A novel hybrid module of skin detector using grouping histogram technique for Bayesian method and segment adjacent-nested technique for neural network

A. A. Zaidan\textsuperscript{1}, H. Abdul Karim\textsuperscript{1}, N. N. Ahmad\textsuperscript{1}, Gazi Mahabubul Alam\textsuperscript{2*} and B. B. Zaidan\textsuperscript{1}

\textsuperscript{1}Faculty of Engineering Multimedia University, 63100 Cyberjaya, Selangor Darul Ehsan, Malaysia.
\textsuperscript{2}Faculty of Education, University of Malaya, Kuala Lumpur, Malaysia.

Accepted 21 October, 2010

Skin detection is a common ancient image processing applications for detecting human images. The applications include video surveillance, naked image filters within unit-spam systems and face detection. Skin color is considered as a useful and discriminating spatial feature for many skin detection related applications, but it is not robust enough to deal with complex image environments. Skin tone ranges from dark (some Africans) to light white (Caucasians and some Europeans). Other factors like light-changing conditions and the presence of objects with skin-like colors could create major difficulties in face pixel-based skin detector when color feature is used. Thus, this paper proposed a novel hybrid module using grouping histogram technique for Bayesian method and back propagation neural network with segment adjacent-nested (SAN) technique based on YCbCr and RGB color space in improving the skin detection performance. The researcher was able to increase the classification reliability in discriminating human skin color and regularizing the skin detection that is exposed to different light conditions. This novel skin detector method depends on three factors. The first part of the method involves the Bayesian part that is applied to a novel grouping histogram technique which uses 600 non-skin images in the processing and then calculates the probability density for each pixel. The second part involves applying the adjacent-nested technique in the preprocessing and calculating the probability density for each pixel in the neural part. Then a combination of the neural part and normalization technique is used to normalize the inputs and targets, so that the target falls in the interval [-1, 1] for each segment, which is created and trained with the training set of the skin and non skin segments. The third part involves a combination of the Bayesian method with the neural network segmentation methods and novel hybrid method. The study, tested on human images, has an upright frontal skin with any background. As such, the results show that the proposed system is able to achieve high detection rates of 98\% segmentation and low false positives when compared with the existing methods.

Key words: Skin detector, Bayesian method, neural network.

INTRODUCTION

Skin segmentation plays a big role in human computer interaction (HCI) based applications, for example, in analysis of gesture, tracking human face, tracing human emotion and other human related image processing applications in computer vision and multimedia such as filtering of web contents, retrieving multimedia databases, video surveillance, videophone and video conferencing applications. The main purpose of skin detection is to detect skin pixels in images and generate the area of skin region (Wadud et al., 2008). These areas of skin regions are further investigated according to the focus of the

\textsuperscript{*}Corresponding author. E-mail: gazi.alam@um.edu.my, gazimalamb@yahoo.com.
specific application. An effective recognition of the skin regions makes the skin detection process better and easier (Wadud et al., 2008). If the skin regions are detected with errors in the beginning, the later process will generate multiple errors and unsuccessful diagnosis. The detection process that does not accurately detect skin regions or generate regions that are not complete reduces the reliability of the applications. The initial step should be at most accurate and most reliable to maintain the efficiency of the systems which depend on it (Mohamed et al., 2008; Wadud et al., 2008). A skin detector is able to divide an image into two distinct classes, one representing the skin region and the other non-skin regions. Human skins do not have a specific geometric form. According to Dargham et al. (2009) and Peer et al. (2009), the only way to detect skin is by using texture and color attributes.

An earlier study by Wadud et al. (2008) on skin segmentation used grayscale images. The researchers’ approach faces inaccurate detections as well as the rejections of skin pixels since grayscale images lack the appropriate variation in color. As a result of this, follow-up researchers adopted skin color as the vital element in skin segmentation. Skin detection methodologies that are based on color information have received a lot of attention and interest since skin color provides computationally effective and robust information against rotations, scaling and partial occlusions (Kakumanu et al., 2007). Therefore, color has been recognized as an important factor in identifying skin areas if the skin color pixels can be represented, modeled, and classified accurately. Nevertheless, real world skin detection can be a hard task to accomplish as skin appearances in a particular image are influenced by various factors. Such factors include different illumination levels like indoor, outdoor, highlights, shadows and many more. These factors can change the original skin color of the human and represent the color differently in the captured image. This inaccuracy is referred to as color inconsistency problem (Kakumanu et al., 2007). The original skin color of humans can also be altered by the factors associated with the camera itself. When cameras are set with different functions, they produce different characteristics like spectral reflectance, sensor sensitivities and so forth. Skin color also differs for different people and ethnic groups (Yao and Gao, 2001). Other human characteristics such as age, sex and body parts also affect the appearance of skin color. As such, Kakumanu et al. (2007) and Yao and Gao (2001) stated that external factors such as subject appearances (makeup, hairstyle and glasses), background colors and motions may affect skin color.

Many studies have been done to minimize the disadvantage factors in normalizing color so the image is less sensitive to illumination. Some of these approaches include making the color space transformation (Cai and Goshtasby, 1999; Liu and Yang, 1994) or multiple color channels ratios (Vezhnevets et al., 2003; Soriano et al., 2000). Nevertheless, according to Mohamed et al. (2008), transformed spaces of the chrominance and information on the color are often lost in the chrominance separation process. Even after several methods are applied in the skin detection method to improve its effectiveness, there are still a lot of the main problems faced by these proposed methods as previously stated. The problem statement and the research objectives provide reasons for the study.

