A New Bayesian Classifier for Skin Detection

Sajad Shirali-Shahreza
Computer Engineering Department
Sharif University of Technology
Tehran, IRAN
shirali@ce.sharif.edu

M. E. Mousavi
IT Department
Nasr Electronics Research Center
Tehran, IRAN

Abstract

Skin detection has different applications in computer vision such as face detection, human tracking and adult content filtering. One of the major approaches in pixel based skin detection is using Bayesian classifiers. Bayesian classifiers performance is highly related to their training set.

In this paper, we introduce a new Bayesian classifier skin detection method. The main contribution of this paper is creating a huge database to create color probability tables and new method for creating skin pixels data set. Our database consists of about 80000 images containing more than 5 billions pixels. Our tests shows that the performance of Bayesian classifier trained on our data set is better than Compaq data set which is one of the currently greatest data sets.

1. Introduction

Computer vision is one of the active computer science fields during the history of computer science. It attracts researchers from early days of computer history till now.

Automatic skin detection is one of primitive computer task which is used in different applications such as face detection [1] and adult content filtering [2]. During past years, different skin detection methods are proposed. These method can be divided into three main categories [3]: explicitly defined skin regions, non-parametric methods and parametric methods.

In explicitly defined skin detection methods, such as [1] and [4], a series of rules such as "If Red>100 and Green<150, then pixel with color RGB is skin" are created. If a skin pixel color matches one of these rules, it will be marked as skin pixel. The main advantage of these methods is their speed, while these methods usually have poor performance.

The other two categories are machine learning approaches. In these methods, the skin detection problem is defined as a learning problem and the skin detector is trained by a training set. The goal of these methods is usually to calculate p(skin|RGB), which is the probability that a pixel with color RGB be a skin pixel. After learning the p(skin|RGB), these values are used to create a skin probability map for image, such as Figure 1(b). In skin probability map, the probability that each pixel be a skin pixel is shown.

This probability map is used to create the skin map which shows skin and non-skin pixels, such as Figure 1(c). The common and simple way to create skin map from skin probability map is using a threshold. The pixels with value greater than threshold are considered as skin pixels and other pixels are considered as non-skin pixels. The threshold can be constant [5] or calculated adaptively for each image [3].

In non-parametric methods, such as [3] and one of the methods proposed in [5], the p(skin|RGB) is estimated base on training data set without considering any model for probability model. The most common way to do this is using Bayesian classifiers. More details are explained in section 2, where we describe our skin detection method which is a non-parametric method. These methods usually have high true positive and low false positive rates. The main drawback of these methods is that storing p(skin|RGB) table (also known as Look Up Table (LUT)) requires high memory. A solution to this problem is using a smaller space color, such as a color space with 32^3 colors instead of 256^3 colors [5].

In parametric methods such as [2] and one of methods in [5], it is assumed that the desired probability function p(skin|RGB) has a special model. Methods in this category usually use the Gaussian Mixture Model (GMM) to approximate the p(skin|RGB). In the model creation phase, the parameters of the model, such as the mean and variance of Gaussian models, are estimated from training data set using algorithms such as EM [5].

The 3rd Intetnational Conference on Innovative Computing Information and Control (ICICIC'08)
978-0-7695-3161-8/08 $25.00 © 2008 IEEE

For creating the skin probability map, the skin probability p(skin|RGB) is calculated for each pixel using the model created in learning phase. The main advantage of these methods in comparison to non-parametric methods is that they require a small space for storing model. But they require more time to compute the skin probability model. These methods usually have higher true positive than non-parametric methods but their false positive is also greater.

As we see, most of the skin detection methods require a training data set to learn and the performance of them depends on the size and quality of the data set. Unfortunately there is no standard data set for this purpose. The most common data set which is currently available and is used is the data set created in [5] and is freely available to university researchers.

This data set which is known as Compaq data set, consists of more than 13000 images. The images are divided into two sets, about 9000 images without any skin region and about 4500 images which contains skin regions. For each of the images with skin regions, the skin regions are labeled by hand. We must note that the skin pixel labeling is not down completely. It means that they don't try to label all of the skin pixels; instead they try to label most important skin regions in each image. This data set has more than 800 millions pixels in images without skin regions and about 80 millions skin pixels in images with skin regions.

