Human Skin Detection

By Md. Zargis Talukder, Avizit Basak & Dr. Muhammad Shoyaib
Bangladesh University of Engineering and Technology

Abstract - Skin-color modeling is a crucial task for several applications of computer vision. Problems such as face detection in video are more likely to be solved if an efficient skin-color model is constructed. Most potential applications of skin-color model require robustness to significant variations in races, differing lighting conditions, textures and other factors. Given the fact that a skin surface reflects the light in a different way as compared to other surfaces. As the color of human skin is created by the combination of blood (red) and melanin (brown, yellow) which gives it a restricted range of hues. A skin region can be classified by comparing large image content of skin database and non-skin database. The RGB color space is widely used and most effective to detect skin region from an image. The segmentation is used to localize and identify homogeneous regions in a picture by perceptual attributes which include the size, the shape and the texture and/or color information. The probability of each RGB color space of skin and non-skin database is important to detect skin pixels.

Keywords: skin detection, image, processing.

GJCST-F Classification: I.5.4
Human Skin Detection

Md. Zargis Talukder*, Avizit Basak* & Dr. Muhammad Shoyaib

Abstract - Skin-color modeling is a crucial task for several applications of computer vision. Problems such as face detection in video are more likely to be solved if an efficient skin-color model is constructed. Most potential applications of skin-color model require robustness to significant variations in races, differing lighting conditions, textures and other factors. Given the fact that a skin surface reflects the light in a different way as compared to other surfaces. As the color of human skin is created by the combination of blood (red) and melanin (brown, yellow) which gives it a restricted range of hues. A skin region can be classified by comparing large image content of skin database and non-skin database. The RGB color space is widely used and most effective to detect skin region from an image. The segmentation is used to localize and identify homogeneous regions in a picture by perceptual attributes which include the size, the shape and the texture and/or color information. The probability of each RGB color space of skin and non-skin database is important to detect skin pixels. For each pixel of testing image, the RGB value is calculated then the probability ratio of that RGB color space for training Skin and non-skin data base is compare with a threshold variable called $\theta$. The threshold value lies between 0 and 1 but for this analyzing purpose threshold value has been taken as 0.4. If the ratio will be greater than 0.4 then that pixel will be detected as skin pixel else detected as non-skin pixel. After segmenting out skin from images, this can be useful for identifying faces, hand sign recognition, offensive content such porn photography. The performance curve (ROC curve) reflects the overall accuracy of our analysis.

Keywords : skin detection, image, processing.

I. Introduction

Skin detection consists in detecting human skin pixels from an image. The system output is a binary image defined on the same pixel grid as the input image. Skin detection plays an important role in various applications such as face detection, searching and filtering image content on the web. Research has been performed on the detection of human skin pixels in color images by use of various statistical color models. Some researchers have used skin color models such as Gaussian, Gaussian mixture or histograms. In most experiments, skin pixels are acquired from a limited number of people under a limited range of lighting conditions. Unfortunately, the illumination conditions are often unknown in an arbitrary image, so the variation in skin colors is much less constrained in practice. This is particularly true for web images captured under a wide variety of conditions. A good skin classifier must be able to discriminate between skin and non-skin pixels for a wide range of people with different skin types such as white, yellow, brown and dark and be able to perform well under different illumination conditions such as indoor, outdoor and with white and non-white illumination sources. An important step in the image classification process is color segmentation of the image into homogeneous skin-color regions and non-skin-color regions in color space that is relatively invariant to minor illuminant changes.

However, given a large collection of labeled training pixels including all human skin (Caucasians, Africans, Asians) we can still model the distribution of skin and non-skin colors in the color space. Jones and Rehg [3] proposed techniques for skin color detection by estimating the distribution of skin and non-skin color in the color space using labeled training data. The histogram models were found to be slightly superior to Gaussian mixture models in terms of skin pixel classification for this color space. A skin detection system is never perfect and different users use different criteria for evaluation. General appearance of the skin-zones detected, or other global criteria might be important for further processing. For quantitative evaluation, false positives and detection rates have used. False positive rate is the proportion of non-skin pixels classified as skin and detection rate is the proportion of skin pixels classified as skin. The user might wish to combine these two indicators his own way depending on the kind of error he is more willing to afford. The larger the value ensures the larger the belief for a skin pixel. Error rates for all possible thresholding are summarized in the Receiver Operating Characteristic (ROC) curve.

In this research Image dataset is classified as Training Image Dataset and Testing Image dataset. The training Image dataset consist two database called Skin database and non-skin database consisting human skin pixels and non-skin pixels respectively. Whereas the testing image dataset containing real image both skin and non-skin pixels. The training skin database consists of five hundred and fifty five images and so as mask corresponds to the actual image. Each of them is
manually segmented such that the skin pixels are labeled carefully. The training non-skin database consists of Eight thousand four hundred and thirty images having no skin pixels. On the other hand testing image dataset contain One Hundred and three image. We select our image data set with proper illuminating condition as far possible. The goal is to infer a model from this set of data in order to perform skin detection on new images.