**Problem statement**

Automatic skin detection is often used in many security applications such as face detection, e-mail spam application, etc. The applications are classified as security applications and are very important to life (Zaidan et al., 2010, 2010a, 2010b; Raad et al., 2010; Alanazi et al., 2010; Hmood et al., 2010). A number of researchers have addressed the problems associated with skin detection (Fleck et al., 1996; Jones et al., 2002). These studies have found that color is an important cue in human skin detection. According to the studies, color offers good detection performance that allows for fast processing, and as such, provides invariance to geometric transformations and high robustness to the images. Nevertheless, the studies offer several differences in their approach in solving the problems associated with skin detection. The differences are related to the choice of colour space, for example, full colour or chrominance-based and model type, e.g. parametric vs. non-parametric; colour-based skin detection involves projecting an unknown pixel to some skin colour model and retrieving a skin probability. Since the color of the skin can also be the color of non skin, colour-based skin detection will result in some non skin pixels having a non-zero skin probability. A non-skin pixel has a lower skin probability and the probability threshold eliminates the majority. Thus, the non skin area that is not eliminated by the threshold will become fragmented and separated. The subsequent component analyses that are connected and the small region elimination in the skin map can eliminate the fragmentation. Spatial filtering method can also capture skin pixels that were originally missed or falsely eliminated from the thresholding process. Non skin material has a high skin probability threshold, and because it has identical chrominance to skin, it is difficult to eliminate it in the spatial processing method. The non skin material is identical to the skin and can be rarely separated by the available spatial processing because it is connected to the skin. Such an example is the hair. The result of detecting a non skin material as skin is called false detection. False detection can affect the overall performance of an image or video processing system.

In the face, detection and tracking can influence the
correct calculation of the window in front. In image, classification can lead to significant misclassification and enhancement can also produce erroneous results, though some authors had examined this issue directly (Fleck et al., 1996) or indirectly (Jones et al., 2002). In context or in general, increasing the detection of skin color, based on it, has not been adequately resolved. So the automatic detection of the skin is a component of many applications. As such, color-based methods have proved to be suitable for this task, but generally suffer from a type of false detection or are not a reliable detector of the skin, which negatively affects the aforementioned activity, namely the identification of regions confident hair like the human skin. This research presents a new hybrid skin detection module with high performance and reliability of the regions’ hair removal that is falsely detected and which addresses these problems (Bao et al., 2009; Zaidan et al., 2010c).

Another challenge comes from the fact that the appearance of the skin in an image depends heavily on lighting conditions, providing the geometry of illumination and color, where the image was captured, and has high sensitivity in identifying human object colors in a wide range of miniatures. This is called color constancy. Color constancy is a mystery of perception. Therefore, an important challenge in the detection of skin is to represent the color in a way that is invariant, or at least insensitive to changes in lighting. The choice of color space greatly affects the performance of any skin detector and its sensitivity to changes in lighting conditions. This paper takes the view of these factors in order to obtain robust detection pattern of the skin with high performance reliability and elimination of false detection of new hybrid form.

Research objectives

Automatic detection of skin is an essential component of imaging applications as diverse as face detection, tracking system and unite-spam. The color-based methods have proved to be suitable for this task, but generally suffer from a type of false detection or are not a reliable detector of the skin, and as such, the appearance of human skin in different lighting conditions adds to the complexity of skin detection, as discussed in the problem statement. The combination of these weaknesses increases false detection in the background if the environment is controlled. This research seeks to achieve the following objectives (Dargham et al., 2009):

1. To conduct a study in the literature on various methods of detection of the skin and determine the reliability, strength and weaknesses of each method (Zaidan et al., 2010c).
2. To design and implement a collective approach, based on the detector skin histogram clustering technique; for the Bayesian method based on the YCbCr color space and back propagation neural network with adjacent segment-nested (SAN) technology, based on the RGB color space that overcomes the disadvantageous skin current detection methods.
3. To evaluate the performance of the new detection module of the skin and compare it with existing methods (Bao et al., 2009; Zaidan et al., 2010c).

Research questions

The study proposed the implementation of a novel hybrid module using the technique of grouping histogram for the Bayesian method and back propagation neural network with the technique of adjacent segment-based nested YCbCr and RGB color space, respectively, with the aim of improving the performance of skin detection. Thus, this research attempts to highlight and answer the following questions:

1. According to Mohamed et al. (2008), Bao et al. (2009), Chai et al. (2000) and Phung et al. (2003) based on networking neural method or the Bayesian method of pre-processing, here comes the question: Is there a method of pre-processing that can be considered to prepare the data of Nural-Bayesian methods together to make the highest level of data classification become the detector of skin that can detect skin pixels with high accuracy?
2. There are many ways in which skin detector can detect the skin with a high rate, but is not reliable because the current pixel can have the same color skin-tone that is falsely identified as the skin, for example clothing skin color, background, color of skin materials, and so on (Chahir and Elmoataz, 2006; Zaidan et al., 2010c). Thus, the second research question is formulated on the basis of the fact that there is no form in which the skin detector can detect skin pixels with high reliability (Bao et al., 2009).
3. One challenge comes from the fact that the appearance of the skin in an image depends on the factors of enlightenment which provide the geometry of illumination and color, where the image was captured. The third research question states that there is no form for the detector that can detect skin pixels in the skin without the influence of lighting conditions (Bao et al., 2009).
4. Methods of neural network or Bayesian method cannot be used alone for skin detector (Zaidan et al., 2010c, 2009). The fourth research question is having a hybrid approach (composed of Bayesian neural system and method) that is commendable.

