GVIP 05 Conference, 19-21 December 2005, CICC, Cairo, Egypt

www.icgst.com

Adaptive Skin Color Classificator

M. Wimmer, B. Radig
Informatik IX, Technische Universität München, Germany
Boltzmannstr. 3, 87548 Garching, Germany
[wimmerm,radig]@informatik.tu-muenchen.de
http://wwwradig.informatik.tu-muenchen.de

Abstract
A lot of computer vision applications benefit from robust skin color classification. But this is a hard challenge due to the various image conditions like camera settings, illumination, light source, shadows and many more. Furthermore people's tans and ethnic groups also extend those conditions. In this work we present a parametric skin color classifier that can be adapted to the conditions of each image or image sequence. This is done by evaluating some previously know skin color pixels which are acquired by applying a face detector.
This approach can distinguish skin color from very similar color like lip color or eye brow color. Its high speed and high accuracy makes it appropriate for real time applications such as face tracking and mimic recognition.

Keywords: *skin color detection, lip color detection, face tracking, mimic recognition.*

1. Introduction
There is a widespread benefit of skin color detection in computer vision applications such as face detection, mimic recognition [1], person identification, hand gesture detection [2][7], image content filtering [3]. Skin color occupies a large subspace within the color space and the various cameras types and settings, illumination conditions and the people's tans and ethnic groups make that subspace even larger.

The need to distinguish between skin and objects colored in a similar way such as lips makes that challenge is even harder. Most detectors classify lips as skin because the difference between skin color and lip color is small. In other words, the subspace of skin color intersects the subspace of lip color. Figure 2 (middle column) shows some shortcomings of non-adaptive skin color classifiers: parts of the background are incorrectly classified as skin; lip color is often classified as skin color, and parts the persons' faces are often incorrectly detected as non-skin color. This is especially true for the images with colored persons.

Knowing the image conditions (the person's ethnic group, camera, lighting, background) in advance a more specific skin color model can be used. This will shrink the subspace of all possible skin colors and improve the classification result. Those image conditions can easily be figured out if we know some pixels to be skin color pixels. Via foreground detection you can define each foreground pixel to be a skin color pixel. We believe this approach is too vague and apply a face detector in order to acquire the skin color pixels more precisely. This requires each classified image and at least the first image of an image sequence to include a frontal view of a person.

The remainder of the paper is organized as follows: Section (2) gives a short overview of other techniques that focus on the same subject, section (3) explains our work, and section (4) demonstrates our results

2. Related work
Vezhnevets et al. give a comprehensive overview about the work within the area of skin color detection that has been done during the last decade [5]. They describe the main color spaces, the skin color models, and the classification techniques which they categorize into four groups. Our approach contributes to the group "dynamic skin distribution models". This group describes parametric skin color classifiers that can be adapted to the image conditions prior to classifying the image. Most other work in this context uses histogram based techniques in order to evaluate the global image conditions. Since most objects in an image do not influence skin color, we specify the image conditions in a different way. We acquire the a priori knowledge by using a face detection algorithm that is completely independent of skin color.

Rheg and Jones [3] use a common approach for estimating the distribution of skin color: Using labeled training data they created sixteen Gaussian kernels within the color space. Those Gaussians describe the probability of each pixel to be a skin color pixel.

Soriano et. al [4] define their skin color model as a subspace with a distinct shape within the color space. This model can be adapted to the image illumination conditions by evaluating the image's histogram.

Viola and Jones [6] created a face detector that works with gray value images and does not use skin color at all. Their detector uses a boosted cascade of classifiers that evaluate rectangular features summing up the gray values inside. That face detector basically relies on the brighter and darker parts within faces.

3. Our technique

This section explains our approach. It can be divided into three phases. First we create a parametric skin color classifier. Second, we detect some skin color pixels using a stand-alone technique and calculate the illumination conditions and the ethnic group of the person. Third we set up the parametric skin color classifier with this information.

3.1 Parametric skin color classifier

Skin color classifiers often specify some bounds that narrow down the color space (e.g. RGB) to the subspace of skin color. In order to gain resistance towards various illumination conditions such as the color of the light source and shadows a different color space is used. The chromatic color space (also referred to as "normalized RGB" [5]) uses the proportional part of each color:

$$
\begin{aligned}
base &= R + G + B \\
R_c &= R / base \\
G_c &= G / base \\
B_c &= B / base
\end{aligned}
$$

(1) Definition of the chromatic color space.

