www.icgst.com

Adaptive Skin Color Classificator

Matthias Wimmer, Bernd Radig
Fakultät für Informatik, Technische Universität München,
Boltzmannstr. 3, 85748 Garching, Germany,
[wimmerm,radig]@cs.tum.edu,
`http://www9.informatik.tu-muenchen.de`

Abstract

Skin color is an important feature of faces. Various applications benefit from robust skin color detection. Skin color may look quite different, depending on camera settings, illumination, shadows, people's tans, ethnic groups. That variation is a challenging aspect of skin color classification.

In this paper, we present an approach that uses a high level vision module to detect an image specific skin color model. This model is representative for the context conditions within the image and is used to adapt dynamic skin color classifiers to it. This approach distinguishes skin color from very similar color like lip color or eyebrow color. Its high speed and accuracy makes it appropriate for real time applications such as face model fitting, gaze estimation, and recognition of facial expressions.

1 Introduction

As computer systems are becoming more and more complex, we need more sophisticated human machine interfaces. Sensor equipped systems make natural interaction mechanisms more feasible. Cameras perceive the user's face and infer his gaze, mood, focus of interest, intention, and so on. Therefore the interpretation of facial expressions has become such an important research focus in the area of computer vision during the last decade.

Unfortunately, perceiving information about a face in natural scenes is a very challenging task because a widespread set of side conditions concerning the person and the context has to be taken into account. People have different hair, different tans, different ethnic groups, beards, glasses, and so on. Context conditions include illumination, camera type and camera settings. Furthermore, the algorithms must not only work with a small set of persons but with any person who is previously unseen. Referring to this issue Gross et al. [4] show that face models and fitting algorithms that are adapted to a specific face result in a much higher robustness than general ones. Since this adaptation can not be done in natural and unconstrained scenes, the same benefit is achieved via a normalized representation of a face image. That representation must reject the person specific and context specific information and keep the relevant information such as the position of the head and the facial parts. In this paper we propose skin color to be that kind of normalized representation of a human face.

Because of the mentioned person specific and context specific variations, skin color looks very differently throughout the images. In order to robustly determine the skin color pixels we adapt skin color classifiers to those image conditions. A programmer is able to do that adaptation manually by evaluating the image and the visible face, but most face interpretation systems must run autonomously. We present an approach that automatically obtains the image specific conditions of skin color and show how to adapt skin color classifiers.

Figure 1: left: deformable face model, middle: skin color image, right: correctly fitted face model.

Face model fitting approaches that use skin color as the primary feature benefit from accurately determining skin color, see Figure 1. Those algorithms need not be robust towards many kinds of variation because the skin color detector is expected to handle them. They determine the parameter values of the face model that describe the content of the image best. Those parameters represent the position and the deformation of the face model. In a subsequent step those values are used to infer further information, such as the mimic of the person, the focus of attention, the level of fatigue, support for natural language understanding, and so on.

In this paper, we propose to approach the problem of fitting curve models to faces by decomposing it into the problem of adaptively detecting skin color and then use the classified skin color as the primary feature for curve fitting. The main contributions of this paper are the following ones:

1. We propose a mechanism that autonomously perceives a person specific and context specific skin color model. That mechanism makes use of a face detector that detects square regions around human faces. It extracts a small number of skin color pixels within those facial regions in order to set up a skin color model that is representative for the person and context conditions in the face.

2. We show how dynamic skin color classifiers can automatically be adapted to the image conditions using our skin color model. Those skin color classifiers result a skin color image that represents the information from the original RGB image that is relevant for fitting a contour model. Model fitting techniques that use this image don't have to be robust against those side conditions which makes them more accurate.

In the remainder of the paper we proceed as follows. Section 2 categorizes previous approaches on skin color detection. Section 3 explains our approach for image specific skin color detection. Section 4 describes the benefit for applications that rely on face model fitting, and Section 5 depicts the accuracy of our approach.

2 Previous work on skin detection

Since skin color provides an important source of information to various computer vision applications, a lot of research has been done in that area. In this paper we focus on pixel based skin color detection which aims at developing a decision rule that categorizes the color of a single pixel as skin or non-skin. Vezhnevets et al. [10] give a comprehensive overview of recent work within that area. They describe the utilized color spaces and skin color models and categorize the detection techniques.