Related work

Chahir and Elmoataz (2006) stated that there are many
methods and processes of skin detection. The methods depend on the target system and the type of goals the system should achieve. Mohamed et al. (2008, 2007) proposed a method to detect skin in a soft computing, especially on the nervous system and the Bayesian method (Bao et al., 2009). Chai and Bouzerdoum (2000) address a technique for image classification using the Bayes’ decision rule for the minimum cost to classify the pixels in the color of their skin color and non-skin color. As such, color statistics are collected from a YCbCr color space. The Bayesian approach to the classification of skin color is discussed together with an overview of the YCbCr color space. Experimental results show that this approach can achieve good classification, and is robust against different skin colors (Gomez and Morales, 2002). Phung et al. (2003) presented a segmentation algorithm which combines skin color and edge information. In the proposed algorithm, regions of skin color are detected using a highly accurate model of skin color according to the Bayesian decision rule for the minimum cost and nonparametric density estimation. The regions identified are further processed using the homogeneity properties of human skin. The proposed segmentation algorithm is shown to be able to reduce false detection background caused by skin color detection, and effective separation of the skin regions from the false detection. The author also reported a complete analysis of the Bayesian model of skin color, along with some interesting results (Elgammal et al., 2009; Kakumanu et al., 2004).

Kakumanu proposed a neural network (NN) approach to estimate the skin color. The formation of NN to stabilize the color of the skin produced significantly, better results than the gray world algorithm or algorithms trained on white spot. A simple method, presented for detection of skin, indicates that the color matching NN can be used for adaptation of skin color. The document presented a strategy for image chromatic adaptation using neural networks (NN) with the application for adjustment of the detection of human skin color. The network consists of randomly chosen color images containing human subject in different lighting conditions, allowing the model to dynamically adapt to changing lighting conditions. The proposed network includes the estimated direct illumination of the image to fit the color of human skin. It is clear that the color matching NN produces better performance in order to stabilize the color of the skin. In the method, when the area of skin is used as a reference, the stabilization assigns values to the gray area of the skin. This property can be used to detect skin regions in the image (Chin, 2008; Wadud et al., 2008).

Wadud proposed an approach for detection of the skin, and as such, the author used a color map distance, which can be implemented using any explicit skin cluster classifier threshold-based in any color space. Although the algorithm operates mainly on an image to grayscale (DM), the treatment is actually done on the basis of color information. The map contains information of the scalar distance vector (color). This makes the method easy to implement. Experimental results show that the proposed approach is better than applying the traditional threshold for cluster-based skin classifier itself. He was pretty confident about its performance in different color spaces. In addition, he had not used any method in the strict threshold, making it applicable in a variety of shooting conditions. He also makes use of the information region, which makes it robust against noisy pixels and generates solid area of the skin (Mohamed et al., 2008; Boussaid et al., 2005).

Boussaid presented a skin detector that is suitable for VLSI implementation. On the basis of multilayer feed forward neural network architecture, the proposed skin detector can be processed in real time and as such, eliminate the need for on-chip memory elements. The proposed skin detector offers a good compromise between performance of skin detection and deployment complexity. The skin detector achieves a classification accuracy of 88.76%. Boussaid et al. (2005) experiments have shown that the use of MLP neural networks, proposed by Smach for face recognition, is a very promising approach for designing hardware. Robustness of the model was obtained by a back propagation learning algorithms and the activation function. The proposed approach is that no pre-treatment is necessary because the normalization is embedded directly in the weights of the network input (Chahir and Elmoataz, 2006).

Mohamed et al. (2008) suggested a new algorithm for face recognition with a compressed domain. Their study used skin color information Cb and Cr color space with back propagation classifier neural network, in the feature extraction of the vector of DCT coefficients after the segmentation (Mohamed et al., 2008, 2007). Uniform face recognition process was classified into three phases. Pre-processing of segmentation and classification using neural networks back propagation (Mohamed et al., 2008, 2007) which explores the use of color images tested dataset upright frontal face from the internet; and studies have shown high detection rate. However, the methods have been criticized for their reliability because of the inability of the detector to fully differentiate the pixels that have the same skin tone or color. Thus, clothing color, materials, color, background, etc. are the challenges present in detecting true skin. Moreover, the skins were still affected by lighting and its variables (Bao et al., 2009; Zaidan et al., 2010c).

Materials and Methods

System overview

The proposal to study the detection system of the skin using the technique of grouping histogram for the Bayesian method and back propagation neural network with the technique of adjacent
segment-based nested YCbCr and RGB color space, respectively is shown in Figure 1. First, researchers have to prepare the database of images. To do this, the researcher collected 600 images from the internet and the human skin was manually cropped. The process is to remove the non-skin pixels from the image. The researcher has used Adobe Photoshop to complete this operation. Figure 2 is an example of an image that had been cropped. The cropped image database with the original image (first cut) database is the database of images that will be used in the hybrid module mapping study of the skin, and as such, the output from the Bayesian neural training will be taken as an input for the function of the threshold's second phase.

In this work, the study uses 10,715 pixels of all skin tones and 2,285 pixels of skin pixels for training in the Neural; 34148312 pixels of all skin tones and 29713469 pixels from pixels of the skin for training in the Bayesian. In the process of testing, given the new input function in four parameters, one of these parameters is the output from neural training and the other is the original image, which is required by this system as a segment of skin pixels. The last two, which are output of the Bayesian training, are probability for skin and non skin. Based on the positions of pixels as correspondence with the original photo and results of Bayesian neural training, the study explains all the pixels on the threshold (that is, neural training <0.2 | training Bayesian |> 0.9). If the condition is satisfied, the pixel refers to the skin; otherwise the pixel refers to the skin. The study concludes that these decisions have contributed to the study of hybrids to increase the reliability of the proposed threshold detector. The skin was chosen on the basis of studies that were conducted in the construction phases of the program, and as such, the proposed method achieved high reliability (98%) detection rate of the skin.

