

# A Comprehensive Survey on Human Skin Detection

**Mohammad Reza Mahmoodi**

Department of Electrical and Computer Engineering, Isfahan University of Technology, 8415683111 Isfahan, Iran  
E-mail addresses: mr.mahmoodi@ec.iut.ac.ir

**Sayed Masoud Sayedi**

Department of Electrical and Computer Engineering, Isfahan University of Technology, 8415683111 Isfahan, Iran  
E-mail addresses: m\_sayedi@cc.iut.ac.ir

**Abstract**—Human Skin detection is one of the most widely used algorithms in vision literature which has been numerously exploited both directly and indirectly in multifarious applications. This scope has received a great deal of attention specifically in face analysis and human detection/tracking/recognition systems. As regards, there are several challenges mainly emanating from nonlinear illumination, camera characteristics, imaging conditions, and intra-personal features. During last twenty years, researchers have been struggling to overcome these challenges resulting in publishing hundreds of papers. The aim of this paper is to survey applications, color spaces, methods and their performances, compensation techniques and benchmarking datasets on human skin detection topic, covering the related researches within more than last two decades. In this paper, different difficulties and challenges involved in the task of finding skin pixels are discussed. Skin segmentation algorithms are mainly based on color information; an in-depth discussion on effectiveness of disparate color spaces is elucidated. In addition, using standard evaluation metrics and datasets make the comparison of methods both possible and reasonable. These databases and metrics are investigated and suggested for future studies. Reviewing most existing techniques not only will ease future studies, but it will also result in developing better methods. These methods are classified and illustrated in detail. Variety of applications in which skin detection has been either fully or partially used is also provided.

**Index Terms**—Skin Detection, Skin Classification, Skin Segmentation, Face Detection, Standard Skin Database.

## I. INTRODUCTION

With the advancement of technology and emersion of automatic systems, image processing plays an indispensable role as countless number of image-analysis-based electronic systems is fabricated diurnal. Automatic door control systems are often designed based on motion detection, face recognition algorithms are employed extensively in surveillance, object tracking algorithms are directly involved in anti-accident systems and driverless

cars and finally, human computer interaction (HCI) technology hugely rely on human segmentation and biometrics; these are all witnesses of this fact. Among these applications, extracting a feature in human's body is one of the most fundamental under interest topics mainly due to its enormous number of applications.

A very shocking glance at actuarial of sleepy people driving cars reported by National Sleep Foundation's 2005 Sleep in America poll call up for serious reactions. Accordingly, 60% of adult drivers (about 168 million people) have admitted that they have driven a vehicle while feeling drowsy in the past year, and more than one-third (37% or 103 million people) have actually fallen asleep at the wheel [1]! As a matter of fact, more than 1 tenth of those who have nodded off, say they have done so at least once a month and four percentages of them avouched that they have had an accident or near accident since they were too tired to drive. In the following, the National Highway Traffic Safety Administration have estimated that 100,000 police-reported crashes are the direct result of driver fatigue each year resulted in an estimated 1,550 deaths, 71,000 injuries, and \$12.5 billion in monetary losses [1]. Fatigue directly affects the driver's reaction response time and drivers are often too tired to realize the level of inattention. Thus, a driving monitoring system [2,3] shall be a promising solution to this dreadful problem. In such systems, several features including different signs and body gestures in the eyes or head such as repetitious yawning, heavy eyes and slow reaction is employed as a drowsiness cue. However, an essential step before facial feature extraction in variety of such systems is detection of human face or skin. Here, as the head pose or rotation may vary frequently, using skin color cue to detect facial features and face could be a promising choice. Skin segmentation algorithm is also a major concern in many other number of applications including content base retrieval [4], identifying and indexing multimedia information [5], gesture recognition [6,7], robotics [8], sign language recognition [9,10], gaming interfaces [11] and human computer interaction [12,13], and filtering objectionable URLs [14].

Skin classification is the act of discriminating skin and non-skin pixels in an arbitrary image. Detection of pixels related to human skin is a prevalent task as it has potential

of high speed processing [15,16] and not sensitive to the changes of posture and facial expression. It is invariant against rotation [15,17], geometrics [16,18], stable against partial occlusion [15,17], scaling [19] and shape [20] and

it is somewhat person independent [18]. Skin detectors are easily implement-able and they often impose low level of computation cost [21,22].

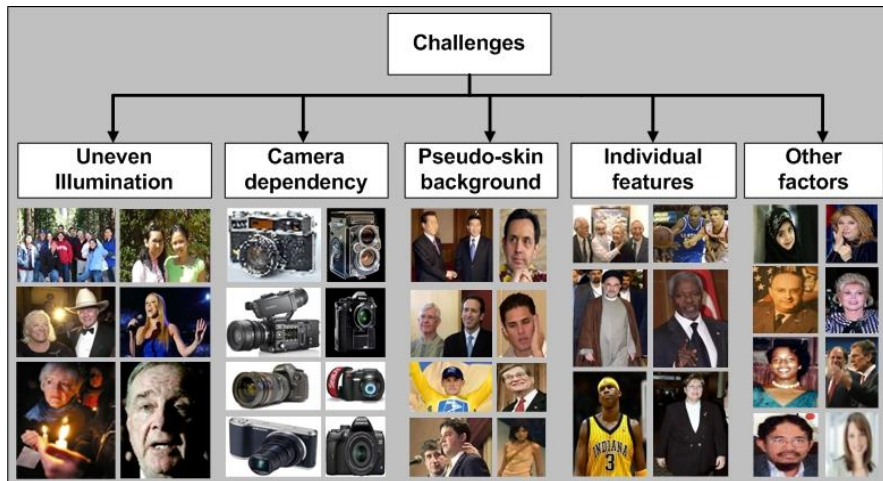


Fig.1. Challenging Factors Involved in Skin Detection in One Glance

Skin detection though seems to be an effortless task for human; there are performance limiting factors which some of them are common in different image processing especially biometrics and object analysis. In Fig. 1, these factors are presented with examples.

#### A. Uneven, inconsistent and nonlinear illumination

Without any exaggeration, the most performance deprecatory factor in skin detection is color constancy problem [17,23]. To be more specific, the intensity of a pixel in an image depends on both reflection and illumination. Reflection is related to the imaging scene while illumination is the effect of light source i.e. light fixtures or daylight natural illumination. The appearance of a scene varies when spectra, source distribution and lighting conditions changes [24] and this imposes difficulty as this behavior is nonlinear and unpredictable (controllable). Handling this problem has been investigated by two strategies. Using color correction and illumination cancelation algorithms prior to segmentation task significantly improve the performance [25, 26]. Dynamic adaptation approaches [27, 28] have been also employed for this problem, however in many cases, the generality of these system is doubted.

#### B. Complex, pseudo skin background

One important challenge in skin segmentation is the fact that a pixel with certain intensity in components of a color space can be perch into both skin and non-skin classes in different conditions [15,29]. There are numerous skin-like objects in the background such as wall, wood and brick which separating real skin pixels and pseudo skin pixels associated with them is not an straightforward task. A good skin detector algorithm shall be able to consider this effect unless the output will not be satisfactory. Most methods which are color and pixel based often have difficulty in dealing with such phenomenon. Complex backgrounds also impacts mostly

on the performance of those approaches based on shape, texture and spatial features.

#### C. Imaging equipment and camera dependency

Camera characteristics (sensor response, lenses, device settings) [24] are very influential as the distribution of human skin color varies device to device [17,30] and this degrades the performance of detectors designed based on distribution of skin color cluster in one specific color space. Some methods attempts to use non-linear color space transformations [31] while some methods use several regression analyses to estimate device functions e.g. [32]. However, the former is not applicable for all devices and the latter has some pre-assumptions which are not always true.

#### D. Individual and intra-personal characteristics

Skin color is not stable across individuals though seems to be a discriminator tool [17,30]. Several factors are officious such as age since human skin color changes slowly with passing the life, sex; the distribution of human skin color is quite different across males and females (cosmetology makes it more disparate), and health; human's skin color changes with specific illnesses. Ethnicity is also one of the most important factors as skin cluster varies slightly among black, yellow and white skin colors. It has been frequently observed that those methods designed based on specific training images (i.e. related to one particular ethnicity) do not deal with other ethnicities very well.

#### E. Other factors

Other less important but effective factors are quality of images [23], traditional scarf or wear [33], Computation inaccuracies and continuous nature of color transformations [34], movement of an object and blurring [25], significant overlap between skin and non-skin cluster [29] and reflection from water and glass [35]. It is worth

to note that aforementioned challenges are related to those algorithms which are based on visible spectrum imaging. For infrared systems, expensive systems [17], tedious setup procedures, and temperature dependency are most common problems. Nevertheless, with all existing limitations and challenges, new algorithms, different variations of traditional methods and mixed systems are put in proposal each year.

A robust skin detector has a couple of features. First of all, as these systems are typically employed for filtering (reduction purposes), it is necessary to have a skin detector with not only high detection rate, but also low false rejection/detection [36]. A robust system should be general i.e. not to be designed for particular condition or specific ethnicity. A relatively fast system ought to be designed which can operate with un-calibrated cameras, arbitrary viewing geometry and unknown but commonly encountered illumination.

While a huge number of methods have been proposed, to the best of author's knowledge, there are only two surveys in this topic. The first one authored by Vezhnevets et al. [38] covered a remarkable number of published works before 2003. The second one documented by Kakumanu et al. [17] in 2006 contains relatively more information regarding different types of methods and previously published works. The first though is coherent and spacious excludes new methods developed in last decade. Vezhnevets et al. have mainly discussed pixel-based methods and algorithms operating on each pixel of an image individually and independent of other pixels leaning on color feature. This paper naturally (most methods were based on statistical analysis of training datasets) emphasizes on statistical methods including parametric and non-parametric ones. Authors have not discussed standard evaluation and training datasets and evaluation metrics. However, at the end, an evaluation comparison is embedded. They have concluded that parametric methods are better suited for classifiers with limited data sets. However, in large data sets, methods less dependent on the skin cluster shape and automatic classification rules are more promising. One other important result was that exclusion of luminance from classification process will not lead to better performance. This was also mentioned and mathematically and practically investigated in [39,40,41]. Unfortunately, this has not been considered in many developed methods after then [42,43,44]. In addition to all previous points, Vezhnevets stated that one cannot evaluate how good a color space is based on the shape of skin's cluster (as also claimed in [45]). This is mainly due to the fact that performance of different skin modeling schemes is too variable in different color spaces.

Kakumanu et al provided a much comprehensive survey as they classified and illustrated different color spaces; explained features, advantages and disadvantages of most of them. In addition, a couple of skin classifiers such as MLP, SOM and MaxEnt classifiers and previously developed statistical based models were also investigated. A very new approach in this paper was discussing illumination adaption algorithms which were very

effective and directional for future studies since many authors began using color correction algorithms after that. Skin color constancy approaches such as Gray world and white patch and a brief comment on Neural Network based illumination adaptation methods were also included in their paper. Kakumanu et al. not only did evaluate skin detection modeling approaches, but they also investigated two standard datasets.

Many novel methods have been developed in recent years which are not covered in [17,38]. These include high performance spatial based methods, multispectral systems, numerous local and global adaptation techniques plus many evaluation and comparison papers. In this paper, a number of issues some missing from former works and some for the sake of completeness are represented. A discussion on skin segmentation challenges was provided. Understanding difficulties when finding a solution to a problem is one of the first steps in inception. It is very critical to investigate different challenges because this will direct the algorithms (which are going to be developed in future) to the way of finding an effective solution rather than a just-solution. Furthermore, variety of color spaces and their impression on the performance of models are also discussed. This is a talk on the choice of color space which will be an answer to many researchers claiming that color space X has the best performance or Y has the least (and they are often repugnance with each other). Color spaces are evaluated and compared based on their performance when using variety of methods. The other contribution of this paper is to provide the most up to date report on almost all (as much as possible) published methods related to skin modeling techniques. Most of these works are developed at recent five years after both of former surveys. As already mentioned, illumination is one of the most important factors in diminution of detection rate. However, recently a couple of plenary review papers have been published discussing illumination correction techniques in general [46-50]. Words are saved here and interesting readers are encouraged to recourse those papers. The other contribution of this paper is paraphrasing skin detection applications.

The rest of the paper is organized as follows. In next section, standard data bases for both training and evaluation phases are followed by superscription of evaluation metrics. In section 3, different color spaces are explained, and then, their performance is compared based upon literature results. In section 4, skin segmentation techniques are explicated in detail and in the following, a fair performance evaluation is discussed in section 5. Finally, skin segmentation applications and directive conclusions are elucidated in subsequent sections 6 and 7 respectively.

## II. STANDARD DATASETS AND EVALUATION METRICS

A complete skin database is required in both training and evaluation phases. Large number of skin and non-skin pixels ought to be used in order to develop an optimum

classifier in training phase. In evaluation also, an exhaustive dataset of images are required in conjunction with their manually annotated ground truth images in order to effectively appraise the performance of a system. Thus, it is essential to investigate characteristics of current available datasets to motivate future studies to utilize them. Most of skin detection datasets are those originally developed for face detection, hand tracking and face recognition problems. There are some important issues which barricade fair and accurate evaluation of different methods. Firstly, although a great number of skin detectors have been proposed, many methods have not been experimented based on standard datasets. In fact, most papers are published based on random collection of personal or online public images. In addition, current experimental results are also based on different databases. Furthermore, many of available databases are neither standard nor specifically designed to measure how effective is an algorithm in dealing with particular condition or challenge. Finally, different datasets are used to train skin classifiers which for obvious reasons, this significantly affect the performance of detectors. Nonetheless, recent and common compiled benchmark datasets are discussed in this section.

The most important group of skin databases are those designed especially for training and assessment of skin classifiers. Compaq [45] is the first large skin dataset and perhaps the most widely used one which consists of 9,731 images containing skin pixels and 8,965 images not containing any skin. The entire database includes approximately 2 billion skin and non-skin pixels collected by crawling Web. Skin regions of 4,675 skin images have been segmented which in conjunction with non-skin images leads to 1 billion pixels. Many skin classifiers have been trained and evaluated based on this database [16,51,52,53]. An automatic software tool is also developed in order to generate ground truth images leading to imprecise results. Fig. 2 represents a set of images from Compaq dataset with GTs. This database is no longer available for public use [54,55].

In contrast with poor image quality in Compaq and its semi-supervised ground truth, ECU skin and face datasets [41] is compiled based on near 4,000 high quality color images and relatively accurate ground truth. ECU images ensure the diversity in terms of the background scenes, lighting conditions, and skin types. The lighting conditions include indoor lighting and outdoor lighting and the skin types include whitish, brownish, yellowish, and darkish skins. Skin dataset provided by Schmutge et al. [56] consists of 845 images. The dataset is composed of 4.9 million skin pixels and 13.7 million non-skin pixels. This dataset is very general as it contains images with different facial expressions, illumination levels and camera calibrations. MCG-skin database [57] contains 1,000 images randomly sampled from social network websites captured in variable ambient lights, confusing backgrounds, diversity of human races and also various resolutions and visual quality. This dataset contains 38,868,720 skin pixels and 139,091,233 non-skin pixels. Ground truth images in this dataset are not accurately

labeled as eyes, eyebrows, etc are also considered as skin and pixels around edges are not also marked with charity. Ling et al. [58] used MCG-skin dataset and additional web collected images to construct a dataset of 37.5 million skin pixels and 135.58 million non-skin pixels. They used half of this dataset to train their SOM-based classifier and the other half for evaluation.

Kawulok et al. compiled HGR [59] database of images for hand gesture recognition (HGR). The dataset is organized into three series acquired in different conditions, individuals, size and background totally including 1558 images. UCI Machine Learning Repository [60] consists of skin pixels collected by randomly sampling B,G,R values from images of various age groups, races, and genders derived from face detection databases. Total learning sample size is 245,057; out of which 50,859 are the skin samples and 194,198 are non-skin samples. This database is only applicable for training purposes. Castai et al. [61] have also created SFA dataset based on FERET (876 images) [62] and AR (242 images) [63] databases. SFA consists of 3,354 samples of skin pixels and 5,590 non-skin samples with different 1 to 35×35 dimensions. A comparison between UCI and SFA based on the best topology and threshold of a neural network based skin classifier yielded to the fact that SFA is slightly more accurate than UCI in evaluation of skin detectors [61]. However, the SFA mainly include semi passport images which are not suitable for evaluation purposes. In Fig. 3, several GT images in SFA are depicted and red arrows shows false points.

Db-skin dataset [64] contains 103 skin images annotated by human with relative charity. In some images, eyes and skin and non-skin boundaries are not marked precisely, but they are taken in different lighting conditions and complex backgrounds. In a recent study, Montenegro et al. [65] compiled a dataset in order to compare the performance of 5 common color spaces. This dataset contains 705 RGB images from 47 Mexican subjects with different ages and distinct skin tones. Apart from former databases, images have been all acquired using a single Kinect sensor in completely controlled conditions. The dataset is not available for public usage. R. Khan et al. [15] also developed Feeval dataset consisting of 8991 images from 25 online videos plus per pixel manually generated ground truths. Videos mostly lack enough quality, and the diversity of only 25 videos is questionable. TDSD [66] and Sigal dataset [67] have been also developed for skin segmentation where ground truth labeling for both databases has not been performed with hand. The former contains 554 images including 24 million skin pixels and 75 million non-skin pixels. Table. I summarizes the characteristics of aforementioned skin datasets.

Face detection/recognitions datasets are also utilized for skin detection purposes. Razmjoo et al. [68] used Bao test bed for training and evaluation. Ding et al. [69] used Caltech with their own ground truth to measure the effectiveness of their classifier. Miguel et al. [70] used near 290 images from each public datasets for human activity recognition such as EDds, LIRIS, SSG, UT, and

AMI as evaluation sets. These 87,000 skin pixels cover a wide range of situations and illumination levels. Naji et al. [71] constructed a database of 125 images collected from

LFW, CVL and Web images. Tan et al. [72] used Pratheepan and ETHZ PASCAL datasets to present quantitative results.



Fig.2. A Set of Images from Compaq Dataset and Corresponding GTs

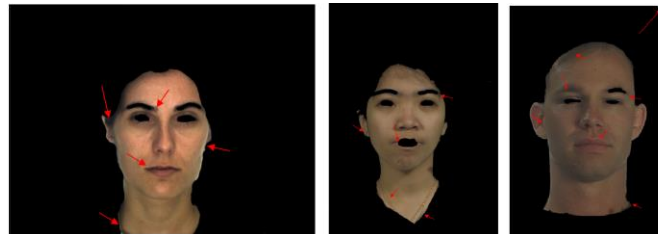


Fig.3. Ground Truth Images in SFA

Table 1. Common Datasets in Skin Segmentation

Dataset	No. S images	No. NS images	No. S pixels	No. NS pixels	GT qualification
Compaq [45]	4,675	8,965	1 billion		Annotated imprecisely
ECU [41]	4,000	2,000	209.4 million	901.8 million	Relatively Precisely annotated
MCG [57]	1,000		39 million	139 million	Annotated imprecisely
Schmugge [56]	845		4.9 million	13.7 million	Ternary marking charily
Ling et al. [58]	1,000		37.5 million	135.58 million	Unavailable
HGR [59]	1,558		-	-	Precisely annotated
UCI [60]	-		50,859	194,198	Unknown
SFA [61]	1,118		-	-	Precisely annotated
Db-skin [64]	103		-	-	Relatively annotated precisely
Montenegro [65]	705		-	-	Unavailable
Feeval [15]	8,991		-	-	annotated imprecisely
TDSD [66]	554		24 million	75 million	Annotated imprecisely

#### A. Evaluation measurement

Evaluation metrics are critical in assessment of a classifier's power. Several performance measures are employed in a general classification problem including Recall, Accuracy, Precision, Kappa, F-measure, AUC (Area Under Curve),  $G_{\text{mean}}$ , Informedness, and Markerndness [73,74]. Note that in different applications, one of these metrics may be more suitable than another one.  $2 \times 2$  confusion (contingency) matrix for an skin detection problem is provided in Table .II.

Table 2. Confusion Matrix

-----	Classified Skin	Classified Non-Skin
Ground Truth skin	True Positive (TP)	False Negative (FN)
Ground Truth N-skin	False Positive (FP)	True Negative (TN)

Empirical and theoretical evidences demonstrate that these measures are biased with respect to data imbalance and proportions of correct and incorrect classifications [75]. These shortcomings have motivated a search for new metrics based on simple indices, such as the True Positive Rate (TPR) which is the percentage of correctly classified pixels and the True Negative Rate (TNR) is the percentage of pixels negatively classified. Precision and Recall are also two other important metrics to evaluate a skin detection algorithm [16,70]. Precision (P) is the percentage of pixels correctly labeled as positive whereas Recall (R) indicates true positive rate;

$$P = \frac{TP}{TP + FP}$$

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

The F-score integrates both Recall and Precision into a measure. It is defined as:

$$F_{score} = \frac{(1+b^2) \times (R \cdot P)}{b^2 P + R} \quad (2)$$

The non-negative real  $b$  is a parameter to control the effect of the Recall and Precision separately. Typically,  $b$  is set to 1, thus obtaining the  $F_{score}$  as a measure in range of  $[0,1]$ , which can be viewed as a harmonic mean of the Recall and Precision [73]. Receiver Operating Characteristic Curve (ROC) is another popular technique for assessment of classifiers considering both Recall and Precision [76,77]. Precision and Recall are often in trade-off with each other in skin segmentation and ROC is a very useful tool to indicate this fact. ROC is the right tool to represent effectiveness of a classifier independent of tunable parameters (e.g. a threshold value or number of training samples, etc). The Area Under Curve (AUC) is the quantitative representation of the ROC [78]. In [56,79], authors employed ROC and AUC concepts as a measure to experiment some of in doubt premises in skin segmentation. For skin cluster boundary models and whenever there is only one run, AUC is calculated as the average of TPR and TNR. CDR (Correct Detection Rate) (or Accuracy) represents the probability of a pixel to be correctly detected, while FDR (False Detection Rate) demonstrates the probability of a pixel to be wrongly detected [80,81]:

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

$$FDR = 1 - CDR \quad (4)$$

Aibinu et al. [82] utilized Mean Square Error (MSE) and Correlation Coefficient (CC) in addition to previous measures to represent the performance of their skin detector. Both MSE and CC are calculated using difference between the binary mask and the ground truth. The former indicates the closeness of the output image to the ground truth considering the same penalty for false detection and false rejection. However, this cannot be used generally as a comparison factor since the result depends on the size of the image. The latter also quantifies the same idea but with a normalized value in the range of  $[-1,+1]$  which  $-1$  shows the extreme dissimilarity and  $+1$  is interpreted as a perfect detection. Considering  $S_M$  as the binary mask and  $S_G$  as the ground truth:

$$CC = \frac{\sum_{x=1}^M \sum_{y=1}^N (S_M(x, y) - \bar{S}_M(x, y))(S_G(x, y) - \bar{S}_G(x, y))}{\sqrt{\sum_{x=1}^M \sum_{y=1}^N (S_M(x, y) - \bar{S}_M(x, y))^2 \sum_{x=1}^M \sum_{y=1}^N (S_G(x, y) - \bar{S}_G(x, y))^2}} \quad (5)$$

$$MSE = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N (S_M(x, y) - S_G(x, y))^2 \quad (6)$$

Matthews's correlation is a metric which takes FP, FN, TP and TN into account [77]. Montenegro et al. [65] utilized this measure to compare the performance of different color spaces using Gaussian model.

$$C = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TN + FP)(TN + FN)(TP + FN)}} \quad (7)$$

$C$  is a measure in the range of  $[-1,+1]$  which  $+1$  express identity and  $-1$  indicates total false detection and rejection. This coefficient is not very useful for assessment purposes in skin segmentation particularly in cases with large background areas. Correct detection of such regions which are often easily segmented by most classifiers will be misinterpreted. In other words, this coefficient is not capable of providing fair evaluation for imbalanced classification problems in which one of the classes often outnumbered other class. In addition, when an algorithm gives very low FP and at the same time very low TP,  $C$  will be unfairly high [77].