Non-parametric skin color distribution modeling (NSDM) defines the skin category (*skin* or *non-skin*) for every color in color space separately. It is often referred to as Skin Probability Map (SPM), assigning a probability value to each point of a discretised color space. It performs at very high speed but it requires a vast amount of memory (e.g. RGB: 256^3 entries). Its accuracy relies on a huge and comprehensive training set. In order to reduce memory requirements common approaches use a two-dimensional color space or subdivide the color space into clusters [1, 2].

Parametric skin color distribution modeling (PSDM) uses a cluster with a distinct shape within color space. Skin color is expected to be located within this cluster. The required memory is limited to the parameters that specify the shape. Usually execution time rises and accuracy lowers, because it depends both on the chosen shape and the chosen color space. Common approaches model skin color distribution via a single Gaussian or a mixture of Gaussians [6, 5].

Explicit definition of the skin color cluster (EDSC) uses a set of rules that explicitly define the skin color cluster. Memory requirements are limited by the chosen rules but the challenge is to find adequate decision rules. Rule induction algorithms often rely on appropriate features that are well associated to the skin category. Gomez et al. [3] propose a rule deduction mechanism that uses a converted color space.

Assuming a huge and representative training set NSDM is the most accurate approach because it is suitable for any distribution of skin color in any color space. PSDM and EDSC are only able to approximate the real distribution via their shape or their rules. Those two approaches can be converted into NSDM for adopting NSDM's other advantage, its runtime performance. Note that this conversion does not improve the accuracy of skin color detection.

A further approach, *dynamic skin color distribution modeling* (DSDM), extends PSDM and EDSC by adding further dependency to the classification process. Those classifiers additionally depend on the context conditions of the processed image rather than on the pixel's color only. Image conditions consist of the camera type, camera settings, illumination, the visible person, etc. Those classifiers adapt the shape of PSDM or the rules of EDSC to the processed image, which improves the skin detection accuracy. Soriano et al. [9] define their skin color cluster within the chromatic color space. Its shape looks like the crescent of the moon and they call it *skin locus*. It is camera specific but they do not provide a mechanism that automatically determines the skin locus. Furthermore the skin locus does not take person specific and context specific side conditions into account. In this paper we demonstrate how classifiers of type PSDM and EDSC are adapted to the conditions of the processed image autonomously, see also [12]. We show which information to extract from the image, how it is obtained automatically, and how the classifiers are adapted.

3 Skin color detection

Skin color can not be modeled properly in the RGB color space [10]. In this representation, the skin color cluster is very large and incompact. A common way to cope with that is to switch over to the normalized RGB color space (NRGB), which uses the proportional part of each component, see (1). b is omitted because it can be calculated from the other components. The color of a pixel p is represented by the corresponding color vector $c_p = (r, g, base)^T$.

$$base = R + G + B, r = \frac{R}{base}, g = \frac{G}{base}, b = \frac{B}{base} \quad (1)$$

Detecting skin color is challenging, because it occupies a large cluster within color space. Camera types and settings, illumination conditions, as well as people's tans and ethnic groups make skin color vary significantly. However, within one image skin color looks similarly, because most of the above conditions are fixed. Our image specific skin color model describes those conditions. Using this model, dynamic skin color classifiers are adapted to the image conditions, which improves their classification accuracy.

Previous work focuses on detecting image specific skin color models via low level techniques such as color segmentation, background subtraction, or histogram prediction [8]. In order to improve the accuracy of the obtained image specific skin color model we use a more sophisticated vision module. It is a combination of the commonly known and widely accepted face detector of Viola and Jones, see Section 3.1, and an empirically obtained skin color mask, see Section 3.2. Since the face detector works with gray value images it is not influenced by the color distribution. The entire module is capable of extracting a small number of skin color pixels from the image, which are used to set up the image specific skin color model.

3.1 The face detector by Viola and Jones

Viola and Jones [11] propose a visual object detection framework that processes images very quickly while achieving high detection rates. Their approach detects previously trained objects within rectangular regions of interest (ROI) by evaluating simple features within those rectangles. Two of those features are depicted in Figure 2. The features are called haar-like features because they are computed similarly to the haar wavelet parameters. The feature value is the sum of the gray values within the black area subtracted from the sum of the gray values within the white area. Viola and Jones introduce an image representation called the *Integral Image* which allows any of their features to be computed very quickly and independently of its size. They utilize simple yet fast classifiers that evaluate those features for object detection. Through a combination of those simple classifiers in a boosted cascade high accuracy is achieved.