The novel hybrid module for skin detector using the technique of histogram grouping for the Bayesian method and back propagation neural network, with the technique-nested adjacent segment, consists of four major phases, namely pre-treatment, training, testing and novel hybrid methods.
Novel pre-processing phase

We have two folders that are keeping the images. First, the contents of the folder, 600 images of the skin and the contents of the folder, 600 images in the second place with no skin have two new factors for this method of pre-processing. The first involves pre-processing for the part using the Bayesian technique of grouping histogram directly to these images: this part of the pre-processing is performed only once in the first phase of the training. We begin by converting the image into the YCbCr color space and then take only CbCr for each pixel and measure the histogram for each pixel in the image for all images and save the result in the same matrix $256 \times 256$ which is primarily two dimensions to one for Cb Cr. When the histogram is calculated for each pixel of the image (and this has been done on any images that are saved in the folder), histogram arrays size within two employees of the gradient, which is between 00 and 255 * 255 pixel image. With that image gradient, and in addition to an array of gradient matrix in the same position after finishing all the pixels of the image, it goes directly to the second image through the end of all the images to the folder.

The swimming pool which is more than one image in the same matrix, enabling the histogram is calculated for all 600 photographs in the same matrix of this new technique. When we speak of grouping the histograms, we apply this in two stages, one for a skin folder to another leather folder as shown in Figures 3 and 4.

The second involves pre-processing for the part applying neural techniques, an adjacent segment-nested. We started by converting the image in RGB color space and then R, G and B for each pixel and converted the string. Then they became string concatenation (RGB) and then we converted the new number, collected the pixels in the same position of the original image. So they became the same image but in the position of each pixel concatenation (RGB). And then we played for the number of pixels between each 000-255255255 and then applied the technique of adjacent segment-nested for each image using 3x3 sliding window. This indicated that the pre-processing for the next step in this technique and for each window is to reach the nested. We then arranged the pixels for each window in the same column from one to nine, with different objectives $[1 -1]$; in Nural, 1 indicates the skin and -1 indicates that the skin does not pass all the images you are holding in these folders. First folder of your skin shows after you have completed all pictures of this same procedure to automatically transfer the folder that contains images of the skin as shown in Figure 5. As a result, it appeared that the pixels to the skin are not in the same matrix within the target difference. At the pre-processing, it is performed several times in the first phase. In this, pre-processing is sufficient to apply the technique of adjacent-segment nested within each image to be tested in the same objective which is suitable for skin that is necessary to deduce the image as shown in Figure 6.

Training phase

This phase of training is done once in the first stage. The new proposed skin detector depends on two factors on the new phase of training, partly by applying Bayesian clustering techniques histogram for 600 images of the skin and 600 images of no skin; and then calculates the probability density histogram of the results; and also partly by Nural technical standards for normalizing the inputs and targets so that they fall outside the range $[-1 1]$. The most important things at this stage is how accurate is the data set. So the proposed method covers a wide range of skin color. Two sets of data for the training phase can be classified as non-leather and leather pictures or images.

Non-skin images database

The image database consists of the skin of the researcher who collected 600 images from the Internet. All pixels of the images will be considered in the methods of the study. Figure 7 shows an example of non-skin sample images used for training.

Skin images database

The database of photos of skin consists of 600 images that were collected from the internet. Not all the pixels of the images will be taken into account by these images. The researcher used only skin
Figure 3. Results from grouping histogram technique for skin pixels.

Figure 4. Results from grouping histogram technique for non skin pixels.
Figure 5. Results from segment adjacent-nested technique for skin and non skin pixels in the same matrix.

Figure 6. Results from segment adjacent-nested technique for image which need the system to detect the skin from it.
pixels and neglected the white pixels in the training. Figure 8 shows examples of skin sample with different body shapes with various human skin without any noise used for the training phase.

Training using neural network

The neural network was successfully used in many detection system of the skin (Zakaria et al., 2009). The neural network has been the subject of increasing interest in recent years (Messaoudi et al., 2007). The neural network is considered appropriate to deal with the nature of uncertainty in the system and human error (Ozcep et al., 2010). The system has been applied in the field of model awards watch in optical character recognition, object, self-driving robot (Celik, 2010; Gencel, 2009). The system has many advantages including the feasibility of establishing the system that can capture the complex class conditional density models face. The researchers also noted that the use of neural networks can produce high accuracy in face detection (Rowley, 1999). The system also has some disadvantages as stated by Zakaria et al. (2009). The neural network needs a longer processing time than other methods such as fuzzy system (Bahari et al., 2009; Ardila and Sandhu, 2010) or Baysian method. Artificial Neural Networks (ANN) is massively parallel systems composed of many processing elements connected by links of variable weight (Yenigün et al., 2010). The formation of back-propagation algorithm is the most commonly used methods and neural network method used in this document (Pradhan and Lee, 2009). The formation of back-propagation algorithm is trained with a number of examples of associated input and output values (Pradhan and Lee, 2009).

In the proposed method, the weakness is that the back-propagation network Nural must be optimized (number of layers, number of nodes, threshold, the algorithm training, and etc.) to get the best performance (Zakaria et al., 2009). In this study, neural network model consists of three layers (1 - inputs, 2 hidden and 3-output). Figure 9 shows the architecture of the neural network back...
Figure 8. Samples of skin pictures used for training.

Figure 9. Neural network architecture.
propagation neural network with four neurons in the hidden layer and one neuron in output layer for the task of segmentation (Bao et al., 2009).