II. IMAGE PREPARATION

Most of the research efforts on skin detection have focused on visible spectrum imaging. Skin-color detection in visible spectrum can be a very challenging task as the skin color in an image is sensitive to various factors such as:

a) **Illumination**

A change in the light source distribution and in the illumination level (indoor, outdoor, highlights, shadows, non-white lights) produces a change in the color of the skin in the image (color constancy problem). The illumination variation is the most important problem among current skin detection systems that seriously degrades the performance.

b) **Camera Characteristics**

Even under the same illumination, the skin-color distribution for the same person differs from one camera to another depending on the camera sensor characteristics. The color reproduced by a CCD camera is dependent on the spectral reflectance, the prevailing illumination conditions and the camera sensor sensitivities.

c) **Ethnicity**

Skin color also varies from person to person belonging to different ethnic groups and from persons across different regions. For example, the skin color of people belonging to Asian, African, Caucasian and Hispanic groups is different from one another and ranges from white, yellow to dark.

d) **Individual Characteristics**

Individual characteristics such as age, sex and body parts also affect the skin-color appearance.

e) **Other Factors**

Different factors such as subject appearances (makeup, hairstyle and glasses), background colors, shadows and motion also influence skin-color appearance.

Many of the problems encountered in visual spectrum can be overcome by using non-visual spectrum such as infrared (IR) and spectral imaging. Skin-color in non-visual spectrum methods is invariant to changes in illumination conditions, ethnicity, shadows and makeup. However, the expensive equipment necessary for these methods combined with tedious setup procedures have limited their use to specific application areas such as biomedical applications [1].

III. SKIN-COLOR MODELING AND ANALYSIS

a) **Color Spaces**

The choice of color space can be considered as the primary step in skin-color classification. The RGB color space is the default color space for most available image formats. Any other color space can be obtained from a linear or non-linear transformation from RGB. The color space transformation is assumed to decrease the overlap between skin and non-skin pixels thereby aiding skin-pixel classification and to provide robust parameters against varying illumination conditions.

Figure 1: RGB Color Space

The RGB color space consists of the three additive primaries: red, green and blue. Spectral components of these colors combine additively to produce a resultant color. The RGB model is represented by a 3-dimensional cube with red green and blue at the corners on each axis (Figure 1). Black is at the origin. White is at the opposite end of the cube. The gray scale follows the line from black to white. In a 24-bit color graphics system with 8 bits per color channel, red is (255, 0, 0). On the color cube, it is (1, 0, 0). The red, green and blue color components are highly correlated. The RGB model simplifies the design of computer graphics systems but is not ideal for all applications.

b) **Explicit skin-color space thresholding**

Image segmentation is a process of pixel classification into classes. Two classes or bi-level segmentation is the simplest case. The real world consists of many objects and so the number of classes. There are primarily four types of segmentation techniques: thresholding, boundary-based, region-based, and hybrid techniques. All these four methods can be implemented in grayscale and color images.

Thresholding is based on histogram of the image and applying assumption that peaks in the
histogram correspond to either background or objects of interest that can be extracted by separating histogram valleys. In segmentation, an image is partitioned into different homogeneous regions, where the homogeneity of a region may be composed based on different criteria such as gray level, color or texture. Region of interest segmentation plays a crucial role as a preliminary step for high-level image processing such as image analysis and pattern recognition.

Jones and Rehg [3], built a 3D RGB histogram model with two billion pixels collected from 18,696 web images. They reported that 77% of the possible RGB colors are not encountered and most of the histogram is empty. There is a marked skew in the distribution towards the red corner of the color cube due to the presence of skin in web images, though only 10% of the total pixels are skin pixels.

This suggested that skin colors occur more frequently than other object colors. By computing two different histograms, skin and non-skin histograms, the probability that a given color belongs to skin and non-skin class (also called class conditional probabilities) is defined as

\[
P(\text{rgb} \mid \text{skin}) = \frac{S[\text{rgb}]}{T_s}
\]

\[
P(\text{rgb} \mid \text{non-skin}) = \frac{N[\text{rgb}]}{T_n}
\]

Where \(S[\text{rgb}]\) is the pixel count contained in bin “rgb” of the skin histogram, \(N[\text{rgb}]\) is the equivalent count from the non-skin histogram, and \(T_s\) and \(T_n\) are the total counts contained in the skin and non-skin histograms, respectively. From the generic skin and non-skin histograms a reasonable separation can be obtained between skin and non-skin classes. This fact can be used to build fast and accurate skin classifiers even for images collected from unconstrained imaging environments such as web images, given that the training dataset is sufficiently huge. A larger training set can lead to better probability density function estimations. Given the class conditional probabilities of skin and non-skin-color models, a skin classifier can be built by which a given image pixel can be classified as skin, if

\[
\frac{P(c \mid \text{skin})}{P(c \mid \text{non-skin})} \geq \theta
\]

Where \(0 \leq \theta \leq 1\) is a threshold value which can be adjusted to trade-off between true positives and false positives. This threshold value is normally determined from the ROC (receiver operating characteristics) curve calculated from the training data set. The ROC curve represents the relationship between the true positives and false positives as function of the detection threshold.

