Chapter 9
Skin Color in Face Analysis

J. Birgitta Martinkauppi, Abdenour Hadid, and Matti Pietikäinen

9.1 Introduction

Color is a common feature used in machine vision applications. As a cue, it offers several advantages: easy to understand and use. Implementations can be made computationally fast and efficient, thus providing a low level cue. Under stable and uniform illumination, color cue remains robust against geometrical changes. Its ability to separate the targets from background depends on the color dissimilarity between targets and background. In some scenes, the color itself is enough for object detection.

The main difficulty in using color in machine vision applications is that the cameras are not able to distinguish changes of surface colors from color shifts caused by varying illumination spectra. Thus, color is sensitive to changes in illumination which are common under uncontrolled environments. The changes can be due to varying light level, for example, shadowing, varying light color due to changes in spectral power distribution (like daylight and fluorescent light source), or both. Cameras and their settings may produce different appearances which are different from the perception of human vision system.

Several strategies have been employed to reduce the illumination sensitivity. In one strategy, the color information is separated into two components, color inten-

J.B. Martinkauppi (✉)
Department of Electrical Engineering and Automation, University of Vaasa, Wolffintie 34, 65101 Vaasa, Finland
e-mail: birmar@uwasa.fi

A. Hadid · M. Pietikäinen
Machine Vision Group, Department of Electrical and Information Engineering, University of Oulu, P.O. Box 4500, 90014 Oulu, Finland

A. Hadid
e-mail: hadid@ee.oulu.fi

M. Pietikäinen
e-mail: mkp@ee.oulu.fi

S.Z. Li, A.K. Jain (eds.), *Handbook of Face Recognition*,
DOI 10.1007/978-0-85729-932-1_9,

sity and color chromaticity. Use of color chromaticity component reduces the effect of varying light levels. To cancel the effect of illumination color and thus different spectral power distributions, numerous color constancy algorithms have been suggested, but their success has been limited [6]. A different strategy to these is to tolerate or adapt the model to the illumination changes. This strategy can produce promising results even under drastic variations in target colors as shown in this chapter for facial recognition.

It is often preferable to get rid as much as possible of the dependencies on lighting intensity. The perfect case would be to also cancel-out the effect of the illuminant color (by defining a color representation which is only a function of the surface reflectance) but, thus far this has not been achieved in machine vision. The human visual system is superior in this sense, since human visual perception in which the color is perceived by the eye depends quite significantly on surface reflectance, although the light reaching the eye is a function of surface reflectance, illuminant color and lighting intensity.

For face detection, color has been an intriguing and popular cue. It is often used as a preprocessing step to select regions of interests for further, more computationally demanding processing. For instance, with the appearance-based face detection, an exhaustive scan (at different locations and scales) of the images is conducted when searching for the faces [54]. However, when the color cue is available, one can reduce the search regions by pre-processing the images and selecting the skin-like areas only.

This chapter deals with the role of color in facial image analysis such as face detection and recognition. First, we introduce the use of color information in the field of facial image analysis in particular (Sect. 9.2). Then, in Sect. 9.3, we give an introduction to color formation and discuss the effect of illumination on color appearance, and its consequences. The skin data can come from different sources like real faces, photos or print. Separating the sources of skin data is presented in Sect. 9.4, and skin color modeling is discussed in Sect. 9.5. Section 9.6 reviews the use of color in face detection, while the contribution of color to face recognition is covered in Sect. 9.7. Finally, conclusions are drawn in Sect. 9.8.

9.2 Color Cue and Facial Image Analysis

The properties of the face pattern pose a very difficult problem for facial image analysis: a face is a dynamic and nonrigid object which is difficult to handle. Its appearance varies due to changes in pose, expressions, illuminations and other factors such as age and make-up. As a consequence, most of the facial analysis tasks generally involve heavy computations due to the complexity of facial patterns. Therefore, one may need some additional cues, such as color or motion, in order to assist and accelerate the analysis. These additional cues also offer an indication of the reliability of the face analysis results: the more the cues support the analysis, the more one can be confident about the results. For instance, with the appearance-based face

