Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 4, Issue. 3, March 2015, pg.442 – 450

RESEARCH ARTICLE

Single Face Detection Based on Skin Color and Edge Detection

Haidar R. Shakir[1], Loay E. George[2], Ghadah K. Tuama[3]

[1,2,3] College of Science, University of Baghdad, Iraq

[1] haidar.raad@scbaghdad.edu.iq, [2] loayedwar57@scbaghdad.edu.iq, [3] hgkta2012@yahoo.com

Abstract— In this paper a method is proposed for human faces detection in color images. Most of the current facial detection systems are based on learning algorithms to scan the whole image so as to identify the human face. However, such systems are subject to problems of pose variation, scale variation, illumination changes, among others. Rather than being dependent on illumination, the proposed detection method uses face outer edge and the skin tone which may appear in the tested color image. The first stage in the introduced method involves enhancement of the image; it is a vital stage particularly if the image has been obtained under unconstrained illumination condition. The next step involves segmentation of the skin in HSV space. The hole filling algorithm is used to refine the skin segmentation results. In order to ensure that all non-face regions are removed from candidate faces, the input image edges and skin tone are fused. Primitive shape features of the candidate's face are used in the confirmation process to verify the segmented candidate facial region. Some of the merits of the proposed method include detection of faces of different sizes, pose variation as well as faces that are making varied expressions under different illumination conditions.

Keywords— Face Detection, Facial Features, K-means Clustering, Gray World Assumption, Robert's Cross Edge

I. INTRODUCTION

There are myriad computer vision and biometric applications which draw heavily on human face detection. Such applications include large-scale face image retrieval systems, video surveillance, automatic face recognition, human computer interaction (HCI) among others. The initial stage of any face processing system involves detecting the presence of a human face followed by the position the face using a video or image [1].

The most prevalent challenge that may impede the face detection process is coping with a myriad of human face variations such as face orientation, face pose and scale, facial expression, skin color and ethnicity. The process may also be impeded by external factors such as complex backgrounds, occlusion, inconsistency in the illumination conditions and image quality [2].

Research in the field of face detection has in the last ten years immensely developed especially in the profusion of methods and approaches. The different methods of face detection in the literature have been broadly reviewed in recent surveys [3] [4]. Face detection that is based on color images has in the past few years also gained much attention. The most proficient way of extracting skin regions is through the use of color cues. This is because it allows easy localization of the potential facial regions without considering the texture and geometrical attributes of the face. Most current skin detection methods that draw on color properties are pixel-based [5], where every pixel is categorized individually and independently as skin or "non-skin". Some methods, on the other hand, use different statistical color models, for example the single Gaussian model [6][7], the Gaussian mixture density model [8], as well as the histogram-based model [9].

In color spaces where the skin pixels and non-skin pixels have higher discrimination at different illumination conditions, the luminance component is usually separated from the chromatic component. The most effective skin color models when making distinctions between different human skin colors are those that only function on chrominance subspaces; for example C_bC_r [10] and HS [11]. When the parametric decision rules are used to explicitly model the skin distribution on certain color spaces, it becomes possible to achieve different skin classifications. The rules of describing a skin cluster in RGB space were developed by Kovac et al. [12], while the classification of skin regions on both HSV and YC_bC_r spaces was done by Azad and Boroujeni [13].

A new color-based HSV skin clustering model that can be used in the detection of human faces is provided in this paper. The proposed model uses some additional information of the image properties that pertains to hue and chrominance to enhance the skin and non-skin pixel discrimination. K-means clustering in HSV, which is based on centroids, is used for the classification of skin regions in this model. The construction of these centroids is done using the distribution of skin color acquired from the training images. In order to further refine the classification of the regions that are extracted, morphological operations are combined.

The organization of this paper's is as follows: Section 2 gives a brief description of the steps involved in the proposed face detection system. Section 3 gives a description of how the HSV skin color model is constructed. In section 4, the use of morphological operations in the proposed face detection system's algorithm is described. Section 5 provides the findings and discussions of the experimental results. The conclusion and future work is provided in section (6).