## B. Discussion

In this section, a review on the most existing and common datasets and evaluation metrics in skin detection literature was provided. Note that the impact of both the number and variety of training and test sets on the performance of classifiers is not hidden from any one. However, this factor is often ignored in evaluation of different works. Most of existing works report their statistical results based on one of abovementioned datasets or a self-made images without considering the effect of both size and diversity of photo sets. The second factor is the execution and training time. Though, few works have reported the training and evaluation time, most have not considered the importance of evaluation time. The training time is an indispensable factor particularly for dynamic (online) approaches. Skin detection is often employed as a preprocessor in most applications which makes them necessary to be real-time. Furthermore, an exact definition of skin is not still ascertained among authors. For example, in labeling ground truth images, some authors have considered lips, mouth, etc as skin while some have excluded them and this definitely affects the results. From another point of view, many authors have reported their performance rates using TP, FP, P, R and other non-graphical measures. However, skin detection is somehow a fuzzy problem in which there is an optimal threshold for each method and it is not constant for different images. In some cases, F-score, ROC curve and AUC have been reported to be strong criteria in describing the performance of an algorithm. In fact, ROC represents the pure stability of a detector independent on the choice of threshold. A similar associated numerical measure is AUC which has been illustrated before. Unfortunately, a few number of authors utilized AUC though it is very simple to calculate by using curve fitting techniques.

In summary, there are two major steps in fairly measuring performance of an skin classifier. For evaluation purposes, ROC and AUC seems to be very promising for comparison goals. The other one is related

to the training and test sets. In order to address so called issues, recently, a new skin detection dataset (SSD) has been compiled to be publicly available for future studies. In developing SDD, limitations of former databases are obviated from different points of view. SDD contain 21,000 images with which 15,000 of them are photos without any skin pixel. All images are divided into 4 sets. Images are captured in different illumination conditions, using variety of imaging devices, from diversity of skin tones from people all around the world. Some images are snapshots from online videos and movies while some are static images acquired from popular face recognition/tracking/detection datasets and already cited skin datasets. Table. III shows the statistics of the SDD. First portion of the database is particularly considered for training purposes mainly comprise of single face images in different lighting conditions. In the table sections 2-4 include evaluation photos with manually annotated GTs.

GTs have been marked with careful attention to the fact that no pixel is misclassified i.e. all skin pixels annotated as skin are actual skin pixels and non-skin pixels are not. There are certain pixels in all images in which either there is a question on the skinness of them (such as those regions around eyebrows and eyes, lips and nose holes, etc) or they are located at the boundary of skin and non-skin pixels in which in some images it is difficult to discriminate their skinness and in some images, it takes a lot of time to annotated them. Thus, GT images are divided into 3 non-overlapping regions; skin pixels, non-skin pixels and pixels which are not going to be considered both in evaluation and training. In Fig. 4 and Fig. 5, a set of training and evaluation test images are located. Obviously, black points indicate non-skin regions, red ones are associated with skin pixels and blue pixels portend ignored pixels. Using this database allows precise measurement of performance of skin detection methods. In addition, authors are capable of using different test sections in order to check if their results are independent of the test set. Also, as the number of pixels both in training and evaluation proliferate, the results would be more reliable.

Table 3. Statistics of the Sdd

Dataset sections	Goal	No. S images	No. NS images
1	Training	2000	4000
2	Test	1000	4000
3	Test	1000	4000
4	Test	1000	4000

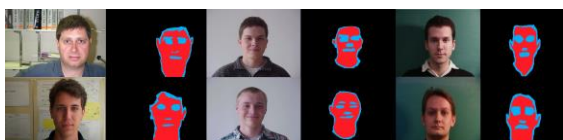


Fig.4. A Set of Training Images in SDD



Fig.5. A Set of Test Images in SDD

### III. STUDY OF DIFFERENT COLOR SPACES

Color is the perceptual result of light in the visible region of the spectrum, having wavelengths between 400 nm to 700 nm, incident upon the retina [83]. Color space is a three-dimensional geometric space with axes appropriately defined so that symbols for all possible color perceptions of humans fit into it in a psychological order [84]. One big group of color spaces are those directly depend on the imaging device and not the way human sees the color! They are often exploited for digital applications, representations and computations. Imaging equipments often use disparate color spaces with specific response to the environmental conditions due to their particular features e.g. printers often use CMY whereas TVs utilize popular orthogonal (YCbCr) color spaces. However, to ensure a proper color rendition in various devices, device-independent color spaces, such as CIE (Commission Internationale de l'Éclairage) color models are required to serve as an interchange standard. CIE colors are the colorimetric spaces that are device-independent.

Definitely, though color is not directly utilized in some skin detection approaches, it is one of the most decisive tools affecting the performance of skin algorithms. Both device dependant and independent color spaces have been used in skin classification tasks with different frequencies. Albiol et al. [5] mathematically proved that the optimum performance of skin classifiers is independent of the selected color space, though, the performance of most skin detectors are significantly in direct relation with the color space choice. Therefore, widely used color spaces are discussed in this section.

#### A. Device dependant color spaces

Device dependant color spaces have been extensively gained popularity for skin detection. RGB as the most common color space is expressed in terms of primary colors i.e. colors are encoded as a linear weighted addition of red, blue and green components. Several algorithms such as [7,78,85,86] are proposed based on RGB color space. In addition, in [40] it has been stated that RGB color space is the best choice for human skin detection based on particular criterions. However, there are several issues associated with RGB color spaces. First of all, R,G and B components are highly correlated, in addition to the fact that luminance (intensity) is in a linear relation with these components which means changing the luminance

that easily happens leads to high variation in the RGB values [38,76]. This can be observed in the stretched skin cluster in RGB color space. Normalized RGB is the normalized version of the RGB color space so that sum of r, g and b components is unity. This color is often considered to be more robust against lighting variations and ethnicity in specific conditions and exhibit tighter cluster [38,17]. Normalized rgb is a popular color space utilized in [87,88,89] for skin segmentation in different approaches.

One big branch of device dependant color spaces are utilized in TV transmission and digital photography systems. The orthogonal color space family includes well-know color spaces such as  $YC_bC_r$ ,  $YCgCr$ ,  $YDbDr$ ,  $YPbPr$ ,  $YIQ$ ,  $YUV$ , and  $YES$ .  $YIQ$  is employed in NTSC TV broadcasting [90] whereas  $YC_bC_r$  is utilized for JPEG, H.263 and MPEG compression tasks since it reduces the redundancy in RGB channels [20,91]. In addition,  $YUV$  color model is exploited by PAL (Phase Alternating Line) and SECAM (Séquentiel Couleur À Mémoire) color TV system [76]. In these color spaces, unlike the RGB, the luminance channel is separated from chrominance ones yielding to a very tight pseudo-ellipse skin cluster. Due to the particular features of these color spaces such as separation of luminance-chrominance channels, relatively simple RGB conversion and relatively tight skin cluster, orthogonal color models have been frequently used in skin segmentation [92-95].

Perceptual color spaces such as HSI, HSV, HSB, HSL and TSL are very attractive color spaces in skin detection literature. In these color models, each pixel is presented by Hue (tint or color), Intensity (lightness) and Saturation associated with physiological feeling of human [96]. Similar to the orthogonal color spaces, luminance and chrominance properties are separated in addition to the intuitiveness of the color space and invariance in white light sources which make them favorable for skin segmentation [97,98]. However, nonlinear RGB conversion of these color spaces are relatively computational, and hue component is discontinued. Furthermore, the cyclic nature of H component disallows usage of color parametric skin models [38]. Of course, HS components can be represented in Cartesian coordinates rather than polar one which not only solves this so called problem, but it also forms a tighter skin cluster. Similar to HSV, TSL is a normalized chrominance-luminance alteration of the normalized RGB with more intuitive values [38]. Baskan et al. [99], Zahir et al. [100], Sanmiguel et al. [70], Vishwakarma et al. [96] and Prasertsakul et al. [81] made use of perceptual color spaces for skin segmentation and illumination compensation.

There are also less frequently used color models which are obtained through either linear or nonlinear transformations of other color spaces. In addition, there are also variations in which different components of disparate color spaces are employed for skin segmentation. Several logarithmic variations of RGB are applied in skin detectors since they render illumination change [72,101,102,103]. In [23], RGB, HS and  $CbCr$  were

combined whereas  $YC_gC_r$  and  $YIQ$  were used simultaneously in [104].

### B. Device independent color spaces

CIE-XYZ, CIE-xy, CIE-Lab, CIE-Luv, CIE-Lch are colorimetric models defined and standardized by CIE committee. CIE-XYZ as both the foundation of other spaces and the oldest model is based on measurements of human visual perception. It is obtained using a linear transformation from RGB [90]. The Y represents the luminance component while X and Z both are related to the chrominance values. CIE-xy is the chromacity normalized version of the CIE-XYZ. In both of them, color difference is not perceived equally through the entire of the space. CIE-Lab attempts to make the luminance scale more perceptually uniform. It describes all the colors visible to the human eye. The chromacity of CIE-Luv is different from CIE-Lab, but the luminance i.e. L is identical. Since RGB is device dependant, there is not straightforward conversion to device-independent color spaces; however, using an absolute color space (e.g. sRGB which is used in many printers), RGB will be initially transferred using a device-dependant adjustment and then CIE-Lab can be obtained. This complexity of transformation is a noticeable disadvantage for device independent color models making them less eligible color model in skin classification algorithms. Nevertheless, Lee et al. [105], Shin et al. [40], Ravi et al. [106] and Lindner et al. [107] built skin classifiers based on device independent color spaces.

### C. Color space comparison

Color spaces undoubtedly affect the performance of skin detectors and different authors derived different results for the influence of color space choice on the performance of skin segmentation algorithms. In [108], it is claimed that TSL is the best color space for skin detection particularly when applied in conjunction with Gaussian and Mixture of Gaussian models. Khan et al. [15] stated that IHLS has the best performance in almost all methods except histogram LUT (look-up-table), and RGB, independent of the method, has the worst performance. Comparing the performance of variety of color spaces on single Gaussian model has been yielded into the fact that  $YC_bC_r$  has the best performance [76]. On the other hand, Montenegro et al. compared RGB, HSV,  $YC_bC_r$ , CIE-Lab and CIE-Luv in SFA data base. They used MCC as a metric for evaluation, and Gaussian classifier as the method, concluding that CIE-Lab is the winner while normalized RGB has the lowest performance [65]. In [109], Based on a comparative study, HSV seems to be more robust skin detector compared with other color spaces. Moreover, it has been concluded in [82] that performance of RGB based techniques lessens when the distinction between skin-pixels and objects increase, but in overall, using  $YC_bC_r$  yields better results.

Littman et al. [110] compared the performance of RGB,  $YIQ$ , and  $YUV$  in their hand recognition and detection approach on a small subset of evaluation images, resulting that the system is independent from the choice of color



space. In [40], they also evaluated the dependency of skin detectors on the selected color model. They performed an exhaustive experiment on different color spaces reporting that RGB color space provides the best separability between skin and non-skin patches. In another work [111], a comparison study is done using Gaussian and histogram approaches on a dataset of 805 color images. It is claimed that the choice of color space significantly changes the performance, and HSI color space accompanying with histogram model outperform other cases. In addition, they stated that skin color modeling has greater impact than the color space transformation.

The work presented in [41] compares the performance of HIS, RGB, CIE-Lab and YCbCr. The best performance in terms of classification error is obtained in HIS and  $YCbCr$  models. However, based on [56], HIS should be employed with histogram technique to produce acceptable results. Authors in [112] stated that HIS is the best color space and used it in an explicitly defined method. In a recent study [106], using of both explicitly defined models and single Gaussian distribution resulted that YPbPr outperforms other models. Gonzalez et al. [114] compared the performance of 10 common color spaces based on k-means clustering algorithm in 15 images from AR database with manually annotated ground truths. They have concluded that HSV, YCgCr, and YDbDr are the best color spaces. Nalepa et al. [115] showed that statistically combination of RGB and HSV outperforms other color spaces in a recall oriented system using an explicitly defined method. In [116] by using a different method, artificial neural network has been subjected for skin segmentation in variety of color spaces yielding to the conclusion that YIQ is the best choice. The idea to adapt an optimal chrominance color space rather than a segmentation model has been suggest in [117]. Here a non-linear transformation between YUV and new TSL\* color space is used to boost the segmentation.

#### D. Discussion

A review on the utilization of different color space in skin detection was presented. Furthermore, a comparison based on literature results was included. To make the comparisons realistic, some points should be considered. First, as it was stated before, there are certain factors which are effective on the performance of detectors, but are missed in the studies, e.g. the size of training database. In some cases the method which is used and the evaluation set are also very decisive. In neural networks, the performance of classifier is directly depend on the number of neurons, as well as the initial guess of the weights. A similar logic exists in other approaches. Some databases are based on general images with different illumination and imaging conditions, but some are compiled based on a set of images in controlled conditions that can change the results in favor of some specific color spaces. In addition, some methods [8,94,95,118,119] have

pretermitted luminance component in order to diminish the role of illumination. Several authors investigated integrity of this notional yielding to the fact that using luminance always boost the performance of detectors [36,39,40, 41,111]. Nevertheless, a very straight conclusion is that the performance of a color space is associated with several factors including the method used. In order to compare the performance of one classifier using different color models, all other factors which may influence the performance should be considered. Also, by looking from a wider perspective, it is tangible that different authors utilizing similar methods have reported inconsistent results. This obviously corroborates the fact that training sets, evaluation datasets and calibration factors used in each method are also decisive. Effort to find the optimum color space for the task of skin segmentation has been redirected into adapting optimal skin detection models and illumination compensation algorithms as performance comparisons show that the effect of segmentation technique is not even comparable with the efficacy of color model.

#### IV. SKIN DETECTION METHODS

In this section, most techniques used in segmentation of skin pixels are elucidated. Skin detection approaches can be classified into 8 groups which are not yet mutually exclusive. Explicitly defined methods are based on observation of the skin cluster in a particular color space and the idea is to design segmentation rules based on the boundary of the skin cluster. Statistical methods are based on actuarial information extracted from histogram of training skin and non-skin pixels. Non-parametric or parametric models are developed with the aim of obtaining the probability that a pixel belongs to skin or non-skin class. Artificial neural networks (ANN) are very useful tools in either estimation of skin distribution or direct classification of the pixels. ANNs with variety of structures have been utilized in skin segmentation of color images based on both color and texture information. Spatial analysis methods which have been recently developed not only concentrate on in-pixel information, but they often exploit neighboring pixels to boost the performance of classification. Adaptive methods often reach higher accuracies with the cost of computation mostly in images and videos with specific features. These methods often improve the performance of former methods by dynamically updating classifiers' parameters based on global or local information of the test image. Non-visible spectral models are based on relatively expensive tools working in different spectral bands. SVM-based systems are also used for classification of skin pixels in specific applications. And finally, the mixture techniques are based on combination of former methods. Fig. 6 shows classification of skin detection methods.

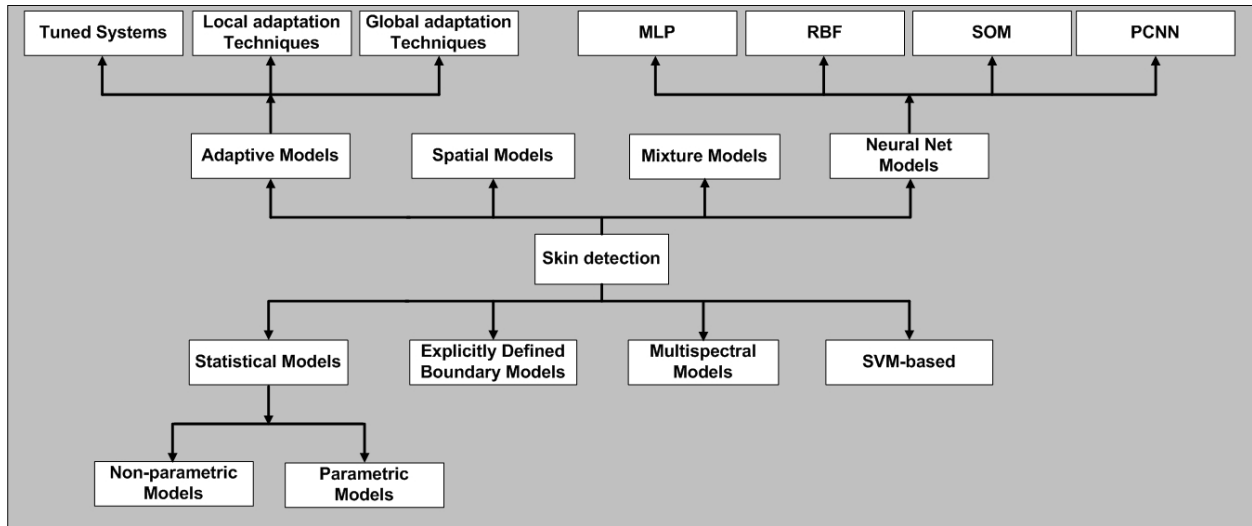


Fig.6. Skin Detection Approaches

### A. Explicitly Defined Boundary Models

Among different people, the main difference in skin color is in its intensity (brightness) rather than the chrominance (color) [24,114,116]. This observation has been a leading factor for developing a very common and simple approach in skin classification. Explicitly defined methods are based on a set of rules derived from the skin locus in 2D or 3D color spaces. These methods have a pixel-based processing scheme in which for any given pixel, rules are investigated to decide on the class of that pixel. They are very popular mainly due to their simple and quick training, low cost implementation and fast processing. However, several parameters are involved in degrading the performance of explicit classifiers, including their static nature, high dependence on training images, effectiveness of rules and inability to deal with most of skin detection challenges. If the training set is selected too general, the classifier gives a high false positive rate and if it consist of images of specific conditions, then the classifier is not applicable for many images. Nevertheless, explicitly defined boundary models with low cost and fast computational requirements are suitable for many applications in which lower accuracy is acceptable.

Kovac et al. [120] proposed a method of explicitly defined boundary model using RGB color space in two of daylight and flashlight conditions which has been reutilized in [16,85,86,119]. Here, a pixel is considered skin in uniform daylight conditions if it satisfies  $R > 95$ ,  $G > 40$ ,  $B > 20$ ,  $\max\{R, G, B\} - \min\{R, G, B\} > 15$ ,  $|R - G| > 15$ ,  $R > G$  and  $R > B$  conditions. In flashlight illumination, however, a pixel is assumed to be skin if  $R > 220$ ,  $G > 210$ ,  $B > 170$ ,  $|R - G| < 15$ ,  $R > G$ ,  $R > B$ . Zhang et al. [121] transform the 3D color space to polar 2D color space, and by means of an explicit defined method on the radius and phase of any pixel in the polar space, the image is segmented to skin and non-skin regions. The boundary defined in [122,123] is based on normalized rgb color space. For a given pixel, upper and lower boundaries for "g" channel is determined based on value of "r"

component. Also, sRGB color space has been employed in [124,125] using very simple rules on sR, sG and sB values. Other variations of this approach based on RGB and relative color spaces are proposed in [87,101].

Orthogonal color spaces are frequently used in case of explicitly defined boundary models. Sagheer et al. [33] exploited two cases of normal lighting ( $C_b \in \{110-120\}$  and  $C_r \in \{135-150\}$ ) and different lighting conditions ( $C_b \in \{110-160\}$  and  $C_r \in \{120-150\}$ ). In order to detect faces, in [126,127], rules are dynamically reconfigured based on pixel's value. A pixel's skinness is investigated based on its luminance value:

$$\begin{aligned} \text{if } (Y > 128) &\rightarrow \theta_1 = -2 + \frac{256 - Y}{16}, \theta_2 = 20 - \frac{256 - Y}{16}, \theta_3 = 6, \theta_4 = 8 \\ \text{if } (Y \leq 128) &\rightarrow \theta_1 = 6, \theta_2 = 12, \theta_3 = 2 + \frac{Y}{32}, \theta_4 = -16 + \frac{Y}{16} \end{aligned} \quad (8)$$

In this case, a pixel is classified as skin if it satisfies following conditions:

$$\begin{aligned} Cr \geq -2(Cb + 24), Cr \geq -(Cb + 17), Cr \geq -4(Cb + 32), Cr \geq 2.5(Cb + \theta_1) \\ Cr \geq \theta_3, Cr \geq 0.5(\theta_4 - Cb), Cr \leq \frac{220 - Cb}{6}, Cr \leq \frac{4}{3}(\theta_2 - Cb) \end{aligned} \quad (9)$$

In [31,34,128,129], different elliptical boundaries are estimated based on the fact that skin locus in  $C_b C_r$  is similar to ellipse, and then in evaluation, only pixels surrounded by the ellipse are considered as skin. An FPGA implementation of a face detector based on explicitly defined boundary model in  $Y C_b C_r$  color space is proposed in [21]. Similar classification works (using orthogonal color models) based on boundary model with variation in parameters originating from different training sets are presented in [12,118,130,131,132].

Many works employed perceptual color spaces such as HSV to define explicit rules. Zahir et al. [100] proposed a simple boundary model using HSV color space for indoor and outdoor conditions. Pitas et al. [133] utilized HSV

color space with following rules to detect skin pixels:

$$V \geq 40, 0.2 < S < 0.6, (0^\circ < H < 25^\circ \vee 335^\circ < H < 360^\circ) \quad (10)$$

Garcia et al. [134] designed more sophisticated rules based on identical color space:

$$\begin{aligned} V \geq 40, H \leq -0.4V + 75, 10 \leq S \leq (-H - 0.1V + 110) \\ \text{if } (H \geq 0) \rightarrow S < 0.08(100 - V)H + 0.5V \\ \text{if } (H < 0) \rightarrow S < 0.5H + 35 \end{aligned} \quad (11)$$

In [112,135,136,137], boundary models based on skin locus in other perceptual color spaces are proposed.

Combination of color spaces has been also effective to boost the performance of explicitly defined boundary models. Thakur et al. [23] employed RGB, CbCr and HS color models. Rules are set independently for triple color spaces and results are simply fused for taking final decision. An algorithm for tattoo detection is proposed in [138] based on combination of  $YC_bC_r$  and HSV. Kim et al. [139] proposed a technique based on fusion of RGB, HSV and  $YC_bC_r$  color spaces. A skin candidate region  $T(x,y)$  is calculated as:

$$T(x,y) = D_1W_1(x,y) + D_2W_2(x,y) + D_3W_3(x,y) \quad (12)$$

where  $W_i = TP_i/FP_i$  (for a set of 50 training images) and  $D_i(x,y)$  represents the detected skin pixel in color space  $i$ . Subsequently, a threshold is adapted to decide whether or not the pixel located at  $(x,y)$  is skin. Furthermore, fusion of the result of different color spaces to reduce false positives has been a common procedure; some examples are: RGB and  $YC_bC_r$  [140],  $YC_bC_r$  and YUV [8], HSV and YUV [109], HSV and  $YC_bC_r$  [141,142], RGB and YUV [143] and HSV and  $YC_gC_r$  [144]. Gasparini et al. [98] used genetic algorithm to find accurate boundaries of skin cluster. He compared the method with some former methods including [120, 133]. The results showed better performance for genetic algorithm.