Although the Compaq data set is a good data set, it has some problems. We will discuss some of them in section 3. In this paper, we try to build a new huge data set and avoid problems exist in Compaq data set.

We will describe our new skin detection algorithm in section 2. In section 3, we discuss some of the problems of Compaq data set and explain how we create our data set to solve these problems. Our experimental results are mentioned in section 4 and the final section is the conclusion.

2. Our skin detection method

The skin detection method which we porpose in this paper is a Bayesian skin detection method. In probability based skin detection methods, we try to calculate the p(skin|RGB) which is the probability that the pixel is a skin pixel if the pixel color is RGB. In the Bayesian skin detection methods, the p(skin|RGB) is calculated using the Bayes theorem:

$$p(skin \mid RGB) = \frac{p(RGB \mid skin)\, p(skin)}{p(RGB)} \quad (1)$$

To calculate the p(skin|RGB) using equation (1), we need to calculate three probabilities: p(RGB|skin) which is the probability that the color of a skin pixel is RGB, p(skin) which is the probability that a pixel is a skin pixel and p(RGB) which is the probability that the color of a pixel is RGB.

After calculating the skin probability of each pixel, usually a threshold is used to create the final skin map and pixels with skin probability greater than the threshold are set as skin pixels. So the accuracy is independent of p(skin), because if someone else use 2*p(skin) instead of p(skin), then using 2*threshold instead of threshold yields the same output.

In addition we must note that calculating p(skin) is very difficult and we did not find any good method for estimating it in the literature. A simple way to estimate it is to number of skin pixels by total number of pixels in the training data set. The problem of this estimation is that in data sets, the ratio between total number of images and images containing skin is not equal to ratio of images containing skin to all images.

For the reasons we mentioned here, we ignored estimating the p(skin) and the main work of our method is to estimate p(RGB) and p(RGB|skin).

To estimate p(RGB|skin), we can create a data set of sample skin pixels and then use the ratio of pixels with color RGB to the number of pixels in data set as an estimation of p(RGB|skin). This is the usual approach used in Bayesian skin detection methods. For example, we can use the skin pixels data set of Compaq data set and use it to estimate p(RGB|skin), like [5]. We use the same method to estimate p(RGB|skin) in our method.

If we have a large enough data set of different images, we can estimate the p(RGB) by the ratio of pixels with that color to the total number of the pixels. In the method proposed in [5] and other method which uses the Compaq data set such as [6] and [3], the p(RGB) is calculated using equation (2):

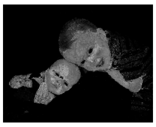
$$\begin{aligned} p(RGB) = \; & p(RGB \mid skin)\, p(skin) + \\ & p(RGB \mid \neg skin)\, p(\neg skin) \end{aligned} \quad (2)$$

In this equation, p(RGB|skin) is calculated based on the skin pixels of the images in skin images data set and the p(RGB|¬skin) is calculated based on the pixels of the images in the non-skin images data set. This approach has two problems. First it needs to know the p(skin). Second, it is calculating p(RGB|skin) and p(RGB|¬skin) on two different data sets and then combines them, while the equation (2) is valid if all of the probabilities are calculated on one data set. We try to solve this problem by creating a skin pixel data set as a part of total data set. We will describe the details of our data set in next section. Our skin pixel data set is part of our full data set. The full data set is used to estimate p(RGB) by calculating the ratio of pixels with color RGB to total number of pixels in data set.

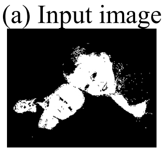
The 3rd Intetnational Conference on Innovative Computing Information and Control (ICICIC'08)
978-0-7695-3161-8/08 $25.00 © 2008 IEEE

A sample output of our skin detection method is shown in Figure 1. Figure 1(a) shows the sample input image to our method. Figure 1(b) shows the probability map (the value p(skin|RGB) for each pixeil). Figure 1(c) is the output skin map, which shown the pixels with p(skin|RGB) greater than defined threshold. Figure 1(d) is pixels of the input image marked as skin pixels.