II. SYSTEM LAYOUT

Figure (1) is an illustration of the proposed face detection system structure. It is involving two stages: (i) training stage and (ii) detection stage. The proposed color-based face detection approach is based on using a set of skin-cropped training images to enhance the initial formulation of the HSV skin model. The construction of the proposed color model is done by applying K-means clustering algorithm. The construction of color centroids is done by analyzing the distribution of skin color acquired during the stage of face detection.

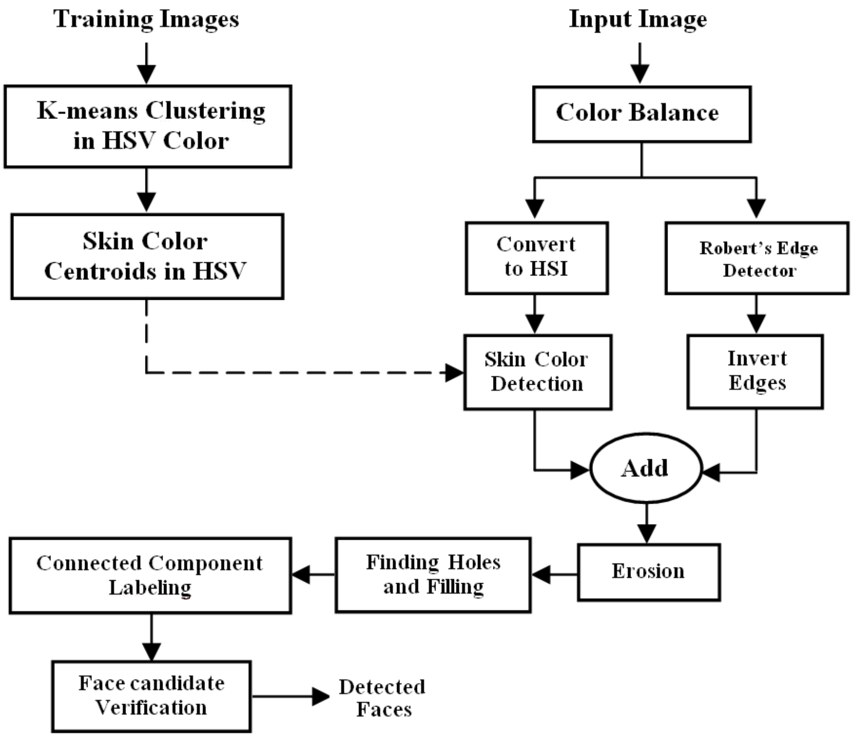


Fig. 1: Layout of the proposed face detection system

The first step of the stage of face detection involves the correction of the input images' color; which comes as a result of illumination condition, using color balance. Afterwards, the extracted skin regions are combined with the generated binary image as well as with the inverted edges images to enable the separation of face skin regions and the background. Then, this is followed by the application of erosion to take out the small regions in the image. Later, a hole filling algorithm is used to fill the holes areas in the image. Labeling of all segments is conducted after the holes are filled. Finally, the detection of the candidate's facial regions is accomplished. The proposed system concentrates on the use of a skin color model to detect human faces. As a result, the system can only be used with color images.

III. HSV SKIN COLOR MODEL

A. Training images

To develop a skin color model, the distributions and attributes of different color subspaces were analyzed using a set of training images. The training image set has 120 patches of skin color of 89 color images that are acquired from the Caltech face dataset [14]. The set therefore covers a huge array of variations (i.e. different ethnicities and skin colors). The images were comprised of skin color regions that were exposed to different forms of illumination; such as daylight illumination (outdoors), maluniform illumination, or flashlight illumination (under dark conditions). The 89 training images were manually cropped in order to examine the color distribution of their skin regions. Some of the samples of the cropped skin training images are illustrated in Figure (2) below.