## B. Statistical Models

Skin detection is a probabilistic problem and many techniques based on general distribution of skin color, each in a particular color model, have been developed. Based on an enormous training set, these methods estimate the probability that an observed pixel is associated with skin. Unlike parametric methods which are based on parameterized families of probability distributions, nonparametric methods are based on descriptive and inferential statistics. In the following subsections, these methods are explained.

### 1) Non-parametric (histogram based) models

There is no any explicit definition of probability density function in non-parametric techniques. Instead, a point to point mapping between color distribution and the quantized color model is provided, i.e., a skinness probability value is assigned to each discrete point of the

entire color model. Single histogram based Look-UP-Table (LUT) model is a common approach in modeling skin color cluster. In this technique by using a set of training skin pixels, distribution of skin pixels in a particular color space is obtained. Considering RGB as a color space with finest possible resolution (i.e. 256 bins per each channel), the 3D RGB histogram (or LUT) is constructed from  $256*256*256$  cells that each representing skin probability of one possible  $R_iG_iB_i$  value. The learning process is simple, but populating the histogram requires a massive skin dataset. The final probability for each possible  $R_iG_iB_i$  value is obtained as:

$$P(R_iG_iB_i) = \frac{\text{No. of occurrence of } R_iG_iB_i}{\text{total counts}} \quad (13)$$

This simply constructs an RGB-based  $256^3$  size LUT. The related huge storage requirement which is histogram model's blemish can be addressed by using both coarser bins and 2D color models.

Several authors have utilized LUT-based skin color modeling. Jones et al. [45] employed this technique for both person detection and adult image recognition. In order to reduce false positives of a face detector, a set of LUTs including different color spaces such as RGB and YUV are used in [145]. Liu et al. [18] utilized HS color model in conjunction with a LUT with  $256*256$  bins to build a skin classifier. They used morphological and blob analysis for further enhancement. Yoo et al. [146] also used a  $32*32$  LUT (total of 1024 bins) in the same color model for face tracking. Ibraheem et al. [147] adopted a mixture of histogram models in RGB,  $YC_bC_r$  and HSV for recognition of hand gestures. The overall model is the linear weighted summation of all 3 histograms. Chen et al. [148] employed the normalized histogram of skin color in HSV color space to filter many non-face regions.

Contrary to former LUT approach, the Bayesian classifier considers two histograms of skin and non-skin pixels. In fact, due to an obvious overlap between skin and non-skin pixels in different color spaces (Jones et al. [45] showed that in their study, 97.2% of colors occurred both as skin and non-skin), in above equation,  $P(R_iG_iB_i)$  is a conditional probability for which it is already assumed the observed pixel belongs to skin class. Hence, according to Bayes rule:

$$\begin{aligned} P(\text{skin} | R_iG_iB_i) = \\ \frac{P(R_iG_iB_i | \text{skin})P(\text{skin})}{P(R_iG_iB_i | \text{skin})P(\text{skin}) + P(R_iG_iB_i | \text{non-skin})P(\text{non-skin})} \end{aligned} \quad (14)$$

Using each of skin and non-skin histograms,  $P(R_iG_iB_i | \text{skin})$  and  $P(R_iG_iB_i | \text{non-skin})$  are calculated respectively. In designing a Bayesian classifier, two assumptions for prior probabilities are possible [25]. Using ML (Maximum likelihood) approach results in  $P(\text{skin}) = P(\text{non-skin})$  while by using MAP (Maximum a posteriori) assumption,  $P(\text{skin})$  and  $P(\text{non-skin})$  are both estimated by employing a training set. Zarit et al. [97]

compared the performance of Bayesian classifiers using both ML and MAP premises in 5 different color spaces. It is concluded that ML based classifier significantly performs better. However, it is possible to use following detection rule to eliminate the effect of priori probabilities [38]:

$$\frac{P(R_i G_i B_i | skin)}{P(R_i G_i B_i | nonskin)} > \theta, \theta = K \left( \frac{P(non-skin)}{P(skin)} \right) \quad (15)$$

where the value of K is tunable parameter to remove the dependency of detector's behavior to priori probabilities. The optimum value for  $\theta$  is obtained by using ROC curve depend on the application. In [45], it is proved that for any choice of priori probabilities, the resultant ROC is the same.

In [149,35], a Bayesian classifier based on YCbCr color space is employed. Erdem et al. [16] employed Bayesian based post filtering method to reduce false positives of Viola-Johns face detector. In [150], the adequate number of quantization levels by minimizing an objective function which is the summation of false acceptance rate and false rejection rate in accordance with  $2^k$  number of bins is found. The Bayesian classifier accompanying with Adaboost method is used for detecting faces. Due to the precision of Bayesian technique in estimating skinness probability, this model has been fostered in several other skin and face detectors [12,51,103].

Zarit et al. [25] conducted a comparison study on the performance of histogram models in Fleck HS, HSV, rgb, CIE Lab and YCbCr. For training, 48 images are used and evaluation was done on 64 other images. It is concluded that Fleck HS and HSV outperform other color spaces. Also by using two resolutions of 128\*128 and 64\*64, the performance did not change significantly. Another work by Phung et al. [36,41] examined the tradeoff between the number of bins per channel and detection rate. The performance of Bayesian classifiers with disparate number of bins in five different color spaces including RGB, CIE XYZ, HSV, YCbCr and CIE Lab using ECU dataset are compared. There was no significant difference in performance when using different color spaces at histogram sizes more than 128 and 256 bins per channel. Contrary to other color spaces in which the performance rapidly degrades when histogram size is less than 64, for RGB color space the performance remains untouched when histogram size is not less than 32. However, performance always lessens if only chrominance channels are utilized, and there are significant performance variations between disparate choices of chrominance channels.

There are several advantage and disadvantages associated with non-parametric techniques. The ease of training and implementation, fast training, and independency on the shape of cluster [38] are the main advantages. In addition, histogram based methods often exhibit higher detection rate compare with parametric ones, as there is no fitting error in their probability estimation procedure. However, one main drawback of these techniques is their dependency on the training set

which requires gathering of a large number of skin pixels. In addition, large memories [38,41,58] are required to implement histogram based models particularly when fine resolutions are needed.

## 2) Parametric models

Parametric models such as single Gaussian models (SGMs), Gaussian mixture models (GMMs), cluster of Gaussian models (CGMs), Elliptical models (EMs), etc, are developed to compensate LUT shortcomings such as high storage requirement. In addition, they generalize very well with a relatively smaller amount of training set [17,38]. Care should be taken in using these models; where the goodness of the fit is a very important parameter in such techniques which specifies how effectively PDF simulates the real distribution [151]. In SGM, there should be a smooth Gaussian distribution around the mean vector. However, in general conditions, the distribution is more sophisticated. In a controlled environment, by using elliptical Gaussian joint PDF, for particular color spaces, multivariate normal distribution of an m-dimensional color vector  $\mathbf{C}$  can be modeled as:

$$P(\mathbf{C}; \mu, \Lambda) = \frac{1}{(2\pi)^{\frac{m}{2}} |\Lambda|^{0.5}} \exp\left(-\frac{(\mathbf{C} - \mu)^T \Lambda^{-1} (\mathbf{C} - \mu)}{2}\right) \quad (16)$$

Distribution parameters i.e. diagonal covariance matrix ( $\Lambda$ ) and mean vector ( $\mu$ ) are calculated based on ML (Maximum Likelihood) approach using N skin pixels of the training dataset in color space  $C_j$ :

$$\mu = \frac{1}{N} \sum_{j=1}^N C_j, \Lambda = \frac{1}{n-1} \sum_{j=1}^N (C_j - \mu)(C_j - \mu)^T \quad (17)$$

Either the probability value [152,153] or Mahalanobis Distance (MD) [154] from the mean vector and observation vector  $\mathbf{C}$  can represent the skinness of the color vector [155]. Using an empirical harsh threshold or a hysteresis (double thresholding) method [52] considering ROC, by comparing it with either of probability or MD, the decision about any given pixel can be taken. MD is defined as [38]:

$$\lambda_c(C_j) = (C_j - \mu)^T \Lambda^{-1} (C_j - \mu) \quad (18)$$

The result of applying the threshold to Eq. 18 is an ellipsoid in 3D color spaces. This will be logically true if the distribution of non-skin pixels is assumed to be uniform. Based on a non-uniform distribution, in [41], the use of unimodal Gaussians for both skin and non-skin distributions is investigated. In this case, the MD from color vector to both centroid of skin cluster and non-skin cluster should be considered:

$$\lambda_c(C_j) = (C_j - \mu)^T \Lambda^{-1} (C_j - \mu) - ((C_j - \mu_{ns})^T \Lambda_{ns}^{-1} (C_j - \mu_{ns})) \quad (19)$$

where “ns” indicates non-skin related parametr. Ketenci et al. [76] compared the performance of SGM in rgb, YCbCr, HSV, YIQ and YUV without considering the goodness of the fit test. However, there are certain color spaces in which the distribution of skin color is somehow follows Gaussian rule. YCbCr color space has been hugely employed in SGM based techniques. Zhu et al. [156] employed SGM in YCbCr color model for both skin detection and lip segmentation. The most significant features of SGM are their training simplicity, negligible storage requirement, and relatively low cost of computation. In [43,93,94], also SGM is employed using  $C=(cb,cr)^T$  as the color vector. Subban et al. [106] compared the performance of single Gaussian models in different color spaces reporting that YPbPr outperforms CIE-XYZ, YCC and YDbDr. In [65], a single Gaussian model is developed in RGB, HSV, YCbCr, CIE-Lab, and CIE Luv color spaces based on their own data base. Using MCC as a metrics for evaluation, they conclude that CIE-Lab has the best performance while normalized rgb has the lowest. In [157,158], the SGM classifier is constructed based on rgb color space, and it is applied to remove probable non-regions of interests.

GMM is developed to model more complex distributions as a normalized weighted sum of Gaussian PDFs [159]. It compensates the inability of SGM in dealing with uncontrolled conditions in a general skin segmentation problem, in addition to the fact that SGM is not capable of approximating the actual distribution due to the asymmetry of distribution in its peak [105]. The mixture PDF is the summation of weighted Gaussian kernels defined as:

$$P(C; \mu, \Lambda, w, N) = \sum_{j=1}^N w_j G_j(C; \mu_j, \Lambda_j) \quad (20)$$

where each  $w_j$  is the weight of each kernels and  $N$  is total number of single Gaussian components. The learning process in GMMs is quite different from that of SGMs; here, an iterative method called EM (Expectation Maximization) is often applied to estimate PDF's parameters. EM is a two stage algorithm which tries to find the maximum Log likelihood function that estimates the real distribution as much as possible. In order to utilize it, the number of Gaussian components i.e.  $K$  and other parameters should be initialized that for which  $k$ -means clustering algorithm can be employed [160]. Evaluation process is also similar to that of SGMs that probability itself or Bayesian rule may be exploited for segmentation.

The most appealing features of GMM models are their simple evaluation process and low memory cost. The training process is however, much longer than the former methods. A comparative study on the performance of single and mixture of Gaussian distributions in [161] showed that mixture models improve the performance only in a relevant operating region (high True Positive rates). In addition, they concluded that increasing the number of kernels is not very effective.

Xie et al. [68] employed GMM in RGB color space for hand detection. In [45], 16 kernels, and in [159], 2

Gaussian components are employed for construction of GMM. Terrillon et al. [108] claimed that  $k=8$  is an appropriate tradeoff between computation cost and goodness of fit. GMM technique was also utilized by Hossain et al. [162] who adopted different GMMs for different values of luminance in two bright and normal conditions.

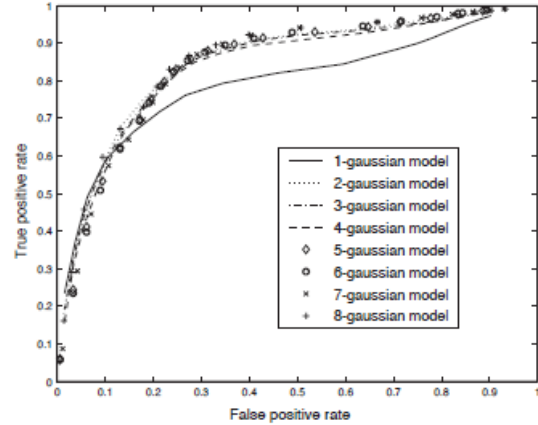


Fig.7. ROC for SGM and GMM [161]

Other less common approaches are cluster of Gaussians and Bivariate Pearson Mixture Model (BPMM). Cluster of Gaussians have been employed for estimation of distribution with the idea that skin cluster for different ethnicities (whitish, darkish and yellowish) or in different illumination conditions (high, medium and low) forms different separate Gaussian patches in particular color space. Phung et al. [39] used 3 Gaussian clusters each characterized by its centroids and covariance matrix, obtained using K-means clustering. Each of clusters is related to one level of luminance defined as low, medium and high. In evaluation, they utilized minimum MD distance of the color vector to the clusters. Zou et al. [164] also applied 3 Gaussian clusters to model different ethnicities. BPMM has been also addressed for the task of skin segmentation [165,166]. Here, BPMM type  $\Pi\alpha$  (Bivariate Beta) is used to model skin color in HS color space. They claimed that this model is more precise, as in compare with GMMs the false rejection rate of BPMM was 5% less than that of GMM.

Another alternative way to redress SGM and GMM limitations was proposed in [105]. Here, skin distribution is estimated based on observation of semi-elliptical shape of skin cluster in some color spaces. For this case, Lee et al. [298] defined the elliptical boundary model as  $P(C; \mu, \Lambda)$  where:

$$P(C; \mu, \Lambda) = (C - \mu)^T \Lambda^{-1} (C - \mu) \quad (21)$$

From the training point of view, this model is less computational and faster than GMM and parameters are simply obtained as:

$$\mu = \frac{1}{N} \sum_{j=1}^N C_j, \Lambda = \frac{1}{N} \sum_{j=1}^N f_j (C_j - \Omega)(C_j - \Omega)^T, \Omega = \frac{1}{N} \sum_{j=1}^N f_j C_j \quad (22)$$

In above equations,  $N$  is total number of samples, and  $\Omega$  is mean of chrominance vectors. Given an input chrominance vector  $C$ , the decision on skinness of the pixel can be calculated using harsh thresholding [105]. The model has been tested under rg, CIE-uv, CIE-ab, CIE-xy, IQ, and CbCr color models concluding that it outperforms SGM and 6 kernel GMM model from detection rate point of view in all color spaces. Computationally, elliptical boundary model is almost as quick as SGM, but it is faster than GMM. Xu et al. [167] also leveraged both elliptical boundary model and depth data (by using Kinect sensor) to perform hand and face localization.

Entropy Based Models have been also utilized for skin segmentation as well as several other image processing task including face detection [24], speech recognition [168], etc. However, they lost their attraction soon in skin classification mainly due to their high volume of computation, and intangible performance improvement in compare with former parametric and non-parametric models. Generally, a model is inferred of a set of training data to be used for classification. Choosing several relevant features, their histogram is calculated on the training pixels, and model parameters are obtained based on MaxEnt model [169]. Varieties of models have been developed based on maximum entropy model (MaxEnt), with constraints concerning marginal distributions. In one model, it is assumed that pixels are independent in color (baseline model) [170]. The MaxEnt solution to this model was obtained using Lagrange multipliers. The second model, Hidden Markov Model (HMM) provided better detection results as it is not as loose as the baseline one [171]. This model is obtained by constraining the baseline to a model in which skin zones are not considered thoroughly random but are made of pact patches. Thus, a constraint is added to the prior  $P(Y)$  ( $Y$  is the skin probability map (SPM) of the image) using 4 adjacent neighbors of each pixel. In this case, the MaxEnt model follows Gibbs distribution [172]. In First Order Model (FOM), there is another constraint imposed on the two-pixel marginal of the posterior, i.e.,  $P(y_s, y_t | x_s, x_t)$  in which  $s, t$  are neighbor pixels,  $x$  is the color vector and  $y$  is the skinness probability. Using Lagrange multipliers, solution to MaxEnt problem in FOM model is [172]:

$$P(Y | X) \approx \exp\left[\sum_{\langle s,t \rangle} \lambda(s,t,x_s,x_t,y_s,y_t)\right] \quad (23)$$

where  $\lambda(y_s, y_t, x_s, x_t) > 0$  are parameters of the distribution. The total number of parameters is  $256^3 \times 256^3 \times 2 \times 2$  for a 24bit RGB color image.

Markov Random Fields (MRF) and Conditional Random Fields (CRF) are also other statistical approaches successfully applied to segmentation tasks with similar concepts to previous methods. The former is a graphical solution that models joint probability distribution based on Bayesian framework which takes the spatial connection between pixels into account, while the latter is an approach directly models conditional PDF [173]. For MRF first spatial relationship between pixels is directly

integrated and then by using Bayesian concept, the model is inferred. The label distribution is computed by maximizing the probability of the MRF model as:

$$x^* = \arg_x \max\{P(x | y)\} \quad (24)$$

Using Bayesian method and the fact that prior probability of  $y$  i.e.  $P(y)$  is independent of  $x$ , the above equation is rewritten as:

$$x^* = \arg_x \max\{P(y | x)P(x)\} \quad (25)$$

In [79,340], an initial segmentation is performed by using elliptical boundary model, and then by means of an iterative method, other skin pixels are annexed to the initially annotated ones. The repetitive algorithm exploits Ising model (a two class problem with interacting particles where pixels arranged in a planar grid) and Hammersley-Clifford Theorem (HCT) to compute the probability in terms of total energy. In each iteration, based on Ising model, the total energy of the system is computed and then by using HCT, the probability of configuration is calculated. The process continues until either the temperature reaches to a small value, or the difference in number of detected skin pixels between two successive results is less than 1 percentage of the total size of the image. In Gibbs distribution the temperature is a parameter which is updated in each replicate. The experimental results in [79,174] show the effectiveness of the method in compare with elliptical boundary model proposed in [105]. In some semi-skin color regions, however the algorithm exhibited poor results.

CRF model primary represented in [173,175] also combines information of different color spaces and model the spatial relationship between image pixels. Here, Eq. 24 is rewritten with an assumption that the detection results of  $K$  classifiers i.e.  $D(i)$  are available. The aim is to combine the high-level information obtained by each  $D(i)$  with the low-level information in image to take the final decision:

$$x^* = \arg_x \max\{P(x | y, D^{(1)}, D^{(2)}, \dots, D^{(k)})\} \quad (26)$$

CRF is a discriminative model which defines the conditional probability as:

$$P(x | y, D^{(1)}, D^{(2)}, \dots, D^{(k)}) = \frac{1}{Z(y, D^{(1)}, D^{(2)}, \dots, D^{(k)})} \times \quad (27)$$

$$\prod_{C_p \in C} \prod_{\psi_c \in C_p} \exp\left(\sum_{q=1}^{Q(p)} \lambda_{pq} f_{pq}(x_{\psi_c}, y, D^{(1)}, D^{(2)}, \dots, D^{(k)})\right)$$

where  $Z(y, D^{(1)}, D^{(2)}, \dots, D^{(k)})$  is normalization constant,  $C_p$  is a clique template from the clique set  $C$ ,  $f_{pq}$  is the  $q^{\text{th}}$  real-valued feature function defined on the clique template  $C_p$  and  $\lambda_{pq}$  is the model parameter. Ahmadi et al. [173,175] utilized a pseudo-likelihood algorithm for parameter estimation as an approximation to ML estimation. For simplicity, in [175], they consider two

simple explicit boundary model as  $D^{(1)}$ ,  $D^{(2)}$  and in [173], they boosted the performance of their detection by using the explicit boundary model in more color spaces.

A Bayesian network (or probabilistic directed acyclic graphical model) is a probabilistic graphical model that encodes a set of random variables and their conditional relationships (dependencies) [176]. Formally, Bayesian networks are directed acyclic graphs (DAG) in which edges represent conditional dependencies, and nodes which are not connected are conditionally independent. Each of the nodes represents random variable  $X_i$  with class conditional probability  $P(X_i|\Pi_i)$  where  $\Pi_i$  is the parent of  $X_i$  in graph. Variables could be observable quantities, latent variables, unknown parameters or hypotheses associated with a probability that takes a particular set of values as inputs for the node's parent variables and gives the probability of the variable represented by the node [176]. The structure for a network,  $S'$ , is correct when it is possible to find a distribution  $P(C, X|S')$  that matches the actual distribution; otherwise, it is incorrect. ML estimation is utilized to learn the network parameters [177]. Two approaches are often used for construction of Bayesian classifiers; either selecting a structure and specify the dependency among variables or determining the distribution of features. Both impose the parameter set that is required to calculate the decision function. Popular Bayesian network classifiers are the Naive Bayes (NB) classifier, in which the features are assumed independent from the given class, and the Tree-Augmented Naive Bayes classifier (TAN) which was proposed to enhance the performance over the simple Naive-Bayes classifier. In the structure of the TAN classifier, the class variable is the parent of all the features and each feature has at most one other feature as a parent, such that the resultant graph of the features forms a tree. For learning the TAN classifier, the structure that maximizes the likelihood function out of all possible TAN structures is found in training data. The learning method proposed by Sebe et al. [177] is stochastic structure search (SSS) which improved the performance on a piece of Compaq dataset they used.

### 3) Discussion on statistical methods

Several statistical skin detection techniques have been elucidated in this subsection; all categorized in two minor classes of non-parametric and parametric approaches. Non-parametric methods are based on histogram of skin and non-skin pixels in a predefined training set. The size of training set has a direct impression on the performance of detectors. Jones et al. [45] employed a database of two times more than  $256^3$  size to train their LUT system. This enormous training set is required to produce remarkable results. They were also unable to generalize and interpolate the training data [17,38]. The large storage requirements make the methods unfavorable choice particularly in embedded micro-system platforms. However, in compare with parametric models, histogram models are independent of the shape of cluster, trained simply and fast. In evaluation also, they can be utilized in several clock cycles per pixel needed for accessing

memory. In parametric methods, the simulation results of real distribution needed to be approved using goodness of fit measures. In addition, the choice of color space is also determinative as it is strongly dependant on the shape of skin cluster. Training complexity varies in different parametric methods, but in overall, both the training time and its complication are not copacetic in compare with histogram models. Parametric and non-parametric models have been compared throughout this section and generally, it was concluded that non-parametric models often outperform parametric ones with the cost of high storage requirement.

### C. Neural Net Models

Artificial neural networks (ANNs) are mathematical models representing function  $F: X \rightarrow Y$ , with a distribution over  $X$  or both  $X$  and  $Y$ , that simulate the function of brain inspired by human nervous system. They are generally presented as systems of interconnected neurons. Like other machine learning methods, neural networks have been used to solve a wide variety of tasks. In skin detection, ANNs have been utilized for different purposes and structures. In illumination compensation, dynamic models, in combination with other techniques, and direct classification, variety of ANNs such as MLP, SOM, PCNN, etc, are exploited.