(a) Input image



(b) Skin probability



(c) Skin map

(d) Skin regions

Figure 1. Sample outputs of our method

3. Our data set

The main application of our skin detection algorithm is adult content filtering. So we create a huge data set of adult images. We run a crawler to gather adult images from web. Among the outputs of this crawler, we select about 80000 images. The image size is usually about 800x600 or 1024x768. All of the images are true color (24 bit) JPEG images. This data set contains more than 5 billions of pixels which is 5 times greater than Compaq data set [5] which is one the greatest data sets currently available.

We use a new method to create the skin pixels data set. The method which is usually used for creating skin pixels data set is to label all of the skin pixels of images by hand and create the skin map. For example, the Compaq data set is created using this method.

One of the problems of this approach is that this is a difficult work and need a lot of time. In addition, using this approach, usually the main skin regions of images are marked and skin regions with certain features, such as skin regions which are in the shadow of another object are ignored. This problem is visible in the Compaq data set and the skin detection methods which use this data set such as [6].

To solve these problems, we propose a new approach for creating skin pixel data sets. The main object of our skin pixel data set is to have samples of different types of skins. Usually it is more important to have more samples of different types of skin instead of having skins samples according to color probabilities.

We usually calculate the p(skin|RGB) to decide whether the pixel is skin or not using a threshold. So we don't need the exact probability of p(skin|RGB), instead we can have an estimation of it such as q(skin|RGB) such that if p(skin|RGB)>threshold, then q(skin|RGB)> threshold and vice. If we can calculate such an approximation, then the results will be similar.

When we are defining q(skin|RGB), we can define it so that we can calculate it more accurately. For example, in equation (1), we need p(RGB|skin) and p(RGB) to calculate p(skin|RGB). If p(RGB) is very low and p(skin|RGB) is high, it means that the RGB color is a rare color, but pixels with this color are usually skin pixels. For such color, the probability that such colors occur in the skin pixel data set in low. In addition, they are more sensitive to labeling error, for example they are only 100 such pixels in data set, if 30 of them is missed in an image, then our error on labeling them is high, while the error resulted of missing 1000 pixels among 100000 pixels of a frequent color is less severe. Based on these assumptions, we create our skin data set so that it contains different types of skins. For example if there is a large region of skin in an image, instead of selecting all of it, we try to select different parts of it with special features.

To create such samples, we design a program which shows the image to the user and he must select rectangular regions of image as skin parts. A sample screenshot of our program is shown in Figure 2. The black rectangles are skin regions selected by the user.

We randomly select 200 images among the entire images in the data set. Then more than 2700 skin regions are extracted from these images, similar to regions in Figure 2. These skin regions contain more than 3.6 millions of skin pixels. In selecting skin regions, we try to select skin regions with different illuminations and conditions, such as the skin of a limb which is in the shadow of another limb.

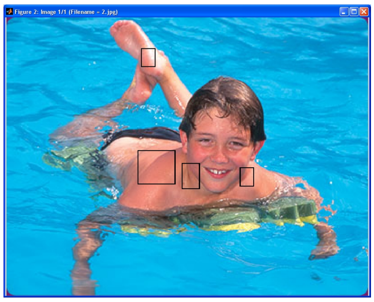


Figure 2. A screenshot of our skin data set creation program.

4. Experimental results

In this section we describe the experimental results of our skin detection algorithm. To test the performance of our method and comparing it with

The 3rd Intetnational Conference on Innovative Computing Information
and Control (ICICIC'08)
978-0-7695-3161-8/08 $25.00 © 2008 IEEE

Compaq data set, we collect 20 images containing large skin regions from internet. We try to collect image with different illumination conditions and human races. Then the correct skin map of these images is created by hand.

The skin probability map (similar to Figure 1(b)) is calculated for each image using the desired method. Then the performance of the method is calculated as a Receiver Operation Characteristics (ROC) chart. In order to create the ROC chart, we calculate the True Positive and False Positive for each image. The True Positive is the ratio of identified skin pixels to the total number of the skin pixels. The False Positive is the ratio of pixles incorrectly identified as skin pixels to the total number of non-skin pixels. The different values of True Positive and False Positive are obtained by using different values for the threshold. For each value used for threshold, the True Positive and False Positive result for each image is calculated and the mean of these results is used to draw the ROC chart.