Usually B_c is omitted because its value can be calculated due to this equation: $R_c + G_c + B_c = 1$ This color space provides a commonly known way to quickly detect skin color pixels via the following classifier:

$$
\begin{aligned}
skin\_color = {} & (R_c > 0.35) \wedge (R_c < 0.5) \wedge \\
& (G_c > 0.2) \wedge (G_c < 0.7) \wedge \\
& (base > 200)
\end{aligned}
$$

(2) Non-adaptive skin color classifier.

We adapt this classifier to the illumination conditions of each image which are determined by camera settings, light source, shadows and many more parameters. Furthermore we adapt the classifier to the ethnic group of the visible person. That means we have to make the upper and lower bounds variable as you can see in (3).

$$
\begin{aligned}
skin\_color = {} & (R_c > lb_R) \wedge (R_c < ub_R) \wedge \\
& (G_c > lb_G) \wedge (G_c < ub_G) \wedge \\
& (base > lb_{base})
\end{aligned}
$$

(3) Adaptive skin color classifier.

3.2 Acquiring image conditions

The bounds of the adaptive skin color classifier can be precisely calculated if we are aware of the image conditions. That means we need to know how skin color usually looks like in this image. In order to get some skin color pixels we apply a stand-alone face detector [6] which gives us the rough position and size of the frontal view face. But since that face detector provides us with a square region around the face, we can not just assume each pixel within the entire square to be a skin color pixel. Therefore we apply a mask that indicates the probability of skin color within that square.

We previously calculated that probability mask by manually annotating skin color pixels within several images and applying the face detector. The resulting mask for skin color within the squares can be seen in figure 1. For a fast application of the mask we divided it into 24 x 24 grid and thresholded it which results in only 131 pixels that need to be considered.

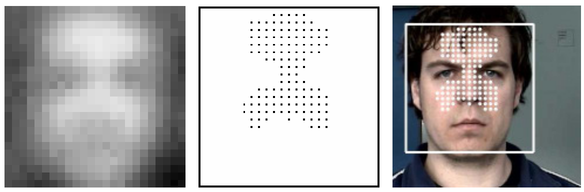

Figure 1. Left: Automatically calculated mask indicating the probability of skin color within a detected face region.
Brighter color means higher probability
Middle: Discretizing and thresholding the mask for faster evaluation
Right: Application of the mask.

Now that we have detected some skin pixels inside of the image we can calculate the image conditions (mainly illumination and ethnic group). Since our parametric skin color classifier works with chromatic coordinates we calculate the mean and standard deviation of the chromatic coordinates of the acquired skin color pixels.

$$
\begin{aligned}
& face\_R_{c,mean},\ face\_G_{c,mean},\ face\_base_{mean} \\
& face\_R_{c,std\_dev},\ face\_G_{c,std\_dev},\ face\_base_{std\_dev}
\end{aligned}
$$

(4) Variables describing the image conditions.

3.3 Setting up the classifier

Now that we have calculated how the average skin color pixels look like within the current image, we can calculate the classifier's bounds. We provided linear regression with the face_xxx values and the expected upper and lower bounds (lb_{XXX} and ub_{XXX} values) which we manually figured out for several images and got the formulae in (5).

$$lb_R = 0.055 + 0.75 * (face\_R_{mean} - face\_R_{std\_dev})$$

$$ub_R = -0.098 + 1.385 * (face\_R_{c,mean} + face\_R_{c,std\_dev})$$
$$lb_G = -0.597 + 2.857 * (face\_G_{c,mean} - face\_G_{c,std\_dev})$$
$$ub_G = -0.17 + 1.6 * (face\_G_{c,mean} + face\_G_{c,std\_dev})$$
$$lb_{base} = -45.26 + 0.79 (face\_base_{c,mean} - face\_base_{c,std\_dev})$$

(5) Formulae for adaptively setting up the bounds of the classifier.

4. Results

Our skin color classifier must be executed in real time because we intend to apply our classifier for face tracking, mimic recognition, and person tracking. The first online step (calculating the image conditions and adapting the classifier) has a very fast execution time which is constant. Using the discretized face mask the mean and standard deviation values have to be calculated out of about 120 pixels. For the second online step (classifying the entire images) we achieved 107 fps evaluating a 480 x 360 image sequence and 44 fps evaluating a 640 x 480 image sequence on a 1800 MHz Pentium 4 processor using OpenCV[1].