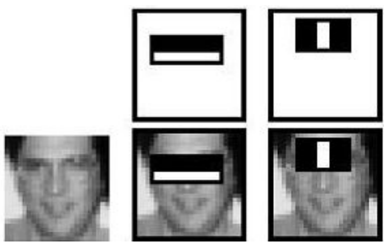

Figure 2: Evaluating two haar features within a face.

Viola and Jones demonstrate the benefit of their approach in the domain of face detection. The trained frontal face detector runs at 15 frames per second on a standard desktop computer which makes it suitable for real-time applications. We use this implementation for obtaining a ROI around a face within the processed image.

3.2 The skin color mask

A skin color mask enables a face detector to extract skin color pixels from the ROI. It is a two-dimensional matrix that specifies the probability of skin color for each pixel within the ROI after being scaled to its size. Runtime performance is increased by only taking those probability entries into account that exceed a given threshold value. A skin color mask exists for any face detector; however, we show its benefit for the Viola and Jones face detector, see section 3.1. Skin color masks are learned via training images whose skin color pixels are previously known. We use a set of K training images that show various faces which originate from well known face detectors, the Boston University skin color database [7], and various web pages.

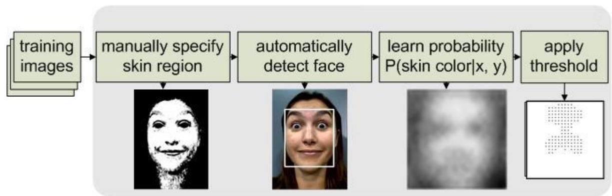

Figure 3: Acquiring a skin color mask.

A skin color mask M is an $n_1 \times n_2$ matrix whose entries are called $m_{i,j} \in [0..1]$. In this work we take $n_1 = n_2 = 24$ as a reasonable compromise between accuracy and runtime performance. We apply the face detector to each image k and receive a region of interest roi_k which is divided into $n_1 \times n_2$ cells $f_{k,i,j}$ with $1 \le i \le n_1$ and $1 \le j \le n_2$ The likelihood for skin color within cell $f_{k,i,j}$ is expressed by $s_{k,i,j}$. Finally, we calculate the entries of the skin color mask $m_{i,j}$, see (2) and (3). The entire procedure is depicted in Figure 3.

$$s_{k,i,j} = \frac{number\ of\ skin\ pixels\ in\ f_{k,i,j}}{total\ number\ of\ pixels\ in\ f_{k,i,j}} \quad (2)$$

$$m_{i,j} = \frac{1}{K} \sum_{k=0}^{K-1} s_{k,i,j} \quad (3)$$

3.3 The image specific skin color model

Skin color within one image occupies a compact cluster within NRGB but its position, size, and shape vary from image to image because of the person and context specific side conditions. In this paper we propose an image specific skin color model that is aware of those side conditions and represents the skin color pixels of an image properly.

Our image specific skin color model consists of an ellipsoid within NRGB. This ellipsoid is defined by the mean μ and the covariance matrix S of the color vectors of all skin color pixels within the image. The maximal Mahalanobis distance t between any color c_p in color space and μ defines the size of the ellipsoid. Formulae (4) and (5) depict the skin color model, where I stands for the number of skin color pixels within the image.

$$\mu = \frac{1}{I}\sum_{i=1}^{I} c_i = \begin{pmatrix} \mu_r \\ \mu_g \\ \mu_{base} \end{pmatrix} \quad (4)$$

$$\begin{aligned} S &= \frac{1}{I-1}\sum_{i=1}^{I}(c_i-\mu)(c_i-\mu)^T = \\ &= \begin{pmatrix} var_r & cov_{r,g} & cov_{r,base} \\ cov_{g,r} & var_g & cov_{g,base} \\ cov_{base,r} & cov_{base,g} & var_{base} \end{pmatrix} \end{aligned} \quad (5)$$

Our approach obtains the image specific skin color model in the following way: First, the face detector delivers a ROI around the detected face. Then the skin color mask is applied to that ROI and extracts a small number of skin color pixels. The image specific skin color model is calculated via (4) and (5) using those pixels only. Section 5.1 shows the reliability of this approach.

