HUMAN SKIN DETECTION USING IMAGE FUSION

Pallavi R. Patil
Research Scholar, Department of Electronics & Telecommunication Engineering, SSBT COET, Bambhori, Jalgaon, India

A. H. Karode
Assistant Professor, Department of Electronics & Telecommunication Engineering, SSBT COET, Bambhori, Jalgaon, India

Dr. S. R. Suralkar
Professor, Department of Electronics & Telecommunication Engineering, SSBT COET, Bambhori, Jalgaon, India

ABSTRACT

Skin detection is a process of finding skin colored pixels and regions in an image or a video. Skin detection is a primary step in variety of image processing applications right from face tracking, gesture analysis, video surveillance, image content filtering, and video annotation to various human-computer interaction domains. Effective skin detection is the process of identifying human skin colors of different ethnic and under different illumination conditions. Different skin detection however has been successfully applied but they are prone to false skin detection and also not able to cope up with the variety of human skin colors across different ethnic. Additionally requires high computational costs. In this project a dynamic threshold approach is proposed which would reduce computational costs by eliminating training stage to improve the accuracy of skin detection despite wide variation in ethnicity and illumination. An eye detector would be used to refine the skin model for a specific person. Additional feature would be implemented i.e. fusion of a smoothed 2-D histogram and Gaussian model for automatic human skin detection in color image(s) to improve the results of skin detection.

Key words: Color extraction, Fusion technique, Skin Detection, Count Skin & Non Skin Pixel value.

http://www.iaeme.com/IJECET/issues.asp?JType=IJECET&VType=8&IType=4
1. INTRODUCTION

With the progress of information society today, images have become more and more important. Among them, skin detection plays an important role in a wide range of image processing applications as skin is the most widely used primitive in human image processing. Skin detection is a process of finding skin colored pixels and regions in an image or a video. Skin detection is a primary step in variety of image processing applications right from face tracking, gesture analysis, video surveillance, image content filtering, and video annotation to various human-computer interaction domains.

Skin detection is not as easy task as the skin appearance in images is affected by various factors such as illumination, complex backgrounds, camera characteristics and ethnicity. Thus effective skin detection is the process of identifying human skin colors of different ethnic and under different illumination conditions.

The approaches to classify skin in images can be grouped into three types: parametric, non-parametric and explicit skin cluster definition methods. The parametric models use a Gaussian color distribution whereas non-parametric methods estimate the skin-color from the histogram obtained from training data. Skin clustering explicitly defines the boundaries of skin in a given color space, generally termed static skin filters. The main drawback of skin clustering is a high number of false detections.

In general skin detection process has two phases:

1. Training Phase
2. Detection Phase

Training a skin detector involves three basic steps –

1. Collecting a database of skin patches from different images. Such a database typically contains skin colored patches from variety of people under different illumination conditions.
2. Choosing a suitable color space. Since skin color occupies a part of such a space, which might be a compact or a large region in the space
3. Learning the parameters of a skin classifier

Detection phase involves-

1. Converting the image (under skin detection) into the same colour space that was used in the training phase.
2. Classifying each pixel using skin classifier to skin or non skin
3. Post processing using morphology to impose homogeneity on the detected regions

One of the simplest and common methods in detecting human skin is to define a fixed decision boundary (threshold) i.e. the training phase. The pixel values that lies between these boundaries are selected as skin pixels else as non-skin pixels.

2. PROPOSED SYSTEM

In the proposed method, to overcome the above mentioned drawbacks by using dynamic approach and the fusion strategy to improve the result of skin detection. Figure 1. shows the block diagram of the proposed system.
Initially first, an approach similar to that of Fusel et al. is adopted to obtain the face(s) in a given image. Second, a dynamic method is employed to calculate the skin threshold value(s) on the detected face(s) region. Third, two features—the 2-D histogram with smoothed densities and Gaussian model—are introduced to represent the skin and non-skin distributions, respectively. Finally, a fusion framework that uses the product rule on the two features is employed to obtain better skin detection results. In this paper, the RGB color space is converted to the HSV color space to mimic visual human perception.

2.1. Acquire Image
The input image is live image taken from camera. Camera image size width*height (320*240). Taking the image from camera flip is used to avoid confusion left & right of image. When getting image first convert the image into grey scale because haar classifier is used for detection of skin for that purpose. I can require first face from the image then haar classifier can create a database for skin tone detection after that this database is compared with criteria of 2D histogram & GMM.

