Study on face detection based on an improved Gaussian skin color model

Lingli Tan
School of Computer Science, Wuhan Donghu University, Wuhan, China
Email: tll@wdu.edu.cn

Abstract: An improved Gaussian skin color model is proposed for face detection applications. It overcomes the issue that the traditional models become ineffective in extracting the skin color region in the presence of diverse skin colors and illuminations. It is a two-dimensional gaussian mixture model based on H-SV (i.e., an improved HSV) and YCbCr color spaces. Experimental results from several face image databases show that the detection accuracy is improved by 35.8% and 6.3%, respectively, compared to a traditional gaussian skin model and Adaboost algorithm.

1. Introduction
In recent years, skin color information has been increasingly used for face detection and tracking. It is useful as the skin color is a relatively constant face feature. In [1], a gaussian model is applied for the feasibility study to detect human body skin color area. The study focuses on the color space distribution, computation speed, and especially for face images with a low resolution taken from a distance. AdaBoost algorithm is used as well. It reveals that the method is effective under limited variations of skin color and lighting environment. A gap still exists for applications with diverse skin colors from different race groups and various lighting conditions. The detection accuracy is affected under these more realistic scenarios.

This paper proposes a method based on improved HSV and YCbCr color spaces and a two-dimensional gaussian mixture model to overcome the above challenge. The conventional HSV space and the YCbCr space mainly use chrominance components for detection and tracking, greatly influenced by the light luminance component. The improved HSV can adapt to varying brightness and eliminate the differences in different ethnic groups' skin color. To best take advantage of the color spaces, a gaussian mixture model is applied to enable remote face detection from low-resolution images.

2. The Proposed Skin Color Model

2.1. H-SV skin color model
The human eye perception of color mainly focuses on three key elements: Hue, Saturation, and Value (or brightness). HSV color space is proposed in 1978 by Alvy Ray Smith, a pioneer computer scientist. It is an alternative representation of RGB color space with a nonlinear transform. H represents the chromaticity, with an angle from 0 to 360 degrees between the horizontal axes. S and V, represents the saturation and brightness, respectively. The conversion relationship between RGB and HSV is shown below.

\[ H(i) = \left[ F(i) + H(i) \right] \times 60 \]  

\[ F(i) = \begin{cases} 
\frac{\text{Max}(i) - \text{Mid}(i)}{\text{Max}(i) - \text{Min}(i)}, & \text{(K(i) is an odd number)} \\
\frac{\text{Mid}(i) - \text{Min}(i)}{\text{Max}(i) - \text{Min}(i)}, & \text{(K(i) is an even number)}
\end{cases} \]

If R>G>B, K(i)=0; when G>R>B, K(i)=1; when G>B>R, K(i)=2; when B>G>R, K(i)=3; when B>R>G, K(i)=4; when R>B>G, K(i)=5.

\[ S(i) = \begin{cases} 
\text{Max}(i), & \text{(Max}(i)=0) \\
\frac{\text{Max}(i) - \text{Min}(i)}{\text{Max}(i)}, & \text{(Max}(i)>0)
\end{cases} \]

\[ V(i) = \text{Max}(i) \]

where Max(i)=Maximum (R, G, B); Min(i)=Minimum (R, G, B); Mid(i)=Median (R, G, B).

It can be noted from the above transformations that H and S components in HSV color space, respectively, express the information of color difference and color purity. It is more in line with the human eye to color information of physical awareness, therefore, suitable for biometrics such as face detection. However, illumination change will directly affect the efficiency of skin color detection, relevant to the presented study. The literature [2] proposed an improved HSV model that can adapt to illumination changes, called the H - SV model. It can effectively detect the face skin color area under different light conditions.

It should be noted that it is found in [3] that the V component is not necessary for the face regions segmented from the background environment. The method with H-S can effectively enhance the anti-jamming of the external environment. As long as the face color is with suitable H and S components, it can be used to detect the face effectively.