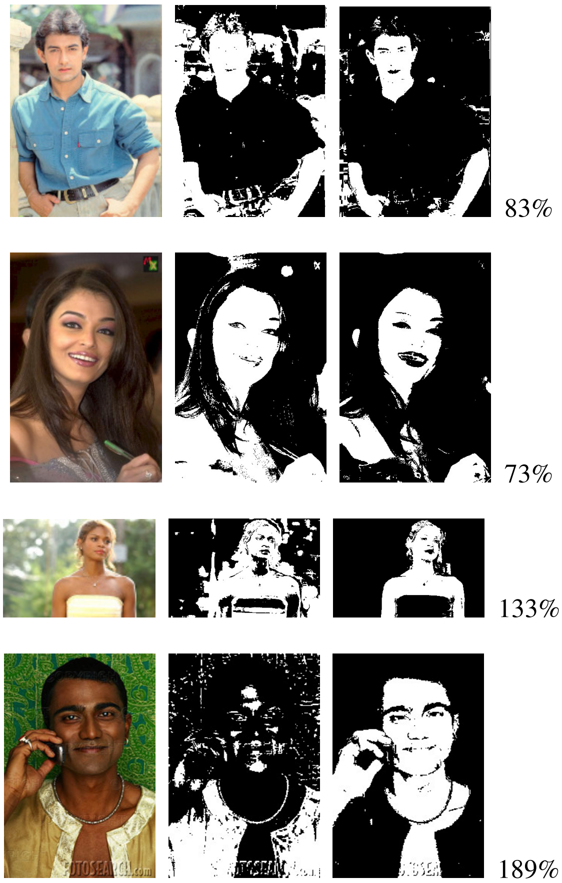

Figure 2. Comparing skin color classifiers
Left: original image,
Middle: non-adaptive classifier,
Right: adapted to illumination conditions and ethnic groups,
Numbers: improvement

Figure 2 shows a few example images of our experiments in which we applied our adaptive skin color classifier to images with various illumination conditions and people out of various ethnic groups. The number next to each row shows the decrease of incorrectly classified pixels compared to the absolute amount of skin color pixels. The classifier correctly classifies facial parts such as lips and eye brows as non-skin.

On the web there are some original images and some skin color classified images that we compared.[2]

Figure 3 shows our first attempt to classify lip color pixels. We used the same classifier assigning the upper and lower bounds with different formulae. We only applied the lip color classifier within the face region.

Figure 3. Lip color detection. The white pixels indicate the detected lip color pixels

5. Conclusion and Future Work

Skin color classification is difficult due to many issues. We created a parametric classifier that can be adapted to the image conditions and the ethnic group of the person on the image. We demonstrated that its result increases dramatically compared to a non-adaptive classifier, especially in the case of poor lightning conditions and colored people. As a special benefit facial parts such as eye, brows, lips, and teeth are detected correctly as non-skin color objects. Because of its real time capability this technique can support head and person tracking purposes.

We are currently working on extending our work towards classifying lip color, eye brow color, and other facial parts. In the future we plan to integrate this approach within our applications for face tracking and mimic recognition.

6. References

[1] S. Fischer, S. Döring, M. Wimmer, A. Krummheuer. Experiences with an Emotional Sales Agent. Workshop on Affective Dialogue Systems (ADS04) S. 309-312, Kloster Irsee, Germany, Juni 2004.

[2] S. Hämmerle, M. Wimmer, B. Radig, M. Beetz. Sensor-based Situated, Individualized, and Personalized Interaction in Smart Environments. Workshop on Situierung, Individualisierung, and Personalisierung, Informatik2005, Bonn, Germany, 19.9.2005.

[3] M. J. Jones, J. M. Rehg. Statistical color models with application to skin detection. Proc. IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, 1999, 274-280.

[1] www.intel.com/research/mrl/research/opencv/

[2] www9.in.tum.de/people/wimmerm/skin_color_images/

[4] M. Soriano, S. Huovinen, B. Martinkauppi, M. Laaksonen. Skin detection in video under changing illumination conditions. In Proc. 15th International Conference on Pattern Recognition, vol. 1, 839–842. 2000.

[5] V. Vezhnevets, V. Sazonov, A. Andreeva. A Survey on Pixel-Based Skin Color Detection Techniques, Graphics and Media Laboratory, Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow, Russia. 2003.

[6] P. Viola, M. Jones. Rapid object detection using a boosted cascade of simple features. CVPR, 2001.

[7] M. Zahedi, D. Keysers, H. Ney. Appearance-Based Recognition of Words in American Sign Language. RWTH Aachen, Germany. 2005.