The phases of training consist of nine major steps:

1. Loading data which are prepared from pre-processing and are then stored in the file mat.
2. Processing row 1 to 13000, columns 1 to 9 as input (1:13000, 1:9);
3. Processing line 6000, from the tenth column as targets (1:13000, 10);
4. Normalize the premnmx involving pre-processing data so that the minimum is -1 and the maximum is 1. \[ \text{premnmx} = (p, t) \text{; Syntax. Normalized inputs and targets so that they fall outside the range } [-1 1] \]
5. Using the train FCN = “trainb algorithm for training.
6. Using the minmax (PN), [4 1], {'tansig' 'PURELINE'}, 'trainlm') for training:
   a. MinMax (p) is used for large data to defend the minimum and maximum for the carrier.
   b. Trainlm creates a neural network of one hidden layer with 4 neurons and Levenberg-Marquardt back propagation c. {'tansig', 'PURELINE'}, (tansig), using a nonlinear function between the inputs and hidden layer because they have the treatment and there is no relationship between the input and hidden layer, and then using PURELINE function, the line between the hidden layer and output layer because there is no other treatment in this sector. As soon as the transfer is out, this non-linear function of the study helps reduce the time for the neural process.
   d. The process consists of four neurons in one hidden layer [4 1].
7. The number of training is 150, the time between reports on the status of the training is 100, the learning rate is 0.30, and the time is 0.6.
8. The weight of this network is initialized net = init (net).
9. For the formation of the lots, the study used the function of training (net, PN, TN) as starter for the formation.

The benefits attributable to the neural network are more stable, as shown in Figure 13, where mistakes have been uniformly decreased by the number of 150 iterations (Figure 10). The study found that the data were trained until the error is fixed (Figure 11). The performance of BPNN was also measured using regression analysis between the network outputs (Y) and targets (T). Regression analysis shows that the relationship between Y and T is near, and the correlation coefficient R is equal to 0. 90891. Figure 12 shows a good fit between the output of the network and the destination where the correlation coefficient R is equal to 0. 90891 (Bao et al, 2009).

**Figure 10.** Shown the ideal iteration time for this network.

---

**Training using Bayesian method**

This method uses a clustering technique based histogram approach to segment skin pixels. The color space is quantified in a number of containers, each corresponding to a particular set of pairs of color component value (in 2D case) or triple (if 3D) (Chai and Bouzerdoum, 2000) in which each bin stores the number of times a particular color occurs in the training skin images. The histogram counts are in the post training time -normalized, converting histogram values to discrete probability distribution (Chai and Bouzerdoum, 2000).
Figure 11. Error characteristics during training to reduce BPNN.

Figure 12. Regression analysis for reducible BPNN between two Targets: 1 for skin and -1 for non skin.
Figure 13. Stability of neural network.

\[ P_{\text{skin}}(CbCr) = \frac{\text{skin}[CbCr]}{\text{Norm}} \]  

(1)

Skin CbCr indicates the value in the histogram bin corresponding to the color vector (CbCr). The standard is normalization coefficients (sum of all values of the histogram bin). The normalized values of the search bins table are the probability that the corresponding colors match your skin. The value of \( P_{\text{skin}}(CbCr) \) calculated in equation (1) is (Bouzerdoum and Chai, 2000) actually a conditional probability \( P(CbCr | \text{skin}) \). It is a probability of observing a pixel of color (CbCr), knowing that a skin pixel is seen. A more appropriate sensing skin would be \( P(\text{skin} | \text{RGB}) \), which is calculated by the Bayes rule in equation (2) (Chai and Bouzerdoum, 2000):

\[ \frac{P(CbCr | \text{skin}) \cdot P(\text{skin})}{P(CbCr | \text{skin}) \cdot P(\text{skin}) + P(CbCr | \sim \text{skin}) \cdot P(\sim \text{skin})} \]  

(2)

\( P(CbCr | \text{skin}) \) and \( P(CbCr | \sim \text{skin}) \) are calculated directly from the skin and non-skin histograms. The prior probability \( P(\text{skin}) \) and \( P(\sim \text{skin}) \) can be easily estimated from the total number of skin and skin pixel in the training set, as is shown below in equations (3) and (4) (Chai and Bouzerdoum, 2000):

\[ P(\text{skin}) = \frac{T_s}{T_s + T_n} \]  

(3)

\[ P(\sim \text{skin}) = \frac{T_n}{T_s + T_n} \]  

(4)

Where, \( T_s \) and \( T_n \) are the numbers of skin and non-skin pixels, respectively.

**Testing phase**

**Testing using neural network**

The testing steps for neural network involves

1. Loading data which are prepared from pre-processing for the set of tests that are saved in the file mat
2. One must be the same size p from row 1 to row 13000, column 1 to 9 from the test data in September
3. s = 1-13000 lines, starting from the tenth test data set of columns.
4. Pre-processing of data using the premnmx so that the minimum is -1 and the maximum interval is a, \( \text{mina}, \text{maxa} \), minutes, \( \text{maxa} = \text{premnmx} (a', s') \) and normalize the inputs and objectives so that they fall outside the range [-1,1].
5. Starting the test with \( y_n = \text{sim} (\text{net}, a) \) in the formation
6. Converting the output \( y_n \) to the original scale including the following command: Postmnmx \( (y_n', \text{minutes}, \text{maxa}) \) (Mohamed et al 2008).