We respectively apply H - S color space and color of H - SV space in the experiment validation, as shown in figure 1. Figure 1(a) presents three different lighting conditions of the same piece of facial skin image; Figures 1(b) and (c) respectively show the distribution of human face skin color in H - S and H- SV space under different lighting conditions. It can be noted from the figure, in H - S space, facial skin pixels S component is almost evenly distributed between [0.2, 1], and in the H - SV space, facial skin pixels SV component is distributed within a limited range of [0.2, 0.6]. The experimental results show that, compared with H - S space, H - SV space has more features.

**Figure 1 Distributions of skin pixels under different illumination conditions in H-S and H-SV space.**

(a) skin pixels under different illumination conditions; (b) skin pixel distribution in H-S color space; and (c) skin pixel distribution in H-SV color space.

### 2.2 The traditional gaussian skin color model

YCbCr color space is often used in face detection. The principle of the traditional color Gaussian model [4] is: the different color of skin in YCbCr color space, the luminance component ignores the effects of cases, the chrominance component, the distribution of Cb and Cr consistent approximation of the two-dimensional Gaussian distribution. Based on each pixel point's color image, calculating the similarity to the pixel belongs to the possibility of skin color area size.
Two-dimensional Gaussian skin model $G = (m, C)$, where $m$ is the average, $C$ is the covariance matrix, $X_i = (C_b, C_r) \text{T}$ with the value of training sample color pixels in the $I$. $n$ is the number of skin color pixels in the training sample.

$$m = E(X) = \frac{1}{n} \sum_{i=1}^{n} x_i$$  \hspace{1cm} (4)
$$C = E[(X - m)(X - m)^{\text{T}}] = \frac{1}{n} \sum_{i=1}^{n} (x_i - m)(x_i - m)^{\text{T}}$$  \hspace{1cm} (5)

The Gaussian skin model can obtain each pixel's probability belonging to the color image and the output color likelihood image. Its probability calculation is:

$$p(C_b, C_r) = \exp \left[-0.5(x - m)^{\text{T}} \frac{1}{C} (x - m) \right]$$  \hspace{1cm} (6)

2.3. Two-Dimensional Mixed Gaussian Skin Color Model

According to the above analysis, the YCbCr color model and the improved H-SV skin model are better in terms of face detection performance, especially the H-SV skin color model, to eliminate the influence of illumination change of face detection.

To give full play to the performance advantages of both, we introduce the two-dimensional Gaussian Mixture model (Gaussian Mixture Models, GMMs) [5] for the human face skin color model. It is used to estimate the parameterized modelling. The probability calculation is shown in (7).

$$p(\lambda, \theta) = \sum_{i=1}^{K} \alpha_i \frac{1}{(2\pi)^{d/2}} \exp \left[-0.5(\lambda - m)^{\text{T}} \Sigma^{-1}_{i}(\lambda - m) \right]$$  \hspace{1cm} (7)

where $\theta = \{\alpha_i, m_i, \Sigma_i\}_{i=1}^{K}$; $K$ is the order to mix, which is 2; $\alpha_i$ is the single Gaussian model in each group held by weight, $\sum_{i=1}^{K} \alpha_i = 1 \left(0 \leq \alpha_i \leq 1 \right)$; $m$ is the statistical average of $\lambda$; $d$ is the single Gaussian model number of parameters, with a value of 2; $\Sigma_i$ is the covariance matrix columns of $\lambda$.

Two groups of the single Gaussian model are applied respectively for H-SV and H/SV in YCbCr space components and Cb/Cr component modeling. They are given by (8) and (9).

$$\lambda_{\text{H-SV}}= (H(x,y), SV(x,y))$$  \hspace{1cm} (8)
$$\lambda_{\text{CbCr}}= (Cb(x,y), Cr(x,y))$$  \hspace{1cm} (9)

Accordingly, two groups of statistical mean and covariance matrix of a single Gaussian model can be used ($m_{\text{H-SV}}, \Sigma_{\text{H-SV}}$) and ($m_{\text{CbCr}}, \Sigma_{\text{CbCr}}$).