3.4 Dynamic skin color classifiers

A pixel based skin color classifier is a function that figures out the skin category for any pixel by evaluating a list of features f_i, see (6). The features used by non-dynamic classifiers only consist of the dimensions of the color space, e.g. $f_1 = R$, $f_2 = G$, $f_3 = B$. Therefore the cluster within color space that defines skin color is fixed. Dynamic skin color classifiers are provided with further features and the skin color cluster's position, size, and shape depend on those additional features.

$$\{skin, nonskin\} \in classifier(f_1, ..., f_n) \quad (6)$$

The following three examples show how dynamic skin color classifiers are adapted to image specific conditions. They vary their skin color cluster depending on our image specific skin color model. Their accuracy is evaluated in Section 5.2.

Cuboidal skin color cluster: Skin color is defined to be located in a cuboidal cluster in NRGB. The lower and upper bounds of the cuboid (l_r, l_g, l_{base}, u_r, u_g, u_{base}) are calculated from the parameters μ and S of the image specific skin color model, see formulae below. That approach adapts this classifier to the image conditions. The standard deviation $\sigma_i = \sqrt{var_i}$ is extracted from the covariance matrix S.

$$\begin{aligned} l_r &= \mu_r - 2\sigma_r & u_r &= \mu_r + 2\sigma_r \\ l_g &= \mu_g - 2\sigma_g & u_g &= \mu_g + 2\sigma_g \\ l_{base} &= \mu_{base} - 2\sigma_{base} & u_{base} &= \mu_{base} + 2\sigma_{base} \end{aligned}$$

Ellipsoidal skin color cluster: Skin color is defined to be located in an ellipsoidal cluster in NRGB. The calculation of the ellipsoid bases on the Mahalanobis distance m_p between μ and the color to be classified c_p, see (7). Since we take μ and S from the image specific skin color model, that approach adapts the classifier to the image conditions. If m_p is smaller than a threshold value t, the color c_p is treated to be skin color. Empirical results show that reasonable values for t vary between 6 and 25. The evaluation in Section 5.2 is performed with $t = 9.8$.

$$m_p = (c_p - \mu)^T S^{-1}(c_p - \mu) \quad (7)$$

Skin color cluster via rules: The skin color cluster is defined by a set of rules. Those rules are generated by a rule induction algorithm such as ID3, C4.5, or J4.8. The algorithm evaluates a large set of training data with annotated skin color pixels and non-skin color pixels for that task. The challenge is how to integrate the image specific skin color model into the rule induction process. In order to obtain good rules, the provided features must be decisive for the skin category. Mathematical combinations of features deliver new features which are probably more specific for the skin category. The first rule below defines a fixed skin color cluster, whereas the second rule adapts the cluster to the conditions of the image. This adaptation is given because the rule includes parameters from the image specific skin color model.

$$skin \Leftarrow (r > 0.38) \wedge (g \leq 0.33) \wedge (R > 74)$$

$$\begin{aligned} skin \Leftarrow\ & (|\mu_r - r| / \sigma_r \leq 1.7) \wedge (g \leq 0.33) \vee \\ & (|\mu_r - r| / \sigma_r \leq 1.7) \wedge (g > 0.33) \wedge (\mu_r - r \leq 0.02) \end{aligned}$$

4 Application to face model fitting

A model includes explicit or implicit knowledge about the object to interpret. Because of this knowledge the understanding of images becomes simpler and therefore model fitting is a major issue in computer vision. Model fitting algorithms adjust models correctly to objects visible in images or image sequences.

In this paper we make a distinction between model fitting and model tracking. Model fitting algorithms are applied to single images whereas model tracking algorithms are applied to image sequences. Sophisticated model tracking algorithms use information about the object from the previous image and predict its behavior. Natural constraints restrict that behavior between two subsequent images and tracking algorithms utilize those restrictions for narrowing the search space and making the execution faster and more reliable. In opposite to that, model fitting algorithms have to handle previously unseen objects and to be robust towards all kinds of variation. They have to be general enough to deal with any possible object that might be visible in an image. Often a model fitting step is executed right before a sequence of model tracking steps. Gross et al. [4] created two Active Appearance Models (AAM) to prove the challenge of this issue. One of them is general enough to fit to any person's face, whereas the other has been made specific for one person. Empirical results attested the increase of accuracy from the general AAM to the person specific AAM.

