010). Detecting Skin in Face Recognition Systems: A colour Spaces Study, *Digital Signal Processing*, 806–823.
- [17] Ankit, C. & Neha, S. (2013). A Robust Rapid Approach to Image Segmentation with Optimal Thresholding and Watershed Transform. *International Journal of Computer Applications*. 65 (9), 1-7.
- [18] Ali, A. & Sedigheh, G. (2011). Robust Component-based Face Detection Using Color Feature. *In Proceedings of the World Congress on Engineering, London, UK*, 2, 2-6.
- [19] Ekta, R. & Saroj, K. L. (2013). Comparative Analysis of Skin Color Based Models for Face Detection. *Signal & Image Processing: An International Journal (SIPIJ)*, 4 (2), 69-75.
- [20] Kumar, A. (2014). An Empirical Study of Selection of the Appropriate Color Space for Skin Detection. *In International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*, 725–730.
- [21] Chandrashekar M. B. (2012). Automated Face Detection in Color Images using Skin Region and Adaptive Template Matching. *International Journal of Computer and Electronics Research*. 1 (3), 106-110.