The final skin second-order Gaussian mixture model is given by:

$$p(\lambda) = \alpha_1 \frac{1}{(2\pi)^{d/2} \Sigma_{\text{CbCr}}} \exp \left[-0.5(\lambda - m_{\text{CbCr}})^{\text{T}} \Sigma_{\text{CbCr}}^{-1}(\lambda - m_{\text{CbCr}}) \right] +$$
$$\alpha_2 \frac{1}{(2\pi)^{d/2} \Sigma_{\text{H-SV}}} \exp \left[-0.5(\lambda_{\text{H-SV}} - m_{\text{H-SV}})^{\text{T}} \Sigma_{\text{H-SV}}^{-1}(\lambda_{\text{H-SV}} - m_{\text{H-SV}}) \right]$$  \hspace{1cm} (10)

By using the Expectation Maximization (Expectation Maximization, EM) [5] algorithm to estimate the parameters of the model above, through a large number of human face skin color samples with the color of skin after statistical learning samples, the relevant parameters are obtained as follows:

$$m_{\text{H-SV}}= [0.0617, 0.3211] \text{T}$$
$$\Sigma_{\text{H-SV}}= \begin{bmatrix} 0.0032 & -0.0201 \\ -0.0201 & 0.0084 \end{bmatrix} \text{T}$$
$$m_{\text{CbCr}}= [110.3221, 142.5251] \text{T}$$
$$\Sigma_{\text{CbCr}}= \begin{bmatrix} 130.1621 & 10.1523 \\ 10.1523 & 280.3534 \end{bmatrix} \text{T}$$

The second-order Gaussian mixture skin color model is used to judge whether a pixel belongs to the face skin color points: firstly, the pixels are calculated separately; secondly, the H/SV components and Cb/Cr are fed to obtain a probability value in the model. If the value exceeds a specific threshold, it can detect the pixel points for skin color or as a color of skin. Here, we must go through training and learning to determine the threshold for $P_0 = 0.72$. The color extraction results are shown in figure 2.

3. Face Detection

The most representative face detection algorithm is Adaboost[6]. The algorithm combines the outputs of different weak learners with a weighted sum. Through the training process of iteration, successive
training samples are weighted step by step to get a robust classifier finally. The algorithm has a high recognition rate and fast recognition speed.

After it meets certain conditions of skin color region extraction, we will use the algorithm to detect whether the area is the face region, i.e., the candidate face area.

![Figure 2 Comparison of face skin region detection.](image)

4. Experimental Analysis

The second-order Gaussian mixture model and Adaboost algorithm are combined to form a new face detection framework. It could achieve remote and robust detection under conditions with various distance resolution. The concrete implementation steps are as follows:

1. Perform the skin color detection and segmentation based on the proposed second-order Gaussian mixture model;
2. Determine the candidate face region based on if the aspect ratio between length and width is within \([0.5, 1]\) or not. The expansion and corrosion are also applied during this step to minimize the noises and interferences.
3. Apply the Adaboost classifier for the candidate areas;
4. Judge if it is a human face using the correlation operator \([7]\) for candidate areas that exceed the threshold in terms of length and width. The correlation operator analyzes if there are eyes and mouth and similar holes, which forms a basis for the judgment.