2.2. Pre-Processing
In the preprocessing steps, for any given image(s), S_t where t is the number of images∈{1,2,……,T} we first locate human eyes. Then, an elliptical mask model as illustrated in Fig. 2 would be used to generate the elliptical face region in the image(s). Here, (x_o,y_o) is the center of the ellipse as well as the eyes symmetry point. Minor and major axes of the ellipse are represented by 1.6D and 1.8D respectively, where D is the distance between two eyes. The detected face regions include smooth (i.e., skin) and non-smooth (i.e.eyes, eye brown, mouth, etc.) textures. As we are only interested in smooth regions, Sobel edge detection would be used to remove non-smooth regions. Then, the detected edge pixels would be further dilated using a dilation operation to get the optimal non-smooth regions. Finally, we would obtain a new image(s), S_t that only consist(s) of face regions.
2.3. Color Space
An image can be represented in a number of different color space models (RGB, HSV, YCbCr). These are some color space models available in image processing. Therefore, it is important to choose the appropriate color space for modeling human skin color. This project work proposes the use of the HSV color space. HSV Color Model HSV color model (Hue, Saturation, and Value) is a no lineal transformation of the RGB space color, and the colors are a combination of the three values: the Hue (H), Saturation or color quantity (S), and itself value (V).

2.4. Skin Detection

2.4.1. Dynamic Threshold With Smoothed 2-D Histogram
Human skin color varies greatly between different ethnicity. Nonetheless, skin appearance in color image(s) can also be affected by illumination, background image, camera characteristic, etc. Therefore, a fixed or pre-learned threshold for detecting skin boundaries is not a feasible solution. In our approach, we employ an online dynamic approach as to calculate the skin threshold value(s) on the face images $S_t'$. The assumption is that the face and body of a person always share the same colors. However here we propose to use a 2-D histogram with smoothing densities.

2.4.2. Gaussian Model
The Gaussian model is a sophisticated model that is capable of describing complex-shaped distributions and is popular for modeling skin-color distributions. The threshold skin-color distribution in the 2-D histogram would be modeled through elliptical Gaussian joint probability distribution functions defined as:

$$ P(H|\lambda) = \sum_{i=1}^{k} w_i g(c|\mu_i, \Sigma_i) $$

Where $H$ is the color vector of $(I, B_y)$, $\lambda = \{w_i, \mu_i, \Sigma_i\}$, $\mu_i$ is the mean vector and $\Sigma_i$ is the diagonal covariance matrix respectively. $w_i$ refers to the mixing weights, which satisfy the constraint $\sum_{k=1}^{k} \pi_i = 1$

2.5. Fusion Strategy
In order to increase the effectiveness and robustness of the skin detection algorithm, a fusion strategy is proposed by integrating the two incoming single features into a combined single representation. Both models will vote for classification of skin and non-skin pixels. This can be done by using product rule to both models.

Let $D_{\text{hist}}(S_t, Z)$ and $D_{\text{gmm}}(S_t, \mu, \Sigma)$ denote the matching results produced by the smoothed 2-D histogram $Z$ and Gaussian model $g(c|\mu_i, \Sigma_i)$ respectively. The combined matching results $D(S_t)$ using the fusion rules can be obtained as follows:

$$ D(S_t) = \Gamma\{D_{\text{hist}}(S_t, Z), D_{\text{gmm}}(S_t, \mu, \Sigma)\} $$

Where $\Gamma$ is the selected fusion rule, which represents the product $\otimes$.

3. METHODOLOGY

3.1. Dynamic Threshold With Smoothed 2-D Histogram:
Human skin color varies greatly between different ethnicity. Nonetheless, skin appearance in color image(s) can also be affected by illumination, background image, camera characteristic, etc. Therefore, a fixed or pre-learned threshold for detecting skin boundaries is not a feasible solution. In our approach, we employ an online dynamic approach as to calculate the skin
threshold value(s) on the face images, $S_t$. The assumption is that the face and body of a person always share the same colors.