A multilayer perceptron (MLP) is a feed forward artificial neural network model that consists of several layers of nodes in a cyclic directed graph, each layer fully connected to the next one. Each neuron is a processing element with a nonlinear activation function except for input ones. A common technique to train MLPs is back propagation (BP) which is used in conjunction with optimization methods such as gradient descent. This method calculates the gradient of a loss function with respects to the weights in the network. The gradient is fed to the optimization method to update the weights, in an attempt to minimize the loss function. To find a local minimum of a function using gradient descent, steps are taken proportional to the negative of the gradient of the function at the current point. This is done by simply taking the derivative of the cost function with respect to the network parameters and then changing those parameters in a gradient-related direction. The success of this method is based on the cost function, the model and initial settings (weights). In skin detection, this has been managed by an exploration searching ability in evolutionary algorithms such as ICA (imperialistic competitive algorithms). In this case, the optimal value for the weights is formulated as an exploration searching problem wherein the structure of ANN is reconstructed and fixed during evolution [68].

Wu et al. [179] presented a work inspired by human visual system as skin detection. They combined color and texture information using 3 layers and 3 types of neurons. The decision about each pixel located at  $(x,y)$  depends on the pixels in two receptive fields ( $R_N(x,y)$  and  $R_C(x,y)$ ). The model is depicted in Fig. 8. Three cones are considered in response to red, green, blue channels as well as a rod (Neuron N) in response to brightness and an orientation neuron (T neuron) to detect orientation

textures. The input for color neurons are the mean value in  $R_C(x,y)$  receptive field while the latter neuron extract texture in  $R_N(x,y)$ .

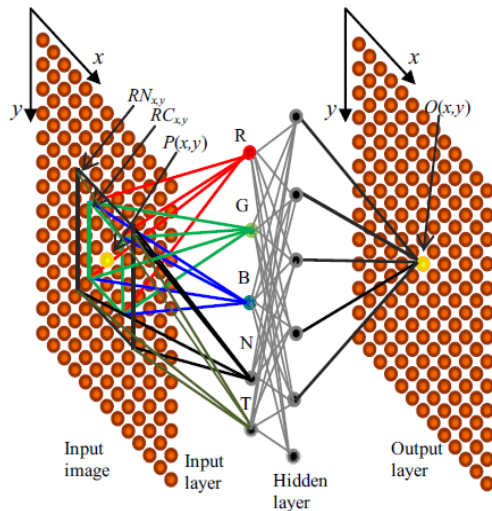


Fig.8. The ANN used in [179]

Al-Mohair et al. [116] employed a 3-layer MLP classifier using different color spaces and different number of neurons in hidden layer. Among common color spaces, they concluded that YIQ gives the highest skin and non-skin separability. In [20], a multi stage technique is characterized by several parallel MLPs trained using a committee machine. The idea is based on the fact that the performance of ANNs depends on the initialized parameters and training data. In this case, several parallel ANNs are trained independently and in evaluation, their outputs are fused to improve the stability of the system. In committee machine, a simple arbitration scheme (median selector) is considered to generate the skinness of a pixel. The authors claimed that this structure improves the skin color detection in compare with a single MLP system. Mitra [52] used a neural network to obtain the probability curve. The Kolmogrov-smirnov goodness of fit test was exploited resulting that the ANN estimate the actual distribution better than a normal fit. In [180], an MLP is trained using back-propagation algorithm to provide an optimized decision boundary and then it is used for both interpolating skin regions and skin classification.

In another approach, Chen et al. [181] exploited rgb color model in an ANN based skin classifier. The ANN is constructed out of 2 input neurons, 4 neurons arranged in two cascade layers in hidden layer, and one output neuron. A logistic sigmoid transfer function  $f(x)=(1+\exp(\lambda x))^{-1}$  is considered where  $\lambda$  represents the steepness of the function. The convergence of the ANN directly depends on the initial value of  $\lambda$ . To address this, they used genetic algorithm to obtain the optimized value of  $\lambda$ . Bhojar et al. [348] proposed an MLP skin classifier with 3 RGB input neurons, 5 neurons in hidden layer and 2 output neurons indicating skinness and non-skinness of a pixel. Final decision about the pixel is based on output values in compare with a harsh threshold. Doukim et al. [182] investigated impact of different color spaces as well as

number of input neurons and hidden neurons in an MLP skin classifier. They used several strategies based on  $YC_bC_r$  deducing that combining Cb/Cr feature by the Cb, Cr color channels improves the performance, and that there is a significant performance variation on selecting different number of hidden neurons.

Duan et al. [184] used synchronous pulse firing mechanism of pulse coupled neural network (PCNN) to simulate the skin detection mechanism of human eyes. PCNN explains synchronous activity among neural assemblies in the cat cortex induced by feature dependent visual activity. In contrary with former neural net composed of rate-coding neuron, PCNN neuron can code information toward time axis. A single layer, two-dimensional network is designed. In the network, the neurons and the pixels are in one to one correspondence i.e. one neuron is equal to a pixel. Pulse output will be delivered to adjacent neurons, they will fire too because of pulse coupled action. In [184], PCNN method considers relationship between neighboring pixels similar to human vision mechanism, that semi colors should be segmented into an area block in presence of illuminations.

A self-organizing map (SOM) or self-organizing feature map (SOFM) is a particular kind of ANN trained using unsupervised learning to produce a low-dimensional (typically two-dimensional), quantized representation of map (the higher dimensional input space of the training samples). Contrary to other ANNs, SOMs use a neighborhood function to preserve the topological properties of the input space which makes SOMs useful for visualizing low-dimensional views of high-dimensional data, akin to multidimensional scaling [185]. In training phase, SOM builds the map using input examples in a competitive process called vector quantization. The arrangement of neurons is a two-dimensional regular spacing in a hexagonal or rectangular grid. The procedure for placing a vector from data space onto the map is to find the node with the closest (smallest distance metric) weight vector to the data space vector.

In the earliest SOM based skin classifier proposed by Brown et al. [186], two different SOMs, a skin SOM and a co skin and non-skin SOM were trained based on more than 500 color images and for evaluation, common color spaces were exploited. The performance of SOM classifier in contrary to the GMMs was almost independent of the choice of color space. In a more recent work [58], the Self-Organizing Mixture Network (SOMN) is modified to improve SOM's performance, stability and applicability. In this case, the non-parametric skin distribution PDF is estimated based on the self-organizing map (SOM) structure, Kullback-Leibler divergence (KLD), and an stochastic approximation method [187]. The Kullback-Leibler divergence (KLD) is the expectation of log likelihood, a suitable measure for density estimation. Furthermore, the SOMN uses a stochastic gradient descent algorithm, so it can easily escape from shallow local minima.

RBF (radial basis function) network is an ANN that uses radial basis functions as activation functions. RBF



networks are typically based on three layers: an input layer, a linear output layer and a hidden layer with a non-linear RBF activation function. In an RBF network with  $N$  neurons in hidden layer, the relationship between input vector of real numbers  $X$  and output, is defined as:

$$\Phi(x) = \sum_{i=1}^N a_i \rho(\|x - C_i\|) \quad (28)$$

where  $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $a_i$  is the weight of neuron  $i$ , and  $C_i$  is the center vector of neuron “ $i$ ” in the output neuron. The norm function is either the Euclidean Distance (ED) or MD and RBF is commonly considered as Gaussian i.e.:

$$\rho(\|x - C_i\|) = \exp(-\beta \|x - C_i\|^2) \quad (29)$$

The Gaussian basis functions are local to the center vector i.e. changing parameters of one neuron has only a small effect for input values that are far away from the center of that neuron. These networks are trained in a two step process; first, center vectors i.e.  $c_i$  are chosen using either a random selection,  $k$ -means clustering or BP. Then, a linear model is fit with coefficient  $w_i$  to the outputs of hidden layer neurons with respect to a set of objective functions such as Least Square Error (LSE). Khan et al. [15] applied RBF networks for skin segmentation in color spaces such as RGB, CIE-Lab, YCbCr, rgb, HIS, IHSL. The performance was varied significantly for different color models and CIE-lab exhibited the best one.

#### D. Spatial based models

In order to address the shortcomings of pixel-based detectors, recently, variety of spatial based (diffusion based) methods have been developed. The strategy in these techniques is extraction of initial skin seed by means of a precision oriented pixel based detector and then annex other skin pixels to the seed. The idea is based on the observation that skin regions are associated with several pact blobs. Thus, it would be reasonable to detect some high probable skin points and then use spatial relationship between pixels to propagate and cover other skin points. Ruiz-de-solar et al. [91] proposed a method of controlled diffusion. In seed generation stage, they utilized a GMM with 16 kernels in  $YC_bC_r$  color space and then the final decision about the pixel’s class is taken using an spatial diffusion process that takes context information into account. In this process, if Euclidean distance between a given pixel and a direct diffusion neighbor (already skin pixel) is smaller than a threshold value, then the propagation occurs. Also, the extension of the diffusion process is controlled using another threshold value, which defines the minimal probability or membership degree allowed for a skin pixel. This process works well in regions where the boundary of skin and non-skin pixels are sharp enough unless the leakage occurs.

Mahmoodi et al. [189] proposed a simpler scheme to generate initial seed. They segmented the skin cluster in  $YC_bC_r$  color model into 3 non-overlapping regions; pixels with high probability to be skin, non-skin pixels and

pixels with unknown status. They used the first category of pixels to generate initial result and then subsequently used a neighboring procedure for each skin pixel to include more skin pixels and finally a window based procedure is used to determine qualified windows for their further face detection [190,191]. In [142], after obtaining skin probability of the image, hysteresis thresholding is employed to segment skin blobs. At first step a “strong” threshold is used to generate high-confidence skin-colored pixels that constitute the seeds of potential hand or face blobs. In second thresholding step, a “weak” threshold is exploited along with prior knowledge with respect to object connectivity.

Abdullah-Al-Wadude et al. [192] also proposed an algorithm that uses color distance map (CDM); a gray scale image robust against variations in imaging conditions and an algorithm based on the property of flow of water which uses spatial analysis to extract skin blobs. For spatial analysis, the edge map of the image is utilized as a height of a pixel and then water is dropped towards both sides of the edges. This water-stream obeys two features of real water; it always flows downwards to the deepest point and if it faces any obstacle, it starts filling up until it overflows the barrier or flows to another direction. Based on this, every stream flows until it either meets with an already defined skin pixel or non-skin pixel or image boundary. Based on each of these conditions, appropriate actions will be taken. Recently, in order to overcome leakage issue in diffusion approaches, Kawulok et al. [115,193] proposed an energy based approach in which the probability of pixels are utilized to determine the skinness. This method is based on cumulative propagation which is less susceptible to leakage. The other work by the same author [194] is a diffusion procedure based on a distance transform for propagating in a combined domain of skin probability, luminance and hue. In this work, after initial seed extraction, a shortest route from the seed to every pixel is determined as:

$$C(x) = \sum_{i=0}^{l-1} \rho(P_i \rightarrow P_{i+1}) \quad (30)$$

where  $\rho$  is the local skin dissimilarity measure between two neighboring pixels,  $p_0$  is a pixel that lies at the seed boundary,  $p_l=x$  and  $l$  is the total length of the path. In geodesic transforms,  $\rho$  is a function of gray level difference between two pixels; however, in [194], they used a combined domain to fuse the information of skin probability map and spatial analysis.

In another method, Mahmoodi et al. [195,196] proposed a new method in which they considered several virtual lines around each skin pixel in initial seed. Afterwards, each skin pixel starts to “propagate” to other non-identified pixel on the virtual lines. In parallel, the edge map of the image is calculated using an empirical threshold to eliminate leakage effect in former methods. The feature used in diffusion is based on the concept of Otsu multi-thresholding. Otsu segmentation is used in several color channels in order to specify the homogenous regions of an image. For diffusion of a master pixel into

an under test pixel, several factors are considered including the group in the initial seed (white, gray or black) and the homogenous class of both pixels. In more recent method [197], not only several features are appended including motion information in local regions, color distance map and skin probability map to the former works, but also an efficient method of conservative extraction of initial seed and multi-step diffusion are used.

#### E. Multispectral skin selection

Hyperspectral cameras provide useful discriminatory data for biometric applications. Electromagnetic spectral bands far below 0.4  $\mu\text{m}$  are extremely dangerous and they are inapplicable for skin detection. Thermal IR imagery has been suggested as an alternative source of information for detection and recognition applications. While visual cameras measure the electromagnetic energy in the visible spectrum range (0.4–0.7  $\mu\text{m}$ ), sensors in the IR camera respond to thermal radiation in the infrared spectrum range at 0.7– 14.0  $\mu\text{m}$  [198]. Multispectral and hyperspectral cameras observe electromagnetic energy across several spectral channels, creating multidimensional clusters of skin and background pixels that are more distinct than in RGB imagery [199]. Skin detection using imaging devices with hyperspectral capabilities are not ordinary approaches as they are only applicable using expensive equipments and in particular conditions. Though, hyperspectral imagery offers a distinct advantage due to the abundance of spectral information. Glass blocks a large portion of thermal energy resulting in a loss of information near the eyes and variations in body temperature also significantly change the thermal characteristics of the object which substantially impact on the performance of multispectral models [198]. Morikawa et al. [200] claimed that using an NIR (Near Infrared) image with a central wavelength of 1050 nm improves the robustness of skin detection.

The overall process of detecting skin regions in non-visible spectral imagery consists of two steps; developing a model which describes skin reflectance and a detection algorithm which is applied on the model. In [201], Nunez et al. proposed a model for justification of skin reflectance near infrared wavelengths. Here, skin is described based on six layers of dermal tissue with varying amounts of water and melanosomes that each is a layer of subcutaneous fat with a large amount of reflectance in the NIR when it is sufficiently thick. Based on optical coefficients of each layer, the thickness of each layer, and the reflectance of subcutaneous fat, a model of skin reflectance is defined as:

$$\hat{I}(\lambda) = I_0(\lambda)F + I_0(\lambda)(1-F)R_1(\lambda) + I_0(\lambda)(1-F)\sum_{m=2}^7 R_m(\lambda)\prod_{n=1}^{m-1} T_n^2(\lambda) \quad (31)$$

where for  $n=1:6$ ,  $R_n$  and  $T_n$  are obtained from the Kubelka Munk equations and  $R_7$  comes from the reflectance of fat diagram, the index of refraction of the stratum corneum i.e.  $n_2$  is estimated 1.5, and the index of refraction of the

atmosphere i.e.  $n_1$  is almost 1. For the air/stratum corneum interface, the amount of reflection  $F$  is considered approximately 4%. The detection algorithm in [201] which is called Normalized Difference Skin Index (NDSI) is described in Eq. 32 where  $\rho_i(\lambda)$  is the reflectance of  $i^{\text{th}}$  pixel at wavelength  $\lambda$ .

$$\gamma_i = \frac{\hat{\rho}_i(1100 \text{ nm}) - \hat{\rho}_i(1400 \text{ nm})}{\hat{\rho}_i(1100 \text{ nm}) + \hat{\rho}_i(1400 \text{ nm})} \quad (32)$$

NDSI values vary in the range of -1 to +1. Experimentally, it is observed that the NDSI values range from 0.641 to 0.742 for the darkest skin to the fairest skin. As such, a very simple thresholding scheme is sufficient to segment skin and non-skin regions. More information on the optical properties of human skin is available in [202]. Recently, Kidono et al. [203] employed a multiband camera, to simultaneously obtain seven spectral images, to perform pedestrian detection (Fig. 9). Two approaches were presented to detect human skin from spectral images in an outdoor environment. In one method, they used an analytical method based on visible color information and subtraction between the NIR spectral images. This is accomplished by fusing the explicitly defined method on visible and non-visible images. Another approach is the statistical method for learning the brightness distribution of human skin in the seven spectral images which is performed using GMM with EM training procedure.

Suzuki et al. [204] also proposed a method based on the subtraction of two NIR images whose central wavelengths are 870 nm and 970 nm. Dowdall et al. [205] also proposed a method of skin detection based on the idea that Human skin exhibits an abrupt change in reflectance around 1400 nm. This phenomenology was employed by taking a weighted difference of the lower band near-IR image and the upper band near-IR image which increases the contrast between human skin and the background in the image. After a simple thresholding, the binary image undergoes a series of opening and closing morphological operations. Angelopoulou et al. [206] investigated the reflectance of skin of different people all around the world in the range of visible spectra (Fig. 10). They analyzed the collected data statistically with goal of finding a particular pattern in the spectrum. Afterwards, a model was constructed that encapsulates the “W” pattern. They investigated the usage of GMM, an orthogonal wavelet approximation and principal component analysis (PCA) to describe the spectral behavior of skin where they opted for the Gaussian representation because it provided a description of filters that are optimal for the detailed reconstruction of the skin’s spectral distribution.



Fig.9. The Multiband Camera used in [203] and Sample Images (From Left to Right: Visible Band, 740nm Band, Three NIR Bands)

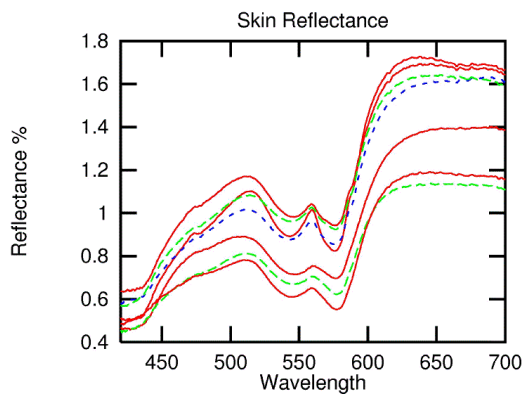


Fig.10. Reflectance Spectra for Different Races. (Caucasian:Red, Asian:Green, East Indian:Blue And African:Magenta)

Schwaneberg et al. [207] also presented a multispectral scanning sensor to classify an object's surface material. The system is meant to detect the presence of limbs and therefore optimized for human skin detection. In [208], skin is characterized by the area of the spectroscopic curve around seven relevant bands which correspond to absorption peaks of specific physiological components. Reddy et al. [209] introduced a touch-less anti-spoofing system which defines an aliveness factor based on the ratio of the reflected radiation in 660 and 940 nm wavelengths. All systems proposed in former methods are based on visible-near infrared or near infrared-short wavelength bands. Ferrer et al. [210] proposed a new approach to increase the security of contactless biometric devices based on a skin detector which is designed using reflectance spectroscopy in visual, near infrared and short wave infrared bands. The experimental results are encouraging in compare with former single band systems.

#### F. Adaptive models

Using online information of the image or sequence of frames has been exploited as an effective idea to counteract non-uniform illumination to some extent. Adaptive models are developed in an effort to present models which are calibrated to given inputs. In some cases, previously defined models are tuned (adapted) for specific conditions i.e. the background, imaging equipment, lighting conditions, and even subject of the image. This approach reasonably yields to high detection rate with the cost of losing generality.

In second category of adapted models, a statistic model is updated based on global information of the image. Chaudhary et al. [28] proposed two automatic color space switching techniques. Initially, an ANN is trained to select the proper color space for an input image and then segmentation is performed using Bayesian classifier. In second approach, after segmentation by means of Bayesian classifier in triple color spaces, the color space with the maximum number of pixels in the largest Blob in all output images is chosen. This is done with assumption that human is the nearest object to the imaging device. Subban et al. [88] constructed a simple adaptive model which changes the upper and lower boundaries of "r" and "g" channels based on their mean

and variance all over the image. Hu et al. [84] proposed a method of updating a Gaussian model by linearly combine of a trained model and color distribution of the input image with weighting factors as:

$$P_{adapt}(x) = w_1 P_{train}(x) + w_2 P_{input}(x) \quad (33)$$

where  $P_{adapt}(x)$  is the adapted model,  $P_{train}(x)$  is the predefined model trained by SFA dataset, and  $P_{input}(x)$  is the new model generated on the input image.  $w_1, w_2$  are considered so that  $0 < w_1, w_2 < 1$  and  $w_1 + w_2 = 1$ . Lee et al. [212] proposed a method of image filtering which exploits an ANN for learning-based chromatic distribution-matching. This specifies the image's skin chromatic distribution online such that it can tolerate the chromatic deviation coming from special lighting without increasing false alarm. Aiding color, coarseness texture is fused to acquire more accurate skin segmentation. Furthermore, low-level geometrical constraints and a mugshot exclusion procedure are employed to examine the skin regions of objects. Yang et al. [213] integrated an ANN to dynamically configure a Gaussian model. The method is based on the fact that distribution of a skin color depends on the luminance level in the  $C_b-C_r$  plane. Thus, firstly, according to statistics of skin color pixels, the covariance and the mean value of  $C_b$  and  $C_r$  with respect to  $Y$  channel is calculated, and then it is used to train a neural network which gives a self-adaptive skin color model based on an online tunable Gaussian classifier. Another work which utilizes an ANN to determine the optimum threshold on Bayesian classifier was proposed by Zhang et al. [387]. In this case, several features extracted from skin probability map are used to help searching candidates of optimum thresholds, and an ANN classifier is trained to select the final optimum threshold.

In [214], skin pixels are detected with the aid of a Bayesian classifier which is bootstrapped with a small set of training data. Then, an off-line iterative training procedure is employed to refine the classifier using additional training images. To make this happen, hysteresis thresholding is applied in which first a relatively high threshold is considered. Non-filtered points constitute potential high blobs. Following that, a weaker thresholding scheme is performed to allow pixels that are immediate neighbors of skin-colored points, being added to each blob recursively. Zhu et al. [66,215] proposed a method in which first, an initial seed is generated based on a generic skin color model. In the second step, a GMM, specific to the image under consideration, is trained using the standard EM algorithm. Then, an SVM classifier is employed to detect the skin distribution from the trained 2-kernel GMM by incorporating spatial and shape information of skin pixels. Oliver et al. [216,217] also employed the same dynamic procedure (using online EM to determine GMM parameters) for tracking, shape description, and classification of the human face and mouth. In [218], Yang et al. addressed dynamically updating GMM parameters using ML algorithm for confronting illumination variations. Online reconfiguration of GMM models using global information

was also exploited in [219,220,221]. Cho et al. [222] presented an adaptive skin color filtering by means of a thresholding box in HSV color space which was updated dynamically based on central gravity of histogram of colors which are more than 1 tenth of maximum color value in the box. In this case, they assume that the size of skin color region is comparable to background region. Afterwards, color vectors inside the box are classified to skin vectors and background vectors, using a multistep clustering analysis.

Sanmiguel et al. [70] performed skin detection by field adaptation where detectors which classify images based on explicitly defined regions in different color spaces are adapted to data within an optimization framework. It selects the best detector's configuration based on agreement maximization (AM). This model is able to adapt its parameters in video surveillance systems with a medium field of view. One of the detectors is employed to detect skin pixels, while the others segment semi-skin pixels. Detectors are mixed through mathematical morphology to effectively keep skin pixels. Additionally, the agreement maximization framework computes the resemblance between the sub-detectors of each detector and maximizes it. This framework is extended to model the relation between the sub-detector parameters, to consider agreement within the expected ranges and to select the optimal channels of color spaces. Sigal et al. [224] proposed a dynamic approach for video sequences based on a second order Markov model to predict evolution of the skin-color (HSV) histogram over time. The first stage of this system is obtaining an initial estimate for the location of the foreground skin and non-skin regions which is achieved by segmenting the first frame, with histogram-based conditional probability distributions for the two classes that have been computed off-line. In the second stage, histograms are dynamically updated based on feedback from the current segmentation and predictions of the Markov model (Fig. 11).