**Testing using Baysian method**

The test states that if the values of \( P(w) \) and \( P(CbCr | w) \) are known for \( W = \{\text{skin} \sim \text{nskin}\} \), \( P(w | CbCr) \) will be determined, which is already a good result that allows the use of the following rule of classification (Chai and Bouzerdoum, 2000): If \( P(\text{skin} | CbCr) > P(\sim \text{Skin} | CbCr) \), then CbCr is classified as skin pixels, otherwise its pixels are classified as non-skin. This is called the maximum, a posteriori probability rule. Using Baye’s formula, the classification rule can be transformed to: IFP \( (CbCr | \text{skin}) / P(CbCr \sim \text{skin}) \cdot P(\sim \text{skin}) / P(\text{skin}) \) and then CbCr is classified as a pixel skin, otherwise it is non-skin pixels (Chai and Bouzerdoum, 2000). We can minimize the cost of the Baye’s decision rule. This rule can be used to classify CbCr in the class of skin color (W1) or non-skin color class (W2). Let \( C_{IJ} \) be the cost to decide CbCr \( \in W_I \) when CbCr \( \in W_J \). \( C_{IJ} \) here represents the cost of correct
The classification problem is how to find the class (skin or non-skin) which gives minimal cost. Equations (7) and (8) can be rewritten as:

\[ C_1(CbCr) = C_{11} \cdot P(skin|CbCr) + C_{12} \cdot P(\neg skin|CbCr) \]  

(5)

\[ C_2(CbCr) = C_{21} \cdot P(skin|CbCr) + C_{22} \cdot P(\neg skin|CbCr) \]  

(6)

The decision rule is thus:

\[ C_1(CbCr) < C_2(CbCr) \rightarrow CbCr \in skin \]  

(w₁)

(7)

\[ C_1(CbCr) > C_2(CbCr) \rightarrow CbCr \notin skin \]  

(w₂)

(8)

The classification problem is how to find the class (skin or non-skin) which gives minimal cost. Equations (7) and (8) can be rewritten as equations (9) and (10):

\[ (C_{12} - C_{22}) \cdot P(skin|CbCr) - (C_{21} - C_{11}) \cdot P(\neg skin|CbCr) > 0 \Rightarrow CbCr \in skin \]  

(9)

\[ (C_{12} - C_{22}) \cdot P(skin|CbCr) - (C_{21} - C_{11}) \cdot P(\neg skin|CbCr) < 0 \Rightarrow CbCr \notin skin \]  

(10)

By applying the Bayesian formula to the above equations, the Bayes' decision rule for minimum cost (Bayes' classifier) are equations (11), (12) and (13) (Chai and Bouzerdoum, 2000):

\[ \frac{P(CbCr | skin)}{P(CbCr | \neg skin)} > \theta \rightarrow CbCr \in skin \]  

(w₁)

(11)

\[ \frac{P(CbCr | skin)}{P(CbCr | \neg skin)} < \theta \rightarrow CbCr \notin skin \]  

(w₂)

(12)

where,

\[ \theta = \frac{C_{12} - C_{22}}{C_{21} - C_{11}} \cdot \frac{P(\neg skin)}{P(skin)} \]  

(13)

\[ \theta = \frac{C_{12}}{C_{21}} \cdot \frac{P(\neg skin)}{P(skin)} \]  

(14)

The threshold was chosen based on multiple experiences that the researchers had previously in the construction phases of the program. If the condition is satisfied, the pixel refers to the skin and not over the vacuum in vacuum [255 255 255], who is black white; otherwise the pixel refers to the skin and also in the same empty white [R (y, x), G (y, x), B (y, x)], which is gray in color from original image. The conclusion of this matter is that the content of the single-pixel image with skin white has been the original image.

This hybrid method has been used to achieve high detection rate with low false positive and false negative. After the detection has been done, the next stage can be taken, which is the application of the unite-spam filter that depends on the skin detector for the first stage. We can extract the features from the skin which entails detecting from the image and then training again in any machine to test if the image is spam or not. This is considered as one of the ways to classify the image spam.

RESULTS AND DISCUSSION

Peer et al. (2009); Tsumura et al. (2003); Zakaria et al. (2009) repeatedly stated in their studies that the neural network and Bayesian methods are popular methods for appearance-based detection system of the skin. Therefore, in this work the study used pre-processing method (neural-Bayesian) together with adjacent segment-nested technique, as the preprocessing technique back propagation neural network method and technique histogram clustering technique of pre-processing method such as the Bayesian method proposed pre-processing to answer the first research question.

Based on the first test and to answer the second research question, the researcher realized that the increases in skin color, such as the rate of false positives and false negatives reduce the true positive. In other words, the skin detector is not reliable. To deal with the wide range of skin problems and skin color in this method, two basic things must be done. First, the researcher has collected more than 600 images with different range of color and used these images in the
training phase of part of the skin to train the system with different types of skin pixels in order to avoid duplication as possible. An example of our skin data set is shown in Figure 14.

Second, the models of the images are shown in Figure 15. The study uses different tones of gray color images. During the formation of part of the skin, that which is not in the same field of hair images for images of the skin is not listed (Figure 16).

The study focuses on the use of gray as this color works similarly with skin color and is used to train a large number of images that contain the color gray. However, the skin does not. Even when using the color gray as part of training, the study can get the form more efficiently with high reliability for skin detector. To answer the research question three, the study presented the results using the RGB, YCbCr color space. For multiple elimination of the effect of changes in illumination, the segmentation of the skin has become more robust for the lighting, when the variations in luminance of the pixel are discarded. As mentioned earlier, this proposed system consists of two parts: first, color spaces YCbCr histogram clustering technique for the Bayesian method. A skin color model is created at the level of YCbCr color
space. The reason for choosing chrominance chrominance red and blue (Cb and Cr) color space instead of YCbCr is to eliminate the effect of changes of lighting using component Y (Bao et al., 2009).