![Figure 3 HSV color model](image)

The three magnitudes can have the following values

- **Hue:**
- **Saturation:**
- **Value:**

Once the transformation of the input image was made, it was observed that the skin tone of a person could be seen in a different color from those seen from different objects within the same image. However, instead of using the 1-D histogram, we introduce a 2-D histogram [see Fig. 3] with smoothing densities. In this paper, the feature vector for the smoothed 2-D histogram, $Z$, is represented by the combination of $I$ and $B$. The smoothed 2-D histogram-based skin segmentation, $D_{hist}$ at pixel, $n$ is given as:

$$D_{hist} (S_t, Z) = \begin{cases} 1 & \text{if } Z (I_n, B_{yn}) > 20 \\ 0 & \text{if } Z (I_n, B_{yn}) \leq 20 \end{cases} \quad (7)$$

∀ $I_n, B_{yn} \in S_t$

![Figure 4 2-D Histogram has two different channels of the same color space on x axis and y axis.](image)

### 3.2 Gaussian Model

The Gaussian model is a sophisticated model that is capable of describing complex-shaped distributions and is popular for modeling skin-color distributions. The threshold skin-color distribution in the 2-D histogram would then be modeled through elliptical Gaussian joint probability distribution functions defined as
Human Skin Detection Using Image Fusion

\[
P(H|\lambda) = \sum_{i=1}^{k} w_i g(c|\mu_i, \Sigma_i)
\]  

(8)

Where \( H \) is the color vector of \((I, B_y)\), \( \lambda = \{ w_i, \mu_i, \Sigma_i \} \), \( \mu_i \) is the mean vector and \( \Sigma_i \) is the diagonal covariance matrix respectively. \( w_i \) refers to the mixing weights, which satisfy the constraint \( \sum_{k=1}^{k} \Pi_1 = 1 \)

The result of Gaussian model-based skin detection, \( D_{gmm} \), can be obtained by using Fig.4.

![Figure 4 Gaussian Model](image)

\( \mu \) is the center of the Gaussian model, while \( T \) is the angle between \( x \)-axis and line \( D \). Let \((I_n, B_{yn})\) be the co-ordinate of pixel \( n \) and is positioned on the red dot along line, \( D \). Distance \((I_n, B_{yn})\), \( d \) and the angle \( T \) are calculated as follows:

\[
d = \sqrt{dx^2 + dy^2}
\]

(9)

\[
T = \tan^{-1}(dy/dx)
\]

Where \( d_x \) and \( d_y \) are the distances between \((I_n, B_{yn})\) and center, \( \mu \) at \( x \)-axis and \( y \)-axis, respectively. And are the coordinate of at \( x \)-axis and \( y \)-axis, respectively. Distance between the boundary and center of the Gaussian model at \( x \)-axis and \( y \)-axis, \( D_x \) and \( D_y \) at given angle, \( T \) are as follows:

\[
D_x = \sum_x \cos(T)
\]

(10)

\[
D_y = \sum_y \sin(T)
\]

(11)

Where \( \sum_x \) and \( \sum_y \) are the variance of \( x \)-axis and \( y \)-axis for Gaussian model. Distance, \( D \) is represented as

Therefore, \( D_{gmm} \) is given as:

\[
d = \sqrt{dx^2 + dy^2}
\]

(12)

\[
D_{gmm}(S_t, \mu, \Sigma) = \begin{cases} 1 & \text{if } D > d \\ 0 & \text{otherwise} \end{cases}
\]

(13)

3.3 Fusion Strategy

In order to increase the effectiveness and robustness of the skin detection algorithm, a fusion strategy is proposed by integrating the two incoming single features into a combined single representation. Both models will vote for classification of skin and non-skin pixels. This can be done by using product rule to both models.

Let \( D_{\text{hist}}(S_t, Z) \) and \( D_{gmm}(S_t, \mu, \Sigma) \) denote the matching results produced by the smoothed 2-D histogram \( Z \) and Gaussian model \( g(c|\mu_i, \Sigma_i) \) respectively. The combined matching results \( D(S_t) \) using the fusion rules can be obtained as follows:

\[
http://www.iaeme.com/IJECET/index.asp
\]

editor@iaeme.com
\[ D(S_t) = \Gamma \{ D_{\text{hist}}(S_t, Z), D_{\text{gmm}}(S_t, \mu, \Sigma) \} \] (14)

Where \( \Gamma \) is the selected fusion rule, which represents the product \( \odot \).