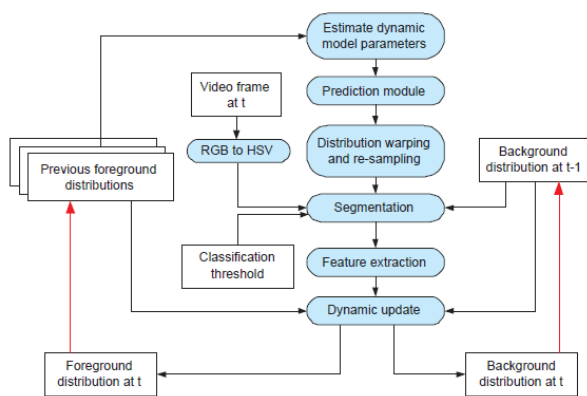


Fig. 11. Online Reconfiguration Model Proposed in [384]

In some dynamic approaches, a model is reconfigured based on local information extracted from pre-detected features. Ibrahim et al. [27] employed VJ face detector [67] to segment face regions. Based on face information, the exact boundary of an explicit method in  $YC_bC_r$  is obtained. Using this dynamic threshold, the skin regions all over the

image is obtained. However, the performance of the method strongly depends on the accuracy of the face detector. In addition, the method is only useful for images with at least one face. Similar face detector was utilized in [89,225,226] where a model is reconfigured based on the skin cluster in face region. In [227,228], authors employed the same idea using PCA-based and SVM-based face detectors respectively. In the latter, the error rate decreases from 26% (when using LUT method alone) to 15% for the face detection plus LUT method. A dynamic model for rapidly changing illumination conditions were proposed by Liu et al. [229] in which first, face detection is applied to sample skin colors online, and a dynamic thresholding technique is used to build and update the skin color model under the Bayesian decision framework. For images where no face is detected, a color correction algorithm is applied to convert the colors of the current frame to those as they appear under the same illuminant of the last model updated frame. The skin color model is still effective for the color corrected images. In [230], Lia et al. presented an edge detector based face detection algorithm in conjunction with extraction of sample pixels from a predefined location (right hand cheeks) of the face window which contains "good" skin pixels. The upper and lower bounds of an explicitly defined skin detector are specified by means of local minimum in hue histogram of those regions.

Tracking based mechanisms were also employed mainly due to their simplicity for video images. Sun et al. [38] proposed an adaptive system in which preliminary skin pixels are extracted globally using a non-parametric model, and then among them high probable skin pixels are exploited for construction of a local GMM model which is trained using k-means algorithm. Yoo et al. [231] proposed a hand tracking algorithm in which the histogram of skin pixels is updated in each frame using former frame histogram and current one. The skin detection in each frame is performed using both motion information (simple frames differencing technique) and histogram of skin pixels in HSV color space.

Yogarajah et al. [232] also used eye as a feature to extract local skin distribution to obtain dynamic threshold values. In [233], Soriana et al. used skin cluster to select training colors for dynamically updating histogram which is used to track faces. The idea in [234] is to utilize texture in order to obtain the optimum threshold in Bayesian classification rule, where the threshold is estimated based on uniformity of local regions. In this case, after an initial segmentations step, a region growing algorithm is applied for each isolated blob where the threshold is increased step by step and then homogeneity of the region is investigated in each iteration. Dadgostar et al. [235] presented a classifier based on adaptive histogram of hue channel in video images. Here, a new histogram is considered for each frame which is linear combination of global histogram and histogram of mobile skin pixels. In [236], Soriano also presented a tracking based local dynamic model in which ratio histogram (the division of skin histogram to histogram of the image) is readjusted in each frame to compensate lighting variations and it is

directly used for classification in an explicitly defined model.

Skin detection based on local models comes with high false negative rate. A possible solution is combination of local and global models. Sun et al. [237] made use of this fact as he proposed a method of tracking skin regions in videos which utilizes the correlation between continuous frames. In this approach, a local skin model shifts a globally trained skin model to adapt the final skin model to the current image. By taking advantage of histogram technique, a set of skin pixels are extracted and then the distribution of these pixels is estimated by means of a GMM subsequently. The final model is combination of dynamic GMM model and the global model for each particular image. Kawulok et al. [238] combined both global and local information of the image to construct a probability map which is used to generate initial seed for an spatial based approach. Here, the final probability  $P_F(C|skin)$  is computed as a weighted mean of the probabilities obtained using the local  $P_L(C|skin)$  and global  $P_G(C|skin)$  models. A recent exhaustive method proposed by Khan et al. [383] employs VJ face detector accompanying with graph cut based segmentation. This systematic approach begins by exploiting the local skin information of detected faces. The detected faces are used as foreground seeds for calculating the foreground weights of the graph. If local skin information is not available, universal seed is selected and to increase the robustness, the decision tree based classifier is used to augment the universal seed weights.

#### G. SVM models

Support vector machine (SVMs) are supervised learning models applied to many pattern recognition tasks as well as human skin classification. Using annotated training set of skin and non-skin pixels, an SVM training algorithm constructs a model which tries to assign pixels into the two classes making it a non-probabilistic binary linear classifier. An SVM model is a representation of the pixels as the points in space, mapped so that skin and non-skin pixels are divided by a clear and as wide as possible gap. New pixels are then mapped into that same space and predicted to a category based on which side of the gap they fall on. The main issue in SVM classifiers is its training complexity as with the increase of learning data and features the size of learning model increases exponentially. This is a drawback particularly for algorithms with large number of training samples and this is a serious issue in applications such as skin detection. However, in some applications of skin detection, where subjects of interests are consistent during training and evaluation, an online training or an adaptive system may be a solution. Han et al. [239] exploited an SVM based on active learning to detect skin pixels for gesture recognition, claiming that in compare with other applications, videos for such systems generally contain very few signers and their skin colors are consistent across the frames. This means it is not essential to get many training samples; and in addition, information from previous frames can be adopted to process the current frame. Hence, for collecting

training data, the pixels from the first couple of frames as the training samples are used. They proposed a multi-stage framework in which in a training stage, a generic skin color model (explicitly defined in RGB) is applied to segment skin areas initially, and then based on that, a binary classifier based on SVM active learning is trained. In evaluation, the classifier is incorporated with the region information to detect final skin mask. Though active learning reduced the accuracy and detection rate in compare with an SVM with general training method, the overall training time significantly improved. However, the performance will be dependent on the generic model. Zhu et al. [66,215] also employed an SMV classifier in an online dynamic system. Genetic algorithm is also applied in [240] to reduce large training sets of SVMs.

#### H. Mixture methods

This category of methods includes the ones designed based on mixture of previously discussed algorithms. Different combination of former methods have been explored. For example, Zaidan et al. [35] incorporate SAN technique (segment-adjacent nested technique) under BP ANN and grouping histogram technique under Bayesian method to detect skin regions. SAN is a procedure in which RGB string value (ranging from 0 to 255255255) of individual R, G and B channels of a pixel is generated. By using 3\*3 sliding windows, a nested vector of these strings are constructed which will be used in neural network. Grouping histogram is algorithm of building skin probability map based on Bayesian rule. Naji et al. [71] builds an explicit classifier in HSV color space for 4 different skin ethnicities (4 layers) working in parallel. After primitive segmentation, a rule based region growing algorithm is performed in which the output of the first layer is used as a seed and then the final mask in other layers is constructed by means of neighboring skin pixels iteratively.

Ng et al. [243] proposed a method which combines both texture and color information. Firstly, by using a 16 kernel GMM classifier, the image is segmented into skin and non-skin pixels and then, the image is segmented into sub-images with the number of skin pixels as the criterion. For each sub-image, 2-D Daubechies Wavelet is calculated. Then, by using Shannon Entropy, Wavelet Energy-vector is obtained that presents texture feature of these pixels. Energy vectors are segmented by k-means clustering algorithm and then by using a skin texture-cluster elimination procedure, skin regions are discriminated. Finally, combined with that of the GMM classifier, final mask is specified. Erdem et al. [16] claimed that blend of explicitly defined boundary model on RGB and Bayesian classifier helps declination of false positives of VJ face detector. Jiang et al. [244] employed not only color information by using Bayesian LUT method; but they have also used a texture filter which is constructed based on features extracted from Gabor wavelet transform. Texture filter will further filter non-skin pixels; and it may also filter some skin pixel. They used a marker driven watershed transform to compensate the loss. Wang et al. [245] also integrated the color and texture information to

boost the performance of skin detection. They used GLCM texture feature which indicates the probability of the simultaneous occurrence of the joint distribution of two gray pixels whose distance is  $d = \Delta x^2 + \Delta y^2$  where  $\Delta x$  and  $\Delta y$  are coordinate differences of pixels. GLCM provides useful information of direction, adjacency, spacing and variation extent of the image. The features they extracted from the GLCM are contrast, angular second moment, entropy, correlation and homogeneity.

Pai et al. [246] presented honeycomb model for skin segmentation. First, possible skin colors are estimated from the pixels of database and the honeycomb structure is built in HSV color space according to the training samples. This model is a clustering model based on the belief that different skin pixels fall within different cells (clusters). Then, the personal skin is captured in one of the honeycomb cells. The structure of this system is depicted in Fig. 12. The initial seed of skin pixels is produced using simple thresholding in  $YCbCr$  and then possible skin pixels are classified in a HSV based honeycomb model. Therefore, pixels will be grouped into  $R_1, R_2, \dots, R_n$  by honeycomb model. Each cell of the honeycomb is defined using an empirically obtained skin cluster in HS color space.

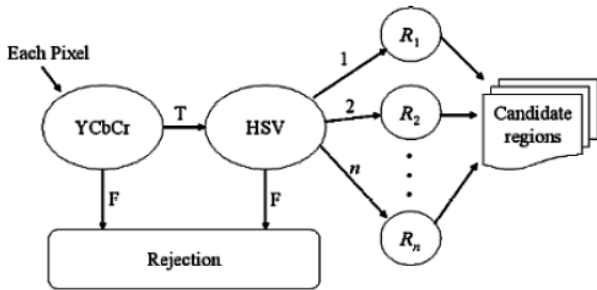


Fig.12. The Method Proposed in [246]

Hai-bo [247] combined explicitly defined boundary method with single Gaussian in HS color space to detect skin pixels. Jmal et al. [248] combined the result of applying boundary rules on RGB space and Bhattacharyya distance between the histogram of these pixels and offline trained histograms.

Conci et al. [250] employed spectral variation coefficient (SVC) texture tool in order to distinguish skin regions. A relatively computationally expensive algorithm is used to calculate the SVC. After obtaining the texture, the k-means clustering algorithm is performed. The centroid vector of each patch is compared with each sample of the training skin centroids. Through this, skin regions are separated, and by applying a region growing algorithm, the final skin-segmented image is obtained. In this procedure, the algorithm cannot discriminate non-smooth skin regions, and also, it will have difficulties in dealing with non-skin smooth surfaces that result in high false positive rate. Kawulok et al. [29] presented a method based on texture based discriminative skin-presence features (DSPF). It extracts textural features from skin probability map rather than luminance channel. DSPF is derived from discriminative textual features (DTFS) claimed to be more effective and faster. First, the skin

probability is transformed using DSPF space and then the spatial analysis based on distance transform combined domain (DTCD) of hue, luminance, and skin probability is applied. The final mask is constructed based on DSPF projection and reference pixel locations in skin probability map using LDA (linear discriminant analysis).

Taqa et al. [251] proposed a filter which is constructed out of three detectors; pre-defined rules of skin color tones, texture features, and a combination of both color and texture features which are estimated by using statistical measures such as range, standard deviation, and entropy. An MLP is then used to learn features and classify any given input. Medeiros et al. [252] employed an offline trained GMM in rgb color space, and then a texture based dictionary is constructed where a stochastic region merging strategy is subsequently performed to segment the image texture regions. Each segment is classified based on the skin color and skin texture models. Zafarifar et al. [253-255] proposed a real-time hardware implementation of a technique which combines histogram-based skin detection, and a color-constrained texture feature suitable for skin detection. The schematic of their system is depicted in Fig. 13. The algorithm first detects skin-colored areas in a color detection procedure and then removes areas that are too textured to be skin. The skin color map is computed using a trained 3D histogram in YUV space. The "constrain map" defines potential skin areas, in which the texture should be further analyzed. Thus, a skin feature called clipped texture is computed based on applying 1D Laplacian operators and the result is post-processed to remove high-frequency variations, and subsequently combined with the color map, it suppresses textured skin-colored areas.

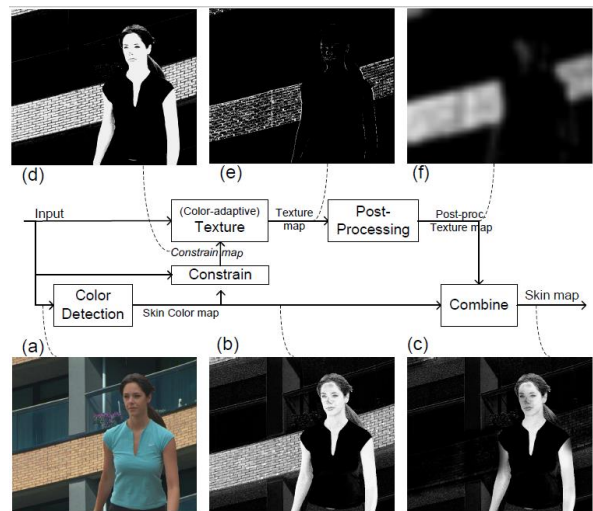


Fig.13. Combination of Color and Texture Features Proposed in [255]

## 1. Discussion

In this section, a literature review of recent publications and methods in skin detection has been provided. Methods were classified into 8 major categories where each group utilized certain features to achieve acceptable rates. Statistical methods use distribution of skin and non-skin

pixels to estimate the probability of an observed pixel to be skin. Different structures of neural networks, the explicitly defined boundary models in disparate color models, are also presented to perform the task of skin segmentation, each with its own pros and cons. Spatial and diffusion models and multispectral methods are also used in skin detection in recent years and results are encouraging. Dynamic models though seems to be computationally expensive, online reconfiguration of models have been reported to be effective and impressive. Methods developed based on SVM classifiers and combination of different features (color, texture,...) and various methods were also discussed in detail.

## V. PERFORMANCE COMPARISON

As it was mentioned before, in order to perform a fair empirical evaluation of skin segmentation techniques, it is vital to use a standard and representative training and test set. Different methods have presented their evaluation results based on different sets and for those using the same test-set, even different photos may have been used. Using a single standard dataset as discussed in section.I may aid the future studies for a more accurate, fair and through collation. Two standard datasets which have been utilized more than others are Compaq and ECU which are considered in this section to evaluate the performance of different systems. Table. IV presents experimental results based on Compaq dataset used in several previously published works.

Table 4. Experimental Results Based on Compaq Dataset

Dataset	Author	Method	Comment	Color-space	Det.rate	Er.rate
Compaq	Platzer et al. [263]	Explicitly	-	RGB-HSV	82.3	11.4
	Brand et al. [51]	Explicitly	-	YIQ	94.7	30.2
	Jones et al. [45]	Bayesian LUT	256 bins	RGB	90	14.2
	Brand et al. [51]	Bayesian LUT	256 bins	RGB	93.4	19.8
	Lee et al. [105]	SGM	6 kernels	CbCr	90	33.3
	Jones et al. [45]	GMM	16 kernels	RGB	90	15.5
	Lee et al. [105]	GMM	-	CbCr	90	37.1
	Doukim et al. [183]	MLP	-	CbCr	83.9	14.9
	Sebe et al. [177]	Bayesian Network	SSS	RGB	98.32	10
	Lee et al. [105]	Elliptical	-	CbCr	90	25
	Jedynak et al. [169]	MaxEnt	TFOM	RGB	72	5
	Brown et al. [186]	SOM	-	TSL	78	32

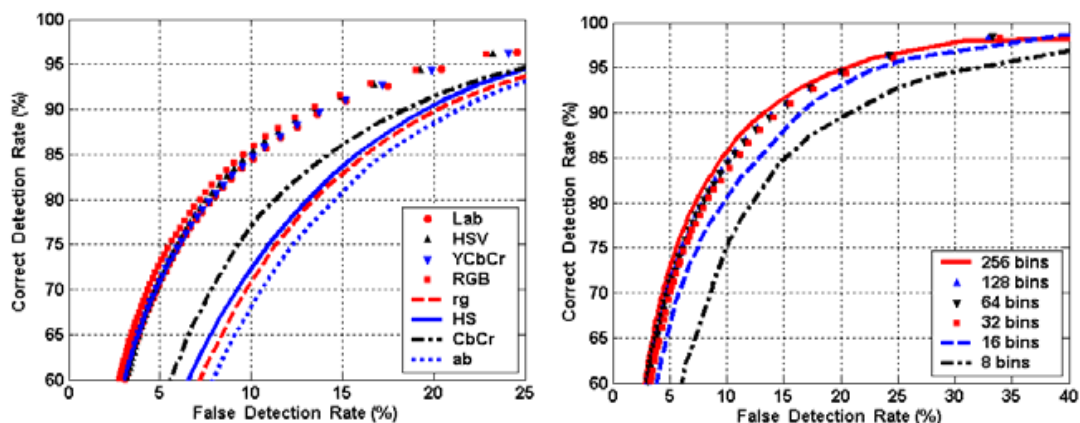


Fig.14. Effect of Number of Bins and Choice of Color Space on the Performance of Bayesian Classifier using ECU Dataset [36]

As Compaq is no more available for public usage, more recently developed methods in literature are not assessed based on this database. Among those presented their results based on this dataset, explicitly defined methods yielded to high detection rate and high false positive rate mainly due to their disability in facing with the overlap between skin and non-skin pixels. In addition, the skin clusters of different ethnicities are not entirely analogous which worsens the conditions. From color space point of view, it seems that for those models in which the skin

cluster is more compact, the explicit rules are simpler to design; however, accuracy is not different since color spaces and consequent rules are convertible. For those applications, which simplicity, area (hardware), cost of computation, and speed are very important and the precision is not the first priority, explicitly defined methods seems to be promising. The performance of Bayesian classifiers has been also remarkable but with relatively high false detection rate. In Fig. 14, the ROC of Bayesian classifier under ECU set for different color

spaces and disparate number of bins is depicted. As it can be seen, the performance of Bayesian LUT in ECU is not significantly different for 3D color spaces and elimination of luminance channel decreased the accuracy substantially. Also the more storage is utilized for construction of table, the better result is achieved. Bayesian models in compare with SGM, GMM, MLP and SOM seems to be better. This has been also certified using ECU database. Phung et al. [41] and Kawulok et al. [159] have compared the performance of a couple of skin detection methods as shown in Fig. 15. As the figure shows, the Bayesian classifier outperforms other systems with a relatively high difference. Comparing to SOM, the performance of LUT models were superior where it may have been as a result of smaller size of training set, or number of neurons used in SOM. A modified SOMN [58] gives a true positive rate of 90.01% with 14.21% false positive, which is equivalent to the detection rate of the histogram model in [45].

Gaussian models were developed to simulate the PDF of skin distribution and they have been quite successful. In compare with Bayesian models, they may be subjected to more false alarm rate, but depend on the application and implementation platform, they seem to be remarkable choice. As it was stated before, GMM models are slightly better than SGM ones in high true positive rates regions of ROC and the number of kernels in GMM has not been effective significantly. Developing MLP models are extremely dependant on different structural factors, training set, etc and this is obvious when comparing ROC curves in Fig. 15. In one of the systems, their performance is negligibly poorer than Bayesian LUT where in the others, the difference is completely clear.

Due to the field evaluation nature of multispectral methods, it is not currently possible to compare the performance of these techniques either among themselves or with other systems. However, the high accuracy of such systems in most of normal situations is not questionable. The algorithms which are based on spatial and dynamic models have been reported to be extremely encouraging with significant higher detection rates in compare with previously discussed methods. The most important point in such systems is their utilization of additional information in pixel processing as well as online information. Often, these methods are more accurate than statistical methods; however, due to the absence of a unified database to compare the performance of all systems; it is not fair to represent a comparison table.

It is really hard to derive a strict and fair conclusion from above discussion. Maybe before, Bayesian classifier with maximum number of bins and large training set accompanying with Bayesian network were the best classifiers in terms of accuracy. From speed, computation and implementation cost, however, there is a definite trade-off between methods. With developing new methods and techniques in recent years, former systems are set aside, when precision is the first priority. But, these methods are much slower than most of traditional methods which makes them indign for real-time applications. However, this is very hopeful to see the tangible progression of the performance of skin detection methods.

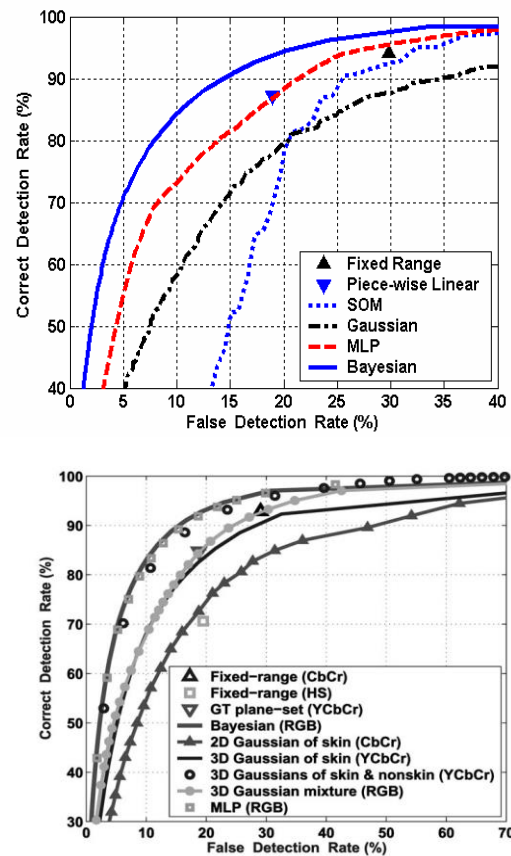


Fig.15. ROC Curve of Different Skin Detection Methods in ECU Dataset [195, 99]

## VI. APPLICATIONS

Skin segmentation technology is useful and sometimes substantial in wide range of biometric systems infolding face detection/tracking/recognition [36,179,234,145,248], pedestrian detection and tracking [131,256,257], gesture segmentation/recognition [6,29,31], content based image retrieval (CBIR) [5,21,72,121], biomedical imaging [125,258], surveillance systems [247], gaming interfaces [259], access control [95], video conferencing [113], human computer interaction (HCI) technology [243,229], detection of anchors in TV news videos for the sake of automatic annotation [72,253], robotics [8], content aware video compression [260], image color balancing [125,229], steganography [27,125], skin color reproduction [139], video phone or sign language recognition [9,109,243], and anti-spoofing [208,210].