Classification using only pixel chrominance Cb and Cr (pure color) segmentation skin may become more robust to lighting variations if pixel luminance is discarded. The classification is also to limit the search and speed up the calculation to identify regions of facial skin. Secondly, the RGB color space to the propagation of neural network backup with the technique of nested-adjacent segment. RGB color space is commonly used in digital images. We will use a new technique for RGB. Take R, G and B for each pixel and convert the string, then concatenation of them becomes string (RGB). And then convert the new number and collect the pixels in the same position with the original image. Note that it becomes the same image but in the position of each pixel concatenation (RGB). After then play for the number of pixels between 000-255255255.

Using these techniques can eliminate the effect of RGB lighting as a condition of using the second part of the multi-space. Also in this section, the study presented the results of back probgation neural network or Bayesian method as the only ones to answer the fourth research question, but the researcher found that they alone are not sufficient to give results of high reliability. Note that the researcher detects the skin pixels and also notes the pixels of the skin.

So the study concludes that this method (or Bayesian methods, neural network back probgation) only in this proposed system will produce a lot of skin-like color. In other words, the skin detector is not reliable. The objective of new methods together with this system is to remedy this problem by using the parallel decision for both the same threshold condition Muti (<tranning Nural 0.2 | | bayesian tranning> 0.9), which is a response to research question four. The hybrid module for the neural network detector leather back probgation and the Bayesian method are considered an appropriate and effective way to detect skin pixels with high precision and reliability.

Our proposed system has achieved a detection rate of 98% as average for classification, as can be seen in (Figure 17) the images 1, 2, 4, 5, 7, 11 and 12 which show that these hybrid methods help us to solve the problem and skin problem such as skin land lighting condition with detection rates of 98.14%.

Images 6 and 10 show that this hybrid method is able to detect more than one person in the same image efficiently with detection rate of 98.37%. Image number 3 shows that this hybrid method is able to detect the skin even if the development of reflective glass in front of any post shows glass skin. Skin with detection rates of 98.35% was taken into account. Images 8 and 9 show that this hybrid method has the ability to detect the skin
even if the development of reflective water in front of the entire skin shows water post. Skin with detection rates of 98.27% was taken into account. This hybrid module for skin detector as shown from the above percentages solved all the problems which are related to skin as detector with some limitations that we faced in this technique. The author used 600 images for skin and 600 images for non-skin as basis for training. This thing made the determinants of the number of pixels trainable a gap, for more number of pixels trained gives better results. Therefore, we had to wait a long time to train a number of pixels to make the system detect the image well, where this training took two consecutive days. Other limitation was that deity data collected from the internet and used in this system were not available. That is there were no images content just the skin pixels with white background in the internet which are used for training process. We collected these images from the internet that contains skin with any background and we segmented them manually using Photoshop. This action took more time. It took us about four months before we were able to segment the skin only from the 600 images deity skin part
of the training.

Result comparison

The classification study shows results will be compared with the simulation results using the fuzzy system based interference detection skin, which is proposed by Taqa and Jalab (2010a). Skin detector is based on combing both texture and color features using neural network which is proposed by Taqa and Jalab (2010b) and our proposed method. The study showed that the proposed techniques for perfect results are not very well and that they still have problems in the reliability of the results of the skin detectors, as shown in Figures 18 and 19. Patterns can be collected in the fifth factors. Primarily, Taqa and Jalab (2010a, 2010b) used the RGB color space, but changed to the method used in this study as RGB color space is not enough. This is because the RGB color space is not separated by luminance and chrominance, and the components R, G and B are highly correlated.

The luminance of a given pixel is a linear combination of RGB values R, G and B. Therefore, it leads to changing the brightness of a skin patch because it affects all components R, G and B. The skin color cluster is extended in space to reflect the difference in light intensity of the patches. Similarly, the clusters of skin color for the patches from different races are in different places in the RGB color space.

Secondly, that this document was proposed by Taqa talked about the problem of skin-like. But we have shown in the results of any image to indicate the problem solved by using the automated system in the same threshold, instead of more than a threshold. So only use the method of fuzzy skin as the representation of objects, that is, the color of the skin-like, which may be present in an image, such as hair, soil, wood materials, sand, drawings, and animals, which are uncontrolled. Because of this, the identification of the skin may contain errors and generate a false detection and therefore fail to identify objects that have non-skin color-like pixels.

Third, this document proposed by Taqa used different thresholds of many pictures to make the perfect result. The researcher can deduce from this that the skin detector can be the basis for any application in the future, as a gesture of analysis of facial recognition, tracking, human motion tracking and other human-related image processing applications in computer vision and multimedia such as web content filtering, recovering in multimedia databases, video surveillance, video, and
Table 1. Comparing between the number of pixels of skin detections.

<table>
<thead>
<tr>
<th>Skin detector</th>
<th>No. of skin pixels</th>
<th>No. of non skin pixels</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIS based skin detector</td>
<td>351,228</td>
<td>428,602</td>
</tr>
<tr>
<td>Skin detector based on combining both texture and color features using neural network</td>
<td>608,129</td>
<td>511,000</td>
</tr>
<tr>
<td>Our proposed module</td>
<td>34,159,027</td>
<td>297,136,954</td>
</tr>
</tbody>
</table>

video conferencing applications. This is because these applications depend directly on the skin where the sensor is exposed. The skin is the first step in the application and then taking the characteristics of this application, depending on the type of skin product. Therefore, see that the detector works as a principal role in these applications, if this requires the existence of skin fixed detector detected by any background of any image without the intervention of the hand, or a change in the threshold. Definitely, the number of pixels that are used in the training skin on both sides of the hips and the sides of the skin is not enough to obtain accurate results with high reliability, as shown in Table 1. Comparison is between our proposed method and the Skin detectors proposed by Taqa and Jalab (2010a, 2010b) in the number of pixels.