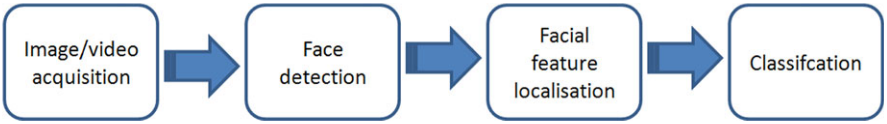

Fig. 9.1 A general block diagram of face analysis which shows different phases of facial image analysis

detection an exhaustive scan (at different locations and scales) of the images is conducted when searching the faces [54]. However, when the color cue is available, one can reduce the search regions by pre-processing the images and selecting only the skin-like areas. Therefore, it is not surprising that the color of skin has been commonly used to assist face detection. Also, in face recognition, it has been argued that color does play a role under degraded conditions by facilitating low-level facial image analysis such as better estimations of the boundaries, shapes and sizes of facial features [56]. As mentioned above, among the advantages of using color is the computational efficiency and robustness against some geometric changes such as scaling and rotation, when the scene is observed under a uniform illumination field. However, the main limitation with the use of color lies in its sensitivity to illumination changes (especially changes in the chromaticity of the illuminant source which are difficult to cancel-out).

Let us consider the general block diagram of face analysis, shown in Fig. 9.1. The color cue is involved at different stages [36]. In the first stage, the color images (or video sequences) are acquired and preprocessed. The preprocessing may include gamma correction, color space transformation, and so on. It is often preferable to get rid as much as possible of the dependencies on lighting intensity.

Among the different stages shown in Fig. 9.1, the use of color in face detection is probably the most obvious. It is generally used to select the skin-like color regions. Then, simple refining procedures can be launched to discriminate the faces from other skin-like regions such as hands, wood, etc. Thus, much faster face detectors are generally obtained when the color cue is considered.

Using the fact that some facial features, such as eyes, are darker than their surrounding regions, holes should then appear in the face area when labeling the skin pixels. Such observation is commonly exploited when detecting facial features in color images [10, 15, 54].

Does color information contribute to face recognition? The answer to this question is not obvious, although some studies have suggested that color does play a role in face recognition as well, and this contribution becomes evident when the shape cues are degraded [56]. Section 9.7 discusses this issue.

9.3 Color Appearance for Color Cameras

9.3.1 Color Image Formation and Illumination

Color cameras reproduce the scene with three components which are typically red (R), green (G) and blue (B). The components are named after the spectral range over which the response was integrated. An example of color camera filters is shown in Fig. 9.2. The spectral filters typically operate in the visible wavelength spectrum range, that is, 400 nm–700 nm. Of course, different filter selections affect the obtainable descriptor set and most likely produce different values for the same input.

The descriptors themselves are obtained by filtering the color signal $C(\lambda)$ with suitable spectral filters and integration over the filtered signal. The color signal is a spectral distribution of electromagnetic radiation, which is the light from an illumination source, light reflected from a surface or a combination of these. This is similar to calculation of human vision responses (see, e.g., [50]).

The following simple model represents camera output with white balancing:

$$D = \frac{\int \eta_D(\lambda) Ip(\lambda) S(\lambda)\, d\lambda}{\int \eta_D(\lambda) Ic(\lambda)\, d\lambda}, \tag{9.1}$$

where D is R, G or B response, λ is the wavelength, p is prevailing (illumination) and c is calibration (illumination), η is the spectral responsivity of a spectral filter, I is the spectral power distribution of the light (SPD), and S is the spectral reflectance of the surface. The nominator of (9.1) alone describes image formation as a sum of the camera sensitivity, the illumination SPD and the reflectance over the wavelength range. Thus, for each pixel, the output value depends on the illumination, reflectance and camera sensitivity. This is a very simplified presentation of the formation but can be used as a basic theoretical estimation of the camera response to the input light. The denominator models the white balance. White balance means adjusting gains of camera so that the cameras response for white (or very bright gray) is equal in every channels. For example, the response of a white is adjusted to (255, 255, 255).