There is a widespread use of exploiting information from human faces and one basic step to achieve this goal is to fit a face model to the image. Applications that fit a face model to the skin color image rather than to the original camera image will improve robustness because of the reduced amount of variation within the face and the image. There is no dependency on illumination, ethnic group, shading, shadows, etc. The edge image of the skin color image represents the facial contour lines better than the camera image, which leads to higher accuracy.

5 Experimental results

Two aspects of our approach need to be evaluated: The acquisition of the image specific skin color model and the correctness of the adapted skin color classifiers. For evaluation purpose we use the Boston University skin color database [7], that consists of 21 image sequences taken from Hollywood movies. Their length varies between 49 and 349 frames and they show persons in natural activities such as talking, walking, or working. They include various illumination conditions and people from various ethnic groups. Each pixel is annotated with *skin*, *non-skin*, or *don't care*. Since our approach works with a face locator for frontal faces we only use those video sequences including frontal faces.

	our approach			color segmentation		
seq	μ_r	μ_g	μ_{base}	μ_r	μ_g	μ_{base}
2	0.8%	0.4%	6.5%	0.9%	0.8%	11.1%
4	1.0%	0.4%	0.5%	7.1%	5.7%	4.6%
6	0.1%	0.0%	2.8%	5.7%	1.4%	16.9%
7	1.0%	0.6%	3.9%	5.3%	5.8%	15.2%
8	1.1%	0.1%	7.3%	5.1%	1.2%	5.5%
9	0.7%	0.4%	1.2%	0.9%	2.2%	10.5%
10	0.4%	0.3%	3.2%	6.0%	1.2%	16.5%
11	0.4%	0.9%	4.1%	3.6%	1.1%	1.5%
15	0.2%	0.1%	3.1%	4.1%	1.3%	1.4%
16	0.5%	0.1%	5.0%	3.3%	1.8%	25.8%
18	0.8%	0.5%	8.1%	4.8%	1.6%	7.3%
avg	0.6%	0.3%	4.2%	4.3%	2.2%	10.6%

Table 1: The relative distance to μ.

5.1 Accuracy of the skin color model

In the following we evaluate two approaches that obtain the image specific skin color model autonomously and compare their results to the corresponding ground truth. We compare our approach to color segmentation, a low level approach. Both of them extract a small number of skin color pixels from the image and assemble the image specific skin color model. Our approach that has been described in Section 3.3 obtains those pixels via the combination of a face detector and the skin color mask. Color segmentation extracts the skin color pixels by applying a simple skin color classifier with a fixed skin color cluster. Table 1 depicts the accuracy for obtaining μ of our approach compared to color segmentation. The table illustrates the relative distances to the ground truth. The ground truth is obtained by evaluating all pixels that are marked as *skin*. Obviously our approach approximates the ground truth up to seven times closer.

5.2 Accuracy of the skin color classifiers

The dynamic classifiers from Section 3.4 are evaluated by comparing three ways of adaptation to the processed image: (a) no adaptation, (b) automatic adaptation via the image specific skin color model, and (c) optimal adaptation. The parameters of (a) are fixed and chosen such that they are optimal for the entire set of images. The parameters of (c) are chosen such that they are optimal for each single image. Those optimal values are obtained via the ground truth. Table 2 shows the accuracy of correctly classifying the skin color pixels (skin) and the non-skin color pixels (bg) in percent. det[C] denotes the determinant of the confusion matrix which is a good measure for a classifier's accuracy. One can clearly see the increase of accuracy between (a) and (b). The result of (c) stands for the upper limit that can be achieved by that kind of classifier. Figure 4 illustrates some resulting skin color images of our experiments.