The complete RGB true color space contains 256^3 colors. The probability p(skin|RGB) can be computed for each color in this color space or for a smaller color space. For example, we can consider the 64^3 color space in which each of color components can have 64 different levels. Smaller color spaces have advantages such as needing less memory to save probability tables and ability to generalize the color model and reduce over fitting. So it is usual to consider different color space sizes in the literature, such as [5].

To evaluate the performance of method in different color space sizes, we calculate the probabilites p(skin|RGB) in 5 different color space sizes: 256^3, 128^3, 64^3, 32^3 and 16^3. The ROC of different color space sizes is shown in Figure 3. As we can se, the result is nearly similar in different color spaces.

We compare our method with Compaq data set. We implement our method as described in section 2 in Matlab. We also implement the method describe in [5] to test Compaq data set. The ROC of two methods is compared in Figure 4. As we see, our data set results are better than Compaq data set.

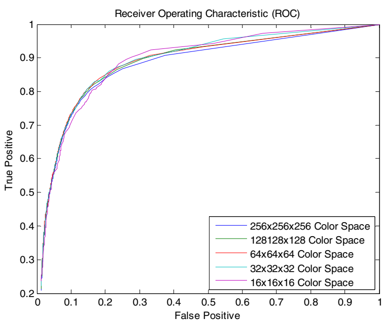


Figure 3. ROC of different color space sizes.

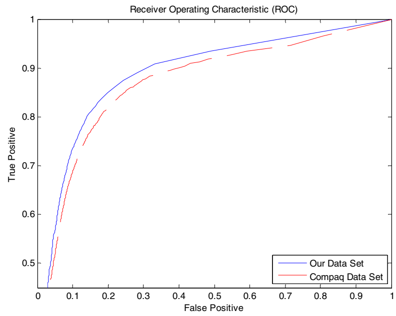


Figure 4. Comparison of our data set and Compaq data set.

5. Conclusion

Skin detection is an important process in computer vision and used in different applications such as human detection and tracking and adult content filtering. Most of the skin detection methods need a large data set to learn on it. In this paper, we present the data set we prepared for skin detection. Our data set consists of about 80000 images with more than 5 billions of pixels and a subset of selected skin pixels with more than 3.6 millions skin pixels. We use new ideas in creating this data set to solve some of problems of previous data sets, such as well known Compaq data set [5]. Our tests show that the Bayesian skin detection method trained with our data set has better performance than the one trained with Compaq data set. Because the Compaq data set used in different skin detection methods as the data set, those methods may be improved by using our data set instead of Compaq data set. This is one of the works which we are planned to do in future.

6. References

[1] R.L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, "Face detection in color images," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 24,. no. 5, pp. 696-706, 2002.
[2] W. Hu, O. Wu, Z. Chen, and Z. Fu, Maybank, S., "Recognition of Pornographic Web Pages by Classifying Texts and Images," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 29, no. 6, pp.1019-1034, 2007.
[3] M.J. Zhang, and W. Gao, "An adaptive skin color detection algorithm with confusing backgrounds elimination," *Proceedings of IEEE International Conference on Image Processing (ICIP 2005),* vol.2, pp. 390-393, 2005.
[4] M.M. Fleck, D.A. Forsyth, and C. Bregler, "Finding naked people," *Proceedings of 4th European Conference on Computer Vision (ECCV'96),* pp. 593-602, April 1996.
[5] M.J. Jones and J.M. Regh, "Statistical Color Models with Application to Skin Detection," *Cambridge Research Laboratory Technical Report,* CRL 98/11, 1998.
[6] W. Zeng, W. Gao, T. Zhang, and Y. Liu "Image guarder-An intelligent detector for adult images," *Proceedings of Asian Conference on Computer Vision,* pp. 198-203, 2004.

The 3rd Intetnational Conference on Innovative Computing Information and Control (ICICIC'08)
978-0-7695-3161-8/08 $25.00 © 2008 IEEE