The fact that color is one of the most principal features in human face analysis algorithms (face recognition, tracking and detection) has been yielded to development of numerous systems that leverage skin segmentation either as a preprocessor, main processor or post processor. In general, assaying face is an intensive task by itself. Thus, it would be essential to use a preprocessor to reduce the size of images (regions of interests) and make the real-time operation feasible. In face recognition, the skin color has potential to be utilized as a discriminating factor to



identify a person among many others. In face detection and tracking however, the skin color can be a very important cue in order to localize the position of the human face in static and video images. Garcia et al. [127] employed color clustering and filtering using  $YCbCr$  and HSV colors in an explicit model to provide quantized skin color regions where then a merging algorithm is utilized to provide a set of potential face areas. Subsequently, several constraints are applied on the shape and size of faces and then a texture based analysis is performed on each face area candidate to detect human faces. Hsu et al. [261] have also embedded skin segmentation in a feature-based face detector. Here, after performing a lighting compensation algorithm, variance based segmentation is performed prior to finding face features (eyes, lips, etc) for further validation of face regions. Baltzakis et al. [262] also proposed visual tracking of hand and facial features in a multi step dynamic procedure. In their system, human skin is segmented using a technique in which global and local information of images are combined to dynamically reconfigure the Bayesian histogram model. Then segmented images are subjected to a multistep tracking algorithm.

The development of digital systems in modern world was not an exception to the world of image processing. The current huge amount of online images and videos lead to construction of massive digital libraries to be used in professional systems for archival and retrieval tasks. CBIR or content based image retrieval is an application of computer vision algorithms in the problem of disquisition of digital images in particular libraries and content recurses to any specific feature such as texture, color, or even high level information such a shape or object. Skin detection is thus counted as a cornerstone algorithm in such systems as complexion could be counted as a vital feature to analyze content of an image based on that. In addition, content-based coding systems are very efficacious and important in video conferencing and video over Internet applications. Conceptually, these systems are designed based on the efficient use of the limited spectrum band for data transmission. Human is the main object of interest in conferencing videos. Therefore, such systems could leverage skin segmentation so that more transmitted data could be exploited to encode human body parts rather than less necessary objects (background). Though the background may not seem as good as skin regions, the regions of interests will be depicted with much more precision. In such applications, the accuracy of the skin detector may not be critical, but, speed is in the first priority. Another application of skin detection in video conferencing systems is the automatic camera guiding in such a way that the focus be always on speaker. In this case, based on the skin, motion, voice or mixture of different feature's information, the camera is automatically adjusted.

Hand gestures are critical communication tool for human computer interaction and their integration with vocal language and facial expressions makes it more interactive and beneficial. The design of algorithms that allow computers to recognize speech, face and hand

gestures, and even emotional state are some of the current challenges that researchers are facing with the ultimate goal of having natural human-machine communications. Gesture segmentation and recognition has been recently utilized in a wide variety of applications, though, shadows, illumination, quality of videos, etc could significantly affect the performance of such systems. Several methods with their own particular pros and cons have been proposed for this task. Building a complexion model is one of the most efficient ones as it has potential of high speed processing as well as robustness against many challenges involved in these systems. One important challenge in HCI systems is the response time of the system in which no appreciable delay between the users and the computer should be sensed [259]. To this end, the role of skin detection which is the first step of most gesture recognition systems will be critical. Here, not only the accuracy is very important in all levels to produce a reliable and robust platform, but it is also substantial for the system to operate in real-time. Mobile robots are also getting a way into many real-applications. It was not until recently that USA army has announced utilization of new generation of robots which are capable of carrying facilities for soldiers by following them step by step. The dynamic nature of the environment and the need to interact with users set limitations that are challenging in robot perception [8]. Robots are capable of acquiring information about the environment throughout variety of sensors embedded in their structures. A reliable embedded skin detection system in conjunction with a multispectral camera in the body of the robot could be exploited to solve such problems to some extent.

Another application of several human skin detection algorithms is in the design of antispooofing systems for personal identification. Physical biometric spoofing attacks are associated with cases where an adversary tries to generate fake samples which are often undertaken by collecting latent fingerprints by force, compromising the personal security. A very common way to confront this issue is aliveness detection [210]. In contactless hand biometrics, aliveness information is available in the outer skin layer (epidermis) without touching it and detection could be performed by reflectance spectroscopy i.e. analyzing the different wavelength of the light reflected by the illuminated skin. The method which is developed in [210] exploits both visual, near infrared and short wave infrared bands skin detection for aliveness detection.

## VII. CONCLUSION

In this paper, a comprehensive survey on human skin detection is presented. The paper covers most topics involved in this subject including challenges, standard metrics and databases, color spaces, variety of techniques and applications. First, factors which could significantly degrade the performance of detectors were illustrated. Only few methods have been developed with a decent consideration and clear solution on an specific challenge. Using standard dataset and metrics introduced in this

paper could be a solution for non-feasibility of comparison tasks in future studies. Performance of color spaces are associated with several factors including the method used, which means generally speaking that one particular color space outperforms others is wrong. In order to compare the performance of one classifier by using different color models, all other factors which may influence the performance should be considered. Effort to find the best color space for the task of skin segmentation has been redirected into adapting optimal skin detection models. Illumination compensation algorithms as performance comparisons show that the effect of segmentation technique is not comparable with the efficacy of color model. The pros and cons of explicitly defined method were illustrated showing that whenever the low cost of computation or speed are bottleneck, these methods can be utilized. Several statistical skin detection techniques have been elucidated, where all categorized into two minor classes of non-parametric and parametric approaches. For non-parametric methods, size of training set has a direct effect on the performance of detectors. A database of several times more than  $256^3$  is required to train an LUT with a finest resolution. These methods are

unable to generalize and interpolate the training data. In addition, the large storage requirements make them unfavorable choice particularly in embedded micro-system platforms. However, in compare with parametric models, histogram models are independent of the shape of cluster, and they are trained simply and fast. In evaluation also, they can be utilized in several clock cycles per pixel needed for accessing memory. In parametric methods, the simulation of real distribution is needed to be approved using goodness of fit measures, but it has been missed from former studies. In addition, the choice of color space is also determinative as they are strongly dependant on the shape of skin cluster. Training complexity varies in different parametric methods, but in overall, both the training time and its complication are not copacetic in compare with histogram models. ANN systems are also as effective as non-parametric methods, but they do not exhibit significant performance improvement. Unlike above methods, spatial based techniques and online adaptation solutions are promising mainly due to the fact that in both, the strategy is toward confronting challenges rather a just-solution.

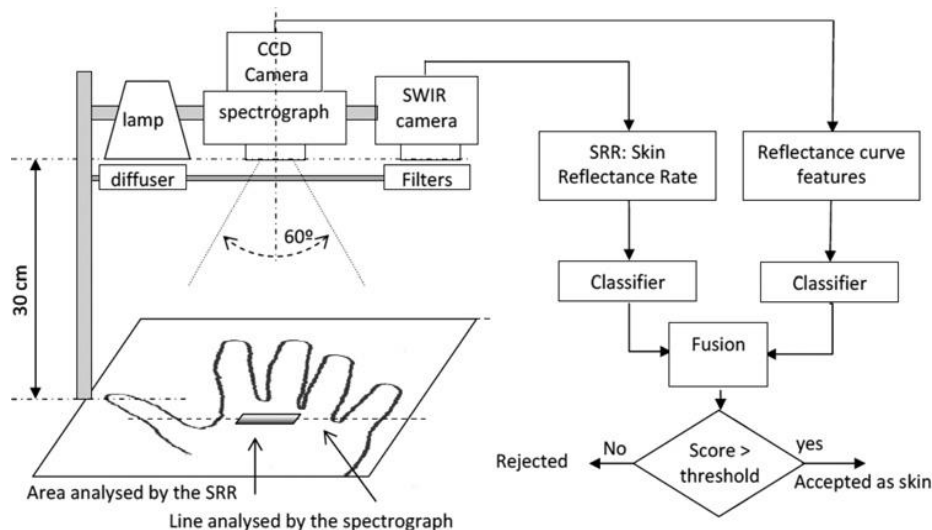


Fig.16. Antispoofing Mechanism Designed in [210]

#### REFERENCES

- [1] M.R. Mahmoodi, and S. M. Sayedi. "Leveraging spatial analysis on homogenous regions of color images for skin classification." In Computer and Knowledge Engineering (ICCKE), 2014 4th International eConference on, pp. 209-214. IEEE, 2014.
- [2] M. S. Devi, and P. R. Bajaj. "Driver fatigue detection based on eye tracking," In Emerging Trends in Engineering and Technology, ICETET'08. First International Conference on, pp. 649-652. IEEE, 2008.
- [3] Wang, Qiong, Jingyu Yang, Mingwu Ren, and Yujie Zheng. "Driver fatigue detection: a survey." In Intelligent Control and Automation, 2006. WCICA 2006. The Sixth World Congress on, vol. 2, pp. 8587-8591. IEEE, 2006.
- [4] W. Ma, and H. Zhang. "Content-based image indexing and retrieval," Handbook of multimedia computing, pp.227-254, 1999.
- [5] A. Albiol, L. Torres, C. A. Bouman, and E. J. Delp. "A simple and efficient face detection algorithm for video database applications," In Image Processing, 2000 International Conference on, vol. 2, pp. 239-242. IEEE, 2000.
- [6] Javed, Ali. "Face recognition based on principal component analysis." International Journal of Image, Graphics and Signal Processing 5.2 (2013): 38.
- [7] Singh, Sanjay, Anil Kumar Saini, and Ravi Saini. "Real-time FPGA based implementation of color image edge detection." International Journal of Image, Graphics and Signal Processing 4.12 (2012): 19.
- [8] P. Vadakkepat, P. L. Liyanage C. De Silva, L. Jing, and L. Ling. "Multimodal approach to human-face detection and tracking," Industrial Electronics, IEEE Transactions on 55, no. 3, pp. 1385-1393, 2008.
- [9] D. Dahmani, and S. Larabi. "User-Independent System for Sign Language Finger Spelling Recognition," Journal of Visual Communication and Image Representation, 2014.

- [10] E. Holden, G. Lee, and R. Owens, "Australian sign language recognition," *Machine Vision and Applications* 16, no. 5, pp. 312-320, 2005.
- [11] H. Kang, C.W. Lee, and K. Jung, "Recognition-based gesture spotting in video games," *Pattern Recognition Letters* 25, no. 15, pp. 1701-1714, 2004.
- [12] Chai, Douglas, Son Lam Phung, and Abdesselam Bouzerdoum. "Skin color detection for face localization in human-machine communications." In *Signal Processing and its Applications, Sixth International, Symposium on*. 2001, vol. 1, pp. 343-346. IEEE, 2001.
- [13] A. Jaimes, and N. Sebe, "Multimodal human-computer interaction: A survey," *Computer vision and image understanding* 108, no. 1, pp. 116-134, 2007.
- [14] Kuo, Yung-Ming, Jiann-Shu Lee, and Pau-Choo Chung. "The nude image identification with adaptive skin chromatic distribution matching scheme." In *Computer Engineering and Technology (ICCET), 2010 2nd International Conference on*, vol. 7, pp. V7-117. IEEE, 2010.
- [15] Khan, Rehanullah, Allan Hanbury, Julian Stöttinger, and Abdul Bais. "Color based skin classification." *Pattern Recognition Letters* 33, no. 2 (2012): 157-163.
- [16] Erdem, C. E., Sezer Ulukaya, Ali Karaali, and A. Tanju Erdem. "Combining Haar feature and skin color based classifiers for face detection." In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pp. 1497-1500. IEEE, 2011.
- [17] Kakumanu, Praveen, Sokratis Makrogiannis, and Nikolaos Bourbakis. "A survey of skin-color modeling and detection methods." *Pattern recognition* 40, no. 3 (2007): 1106-1122.
- [18] Liu, Qiong, and Guang-zheng Peng. "A robust skin color based face detection algorithm." In *Informatics in Control, Automation and Robotics (CAR), 2010 2nd International Asia Conference on*, vol. 2, pp. 525-528. IEEE, 2010.
- [19] Vrânceanu, Ruxandra, Razvan Condorovici, Carmen Patrascu, Foti Coleca, and Laura Florea. "Robust detection and tracking of salient face features in color video frames." In *Signals, Circuits and Systems (ISSCS), 2011 10th International Symposium on*, pp. 1-4. IEEE, 2011.
- [20] Phung, Son Lam, Douglas Chai, and Abdesselam Bouzerdoum. "Skin colour based face detection." In *Intelligent Information Systems Conference, The Seventh Australian and New Zealand 2001*, pp. 171-176. IEEE, 2001.
- [21] Ooi, M. P. "Hardware implementation for face detection on Xilinx Virtex-II FPGA using the reversible component transformation colour space." In *Electronic Design, Test and Applications, 2006. DELTA 2006. Third IEEE International Workshop on*, pp. 6-pp. IEEE, 2006.
- [22] Kumar, Anil. "An empirical study of selection of the appropriate color space for skin detection: A case of face detection in color images." In *Issues and Challenges in Intelligent Computing Techniques (ICICT), 2014 International Conference on*, pp. 725-730. IEEE, 2014.
- [23] Thakur, Sayantan, Sayantanu Paul, Ankur Mondal, Swagatam Das, and Ajith Abraham. "Face detection using skin tone segmentation." In *Information and Communication Technologies (WICT), 2011 World Congress on*, pp. 53-60. IEEE, 2011.
- [24] Yang, Ming-Hsuan, David Kriegman, and Narendra Ahuja. "Detecting faces in images: A survey." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 24, no. 1 (2002): 34-58.
- [25] Zarit, Benjamin D., Boaz J. Super, and Francis KH Quek. "Comparison of five color models in skin pixel classification." In *Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, 1999. Proceedings. International Workshop on*, pp. 58-63. IEEE, 1999.
- [26] Gundimada, Satyanadh, Li Tao, and Vijayan Asari. "Face detection technique based on intensity and skin color distribution." In *Image Processing, 2004. ICIP'04. 2004 International Conference on*, vol. 2, pp. 1413-1416. IEEE, 2004.
- [27] Ibrahim, Nada B., Mazen M. Selim, and Hala H. Zayed. "A dynamic skin detector based on face skin tone color." In *Informatics and Systems (INFOS), 2012 8th International Conference on*, pp. MM-1. IEEE, 2012.
- [28] Chaudhary, Ankit, and Ankur Gupta. "Automated switching system for skin pixel segmentation in varied lighting." In *Mechatronics and Machine Vision in Practice (M2VIP), 2012 19th International Conference*, pp. 26-31. IEEE, 2012.
- [29] Kawulok, Michal, Jolanta Kawulok, and Jakub Nalepa. "Spatial-based skin detection using discriminative skin-presence features." *Pattern Recognition Letters* (2013).
- [30] Abbadi, Nidhal K. El, Nazar Dahir, and Zaid Abd Alkareem. "Skin texture recognition using neural networks." *arXiv preprint arXiv:1311.6049* (2013).
- [31] Tan, Wenjun, Gaoyang Dai, Han Su, and Ziyi Feng. "Gesture segmentation based on YCbCr color space ellipse fitting skin color modeling." In *Control and Decision Conference (CCDC), 2012 24th Chinese*, pp. 1905-1908. IEEE, 2012.
- [32] Pamornak, Burawich, Somchai Limsiroratana, and Mitchai Chongcheawchamnan. "Color correction scheme for different illumination and camera device conditions." In *TENCON Spring Conference, 2013 IEEE*, pp. 430-434. IEEE, 2013.
- [33] Sagheer, Alaa, and Saleh Aly. "An Effective Face Detection Algorithm Based on Skin Color Information." In *Signal Image Technology and Internet Based Systems (SITIS), 2012 Eighth International Conference on*, pp. 90-96. IEEE, 2012.
- [34] Qiang-rong, Jiang, and Li Hua-lan. "Robust human face detection in complicated color images." In *Information Management and Engineering (ICIME), 2010 The 2nd IEEE International Conference on*, pp. 218-221. IEEE, 2010.
- [35] Zaidan, A. A., N. N. Ahmad, Hazerul Abdul Karim, Moussa Larbani, B. B. Zaidan, and Aduwati Sali. "On the multi-agent learning neural and Bayesian methods in skin detector and pornography classifier: An automated anti-pornography system." *Neurocomputing* 131 (2014): 397-418.
- [36] Phung, Son Lam, Abdesselam Bouzerdoum, and Douglas Chai. "Skin segmentation using color and edge information." In *Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on*, vol. 1, pp. 525-528. IEEE, 2003.
- [37] DU, Cui-huan, Hong ZHU, Li-ming LUO, Jie LIU, and Xiang-yang HUANG. "Face detection in video based on AdaBoost algorithm and skin model." *The Journal of China Universities of Posts and Telecommunications* 20 (2013): 6-24.
- [38] Vezhnevets, V, Vassili Sazonov, and Alla Andreeva. "A survey on pixel-based skin color detection techniques." In *Proc. Graphicon*, vol. 3, pp. 85-92. 2003.
- [39] Phung, Son Lam, Abdesselam Bouzerdoum, and Douglas Chai. "A novel skin color model in ycbcr color space and its application to human face detection." In *Image*

- Processing. 2002. Proceedings. 2002 International Conference on, vol. 1, pp. I-289. IEEE, 2002.
- [40] Shin, Min C., Kyong I. Chang, and Leonid V. Tsap. "Does colorspace transformation make any difference on skin detection?." In Applications of Computer Vision, 2002.(WACV 2002). Proceedings. Sixth IEEE Workshop on, pp. 275-279. IEEE, 2002.
- [41] Phung, Son Lam, Abdesselam Bouzerdoum, and D. Chai Sr. "Skin segmentation using color pixel classification: analysis and comparison." Pattern Analysis and Machine Intelligence, IEEE Transactions on 27, no. 1 (2005): 148-154.
- [42] Ibrahim, Nada B., Mazen M. Selim, and Hala H. Zayed. "A dynamic skin detector based on face skin tone color." In Informatics and Systems (INFOS), 2012 8th International Conference on, pp. MM-1. IEEE, 2012.
- [43] Wang, YuanHui, and LiQian Xia. "Skin color and feature-based face detection in complicated backgrounds." In Image Analysis and Signal Processing (IASP), 2011 International Conference on, pp. 78-83. IEEE, 2011.
- [44] Vadakkepat, Prahlad, Peter Lim, Liyanage C. De Silva, Liu Jing, and Li Li Ling. "Multimodal approach to human-face detection and tracking." Industrial Electronics, IEEE Transactions on 55, no. 3 (2008): 1385-1393.
- [45] Jones, Michael J., and James M. Rehg. "Statistical color models with application to skin detection." International Journal of Computer Vision 46, no. 1 (2002): 81-96.
- [46] Zou, Xuan, Josef Kittler, and Kieron Messer. "Illumination invariant face recognition: A survey." In Biometrics: Theory, Applications, and Systems, 2007. BTAS 2007. First IEEE International Conference on, pp. 1-8. IEEE, 2007.
- [47] Barnard, Kobus. "Modeling scene illumination colour for computer vision and image reproduction: A survey of computational approaches." Computing Science at Simon Fraser University 39 (1998).
- [48] Sridharan, Mohan, and Peter Stone. "Color learning and illumination invariance on mobile robots: A survey." Robotics and Autonomous Systems 57, no. 6 (2009): 629-644.
- [49] LIANG, Lin, Wei-ping HE, Lei LEI, Wei ZHANG, and Hong-xiao WANG. "Survey on enhancement methods for non-uniform illumination image [J]." Application Research of Computers 5 (2010): 008.
- [50] Huang, Xiaowei, Zhaohui Jiang, Lili Lu, Chunjie Tan, and Jun Jiao. "The study of illumination compensation correction algorithm." In Electronics, Communications and Control (ICECC), 2011 International Conference on, pp. 2967-2970. IEEE, 2011.
- [51] Brand, Jason, and John S. Mason. "A comparative assessment of three approaches to pixel-level human skin-detection." In Pattern Recognition, 2000. Proceedings. 15th International Conference on, vol. 1, pp. 1056-1059. IEEE, 2000.
- [52] Mitra, S., "A probabilistic method of skin detection," Image Information Processing (ICIIP), 2013 IEEE Second International Conference on , vol., no., pp.73,77, 9-11 Dec. 2013, doi: 10.1109/ICIIP.2013.6707558
- [53] Conci, Aura, E. Nunes, Juan José Pantrigo, and Ángel Sánchez. "Comparing Color and Texture-Based Algorithms for Human Skin Detection." In ICEIS (5), pp. 166-173. 2008.
- [54] Osman, Ghazali, and Muhammad Suzuri Hitam. "Skin colour classification using linear discriminant analysis and colour mapping co-occurrence matrix." In Computer Applications Technology (ICCAT), 2013 International Conference on, pp. 1-5. IEEE, 2013.
- [55] J. Ruiz-del-Solar and R. Verschae, "SKINDIFF - Robust and fast skin segmentation," Department of Electrical Engineering, Universidad de Chile 2006.
- [56] Schmutz, Stephen J., Sriram Jayaram, Min C. Shin, and Leonid V. Tsap. "Objective evaluation of approaches of skin detection using ROC analysis." Computer Vision and Image Understanding 108, no. 1 (2007): 41-51.
- [57] Huang, Lei, Tian Xia, Yongdong Zhang, and Shouxun Lin. "Human skin detection in images by MSER analysis." In Image Processing (ICIP), 2011 18th IEEE International Conference on, pp. 1257-1260. IEEE, 2011.
- [58] Chang, Lin, Leng Jun-min, and Yu Chong-xiu. "Skin Detection Using a Modified Self-Organizing Mixture Network." In Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on, pp. 1-6. IEEE, 2013.
- [59] <http://sun.aei.polsl.pl/~mkawulok/gestures/>
- [60] Rajen Bhatt, Abhinav Dhall, 'Skin Segmentation Dataset', UCI Machine Learning Repository
- [61] Casati, Joao Paulo Brognoni, Diego Rafael Moraes, and Evandro Luis Linhari Rodrigues. "SFA: A human skin image database based on FERET and AR facial images." In IX workshop de Visao Computational, Rio de Janeiro. 2013.
- [62] Phillips, P. Jonathon, Hyeonjoon Moon, Syed A. Rizvi, and Patrick J. Rauss. "The FERET evaluation methodology for face-recognition algorithms." Pattern Analysis and Machine Intelligence, IEEE Transactions on 22, no. 10 (2000): 1090-1104.
- [63] Martinez, Aleix M. "The AR face database." *CVC Technical Report 24* (1998).
- [64] J. Ruiz-del-Solar and R. Verschae, "SKINDIFF - Robust and fast skin segmentation," Department of Electrical Engineering, Universidad de Chile 2006.
- [65] Montenegro, J., W. Gomez, and P. Sanchez-Orellana. "A comparative study of color spaces in skin-based face segmentation." In Electrical Engineering, Computing Science and Automatic Control (CCE), 2013 10th International Conference on, pp. 313-317. IEEE, 2013.
- [66] Zhu, Qiang, Ching-Tung Wu, Kwang-Ting Cheng, and Yi-Leh Wu. "An adaptive skin model and its application to objectionable image filtering." In Proceedings of the 12th annual ACM international conference on Multimedia, pp. 56-63. ACM, 2004.
- [67] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, vol. 1, pp. I-511. IEEE, 2001.
- [68] Razmjoooy, Navid, B. Somayeh Mousavi, and Fazlollah Soleymani. "A hybrid neural network Imperialist Competitive Algorithm for skin color segmentation." Mathematical and Computer Modelling 57, no. 3 (2013): 848-856.
- [69] Ding, Xintao, Yonglong Luo, Liping Sun, and Fulong Chen. "Color balloon snakes for face segmentation." Optik-International Journal for Light and Electron Optics 125, no. 11 (2014): 2538-2542.
- [70] Sanmiguel, Juan C., and Sergio Suja. "Skin detection by dual maximization of detectors agreement for video monitoring." Pattern Recognition Letters 34, no. 16 (2013): 2102-2109.
- [71] Naji, Sinan A., Roziati Zainuddin, and Hamid A. Jalab. "Skin segmentation based on multi pixel color clustering models." Digital Signal Processing 22, no. 6 (2012): 933-940.
- [72] Tan, Wei Ren, Chee Seng Chan, Pratheepan Yogarajah,