IV. Data Preparation and Analysis

The color of skin in images depends primarily on the concentration of hemoglobin and melanin and on the conditions of illumination. It is well-known that the hue of skin is roughly invariant across different ethnic groups after the illuminant has been discounted. This is because differences in the concentration of pigments primarily affect the saturation of skin color, not the hue. For detecting skin a set of 9,088 real images are gathered. These images set have divided into 2 set of data called Training data set and testing data set. The training data have further divided into two dataset called Skin database and Non-skin database. The Skin database consists of 555 real images containing some skin pixels. Also 555 corresponding masked images are contained within this database. These images contain skin pixels belonging to persons of different origins, with unconstrained illumination and background conditions, which make the skin detection task more challenging and difficult. Figure 2 shows some images and their corresponding masked images. The skin pixels within these images have been labeled manually and prepared for skin histogram. The masks are prepared manually with proper caring and consciousness. The Skin database sample consists of 20796952 image pixels and will be validated by a similar number of image pixels randomly selected. Each image pixel has been selected only once.

The Non-skin database consists of 8430 images that did not contain any skin pixel have prepared for the non-skin histogram. These images does not contain any skin pixels with unconstrained illumination and background conditions. The Non-Skin database sample consists of 800927510 image pixels for analyzing purpose. Each data file will be used to train a network during a training run.

The test data consists of 103 images selected at randomly from the remaining image database. The test images are used to evaluate the performance of the skin detection system.

For analyzing skin and non-skin dataset various programming language are available such as “C”, “C++”, “JAVA”, “MATLAB” etc. All has flexibility enough and so as constrains. As JAVA is a platform independent language and has various built-in library function, so we choose JAVA for our analyzing purpose. Without it JAVA is one of the most powerful programming language this is another reason to select JAVA for our analyzing purpose.

a) Skin database analysis

The skin database has contained 555 real images containing skin pixel within them, which we considered as training image. There also have a mask database containing mask corresponding to those actual images. The mask was prepared manually.
During preparing mask the non-skin pixels were marked as white and skin as others. Hence by comparing actual image and corresponding to its mask the actual skin pixels have been detected easily. At the starting of the program we declared a three dimensional array having $256 \times 256 \times 256$ space for skin database to construct a 256 bin histogram. When the skin pixel was detected, corresponding RGB color space have been increased by 1. The total process was continued till the last skin pixel detection in the Skin-database. After that the probability has calculated to each RGB color space which stored in a file named skin probability.

![Figure 2: Training Skin Database Image Corresponding to its Mask](image)

### b) Non-Skin Database Analysis

The Non-skin database has contained 8430 real image which does not contain any skin pixels. At the starting of the program we also declared a three dimensional array having $256 \times 256 \times 256$ space for Non-skin database to construct a 256 bin histogram. For each non skin pixels, the basic RGB combination is detected. Then the array is incremented by 1 to its corresponding array space (RGB combination). The total process continued till to find out the last pixel in the Non-Skin database. After that the probability has calculated for each RGB space which stored in a file named non-skin probability.

The calculated general Information about Skin and Non-skin image are listed below:

<table>
<thead>
<tr>
<th></th>
<th>Total Counts</th>
<th>Total Occupied Bins</th>
<th>Total Unoccupied Bins</th>
<th>Percent Unoccupied</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin Model</td>
<td>20796952</td>
<td>433460</td>
<td>16343748</td>
<td>97.4%</td>
</tr>
<tr>
<td>Non-skin Model</td>
<td>800927510</td>
<td>3237834</td>
<td>13539382</td>
<td>80.7%</td>
</tr>
<tr>
<td>Overlapping skin/non-skin bins</td>
<td>433394</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skin pixels as a percentage of total pixels:</td>
<td>2.53 %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total photos in labeled dataset:</td>
<td>9088</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of photos containing skin:</td>
<td>7.15 %</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### c) Histogram-based Skin Classifier