Fig. 2: Cropped skin samples for training

B. HSV K-means Clustering

The training images' HSV color components are drawn on by the k-means clustering algorithm to identify the input images' skin colors. In order to determine each cluster's initial centroid value, each color band's (i.e. H, S, and V), mean (μ) and standard deviation (σ) is calculated in all skin patches. In this case, the initial color contents of the input image were represented using 8 centroids. The initial centroid vector mean vector and corresponding standard deviation vector are denoted using CentroidHSV, μHSV and σHSV respectively. Table (1) presents the initial centroids values that are produced as the following equation is applied:

$$Centroid_{HSv} = \mu_{HSV} \pm \sigma_{HSV} \quad (1)$$

Table (1) K-means initial centroids for HSV color

Cluster Number (k)	Centroid (C_k)		
	Hue (H)	Saturation (S)	Value (V)
0	$\mu_{Hue} + \sigma_{Hue}$	$\mu_{Sat} + \sigma_{Sat}$	$\mu_{Val} + \sigma_{Val}$
1	$\mu_{Hue} + \sigma_{Hue}$	$\mu_{Sat} - \sigma_{Sat}$	$\mu_{Val} + \sigma_{Val}$
2	$\mu_{Hue} - \sigma_{Hue}$	$\mu_{Sa}t + \sigma_{Sat}$	$\mu_{Val} + \sigma_{Val}$
3	$\mu_{Hue} - \sigma_{Hue}$	$\mu_{Sat} - \sigma_{Sat}$	$\mu_{Val} + \sigma_{Val}$
4	$\mu_{Hue} + \sigma_{Hue}$	$\mu_{Sat} + \sigma_{Sat}$	$\mu_{Val} - \sigma_{Val}$
5	$\mu_{Hue} + \sigma_{Hue}$	$\mu_{Sat} - \sigma_{Sat}$	$\mu_{Val} - \sigma_{Val}$
6	$\mu_{Hue} - \sigma_{Hue}$	$\mu_{Sat} + \sigma_{Sat}$	$\mu_{Val} - \sigma_{Val}$
7	$\mu_{Hue} - \sigma_{Hue}$	$\mu_{Sat} - \sigma_{Sat}$	$\mu_{Val} - \sigma_{Val}$

The purpose of applying k-means clustering algorithm is to establish the training images' minimum variance clusters. The objects should be assembled into k clusters so that k points {mj}(j=1, 2, …, k) can be found in D using the equation below:

$$D = \frac{1}{n}\sum_{i=1}^{n}\left|min_j\left(d^2(x_i, m_j)\right)\right| \quad (2)$$

where $d(x_i,m_j)$ represents the x_i and m_j's Euclidean distance. The cluster centroids are represented by the points $\{m_j\}$ (j=1, 2… k). The clustering of k-means ceases until k cluster centroids are found, to the extent that a reduction of mean squared Euclidean distance between a training image and the closest cluster centroid is achieved. The convergence state is arrived at by repeating the iterations of k-means algorithm up to seven times.

IV. FACE DETECTION

A. Color Balance

In face image light usually affects a person's skin color and this leads to a deviation from the real person's skin color. The correction in color images is enabled through Gray World Theory (GWT) [15], which is an algorithm of lighting compensation. In terms of description the GWT method functions as follows: the magnitude of stimulus of red (R), green (G) and blue (G) in the recorded scenery is identified and the average of the all color channels (i.e. R_{avg}, G_{avg} *and* B_{avg}) is calculated as illustrated in equations (3) & (4), and in Figure (3) shown below.

$$Gray_{avg} = \frac{Red_{avg} + Green_{avg} + Blue_{avg}}{3} \quad (3)$$

$$R' = R \times \frac{Gray_{avg}}{Red_{avg}}, \quad (4a)$$

$$G' = G \times \frac{Gray_{avg}}{Green_{avg}}, \quad (4b)$$

$$B' = B \times \frac{Gray_{avg}}{Blue_{avg}}, \quad (4c)$$

Fig. 3: Original image (left) and result of color balance (right)