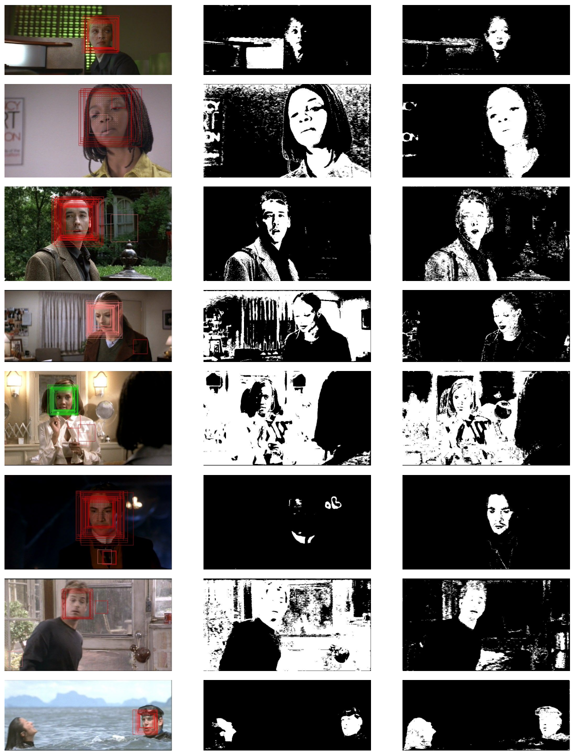

Figure 4: Comparing skin color classifiers: original image (left), fixed ellipsoid (middle), adapted ellipsoid (right)

5.3 Runtime performance

Our approach allows the necessary information to be computed partly offline. Only the following four steps need to be executed online. (S1) detect the face, (S2) apply the skin color mask, (S3) calculate the skin color model, (S4) and compute the skin color image. Nevertheless S1-S3 must only be executed once for an image sequence. S4 is the only step to be executed for each image.

S1 is executed in $O(n)$ where n denotes the number of pixels of the image, see [11]. It runs at an average of 50 ms

	fixed cuboid			fixed ellipsoid			fixed rules			adapted cuboid			adapted ellipsoid			adapted rules			optimal cuboid			optimal ellipsoid		
seq	skin	bg	det[C]	skin	bg	det[C]	skin	bg	det[C]	skin	bg	det[C]	skin	bg	det[C]	skin	bg	det[C]	skin	bg	det[C]	skin	bg	det[C]
2	84.4	86.1	70.5	99.8	26.1	25.9	88.9	92.6	81.5	62.1	99.6	61.6	71.9	99.7	71.7	82.9	59.3	42.3	84.5	95.5	80.0	92.2	94.3	86.5
4	49.2	89.2	38.4	64.4	49.7	14.1	59.6	97.2	56.8	68.5	97.7	66.2	66.9	99.0	65.9	73.2	96.1	69.3	87.1	90.8	77.9	96.6	85.6	82.2
6	80.5	78.6	59.0	99.6	11.2	10.8	87.5	99.2	86.7	82.3	99.8	82.1	89.0	99.6	88.6	90.4	96.1	86.5	88.2	99.5	87.7	97.2	98.1	95.3
7	72.5	93.7	66.2	90.0	71.6	61.7	50.2	99.8	49.9	68.4	97.5	65.8	76.4	95.8	72.2	84.4	85.4	69.8	85.7	91.6	77.3	90.7	92.8	83.6
8	89.7	60.1	49.8	88.6	10.9	-0.6	100.0	89.1	89.1	67.7	98.2	65.9	87.1	98.6	85.6	73.5	88.1	61.6	60.0	99.5	59.5	98.6	97.8	96.4
9	77.4	99.0	76.4	99.8	94.5	94.4	85.9	96.8	82.7	83.7	100.0	83.7	89.1	99.9	89.0	86.9	98.4	85.3	87.1	100.0	87.1	99.5	99.4	98.9
10	60.2	28.4	-11.4	65.8	51.5	17.4	75.9	94.0	69.9	86.9	84.2	71.0	95.8	78.2	74.1	90.3	92.3	82.6	87.1	92.1	79.2	89.6	92.5	82.1
11	6.0	99.2	5.2	37.0	97.9	34.8	73.8	99.4	73.2	64.0	100.0	64.0	69.0	100.0	69.0	85.7	98.5	84.2	87.4	99.8	87.2	97.6	97.8	95.4
15	96.6	44.3	40.9	99.9	10.7	10.6	97.7	94.8	92.5	74.6	97.4	72.0	83.7	93.6	77.2	80.1	93.0	73.0	78.0	98.2	76.2	94.3	96.8	91.1
16	92.5	95.5	88.0	98.7	16.2	14.9	25.2	99.8	25.0	73.2	98.7	71.9	82.3	96.2	78.5	77.5	84.8	62.3	61.9	100.0	61.8	98.7	99.0	97.7
18	97.1	99.7	96.8	79.4	47.8	27.2	83.1	100.0	83.1	95.6	96.8	92.4	92.4	97.0	89.4	94.5	99.7	94.1	85.0	99.9	84.9	98.2	99.3	97.5
21	81.8	69.6	51.4	97.9	39.8	37.7	82.0	95.4	77.4	68.8	93.5	62.3	86.7	89.6	76.3	87.6	92.8	80.4	100.0	17.0	17.0	100.0	4.6	4.6
avg:	74.0	78.6	52.6	85.1	44.0	29.1	75.8	96.5	72.3	74.6	96.9	71.6	82.5	95.6	78.1	83.9	90.4	74.3	82.7	90.3	73.0	96.1	88.2	84.3