### 4. EXPERIMENTAL RESULT

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Live Image</th>
<th>Crop Face Image</th>
<th>2D Histogram</th>
<th>GMM</th>
<th>Proposed Method</th>
<th>Skin Colored</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1.jpg" alt="Image 1" /></td>
<td><img src="image2.jpg" alt="Image 2" /></td>
<td><img src="image3.jpg" alt="Image 3" /></td>
<td><img src="image4.jpg" alt="Image 4" /></td>
<td><img src="image5.jpg" alt="Image 5" /></td>
<td><img src="image6.jpg" alt="Image 6" /></td>
</tr>
<tr>
<td>2</td>
<td><img src="image1.jpg" alt="Image 1" /></td>
<td><img src="image2.jpg" alt="Image 2" /></td>
<td><img src="image3.jpg" alt="Image 3" /></td>
<td><img src="image4.jpg" alt="Image 4" /></td>
<td><img src="image5.jpg" alt="Image 5" /></td>
<td><img src="image6.jpg" alt="Image 6" /></td>
</tr>
<tr>
<td>3</td>
<td><img src="image1.jpg" alt="Image 1" /></td>
<td><img src="image2.jpg" alt="Image 2" /></td>
<td><img src="image3.jpg" alt="Image 3" /></td>
<td><img src="image4.jpg" alt="Image 4" /></td>
<td><img src="image5.jpg" alt="Image 5" /></td>
<td><img src="image6.jpg" alt="Image 6" /></td>
</tr>
<tr>
<td>4</td>
<td><img src="image1.jpg" alt="Image 1" /></td>
<td><img src="image2.jpg" alt="Image 2" /></td>
<td><img src="image3.jpg" alt="Image 3" /></td>
<td><img src="image4.jpg" alt="Image 4" /></td>
<td><img src="image5.jpg" alt="Image 5" /></td>
<td><img src="image6.jpg" alt="Image 6" /></td>
</tr>
<tr>
<td>5</td>
<td><img src="image1.jpg" alt="Image 1" /></td>
<td><img src="image2.jpg" alt="Image 2" /></td>
<td><img src="image3.jpg" alt="Image 3" /></td>
<td><img src="image4.jpg" alt="Image 4" /></td>
<td><img src="image5.jpg" alt="Image 5" /></td>
<td><img src="image6.jpg" alt="Image 6" /></td>
</tr>
<tr>
<td>6</td>
<td><img src="image1.jpg" alt="Image 1" /></td>
<td><img src="image2.jpg" alt="Image 2" /></td>
<td><img src="image3.jpg" alt="Image 3" /></td>
<td><img src="image4.jpg" alt="Image 4" /></td>
<td><img src="image5.jpg" alt="Image 5" /></td>
<td><img src="image6.jpg" alt="Image 6" /></td>
</tr>
</tbody>
</table>

The methodology discussed earlier is simulated using Open CV (Microsoft Visual studio 2010). The results are obtained using programming in c#.net. For experiment purpose and performance analysis dataset and databases that are created by taking live images of human being. For resulting purpose I have considered images with face regions. Result and conclusion are drawn as follows.

**Table 1** Skin & Non skin pixel count of various methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Skin Pixel Count</th>
<th>Non-Skin Pixel Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 D Histogram</td>
<td>11424</td>
<td>65376</td>
</tr>
<tr>
<td>GMM</td>
<td>12366</td>
<td>64434</td>
</tr>
<tr>
<td>Proposed method</td>
<td>14154</td>
<td>62646</td>
</tr>
</tbody>
</table>
Table 2 Comparison between fusion & non fusion approach using live image dataset

<table>
<thead>
<tr>
<th>Sr. no</th>
<th>Image</th>
<th>Method</th>
<th>TPR</th>
<th>FPR</th>
<th>FNR</th>
<th>Fscore</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Live image 1</td>
<td>2-D Histogram</td>
<td>0.7616</td>
<td>0.0578</td>
<td>0.2385</td>
<td>0.8368</td>
<td>0.9294</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GMM</td>
<td>0.8244</td>
<td>0.04262</td>
<td>0.1785</td>
<td>0.8837</td>
<td>0.9508</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Proposed method</td>
<td>0.9436</td>
<td>0.0136</td>
<td>0.0564</td>
<td>0.9641</td>
<td>0.9857</td>
</tr>
</tbody>
</table>

Graph 1 Graph for Various Methods Comparison

6. CONCLUSION
Fusion framework based on histogram and Gaussian model has been proposed to detect human skin automatically in images. In this proposed method haar classifier is used to detect face from the image & it will create a database to detect different human skin tone detection using that database comparison. In this work an approach of color image segmentation which is based on the analysis of 2D histogram using HSV space. The proposed approach reduces computational costs as no training is required, and it improves the accuracy of skin detection despite wide variation in ethnicity and illumination. This is the first method to employ fusion strategy for this purpose. Thus a dynamic approach, with a fusion framework based on 2-D histogram and Gaussian model has been proposed for automatic human skin detection in images. Experimental results show that the proposed method achieved satisfactory performance; it reduces FPR & FNR and increases TPR, Precision & F-score at the same time. The obtained result of live images compared to the other methods.

REFERENCES