B. Skin Color Detection

The image is converted into HSV color space after the application of the color balance. At this stage, the k-means clustering is used for the creation of a binary image which is black and white. The equation below is used to produce the binary image:

$$Minimum\ distance = \lim_{k=0..7}(|I_H - C(k).H)| + |I_S - C(k).S| + |I_V - C(k).V| \quad (5)$$

The input image's pixel is denoted by *I*. The pixel that is identified to have a minimal distance is set as white or black using the threshold T in accordance with the formula below:

$$Bin = \begin{cases} 1 & if\ Minimum\ distance < T \\ 0 & if\ Minimum\ distance \geq T \end{cases}, \quad (6)$$

C. Combining Edge Image with Skin Color Image

The majority of algorithms of face detection that are based on skin color only function well if the images have a non-skin tone background or if the people in the image are wearing non-skin tone clothing. If the image contains skin tone clothing or backgrounds, the algorithm identifies the entire region as a skin region as illustrated in Figure (4). In such case, the skin region of the candidate's face may merge with the background as illustrated in the figure. This necessitates an establishment of a mechanism that can separate the background regions from candidate's face to allow easy localization of the face. This problem is solved by taking the edges of the input image, particularly those of the skin regions, into consideration and combining the edge results with the results of the previous step.

The most common methods that are applied during edge detection include Robert's cross edge, Sobel edge, Prewitt edge, Canny edge, and the Laplacian of Gaussian edge. The method used in this study was Robert's cross edge. Roberts Cross edge wields the ability to conduct simple and quick computations of a 2-D image's spatial gradient. The output images' pixel values correspond to the estimated absolute magnitude of the input images' spatial gradient. As illustrated below, the Robert's cross edge is comprised of two 2×2 convolution kernels. Each kernel is a 90° rotation of the other, see figure (5).

Fig. 4: Original image (left) and result of skin color segmentation (right)

1	*0*
0	*-1*

G_x

0	*1*
-1	*0*

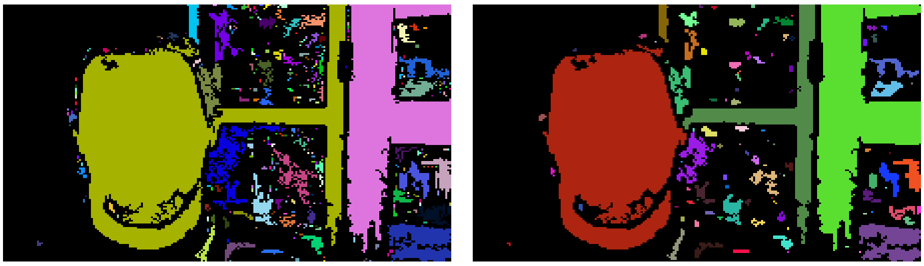
G_y

Fig. (5) Robert's 2×2 convolution kernels

D. Erosion

The implementation of edge detection facilitated the separation of face region from other objects such as the background and clothing. However, in some instances, full separation of the face region from other objects cannot be achieved through fusion with edge image; this failure is due to resemblance of colors among the neighbor objects. For this reason, erosion was also included in the proposed face detection system to increase its separation effectiveness. An illustration of face segmentation results before and after being subjected to erosion is shown in Figure (6).

In erosion, the binary image's segments are shrunk and its pixels are scanned. Detection of a white pixel leads to checkup of its 8-connected neighbors. If at least one of the 8-connected neighbors is found to be black, the white pixel is removed.