Table 2: Evaluation of dynamic skin color classifiers: fixed and adapted to the image conditions.

using an image size of 480 x 360 pixels, whereas S2 and S3 take 0.05 ms independently of the size of the image. S4 is also executed in $O(n)$ at an average of 9.3 ms using the same image size on a 1800 MHz Pentium 4 processor.

6 Conclusion and outlook

Depending on the person and context conditions the color of human skin appears differently in any image which makes automatic skin color classification a hard challenge. Within one image those conditions are fixed and skin color pixels look similarly. We use a face detector extended by a skin color mask and obtain an image specific skin color model that describes the look of skin color within one image. Our experiments prove that this high level approach performs much better than low level approaches such as color segmentation or background subtraction. We show how dynamic skin color classifiers are adapted to the image conditions using our skin color model in order to increase accuracy. This increase is even higher facing poor illumination conditions and colored people. Furthermore, facial parts such as eyes, brows, lips, and teeth are detected correctly as non-skin colored objects. We are currently extending our approach towards classifying lip color and eyebrow color and we will integrate more sophisticated classifiers. Furthermore we are investigating the impact of skin color detection on face model fitting.

References

[1] J. Brand and J.S. Mason. A comparative assessment of three approaches to pixel-level human skin-detection. In *15th International Conference on Pattern Recognition, vol. 1*, pages 1056 – 1059, 2000.

[2] Giovani Gomez. On selecting colour components for skin detection. In *16th International Conference on Pattern Recognition, vol. 2*, pages 961–964, 2002.

[3] Giovani Gomez and E. Morales. Automatic feature construction and a simple rule induction algorithm for skin detection. In A. Sowmya and T. Zrimec, editors, *Proc. of the ICML Workshop on Machine Learning in Computer Vision*, pages 31–38, Sydney, July 2002.

[4] Ralph Gross, Iain Matthews, and Simon Baker. Generic vs. person specific active appearance models. *Image and Vision Computing*, 23(11):1080–1093, November 2005.

[5] R. L. Hsu, M. Abdel-Mottaleb, and A. K. Jain. Face detection in color images. In *IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 5*, pages 696–706, May 2002.

[6] M. J. Jones and J. M. Rehg. Statistical color models with application to skin detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 274–280, Fort Collins, 1999.

[7] Leonid Sigal, Stan Sclaroff, and Vassilis Athitsos. Esitmation and prediction of evolving color distributions for skin segmentation under varying illumination. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2000)*, March 2000.

[8] Leonid Sigal, Stan Sclaroff, and Vassilis Athitsos. Skin color-based video segmentation under time-varying illumination. In *IEEE Trans. Pattern Anal. Mach. Intell., vol. 26*, pages 862–877, 2004.

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

[10] V. Vezhnevets, V. Sazonov, and A. Andreeva. A survey on pixel-based skin color detection techniques. In *Graphics and Media Laboratory, Faculty of Computational Mathematics and Cybernetics*, Russia, 2003.

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

[12] M. Wimmer and B. Radig. Adaptive skin color classificator. In Ashraf Aboshosha et al., editor, *Proc. of the first ICGST International Conference on Graphics, Vision and Image Processing GVIP-05*, volume I, pages 324–327, Cairo, Egypt, Dec. 19–21 2005. ICGST, ICGST.