To verify the effectiveness of the skin color model and face detection algorithm, the experiment with the face of a library for LFW, GTFD, and face image library is conducted. The library consists of 352 face images, including 1508 individuals, downloaded from the internet and taken in laboratories.

| Skin model                      | The Accuracy rate of face skin color extraction (%) |
|---------------------------------|-----------------------------------------------------|
|                                 | LFW       | GTFD      | Multiple faces | Average value |
| Traditional Gaussian Model[2]   | 84.6      | Fail      | 81.0          | 55.2          |
| Proposed skin model             | 93.9      | 87.5      | 91.5          | 91.0          |

The results of the proposed method and the traditional Gaussian model are shown in Table 1. It demonstrates that the proposed method achieves 35.8% higher accuracy for the average skin color extraction. For the case of LFW and images with multiple faces, its extraction accuracy is above 90%.
The main factors affecting the extraction errors are: 1) the color of the neighboring areas of the faces are close to that of the faces; and 2) the illumination conditions for multiple faces in a single image may vary.

In the experiment for face detection, a computer with an i7-8700 3.2GHz CPU is used. The software platform is VC6.0 and OpenCV1.0. A comparison between the proposed method and Adaboost based detection method. A total of 100 single face images in the LFW library are detected. In both methods, the face detection accuracies are above 90%. A total of 354 faces are randomly selected from 40 images in the multi-face gallery. Part of the results is shown in Figure 3. The detailed comparison data are shown in Table 2.

![Fig. 3](image)

Fig. 3(a) and Fig. 3(b) show the results based on AdaBoost algorithm and the proposed method, respectively. There are two false tests in Fig. 3(a), while the false detection area is eliminated by the skin color model proposed in this paper.

| Face detection method | Accuracy rate (%) | False rate (%) | Miss Rate (%) | Average time/ms |
|-----------------------|-------------------|----------------|---------------|-----------------|
| Adaboost              | 90.5              | 3.4            | 9.5           | 155.4           |
| Proposed method       | 96.8              | 3.0            | 3.2           | 160.3           |

As can be seen from the data shown in Table 2, the face detection accuracy of the proposed method is 6.3% higher than that of the Adaboost algorithm. The average time is 4.1% longer, which is reasonably acceptable. The main reason for the missed detection of the Adaboost algorithm is that it is sensitive to the illumination and size changes of the face, while the proposed method can effectively detect the illumination and large size faces. The difference in false detection rate is relatively small, and the detection method in this paper is better than Adaboost.

5. Conclusion
This paper presents an improved Gaussian skin color model for face detection. To a large extent, the parameter selection method overcomes the challenges in the traditional Gaussian skin color model. The proposed model is robust to skin color and illumination change by combing an improved HSV, i.e., H-SV, and YCbCr color space, and a second-order mixed Gaussian skin color model. The experimental results show that the proposed skin color model and the face detection method are effective. The detection accuracy is compared with traditional models with improved performance. It should be noted that there is still room for improvement for the proposed skin color model, especially for images with low resolutions. Future research will focus on studying super-resolution technology for the scenarios with low-resolution face images.
Acknowledgments
2019 Wuhan Donghu University Youth Fund Project

References
[1] ZONG SU. (2011) Application of Gaussian model in human skin color recognition. SME management and technology, 22:273-274.
[2] YOU Y.P., YUAN B.Z. (2017) Face detection under variable lighting. Journal of Information Processing, 3:127-134.
[3] Sobottka K, Pitas I. (2013) A novel method for automatic face segmentation, facial feature extraction and tracking. Processing of the Second International Conference on Automatic Face and Gesture Recognition, 823-829.
[4] LIU CH.J., RO Y.M., PLATANIOTIS K N. (2017) Boosting color feature selection for color face recognition. IEEE Trans. Image Processing, 20(5):1425-1434.
[5] RichardO. Duda, PeterE. Hart. (2013) Pattern classification, China Machine Press, BeiJing pp.201-230.
[6] LIU CH, ZHANG J. D., ZENG H. (2010) Infinite dimension AdaBoost algorithm based on support vector machine and its application. Chinese Journal of Scientific Instrument, 31(4):764-769.
[7] RUAN Q.Q. (2013) Digital image processing, Electronic Industry Press, BeiJing. pp.178-194.