In this stage some binary images have been generated from its actual images. It is the testing stage, so testing dataset have used in this stage. Testing dataset contains 103 real images. The masks correspond to its real image of training database have prepared carefully. In this stage the first goal was to generate binary image from testing image using probability of RGB color space of training skin and non-skin databases, which we generate earlier. An important terminology called “threshold” is magnificent which is denoted by $\theta$ which can vary from 0 to 1. We compared the value of $\theta$ with the ratio of skin probability and Non-skin Probability. So it is important consideration to fix up the value of $\theta$. As the variation of illumination condition, skin color depending on geographical location and many other unavoidable constrain threshold $\theta$ can be vary. So normally it depends on our image dataset. The general relationship between threshold $\theta$ and the ratio of skin and Non-skin probability can be defined as:

$$\frac{P(c/\text{skin})}{P(c/\text{non-skin})} \geq \theta$$
Where $0 < \theta < 1$. Considering our training dataset we have taken 0.4 as the value of Threshold. From testing dataset we have considered each pixel and its relevant RGB values. Depending on that RGB value the skin probability value and Non-skin probability value have chosen from calculated dataset. If the ratio of skin probability to the non-skin probability corresponds to the testing RGB space is greater than 0.4 then that pixel considered as a skin pixel and so that the output marker defined as 1 that is Pure white, otherwise we considered as a non-skin pixel and so that the output marker is defined as 0 that is pure Black. By this process a binary image corresponds to its actual training images is generated. In figure 3 some testing image, mask and generated binary images are shown.

![Figure 3: Testing image, mask and generated binary image](image-url)
The classifier does a good job of detecting skin in most images. In particular, the skin labels form dense sets whose shape often resembles that of the true skin pixels. The detector tends to fail on highly saturated or shadowed skin. The example photos also show the performance of the detector on non-skin pixels. In photos such as the sand or flowers (which having Human skin like color) the false detections are sparse and scattered. More problematic are images with wood or copper-colored metal such as the kitchen scene or railroad tracks. These photos contain colors which often occur in the skin model and are difficult to discriminate reliably. This results in fairly dense sets of false positives. Classifier performance can be quantified by computing the ROC curve which measures the threshold-dependent trade-off between misses and false detections.

V. Performance Analysis

The ROC (Receiver Operating Characteristic) curve is an important trend-off for measuring our performance. In addition to the threshold setting, classifier performance is also a function of the size of the histogram (number of bins) in the color models. Too few bins result in poor accuracy while too many bins lead to overfitting. Figure 4 shows the family of ROC curves produced as the size of the histogram varies from 256 bins/channel. The axis labeled “Probability of correct detection” gives the fraction of pixels labeled as skin that were classified correctly, while “Probability of false detection” gives the fraction of non-skin pixels which are mistakenly classified as skin. These curves were computed from the test data.

![ROC Curve](image)

**Figure 4**: ROC Curve for 256 Bin Histogram Model

By analyzing other size of bin, it can be found that histogram size 32 gave the best performance, superior to the size 256 model at the larger false detection rates and slightly better than the size 16 model in two places. The performance of the skin classifier is good considering the unconstrained nature of web images. The classifier (256 bin) can detect roughly \( \frac{4258578}{4258578+3470391} \times 100\% = 55\% \) of skin pixels with a false positive rate of 45\%. This corresponds to the point on the ROC curve where the probability of false rejection (which is one minus the probability of correct detection) equals the probability of false detection.

In addition to histogram size, classifier performance is also affected by the amount of training data. To do so the list of skin and non-skin images in the training set and divided then into chunks containing approximately 5744138.0 skin pixels and 27375162 non-skin pixels. On each iteration we added one such chunk of new skin and non-skin pixels to the evolving training
A ROC curve was computed at each iteration showing the classifier performance on the full test set.

If more data is added, performance on the training set decreases because the overlap between skin and non-skin data increases. In that case a smooth shaped curve can be obtained which is a measure of good performance and tends to fill up the area near to 1 underlining the curve. Increasing testing data increase the performance characteristic in a similar manner.

VI. Conclusion

Skin detection is very important but very complex procedures are maintained. A single model is can’t give the proper result. But various mixture models can give more appropriate result. Whatever this analysis emphasis on analyze the basic properties of skin and fix up a proper threshold. The testing data set was small so that ROC curve occupied less area in the graph. Using 256 bins Histogram model is another reason to decrement the curve sharpness. Generally 32 bin color model can give more appropriate result. Many objects in the real world have skin-tone colors, such as some kinds of leather, sand, wood, fur, etc., which might be mistakenly detected by a skin detector. Therefore, skin detection can be very useful in finding human faces and hands in controlled environments where the background is guaranteed not to contain skin-tone colors. Since skin detection depends on locating skin-colored pixels, its use is limited to color images, i.e., it is not useful with gray-scale, infrared, or other types of image modalities that do not contain color information. This research hypothesize that taking a minimalistic approach to the software design that can perform image preprocessing, and skin detection in real time.

References Références Referencias

4. E. Ahmed., C. Muang and D. Hu., Skin Detection, Rutgers University, Piscataway, NJ, 08902, USA.