A B

Fig. 6: Illustration of erosion (A) face segment before erosion (B) Face segment after erosion

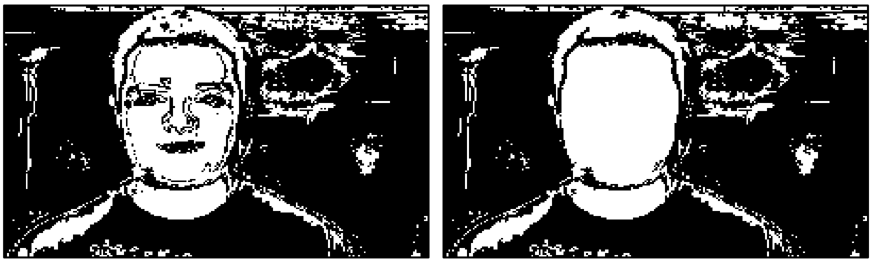
E. Hole Filling

The purpose of holes filling process is to get rid of the major holes may appear in the image after thresholding stage. Holes filling process turns the faces into one connected region devoid of the numerous facial cavities and holes. The hole filling process is contingent on labeling of the background segments. During the process of labeling each connected background is assigned a special ID. The steps followed in the method are highlighted below [16]:

1. The regions that correspond to background intensity are found and labeled.
2. The regions that belong to background and/or holes are separated in accordance with the following condition below:
   Each collected region is regarded as a background if the pixels end with image limit. If not, it is regarded as a hole.

Step 2 is repeated in regions that are identified to have similar intensities with the background.

For purpose of hole filling, the intensity of the collected hole pixels are changed to match the intensity of the surrounding. Figure (7) presents an example for illustrating the effect of hole detection and filling stage.



(A) Image before holes filling (B) Image after hole filling

Fig. 7: Example of a hole filled image

F. Connected Component Labeling

The image generated after the application of holes filling stage is subjected to the stage of search for connected components using the adjacent 8-neighbor pixels test. The adopted procedure for carrying out this task is as follows [17]:

1. The procedure begins with labeling the 1D connected components of every row.
2. Then, the labels of row-adjacent components are merged by means of an associative memory scheme.
3. Finally, the consecutive sets of positive component labels are relabeled.

This procedure helps in the identification of all connected components. The step that follows next pertains to deciding the connected components that correspond to a face and those do not.

G. Candidate Face Verification

At this stage, the remaining components are subjected to a set of shape based connected operators such as *Area*, *Solidity*, *Centroid*, *Orientation*, and *Ellipse area* to make decisions whether the tested shape correspond to a face or not. Through these operators, the basic assumptions regarding the shape of a face are drawn upon. If a component does not correspond to the shape of the candidate's face, it is removed. The criteria used in making these simple and effective decisions is contingent on the combinations of the perimeter P, area A, and D_x & D_y of the connected component's min-max box. Consequently these features are only calculated once for the three operators.

The *Area* represents the number of pixels of a given shape. Small components are eliminated as any one is less than 15% of the total image pixels. *Solidity* refers to the ratio of its area to the min-max box (i.e., the rectangular bounding box) area:

$$Solidity = \frac{A}{D_x D_y} \quad , \tag{8}$$

The value of solidity represents the area occupied by a connected component in its min-max box dimensions.

The face components often have high solidity values. If the tested shape is found to have a solidity values lower than the specified 0.5 or greater than 0.9 thresholds, it can be discarded or preserved so that it can be further analyzed.

O*rientation* refers to the aspect ratio of the min-max box that surrounds the component [18]:

$$Orientation = \frac{D_y}{D_x} , \tag{9}$$

There is an assumption that there is a certain range of the orientation of face components. The determination of this range is done through observations on numerous images. If the orientation of a component is less than 0.5 or larger than 0.9 of the range; it is discarded.

Human faces are usually evenly distributed in the center. As a result, sides of images do not have face components in them. The *centroid* of a face region is usually located in a small window at the central point of the bounding box.

The average centroid location is determined by taking the center of mass of the shape after the application of the above criteria. If a skin regions' centroids (Y coordinate) is extremely below or above by 20% of central windows, it is rejected as a face candidate area.

The equation of calculating the *ellipse area* is as follows [19]:

$$Ellipse\ area\ = \frac{4A}{\pi D_x D_y} \ , \tag{10}$$

The ellipse area criterion stipulates that it is possible to calculate the probability of every skin color of the nominated face region. For a region being tested to be considered as a human face, it must result in a S_e value that is greater than threshold 0.7. After all the above operators are applied, the connected components that are left behind contain faces.