- and Joan Condell. "A fusion approach for efficient human skin detection." *Industrial Informatics, IEEE Transactions on* 8, no. 1 (2012): 138-147.
- [73] Di Martino, Matías, Guzmán Hernández, Marcelo Fiori, and Alicia Fernández. "A new framework for optimal classifier design." *Pattern Recognition* 46, no. 8 (2013): 2249-2255.
- [74] Powers, David Martin. "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." 2011.
- [75] Garcia, Vicente, J. Salvador Sánchez, and Ramón A. Mollineda. "On the suitability of numerical performance measures for class imbalance problems." *International Conference in Pattern Recognition Applications and Methods*, pp. 310-313. 2012.
- [76] Ketenci, Seniha, and Beste Gencturk. "Performance analysis in common color spaces of 2D Gaussian Color Model for skin segmentation." In *EUROCON, 2013 IEEE*, pp. 1653-1657. IEEE, 2013.
- [77] Baldi, Pierre, Søren Brunak, Yves Chauvin, Claus AF Andersen, and Henrik Nielsen. "Assessing the accuracy of prediction algorithms for classification: an overview." *Bioinformatics* 16, no. 5 (2000): 412-424.
- [78] Kurmi, Uma Shankar, Hari Shanker Srivastava, Dheeraj Agrawal, R. K. Baghel, and India Bhopal. "Performance Evaluation of RGB Skin Color Segmentation Based Face Detection Technique." (2014).
- [79] Chenaoua, Kamal, and Ahmed Bouridane. "Skin detection using a Markov random field and a new color space." In *Image Processing, 2006 IEEE International Conference on*, pp. 2673-2676. IEEE, 2006.
- [80] Hazar, Mliki, Hammami Mohamed, and B. Hanene. "Real time face detection based on motion and skin color information." In *Parallel and Distributed Processing with Applications (ISPA), 2012 IEEE 10th International Symposium on*, pp. 799-806. IEEE, 2012.
- [81] Prasertsakul, Pawin, Toshiaki Kondo, Teera Phatrapornant, and Tsuyoshi Isshiki. "A Robust Hand Segmentation Method Based on Color and Background Subtraction."
- [82] Aibinu, Abiodun Musa, Amir Akramin Shafie, and Momoh Jimoh Emiyoka Salami. "Performance Analysis of ANN based YCbCr Skin Detection Algorithm." *Procedia Engineering* 41 (2012): 1183-1189.
- [83] Kang, Henry R. *Color Technology for electronic imaging devices*. Vol. 28. SPIE press, 1997.
- [84] Kuehni, Rolf G. *Color space and its divisions: color order from antiquity to the present*. John Wiley & Sons, 2003.
- [85] Mustafah, Yasir M., Abbas Bigdeli, Amelia W. Azman, and Brian C. Lovell. "Face detection system design for real time high resolution smart camera." In *Distributed Smart Cameras, 2009. ICDSC 2009. Third ACM/IEEE International Conference on*, pp. 1-6. IEEE, 2009.
- [86] NACER, Nadia, Bouraoui MAHMOUD, and Mohamed Hedi BEDOUI. "FPGA ARCHITECTURE FOR FACIAL-FEATURES AND COMPONENTS EXTRACTION." *International Journal of Computer Science* (2013).
- [87] Nacer, Nadia, Bouraoui Mahmoud, and Mohamed Hedi Bedoui. "FPGA Architecture for facial-feature and components extraction." *International Journal of Computer Science*, 2013.
- [88] Bian, Jing, and Wei Du. "Research of face detection system based on skin-tone feature." In *Mechatronic Science, Electric Engineering and Computer (MEC), 2011 International Conference on*, pp. 2407-2410. IEEE, 2011.
- [89] Subban, Ravi, and Richa Mishra. "Rule-based face detection in color images using normalized RGB color space—A comparative study." In *Computational Intelligence & Computing Research (ICCIC), 2012 IEEE International Conference on*, pp. 1-5. IEEE, 2012.
- [90] Taylor, Michael J., and Tim Morris. "Adaptive skin segmentation via feature-based face detection." In *SPIE Photonics Europe*, pp. 91390P-91390P. International Society for Optics and Photonics, 2014.
- [91] Nidhu, R., and Manu G. Thomas. "Real time segmentation algorithm for complex outdoor conditions."
- [92] Ji, Zhichao, and Huabiao Qin. "A vehicle surveillance system for face detection." In *Vehicular Electronics and Safety, 2007. ICVES. IEEE International Conference on*, pp. 1-4. IEEE, 2007.
- [93] Li, Gang, Yinping Xu, and Jiaying Wang. "An improved adaboost face detection algorithm based on optimizing skin color model." In *Natural Computation (ICNC), 2010 Sixth International Conference on*, vol. 4, pp. 2013-2015. IEEE, 2010.
- [94] Wu, Yan-Wen, and Xue-Yi Ai. "Face detection in color images using AdaBoost algorithm based on skin color information." In *Knowledge Discovery and Data Mining, 2008. WKDD 2008. First International Workshop on*, pp. 339-342. IEEE, 2008.
- [95] Mohamed, Aamer SS, Ying Weng, Stan S. Ipson, and Jianmin Jiang. "Face detection based on skin color in image by neural networks." In *Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on*, pp. 779-783. IEEE, 2007.
- [96] Vishwakarma, Anish Kumar, Agya Mishra, Kumar Gaurav, and Abhishek Katariya. "Illumination reduction for low contrast color image enhancement with homomorphic filtering technique." In *Communication Systems and Network Technologies (CSNT), 2012 International Conference on*, pp. 171-173. IEEE, 2012.
- [97] Ginjac, Dominique, Fan Yang, and Michel Paindavoine. "Design, Implementation and Evaluation of Hardware Vision Systems dedicated to Real-Time Face Recognition." *Face Recognition (2007)*: 123-148.
- [98] Gasparini, Francesca, and Raimondo Schettini. "Skin segmentation using multiple thresholding." In *Electronic Imaging 2006*, pp. 60610F-60610F. International Society for Optics and Photonics, 2006.
- [99] Baskan, Selin, M. Mete Bulut, and Volkan Atalay. "Projection based method for segmentation of human face and its evaluation." *Pattern Recognition Letters* 23, no. 14 (2002): 1623-1629.
- [100] Zahir, Nur Baiti, Rosdiyana Samad, and Mahfuzah Mustafa. "Initial experimental results of real-time variant pose face detection and tracking system." In *Signal and Image Processing Applications (ICSIPA), 2013 IEEE International Conference on*, pp. 264-268. IEEE, 2013.
- [101] Paschalakis, Stavros, and Miroslaw Bober. "A low cost FPGA system for high speed face detection and tracking." In *Field-Programmable Technology (FPT), 2003. Proceedings. 2003 IEEE International Conference on*, pp. 214-221. IEEE, 2003.
- [102] Prabhu, K. Edison, and A. Arul Kumar. "Efficient Human Skin Detection Using 2D Histogram and Gaussian Approach."
- [103] Fleck, Margaret M., David A. Forsyth, and Chris Bregler. "Finding naked people." In *Computer Vision—ECCV'96*, pp. 593-602. Springer Berlin Heidelberg, 1996.
- [104] Tao, Luo. "An FPGA-based Parallel Architecture for Face Detection using Mixed Color Models." *arXiv preprint arXiv:1405.7032* (2014).
- [105] Lee, Jae-Young, and Suk I. Yoo. "An elliptical boundary

- model for skin color detection." In *Proc. of the 2002 International Conference on Imaging Science, Systems, and Technology*. 2002.
- [106] Subban, Ravi, and Richa Mishra. "Human Skin Segmentation in Color Images Using Gaussian Color Model." In *Recent Advances in Intelligent Informatics*, pp. 13-21. Springer International Publishing, 2014.
- [107] Lindner, Albrecht, and Stefan Winkler. "What impacts skin color in digital photos?." In *IS&T/SPIE Electronic Imaging*, pp. 901505-901505. International Society for Optics and Photonics, 2014.
- [108] Terrillon, J-C., Mahdad N. Shirazi, Hideo Fukamachi, and Shigeru Akamatsu. "Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images." In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pp. 54-61. IEEE, 2000.
- [109] Wang, Dong, Jinchang Ren, Jianmin Jiang, and Stan S. Ipson. "Skin detection from different color spaces for model-based face detection." In *Advanced Intelligent Computing Theories and Applications. With Aspects of Contemporary Intelligent Computing Techniques*, pp. 487-494. Springer Berlin Heidelberg, 2004.
- [110] Littmann, Enno, and Helge Ritter. "Adaptive color segmentation-a comparison of neural and statistical methods." *Neural Networks, IEEE Transactions on* 8, no. 1 (1997): 175-185.
- [111] Jayaram, Sriram, Stephen Schmutz, Min C. Shin, and Leonid V. Tsap. "Effect of colorspace transformation, the illuminance component, and color modeling on skin detection." In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 2, pp. II-813. IEEE, 2004.
- [112] Petrisor, Daniel, Cristian Fosalau, Manuel Avila, and Felix Mariut. "Algorithm for face and eye detection using colour segmentation and invariant features." In *Telecommunications and Signal Processing (TSP), 2011 34th International Conference on*, pp. 564-569. IEEE, 2011.
- [113] Liu, Zaiying, Jie Sha, and Ping Yang. "Multi-face Detection Based on Improved Gaussian Distribution." In *Intelligent Human-Machine Systems and Cybernetics (IHMSC), 2013 5th International Conference on*, vol. 1, pp. 54-57. IEEE, 2013.
- [114] Chaves-González, Jose M., Miguel A. Vega-Rodríguez, Juan A. Gómez-Pulido, and Juan M. Sánchez-Pérez. "Detecting skin in face recognition systems: A colour spaces study." *Digital Signal Processing* 20, no. 3 (2010): 806-823.
- [115] Nalepa, Jakub, Tomasz Grzejszczak, and Michal Kawulok. "Wrist localization in color images for hand gesture recognition." In *Man-Machine Interactions 3*, pp. 79-86. Springer International Publishing, 2014.
- [116] Al-mohair et. al. "Impact of Color Space on Human Skin Color Detection Using an Intelligent System."
- [117] Dahm, Ingo, Sebastian Deutsch, Matthias Hebbel, and André Osterhues. "Robust color classification for robot soccer." In *7th International Workshop on RoboCup*. 2003.
- [118] Wang, Nai-Jian, Sheng-Chieh Chang, and Pei-Jung Chou. "A real-time multi-face detection system implemented on FPGA." In *Intelligent Signal Processing and Communications Systems (ISPACS), 2012 International Symposium on*, pp. 333-337. IEEE, 2012.
- [119] Powar, Varsha, Amruta Kulkarni, Renuka Lokare, and Aishwarya Lonkar. "Skin detection for forensic investigation." In *Computer Communication and Informatics (ICCCI), 2013 International Conference on*, pp. 1-4. IEEE, 2013.
- [120] Kovac, Jure, Peter Peer, and Franc Solina. *Human skin color clustering for face detection*. Vol. 2. IEEE, 2003.
- [121] Zhang, Qieshi, Sei-ichiro Kamata, and Jun Zhang. "Face detection and tracking in color images using color centroids segmentation." In *Robotics and Biomimetics, 2008. ROBIO 2008. IEEE International Conference on*, pp. 1008-1013. IEEE, 2009.
- [122] Yutong, Zheng, Yang Guosheng, and Wu Licheng. "Fast Face Detection in Field Programmable Gate Array." In *Digital Manufacturing and Automation (ICDMA), 2010 International Conference on*, vol. 1, pp. 719-723. IEEE, 2010.
- [123] Angheliescu, Petre, Ionut Serbanescu, and Silviu Ionita. "Surveillance system using IP camera and face-detection algorithm." In *Electronics, Computers and Artificial Intelligence (ECAI), 2013 International Conference on*, pp. 1-6. IEEE, 2013.
- [124] Hu, Kai-Ti, Yu-Ting Pai, Shanq-Jang Ruan, and Edwin Naroska. "A hardware-efficient color segmentation algorithm for face detection." In *Circuits and Systems (APCCAS), 2010 IEEE Asia Pacific Conference on*, pp. 688-691. IEEE, 2010.
- [125] Chen, Yen-Hsiang, Kai-Ti Hu, and Shanq-Jang Ruan. "Statistical skin color detection method without color transformation for real-time surveillance systems." *Engineering Applications of Artificial Intelligence* 25, no. 7 (2012): 1331-1337.
- [126] Huang, Deng-Yuan, Ta-Wei Lin, Chun-Ying Ho, and Wu-Chih Hu. "Face detection based on feature analysis and edge detection against skin color-like backgrounds." In *Genetic and Evolutionary Computing (ICGEC), 2010 Fourth International Conference on*, pp. 687-690. IEEE, 2010.
- [127] C. Gracia, and G. Tziritas, "Face detection using quantized skin color regions merging and wavelet packet analysis." *IEEE Transactions on Multimedia*, Vol. 1, No. 3, 1999, pp. 264-277.
- [128] Shemshaki, M., and R. Amjadifard. "Lip Segmentation Using Geometrical Model of Color Distribution." In *Machine Vision and Image Processing (MVIP), 2011 7th Iranian*, pp. 1-5. IEEE, 2011.
- [129] Li, Wei, Qinghua Yang, and Xianbo He. "Face detection algorithm based on double ellipse skin model." In *Software Engineering and Service Science (ICSESS), 2011 IEEE 2nd International Conference on*, pp. 335-339. IEEE, 2011.
- [130] Liu, Guangdong, and Zhongke Shi. "Embedded implementation of real-time skin detection system." In *Transportation, Mechanical, and Electrical Engineering (TMEE), 2011 International Conference on*, pp. 2463-2466. IEEE, 2011.
- [131] Johan, Nurul Fatiha, Yasir Mohd Mustafah, Md Rashid, and Nahrul Khair Alang. "Human Body Parts Detection Using YCbCr Color Space." *Applied Mechanics and Materials* 393 (2013): 556-560.
- [132] Wang, Chunyang, and Bo Yuan. "Robust Fingertip Tracking with Improved Kalman Filter."
- [133] Tsekeridou, Sofia, and Ioannis Pitas. "Facial feature extraction in frontal views using biometric analogies." In *Proceedings of the IX European Signal Processing Conference*, vol. 1, pp. 315-318. 1998.
- [134] Chai, Douglas, and Abdesselam Bouzerdoum. "A Bayesian approach to skin color classification in YCbCr color space." In *TENCON 2000. Proceedings*, vol. 2, pp.

- 421-424. IEEE, 2000.
- [135] Hsieh, Ing-Sheen, Kuo-Chin Fan, and Chiunhsun Lin. "A statistic approach to the detection of human faces in color nature scene." *Pattern Recognition* 35, no. 7 (2002): 1583-1596.
- [136] Fooprateepsiri, Rerkchai, and Werasak Kurutach. "A general framework for face reconstruction using single still image based on 2D-to-3D transformation kernel." *Forensic science international* 236 (2014): 117-126.
- [137] Mariappan, Muralindran, Manimehala Nadarajan, Rosalyn R. Porle, Vigneswaran Ramu, and Brendan Khoo Teng Thiam. "A LabVIEW Design for Frontal and Non-Frontal Human Face Detection System in Complex Background." *Applied Mechanics and Materials* 490 (2014): 1259-1266.
- [138] Duangphasuk, Pruegsa, and Werasak Kurutach. "Tattoo skin detection and segmentation using image negative method." In *Communications and Information Technologies (ISCIT), 2013 13th International Symposium on*, pp. 354-359. IEEE, 2013.
- [139] Kim, Dae-Chul, Wang-Jun Kyung, Ho-Gun Ha, and Yeong-Ho Ha. "Selective skin tone reproduction using preferred skin colors." In *Consumer Electronics (ISCE), 2012 IEEE 16th International Symposium on*, pp. 1-4. IEEE, 2012.
- [140] Deshmukh, C. N., and Ms SP Wankhade. "An efficient algorithm for face detection using color segmentation and energy thresholding." *International Journal* 1, no. 5 (2013).
- [141] Shelke, Mrs Kavita. "Efficient Face Segmentation for Recognition in Group Photographs."
- [142] Karishma, S. N., and V. Lathasree. "Fusion of Skin Color Detection and Background Subtraction for Hand Gesture Segmentation." In *International Journal of Engineering Research and Technology*, vol. 3, no. 2 (February-2014). ESRSA Publications, 2014.
- [143] Xiang, Fan Hai, and Shahrel Azmin Suandi. "Fusion of Multi Color Space for Human Skin Region Segmentation." *International Journal of Information and Electronics Engineering* 3, no. 2 (2013).
- [144] Tang, San. "Human Face Detection Method Based on Skin Color Model." *Advanced Materials Research* 706 (2013): 1877-1881.
- [145] Nanni, Loris, Alessandra Lumini, Fabio Dominio, and Pietro Zanutigh. "Effective and precise face detection based on color and depth data." *Applied Computing and Informatics* (2014).
- [146] Yoo, Tae-Woong, and Il-Seok Oh. "A fast algorithm for tracking human faces based on chromatic histograms." *Pattern Recognition Letters* 20, no. 10 (1999): 967-978.
- [147] Ibraheem, Noor A., and Rafiqul Z. Khan. "Novel Segmentation Algorithm based on Mixture of Multiple Histograms."
- [148] Chen, Qian, Haiyuan Wu, and Masahiko Yachida. "Face detection by fuzzy pattern matching." In *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pp. 591-596. IEEE, 1995.
- [149] Choi, Byeongcheol, Seungwan Han, Byungho Chung, and Jaecheol Ryou. "Human body parts candidate segmentation using laws texture energy measures with skin color." *Advanced Communication Technology (ICACT)(2011)*: 13-16.
- [150] Ban, Yuseok, Sang-Ki Kim, Sooyeon Kim, Kar-Ann Toh, and Sangyoun Lee. "Face detection based on skin color likelihood." *Pattern Recognition* 47, no. 4 (2014): 1573-1585.
- [151] Yang, Jie, Weier Lu, and Alex Waibel. *Skin-color modeling and adaptation*. Springer Berlin Heidelberg, 1997.
- [152] Yang, Ming-Hsuan, and Narendra Ahuja. "Detecting human faces in color images." In *Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on*, vol. 1, pp. 127-130. IEEE, 1998.
- [153] Menser, Bernd, and Mathias Wien. "Segmentation and tracking of facial regions in color image sequences." In *Visual Communications and Image Processing 2000*, pp. 731-740. International Society for Optics and Photonics, 2000.
- [154] Mohamed, Aamer SS, Ying Weng, Stan S. Ipson, and Jianmin Jiang. "Face detection based on skin color in image by neural networks." In *Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on*, pp. 779-783. IEEE, 2007.
- [155] Colmenarez, Antonio J., and Thomas S. Huang. "Face detection with information-based maximum discrimination." In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pp. 782-787. IEEE, 1997.
- [156] Zhu, Shiping, and Nan Zhang. "Face detection based on skin color model and geometry features." In *Industrial Control and Electronics Engineering (ICICEE), 2012 International Conference on*, pp. 991-994. IEEE, 2012.
- [157] Chandrappa, D. N., M. Ravishankar, and D. R. RameshBabu. "Face detection in color images using skin color model algorithm based on skin color information." In *Electronics Computer Technology (ICECT), 2011 3rd International Conference on*, vol. 1, pp. 254-258. IEEE, 2011.
- [158] Wang, Wei, and Jing Pan. "Hand segmentation using skin color and background information." In *Machine Learning and Cybernetics (ICMLC), 2012 International Conference on*, vol. 4, pp. 1487-1492. IEEE, 2012.
- [159] Yang, Ming-Hsuan, and Narendra Ahuja. "Gaussian mixture model for human skin color and its application in image and video databases." In *Proc. SPIE: Storage and Retrieval for Image and Video Databases VII*, vol. 3656, pp. 458-466. 1999.
- [160] Redner, Richard A., and Homer F. Walker. "Mixture densities, maximum likelihood and the EM algorithm." *SIAM review* 26, no. 2 (1984): 195-239.
- [161] Caetano, Tiberio S., Silvia D. Olabarriaga, and Dante AC Barone. "Do mixture models in chromaticity space improve skin detection?." *Pattern Recognition* 36, no. 12 (2003): 3019-3021.
- [162] Hossain, Md Foisal, Mousa Shamsi, Mohammad Reza Alsharif, Reza A. Zoroofi, and Katsumi Yamashita. "Automatic facial skin detection using Gaussian mixture model under varying illumination." *Int J Innovative Comput Inf Control* 8, no. 2 (2012): 1135-1144.
- [163] Zou, Li, and Sei-ichiro Kamata. "Face detection in color images based on skin color models." In *TENCON 2010-2010 IEEE Region 10 Conference*, pp. 681-686. IEEE, 2010.
- [164] Aliradi, Rachid, Naima Bouzera, Abdelkrim Meziane, and Abdelkader Belkhir. "Detection of Facial Components Based on SVM Classification and Invariant Feature." In *Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2013 IEEE/WIC/ACM International Joint Conferences on*, vol. 3, pp. 30-36. IEEE, 2013.
- [165] Rao, K. Srinivasa, B. N. Jagadesh, and Ch Satyanarayana. "Skin Colour Segmentation using Fintte Bivariate Pearsonian Type-IV a Mixture Model." *Computer Engineering and Intelligent Systems* 3, no. 5 (2012): 45-55.
- [166] Jagadesh, B. N., K. Srinivasa Rao, and Ch Satyanarayana. "A Robust Skin Colour Segmentation Using Bivariate