Finally in Figure 19 Taqa and Jalab (2010b) had achieved a high rate in detecting skin, but did not take into account the false positive- the real problem which is contact with the skin-like. They achieved 0.8795% comparing with our proposed method which achieved 0.14, as shown in Table 2. The study proposed the hybrid method using multi color space to overcome these problems; this will be followed by accuracy rates comparison in the study.

Evaluation comparisons between our proposed module, the FIS based skin detector and skin detector based on combing both texture and color features using neural network

It was discovered that processing of the skin is not perfect and different people adopt different criteria for performance evaluation. One of the evaluation criteria is the general appearance of the areas of size detected. To
Table 2. True positive and false positive of skin detections.

<table>
<thead>
<tr>
<th>Skin detector</th>
<th>TP (%)</th>
<th>FP (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIS based skin detector</td>
<td>90.00</td>
<td>0.22</td>
</tr>
<tr>
<td>Skin detector based on combing both texture and color features using neural network</td>
<td>95.6176</td>
<td>0.8795</td>
</tr>
<tr>
<td>Our proposed module</td>
<td>98.00</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Table 3. Evaluation metrics of skin detections.

<table>
<thead>
<tr>
<th>Skin detector</th>
<th>Recall</th>
<th>Precision</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIS based skin detector</td>
<td>0.900</td>
<td>0.700</td>
<td>0.910</td>
<td>0.910</td>
</tr>
<tr>
<td>Skin detector based on combing both texture and color features using neural network</td>
<td>0.956</td>
<td>0.916</td>
<td>0.991</td>
<td>0.998</td>
</tr>
<tr>
<td>Our proposed module</td>
<td>0.982</td>
<td>0.995</td>
<td>0.995</td>
<td>0.999</td>
</tr>
</tbody>
</table>

quantify the performance evaluation, true positive (TP), false positives (FP), true negative (TN) and false negatives (FN) were calculated for all pixels in the test set through skin testing detectors. FP is the proportion of pixels incorrectly classified as non-skin skin, while TP is the proportion of skin pixels correctly classified as skin; (Aamer Mohamed et al., 2009; Taqa and Jalab, 2010a).

TN and FN are the additions of FP and TP, respectively. To evaluate the proposed skin detector, a sample of 15 images was used (Table 2). Four parameters in Table 3 were used to evaluate the performance of the two surveys skin. These parameters are: Recalling, accuracy, specificity and precision (Gasparini and Schettini, 2006). The kind of images linked to the Internet, the performance of the proposed skin detector is surprisingly good. The best performance can detect 98% of skin pixels with a very low FP rate of 0.14% by our proposed module. Although, the FIS based Skin detector and Skin Detector Based on Combing Both Texture and Color Features using Nural Network can detect 90 and 95.6176%, respectively of skin pixels correctly, Taqa and Jalab (2010a, 2010b) also have a high FP rate of 0.22 and 0.8795%, respectively.

The first approximation of the model proposed by Wadud et al. (2008) can detect 72% of skin pixels with a rate of 5% FP, while the Bayesian model proposed by Jones and Rehg (2002) can detect 69% FP at the same rate. In the meantime, this model can detect 80% of skin pixels with a rate of 8.5 or 90% correct recognition with 14.2% FP. The detection method recommended by the skin (Zafarifar et al., 2010) can detect more than 83% of skin pixels correctly with a rate of 20% FP. The recall rate of skin color pixel-based classification proposed by Gasparini and Schettini (2006) is 92% and the accuracy rate is 39%. These values indicate that the evaluation metric proposed skin detector outperforms the other methods listed above, especially in terms of reducing the rate of FP (Mohamed et al., 2008; Taqa and Jalab, 2010a).

Conclusion

The study applied a novel hybrid module using the technique of grouping histogram for the Bayesian method and back propagation neural network with adjacent segment-nested (SAN) technology based on YCbCr and RGB color space, respectively. We considered the results satisfactory, as in the first place, the modified form of the classic used hybrid skin detection is only tested on the skin candidate pixels for the skin. So the search space is reduced. Secondary form of a hybrid in the skin color of the Bayesian method is created at the level of YCbCr color space. The reasons for choosing chrominance chrolninance red and blue (Cb and Cr) in YCbCr color space instead of eliminating the effect of changes in lighting using component Y: classification uses only pixel chrominance Cb and Cr (pure color), the segmentation of the skin has become more robust for the lighting, when the variations in luminance of the pixels are discarded, and also to narrow the search to speed up the calculation for the identification of areas of skin. Third, we realized that the color of the skin-like increases the rate of false positives and false negatives, thereby reducing the true positive. In other words, the skin detector is not reliable. To overcome this problem in the skin color of anger, we collected more than 600 images with different range of skin color.

After updating the dataset with a wider range of skin colors, the results are more reliable. Fourth, the author used the technique of an adjacent segment-nested according to the RGB color space in the preprocessing for the Nural part. It helped us to increase the number of pixels that is used during training, where it became 117,000 pixels of skin and skin Nural for better performance.

Our results suggest that the technique of adjacent segment-nested with the color is a strong signal to identify that people in pictures are not bound. Our solution has shown the effectiveness of the classification...
rate of 98% with high reliability. This technique will be improved by taking into account different characteristics as heretical rule should be considered in future work. Using the rules of the heretical, the human object can be extracted from other images that can filter out unwanted noise and the homogeneity region of skin found in the images. With the combination of these methods and proposed heretical rules, it is believed that a better classification performance can be achieved.

ACKNOWLEDGEMENTS

This paper has been a part of PhD research from Multimedia University. The authors would like to acknowledge all those who were involved in this project and had given their support in more than one ways.

REFERENCES