V. EXPERIMENTAL RESULTS

The face images used in this study to evaluate the performance of the proposed method of face detection are acquired from Caltech face database; it is comprised of 450 images that have different levels of illumination, different facial expressions and from different locations. The size of each image is 896×592 and each image contains only one face. Since the database has numerous images that have different levels of illumination, it suitable for the evaluation of the performance of proposed illumination algorithm. There are two performance metrics which are used to evaluate the success rate. One of them is the Non-Face Detection Rate (NDR) which calculates the rate of non-face detections relative to the total number of detections:

$$\text{NDR} = \frac{\text{non} - \text{face detections}}{\text{number of faces}} \times 100\% \ , \tag{11}$$

The other is the Detection Success Rate (DSR) which calculates the rate of correct face detections relative to the total number of detections:

$$\text{DSR} = \frac{\text{correctly detected faces}}{\text{number of faces}} \times 100\% \ , \tag{12}$$

The number of correctly detected faces is obtained by subtracting the number of faces from the number of false dismissals. The obtained results of the conducted face detection system experiment, that is based on a combination of several morphological operations, are tabulated in Table (2). The detection rate of the face detection system that was based on skin-color and Robert's cross edge detector was a remarkable 96.44%.

A comparison was made between the proposed scheme and Viola and Jones face detector/classifier [20]; it is found that the proposed scheme has the ability to achieve comparable results to that of the Jones algorithm (i.e., 93.54%) when the same face dataset is used.

Also, RGB and YUV color models have been used to assess the efficiency of the face detection system and compared with HSV skin color model (Table 3). Some of the proposed face detection system's sample results are illustrated in Figure (8). As evident from the tabulated results, most of the face detections by the proposed method were correct. As as shown in Table (2) the NDR performance was low (i.e., 3.56%) as a result of using the edge detector as well as the image enhancement.

Table (2) Face detection experimental results

Method	**NDR**	**DSR**
Skin-Color Segmentation	32%	68%
Skin-color + edge detection	7.78%	92.22%
Image balance with Skin-color + edge detection	3.56%	96.44%
Viola & Jones	6.46%	93.54%

Table (3) Results using various color models

RGB		**HSV**		**YUV**	
DSR	**NDR**	**DSR**	**NDR**	**DSR**	**NDR**
24.44%	75.56%	96.44%	3.56%	75.33%	24.67%

Regardless of the ability of the HSV skin color model to handle different levels of brightness and illumination, it is subject to detection of non-skin objects that have chrominance levels similar to those of skin color.

Fig. 8: Example of correctly detected faces by the proposed system

VI. CONCLUSIONS AND FUTURE WORK

The detection of human faces in this study was enabled by applying a hybrid method based on three different conditions. They include: normal lighting, different background, and different conditions of illumination. Nonetheless, this method was not able to solve some limitations related to failure in detection, fault detection, over-bounding of the face region if the candidate's hair, clothing or background corresponds to the color of his/her skin. Such problem can be solved by identifying the facial features (i.e. eyes, mouth and nose) on in the face region. The proposed hybrid method of face identification has provided a balance between the accuracy and the rate of detection. Based on the Caltech database, the achieved detection accuracy is 96.44% (Table 2). However, this can further be improved if the suggestions mentioned above are taken into consideration.