- Pearson Type  $II_{\alpha\alpha}$  (Bivariate Beta) Mixture Model." *International Journal of Image, Graphics and Signal Processing (IJIGSP)* 4, no. 11 (2012): 1.
- [167] Xu, Dan, Yen-Lun Chen, Xinyu Wu, Yongsheng Ou, and Yangsheng Xu. "Integrated approach of skin-color detection and depth information for hand and face localization." In *Robotics and Biomimetics (ROBIO)*, 2011 IEEE International Conference on, pp. 952-956. IEEE, 2011.
- [168] Yu, Dong, Geoffrey Hinton, Nelson Morgan, Jen-Tzung Chien, and Shigeki Sagayama. "Introduction to the special section on deep learning for speech and language processing." *Audio, Speech, and Language Processing, IEEE Transactions on* 20, no. 1 (2012): 4-6.
- [169] Jedynak, Bruno, Huicheng Zheng, and Mohamed Daoudi. "Statistical models for skin detection." In *Computer Vision and Pattern Recognition Workshop*, 2003. CVPRW'03. Conference on, vol. 8, pp. 92-92. IEEE, 2003.
- [170] Jedynak, Bruno, Huicheng Zheng, and Mohamed Daoudi. "Statistical models for skin detection." In *Computer Vision and Pattern Recognition Workshop*, 2003. CVPRW'03. Conference on, vol. 8, pp. 92-92. IEEE, 2003.
- [171] Terrillon, J-C., Martin David, and Shigeru Akamatsu. "Automatic detection of human faces in natural scene images by use of a skin color model and of invariant moments." In *Automatic Face and Gesture Recognition*, 1998. Proceedings. Third IEEE International Conference on, pp. 112-117. IEEE, 1998.
- [172] Jedynak, Bruno, Huicheng Zheng, and Mohamed Daoudi. "Skin detection using pairwise models." *Image and Vision Computing* 23, no. 13 (2005): 1122-1130.
- [173] Nadimi, Nahid, Zohreh Azimifar, and Ehsan Ahmadi. "Skin detection using a statistical color spaces fusion model." In *Machine Vision and Image Processing (MVIP)*, 2013 8th Iranian Conference on, pp. 270-274. IEEE, 2013.
- [174] K. Chenaoua, A. Bouridane, "A Markov random field based skin detection approach", 14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP
- [175] Ahmadi, Ehsan, Fahimeh Garmsirian, and Zohreh Azimifar. "A discriminative fusion framework for skin detection." In *Artificial Intelligence and Signal Processing (AISP)*, 2012 16th CSI International Symposium on, pp. 542-545. IEEE, 2012.
- [176] Heckerman, David. "A tutorial on learning with Bayesian networks." In *Innovations in Bayesian Networks*, pp. 33-82. Springer Berlin Heidelberg, 2008.
- [177] Sebe, Nicu, Ira Cohen, Thomas S. Huang, and Theo Gevers. "Skin detection: A bayesian network approach." In *Pattern Recognition*, 2004. ICPR 2004. Proceedings of the 17th International Conference on, vol. 2, pp. 903-906. IEEE, 2004.
- [178] Jan A. Snyman (2005). *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*. Springer Publishing. ISBN 0-387-24348-8
- [179] Wu, Qing Xiang, Rongtai Cai, Lijuan Fan, Chengmei Ruan, and Gang Leng. "Skin detection using color processing mechanism inspired by the visual system." In *Image Processing (IPR 2012)*, IET Conference on, pp. 1-5. IET, 2012.
- [180] Seow, Ming-Jung, Deepthi Valaparla, and Vijayan K. Asari. "Neural network based skin color model for face detection." In *Applied Imagery Pattern Recognition Workshop*, 2003. Proceedings. 32nd, pp. 141-145. IEEE, 2003.
- [181] Chen, Li, Jiliu Zhou, Zhiming Liu, Wei Chen, and Guoqing Xiong. "A skin detector based on neural network." In *Communications, Circuits and Systems and West Sino Expositions, IEEE 2002 International Conference on*, vol. 1, pp. 615-619. IEEE, 2002.
- [182] Bhojar, K. K., and O. G. Kakde. "Skin color detection model using neural networks and its performance evaluation." *Journal of Computer Science* 6, no. 9 (2010): 963.
- [183] Medeiros, R. S., Jacob Scharcanski, and Alexander Wong. "Multi-scale stochastic color texture models for skin region segmentation and gesture detection." In *Multimedia and Expo Workshops (ICMEW)*, 2013 IEEE International Conference on, pp. 1-4. IEEE, 2013.
- [184] Duan, Lijuan, Zhiqiang Lin, Jun Miao, and Yuanhua Qiao. "A method of human skin region detection based on PCNN." In *Advances in Neural Networks-ISNN 2009*, pp. 486-493. Springer Berlin Heidelberg, 2009.
- [185] Kohonen, Teuvo (1982). "Self-Organized Formation of Topologically Correct Feature Maps". *Biological Cybernetics* 43 (1): 59-69.
- [186] Brown, David A., Ian Craw, and Julian Lewthwaite. "A SOM Based Approach to Skin Detection with Application in Real Time Systems." In *BMVC*, vol. 1, pp. 491-500. 2001.
- [187] H. Yin, N. Allinson. Self-organising mixture networks for probability density estimation. *IEEE Transactions on Neural Networks*. 2001, 12(2): 405-411.
- [188] Ruiz-del-Solar, Javier, and Rodrigo Verschae. "Skin detection using neighborhood information." In *Automatic Face and Gesture Recognition*, 2004. Proceedings. Sixth IEEE International Conference on, pp. 463-468. IEEE, 2004.
- [189] M. R. Mahmoodi, and S. M. Sayedi, "Boosting Performance of Face Detection Using an Efficient Skin Detection Algorithm," In *Information Technology and Electrical Engineering (ICITEE)*, 2014 International Conference on, pp.1-6, IEEE, 7-8 Oct, 2014
- [190] M. R. Mahmoodi, and S. M. Sayedi, "A Face Detection Method Based on Kernel Probability Map," *Computers & Electrical Engineering*, Elsevier, 2014.
- [191] M. R. Mahmoodi, and S. M. Sayedi, "A Face Detector Based on Color and Texture," In *Information Technology and Electrical Engineering (ICITEE)*, 2014 International Conference on, pp.1-6, IEEE, 7-8 Oct, 2014.
- [192] Abdullah-Al-Wadud, Mohammad, and Oksam Chae. "Skin segmentation using color distance map and water-flow property." In *Information Assurance and Security*, 2008. ISIAS'08. Fourth International Conference on, pp. 83-88. IEEE, 2008.
- [193] Kawulok, Michal. "Energy-based blob analysis for improving precision of skin segmentation." *Multimedia Tools and Applications* 49, no. 3 (2010): 463-481.
- [194] Kawulok, Michal. "Fast propagation-based skin regions segmentation in color images." In *Automatic Face and Gesture Recognition (FG)*, 2013 10th IEEE International Conference and Workshops on, pp. 1-7. IEEE, 2013.
- [195] M. R. Mahmoodi, and S. M. Sayedi, "Leveraging Spatial Analysis on Homogeneous regions of Color Images for Skin Classification," *Computer and Knowledge Engineering (ICCKE)*, 7th International Conference on, pp.1-6, IEEE, 29-30 Oct, 2014.
- [196] M. R. Mahmoodi, S. M. Sayedi, Z. Fahimi, F. Karimi, and Z. Mannani "SDD: A Skin Detection Dataset for Assessment of Human Skin Classifiers," In *Knowledge-based Engineering and Innovation (KBEI)*, 2015 International Conference on, IEEE, 5-6 Nov, 2015.
- [197] Mahmoodi, M.R.; Sayedi, S.M.; Karimi, F., "Propagation



- from conservatively selected skin pixels using a multi-step multi-feature method," in *Electrical Engineering (ICEE), 2015 23rd Iranian Conference on*, vol., no., pp.219-224, 10-14 May 2015 doi: 10.1109/IranianCEE.2015.7146213
- [198] Kong, Seong G., Jingu Heo, Besma R. Abidi, Joonki Paik, and Mongi A. Abidi. "Recent advances in visual and infrared face recognition—a review." *Computer Vision and Image Understanding* 97, no. 1 (2005): 103-135.
- [199] Beisley, Andrew P. Spectral detection of human skin in VIS-SWIR hyperspectral imagery without radiometric calibration. No. AFIT/GE/ENG/12-03. AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH GRADUATE SCHOOL OF ENGINEERING AND MANAGEMENT, 2012.
- [200] Morikawa, Shohei, Kazuhiko Yamamoto, K. A. T. O. Kunihiro, Yoshikatsu Kimura, and Kiyosumi Kidono. "Decision Method of the Material Characteristics by Using Three Wavelength Images." (2010): 362-367.
- [201] Nunez, Abel S., and Michael J. Mendenhall. "Detection of human skin in near infrared hyperspectral imagery." In *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, vol. 2, pp. II-621. IEEE, 2008.
- [202] Bashkatov, A. N., E. A. Genina, V. I. Kochubey, and V. V. Tuchin. "Optical properties of human skin, subcutaneous and mucous tissues in the wavelength range from 400 to 2000 nm." *Journal of Physics D: Applied Physics* 38, no. 15 (2005): 2543.
- [203] Kidono, Kiyosumi, Yusuke Kanzawa, Takaaki Tagawa, Yoshiko Kojima, and Takashi Naito. "Skin segmentation using a multiband camera for early pedestrian detection." In *Intelligent Vehicles Symposium (IV), 2013 IEEE*, pp. 346-351. IEEE, 2013.
- [204] Suzuki, Yasuhiro, Kazuhiko Yamamoto, Kunihiro Kato, Michinori Andoh, and Shinichi Kojima. "Skin detection by near infrared multi-band for driver support system." In *Computer Vision—ACCV 2006*, pp. 722-731. Springer Berlin Heidelberg, 2006.
- [205] Dowdall, Jonathan, Ioannis Pavlidis, and George Bebis. "Face detection in the near-IR spectrum." *Image and Vision Computing* 21, no. 7 (2003): 565-578.
- [206] Angelopoulos, E., Rana Molana, and Kostas Daniilidis. "Multispectral skin color modeling." In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 2, pp. II-635. IEEE, 2001.
- [207] Schwaneberg, Oliver, Holger Steiner, Peter Haring Bolívar, and Norbert Jung. "Design of an LED-based sensor system to distinguish human skin from workpieces in safety applications." *Applied optics* 51, no. 12 (2012): 1865-1871.
- [208] Pishva, Davar. "Spectroscopic Approach for Aliveness Detection in Biometrics Authentication." In *Security Technology, 2007 41st Annual IEEE International Carnahan Conference on*, pp. 133-137. IEEE, 2007.
- [209] Reddy, P. Venkata, Ajay Kumar, S. Rahman, and T. Mundra. "A new antispoofing approach for biometric devices." *Biomedical Circuits and Systems, IEEE Transactions on* 2, no. 4 (2008): 328-337.
- [210] Ferrer, M. A., A. Morales, C. M. Travieso, and J. B. Alonso. "Wide band spectroscopic skin detection for contactless hand biometrics." *IET computer vision* 6, no. 5 (2012): 415-424.
- [211] Hu, Kexin, Yanli Liu, Qi Dong, Hao Liu, and Guanyu Xing. "Color face image decomposition under complex lighting conditions." *The Visual Computer* 30, no. 6-8 (2014): 685-695.
- [212] Skarbek, Wladyslaw, Andreas Koschan, Technischer Bericht, and Zur Veröffentlichung. "Colour image segmentation—a survey." (1994).
- [213] Yang, Guoliang, Huan Li, Li Zhang, and Yue Cao. "Research on a skin color detection algorithm based on self-adaptive skin color model." In *Communications and Intelligence Information Security (ICCIIS), 2010 International Conference on*, pp. 266-270. IEEE, 2010.
- [214] Argyros, Antonis A., and Manolis IA Lourakis. "Real-time tracking of multiple skin-colored objects with a possibly moving camera." In *Computer Vision—ECCV 2004*, pp. 368-379. Springer Berlin Heidelberg, 2004.
- [215] Zhu, Qiang, Kwang-Ting Cheng, Ching-Tung Wu, and Yi-Leh Wu. "Adaptive learning of an accurate skin-color model." In *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, pp. 37-42. IEEE, 2004.
- [216] Oliver, Nuria, Alex P. Pentland, and Francois Berard. "Lafter: Lips and face real time tracker." In *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pp. 123-129. IEEE, 1997.
- [217] Oliver, Nuria, Alex Pentland, and François Bérard. "LAFTER: a real-time face and lips tracker with facial expression recognition." *Pattern recognition* 33, no. 8 (2000): 1369-1382.
- [218] Yang, Lei, and Alex Waibel. "A real-time face tracker." In *Applications of Computer Vision, 1996. WACV'96., Proceedings 3rd IEEE Workshop on*, pp. 142-147. IEEE, 1996.
- [219] McKenna, Stephen J., Shaogang Gong, and Yogesh Raja. "Modelling facial colour and identity with gaussian mixtures." *Pattern recognition* 31, no. 12 (1998): 1883-1892.
- [220] Bergasa, Luis Miguel, Manuel Mazo, Alfredo Gardel, M. A. Sotelo, and Luciano Boquete. "Unsupervised and adaptive Gaussian skin-color model." *Image and Vision Computing* 18, no. 12 (2000): 987-1003.
- [221] Schwerdt, Karl, and James L. Crowley. "Robust face tracking using color." In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pp. 90-95. IEEE, 2000.
- [222] Cho, Kyung-Min, Jeong-Hun Jang, and Ki-Sang Hong. "Adaptive skin-color filter." *Pattern Recognition* 34, no. 5 (2001): 1067-1073.
- [223] Sigal, Leonid, Stan Sclaroff, and Vassilis Athitsos. "Skin color-based video segmentation under time-varying illumination." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 26, no. 7 (2004): 862-877.
- [224] Cheng, Jian, Haijun Liu, Jian Wang, and Hongsheng Li. "Face Detection Based on Cost-Gentle Adaboost Algorithm." In *The Proceedings of the Second International Conference on Communications, Signal Processing, and Systems*, pp. 263-270. Springer International Publishing, 2014.
- [225] Mittal, Arpit, Andrew Zisserman, and Philip HS Torr. "Hand detection using multiple proposals." In *BMVC*, pp. 1-11. 2011.
- [226] Wimmer, Matthias, and Bernd Radig. "Adaptive skin color classifier." In *Proc. of the first ICGST International Conference on Graphics, Vision and Image Processing GVIP-05*, vol. 1, pp. 324-327. 2005.
- [227] Fritsch, Jannik, Sebastian Lang, Marcus Kleinhagenbrock, Gernot A. Fink, and Gerhard Sagerer. "Improving adaptive skin color segmentation by incorporating results from face detection." In *Robot and*

- Human Interactive Communication, 2002. Proceedings. 11th IEEE International Workshop on, pp. 337-343. IEEE, 2002.
- [228] Kawulok, Michal. "Dynamic skin detection in color images for sign language recognition." In *Image and Signal Processing*, pp. 112-119. Springer Berlin Heidelberg, 2008.
- [229] Shen, Xuanjing, Wei Wei, and Qingji Qian. "A pornographic image filtering model based on erotic part." In *Image and Signal Processing (CISP)*, 2010 3rd International Congress on, vol. 5, pp. 2473-2477. IEEE, 2010.
- [230] Liao, Wen-Hung, and Yu-Hao Chi. "Estimation of Skin Color Range Using Achromatic Features." In *Intelligent Systems Design and Applications, 2008. ISDA'08. Eighth International Conference on*, vol. 2, pp. 493-497. IEEE, 2008.
- [231] Yoo, Kook-Yeol. "Robust hand segmentation and tracking to illumination variation." In *Consumer Electronics (ICCE)*, 2014 IEEE International Conference on, pp. 286-287. IEEE, 2014.
- [232] Yogarajah, Pratheepan, Joan Condell, Kevin Curran, Paul McKevitt, and Abbas Cheddad. "A dynamic threshold approach for skin tone detection in colour images." *International Journal of Biometrics* 4, no. 1 (2012): 38-55.
- [233] Soriano, Maricor, Birgitta Martinkauppi, Sami Huovinen, and Mika Laaksonen. "Adaptive skin color modeling using the skin locus for selecting training pixels." *Pattern Recognition* 36, no. 3 (2003): 681-690.
- [234] Phung, Son Lam, Douglas Chai, and Abdesslem Bouzerdoum. "Adaptive skin segmentation in color images." In *Multimedia and Expo, 2003. ICME'03. Proceedings. 2003 International Conference on*, vol. 3, pp. III-173. IEEE, 2003.
- [235] Dadgostar, Farhad, and Abdolhossein Sarrafzadeh. "An adaptive real-time skin detector based on Hue thresholding: A comparison on two motion tracking methods." *Pattern Recognition Letters* 27, no. 12 (2006): 1342-1352.
- [236] Soriano, Maricor, Birgitta Martinkauppi, Sami Huovinen, and Mika Laaksonen. "Skin detection in video under changing illumination conditions." In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, vol. 1, pp. 839-842. IEEE, 2000.
- [237] Sun, Hung-Ming. "Skin detection for single images using dynamic skin color modeling." *Pattern recognition* 43, no. 4 (2010): 1413-1420.
- [238] Kawulok, Michal, Jolanta Kawulok, Jakub Nalepa, and Maciej Papiez. "Skin detection using spatial analysis with adaptive seed." In *ICIP*, pp. 3720-3724. 2013.
- [239] Han, Junwei, G. M. Award, Alistair Sutherland, and Hai Wu. "Automatic skin segmentation for gesture recognition combining region and support vector machine active learning." In *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on*, pp. 237-242. IEEE, 2006.
- [240] Kawulok, Michal, and Jakub Nalepa. "Support vector machines training data selection using a genetic algorithm." In *Structural, Syntactic, and Statistical Pattern Recognition*, pp. 557-565. Springer Berlin Heidelberg, 2012.
- [241] Khan, Rehanullah, Allan Hanbury, and Julian Stoeftinger. "Skin detection: A random forest approach." In *Image Processing (ICIP)*, 2010 17th IEEE International Conference on, pp. 4613-4616. IEEE, 2010.
- [242] Ho, Tin Kam. "Random decision forests." In *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on*, vol. 1, pp. 278-282. IEEE, 1995.
- [243] Ng, Pan, and Chi-Man Pun. "Skin color segmentation by Texture Feature Extraction and K-mean Clustering." In *Computational Intelligence, Communication Systems and Networks (CICSyN)*, 2011 Third International Conference on, pp. 213-218. IEEE, 2011.
- [244] Jiang, Zhiwei, Min Yao, and Wei Jiang. "Skin detection using color, texture and space information." In *Fuzzy Systems and Knowledge Discovery, 2007. FSKD 2007. Fourth International Conference on*, vol. 3, pp. 366-370. IEEE, 2007.
- [245] Wang, Xiaohua, Xi Zhang, and Jingliang Yao. "Skin color detection under complex background." In *Mechatronic Science, Electric Engineering and Computer (MEC)*, 2011 International Conference on, pp. 1985-1988. IEEE, 2011.
- [246] Pai, Yu-Ting, Li-Te Lee, Shanq-Jang Ruan, Yen-Hsiang Chen, Saraju Mohanty, and Elias Kougiannos. "Honeycomb model based skin colour detector for face detection." *International Journal of Computer Applications in Technology* 39, no. 1 (2010): 93-100.
- [247] Hai-bo, Lin. "A Kind of Human Face Region Detection and Recognition Method Based on Chrominance Information Characteristics." In *Control Engineering and Communication Technology (ICCECT)*, 2012 International Conference on, pp. 469-472. IEEE, 2012.
- [248] Jmal, Marwa, Wided Souidene Mseddi, Rabah Attia, and Anis Youssef. "Classification of Human Skin Color and its Application to Face Recognition." In *MMEDIA 2014, The Sixth International Conferences on Advances in Multimedia*, pp. 126-131. 2014.
- [249] Khan, Rehanullah, Allan Hanbury, Robert Sablatnig, Julian Stöttinger, F. Ali Khan, and F. Alam Khan. "Systematic skin segmentation: merging spatial and non-spatial data." *Multimedia tools and applications* 69, no. 3 (2014): 717-741.
- [250] Yao, Min, Kota Aoki, and Hiroshi Nagahashi. "Segmentation-based illumination normalization for face detection." In *Computational Intelligence & Applications (IWCIA)*, 2013 IEEE Sixth International Workshop on, pp. 95-100. IEEE, 2013.
- [251] Taqa, Alaa Y., and Hamid A. Jalab. "Increasing the reliability of skin detectors." *Scientific Research and Essays* 5, no. 17 (2010): 2480-2490.
- [252] Medeiros, R. S., Jacob Scharcanski, and Alexander Wong. "Multi-scale stochastic color texture models for skin region segmentation and gesture detection." In *Multimedia and Expo Workshops (ICMEW)*, 2013 IEEE International Conference on, pp. 1-4. IEEE, 2013.
- [253] Zafarifar, B., E. B. Bellers, and P. H. N. de With. "Application and evaluation of texture-adaptive skin detection in TV image enhancement." In *Consumer Electronics (ICCE)*, 2013 IEEE International Conference on, pp. 88-91. IEEE, 2013.
- [254] Zafarifar, Bahman, Tim van den Kerkhof, and Peter HN de With. "Feasibility of real-Time skin detection on TV DSP and FPGA platforms." In *Consumer Electronics (ICCE)*, 2012 IEEE International Conference on, pp. 469-470. IEEE, 2012.
- [255] Zafarifar, Bahman, Tim van den Kerkhof, and Peter HN de With. "Texture-adaptive Skin Detection for TV and its Real-time Implementation on DSP and FPGA." *Consumer Electronics, IEEE Transactions on* 58, no. 1 (2012): 161-169.
- [256] Barbu, Tudor. "Pedestrian detection and tracking using temporal differencing and HOG features." *Computers & Electrical Engineering* 40, no. 4 (2014): 1072-1079.

- [257] Kawulok, Michal. "Fast propagation-based skin regions segmentation in color images." In Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on, pp. 1-7. IEEE, 2013.
- [258] Lu, Juan, Ed Kazmierczak, Jonathan H. Manton, and Rodney Sinclair. "Automatic Segmentation of Scaling in 2-D Psoriasis Skin Images." *Medical Imaging, IEEE Transactions on* 32, no. 4 (2013): 719-730.
- [259] More, Prathmesh, Abhijeet Thenge, and Tushar Ghanti. "Neural Network based Dynamic Kinesics Recognition for Gaming Interface."
- [260] Zhu, Jian-qing, and Can-hui Cai. "Region growing based high brightness skin detection." In *Signals, Circuits and Systems (ISSCS)*, 2011 10th International Symposium on, pp. 1-4. IEEE, 2011.
- [261] Hsu, Rein-Lien, Mohamed Abdel-Mottaleb, and Anil K. Jain. "Face detection in color images." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 24, no. 5 (2002): 696-706.
- [262] Baltzakis, Haris, Maria Pateraki, and Panos Trahanias. "Visual tracking of hands, faces and facial features of multiple persons." *Machine Vision and Applications* 23, no. 6 (2012): 1141-1157.
- [263] Platzer, Christian, Martin Stuetz, and Martina Lindorfer. "Skin sheriff: a machine learning solution for detecting explicit images." In *Proceedings of the 2nd international workshop on Security and forensics in communication systems*, pp. 45-56. ACM, 2014.

### Authors' Profiles



**Mohammad Reza Mahmoodi** was born in Isfahan, Iran, in 1989. He received his B.Sc. and M.Sc. degree in Electronics from Isfahan University of Technology (IUT) in 2011. His areas of interest include analog/digital circuit design, computer vision, pattern recognition and image processing.



**Sayed Masoud Sayedi** was born in Maragheh, Iran, in 1960. He received the B.Sc. and M.Sc. degrees in electrical engineering from Isfahan University of Technology (IUT), and the Ph.D. degree in electronics from Concordia University in 1986, 1988, and 1996, respectively. From 1988 to 1992, and then since 1997, he has been with IUT, where he is currently an associate professor in the Department of Electrical and Computer Engineering. His areas of interest include VLSI fabrication processes, low power VLSI circuits, and data converters.

**How to cite this paper:** Mohammad Reza Mahmoodi, Sayed Masoud Sayedi, "A Comprehensive Survey on Human Skin Detection", *International Journal of Image, Graphics and Signal Processing(IJIGSP)*, Vol.8, No.5, pp.1-35, 2016. DOI: 10.5815/ijigsp.2016.05.01