REFERENCES

[1] N. A. bin Abdul Rahman, K. C. Wei, and J. See, "RGB-H-CbCr Skin Colour Model for Human Face Detection", *Faculty of Information Technology, Multimedia University*, 2007.
[2] M. Jain, S. Mitra, and N. Jotwani, "Eye Detection and Face Recognition using Line Edge Map", *National conference on computer vision, pattern recognition, image processing and graphics (NCVPRIPG)*, India, 2008.
[3] S. K. Rath and S. S. Rautaray, "A Survey on Face Detection and Recognition Techniques in Different Application Domain", *International Journal of Modern Education and Computer Science (IJMECS)*, vol. 6, no.8, pp. 34-44, 2014.
[4] M. H. Yang, D. Kriegman, and N. Ahuja, "Detecting Faces in Images: A Survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, pp. 34-58, 2002.
[5] I. Zaqout, R. Zainuddin, and S. Baba, "Pixel-based Skin Color Detection Technique", *Machine Graphics and Vision*, vol. 14, no.1, 2005.
[6] Jr. O. Severino, and A. Gonzaga, "A new approach for color image segmentation based on color mixture", *Machine vision and applications*, vol. 24, no.3, pp. 607-618, 2013.
[7] C. Aiping, P. Lian, , T. Yaobin, and N. Ning, "Face Detection Technology Based on Skin Color Segmentation and Template Matching", *2010 Second International Workshop on Education Technology and Computer Science (ETCS)*,vol. 2, pp. 708-711, 2010.
[8] R. Hassanpour, A. Shahbahrami, and S. Wong, "Adaptive Guassian Mixture Model for Skin Color Segmentation", *World Academy of Science, Engineering and Technology*, vol. 41, 2008.
[9] M. Soriano, B. Martinkauppi, S. Huovinen, and, M. Laaksonen, "Adaptive Skin Color Modeling using the Skin Locus for Selecting Training Pixels", *Pattern Recognition*, vol. 36, no. 3, pp. 681-690, 2003.
[10] S. Gundimada, L. Tao, and V. Asari, "Face Detection Technique based on Intensity and Skin Color Distribution", *ICIP2004*, pp. 1413-1416, 2004.
[11] V. S. Bhat, and D. J. Pujari, "Face Detection System using HSV Color Model and Morphing Operations", *Proceedings of National Conference on Women in Science & Engineering (NCWSE)*, September 2013,
[12] J. Kovac, P. Peer and F. Solina, "Human Skin Colour Clustering for Face Detection", *Proceedings of International Conference on Computer as a Tool*, vol. 2, pp. 144-148, September 2003.
[13] R. Azad and H. S. Boroujeni, "Novel and Tunable Method for Skin Detection Based on Hybrid Color Space and Color Statistical Features", *International Journal of Computer & Information Technologies (IJOCIT)*, vol. 1, no. 3, pp. 211 -219, 2013.
[14] W. Zhang, B. Yu, G. J. Zelinsky, and D. Samaras, "Object class recognition using multiple layer boosting with heterogeneous features", *in Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 323–330, San Diego, CA, June 2005.
[15] P. Kakumanu, S. Makrogiannis, R. Bryll, S. Panchanathan, and N. Bourbakis, "Image Chromatic Adaptation using ANNs for Skin Color Adaptation", *16th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2004)*,pp. 478-485, 2004.
[16] K. Somasundaram, and T. Kalaiselvi, "A Method for Filling Holes in Objects of Medical Images Using Region Labeling and Run Length Encoding Schemes", *National Conference on Image Processing (NCIMP)* , pp. 110-115, 2010.
[17] K. B. Eum, "Effective Face Detection and Robust Face Tracking using Hybrid Filter", *Journal of Korean Society for Imaging Science and Technology*, vol. 18, no. 2, pp. 9-19, 2012.
[18] P. Kuchi, , P. Gabbur, and P. SUBBANNA BHAT, "Human face detection and tracking using skin color modeling and connected component operators", *IETE journal of research*, vol. 48, pp. 289-293, 2002.
[19] G. Yang, H. Li, L. Zhang, and Y. Cao, "Research on a Skin Color Detection Algorithm Based on Self-Adaptive Skin Color Model, *2010 International Conference on Communications and Intelligence Information Security (ICCIIS)*, pp. 266-270, 2010.
[20] P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features", *IEEE Conf. CVPR*, vol. 1, pp. 511-518, 2001.