Equation (9.1) can be used to simulate the effect of illumination. When the prevailing and calibration illumination are the same, then the output image is called as a canonical or calibrated image and colors are canonical colors. This is described in more detail in Sect. 9.3.2. The prevailing and calibration illumination can also be different, and the output image in this case is called non-canonical image. The modeling is, however, more problematic. The problem of normalization can be demonstrated theoretically [32]. Let us assume that the prevailing illumination is originally I_{np} and its normalization factor is the constant factor f_p, and the calibration illumination is in the unnormalized format I_{nc}, which is normalized by the factor constant f_c. For example, if we insert these variables into (9.1), we can derive the following

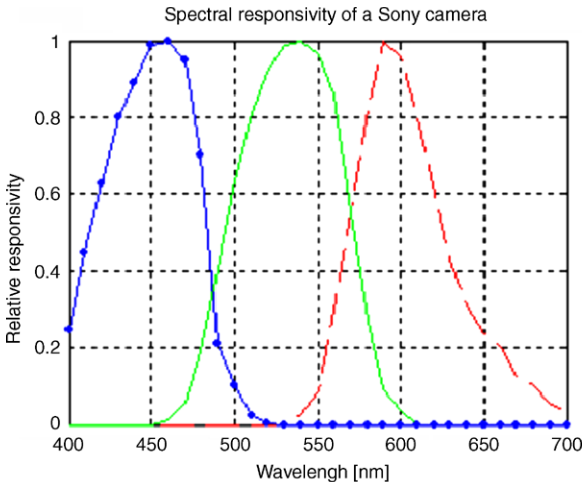

Fig. 9.2 Spectral responsivity curves of a Sony camera, originally obtained from a graph provided by the manufacturer

format:

$$R = \frac{\int \eta_R(\lambda) I_p(\lambda) S(\lambda)\, d\lambda}{\int \eta_R(\lambda) I_c(\lambda)\, d\lambda} = \frac{\int \eta_R(\lambda) \frac{I_{np}(\lambda)}{f_p} S(\lambda)\, d\lambda}{\int \eta_R(\lambda) \frac{I_{nc}(\lambda)}{f_c}\, d\lambda} = \frac{f_c}{f_p} \frac{\int \eta_R(\lambda) I_{np}(\lambda) S(\lambda)\, d\lambda}{\int \eta_R(\lambda) I_{nc}(\lambda)\, d\lambda}. \tag{9.2}$$

The ratio f_c/f_p is 1 only when the illumination conditions are the same. Different choices for normalization methods may produce different results [32].

9.3.2 The Effect of White Balancing

The effect of white balancing on the perceived images and colors is examined in more details in this chapter. White balancing is one of the important factors affecting image quality. The white balancing factor depends on the illumination. Many digital images have been taken under canonical conditions or very near to them to avoid distortions in colors. The color distortions are easily noticed and taken as annoying artifacts. This is especially true for certain colors which humans remember very well; thus, they are referred to as memory colors. One of these memory colors is, quite naturally, skin tone.

Humans are very sensitive to any distortion in skin tones [12, 25], thus, it is not so surprising that these have been investigated a lot. Skin tones refer here to the correct or acceptable colors for skin as perceived by a human. Skin colors refers to all those RGBs which a camera can perceive as skin under different illuminations. Note that human and cameras can perceive skin color differently.

In cameras, white balancing can be done automatically or manually. In manual selection, the user selects the best option for the prevailing illumination, while automatic option provides settings from a program. However, it is not always possible to

Fig. 9.3 The face is illuminated by the nonuniform illumination field and the white balancing partially fails. The color appearance of the face varies at different parts of the light field

select or compute proper white balancing factors. This is especially true under varying, nonuniform illumination, which can cause more drastic color changes. For example, it is common to have more than one light source on a scene. If these sources with different SPDs shine over an object, it is not possible to conduct the correct white balancing for the whole image. This is demonstrated in Fig. 9.3. The face is imaged under a nonuniform illumination field. The camera was balanced under the light of fluorescent lamps on the ceiling and thus the part of the face under only fluorescent illumination field appears in skin tones. However, the daylight from windows causes a bluish color shift on the right side of the face image. The colors are distorted because the white balancing fails partially. The distortion between these two sides varies to a different degree as a function of illumination field. The nonuniform illumination fields are encountered commonly, but they are rarely considered in face detection or recognition applications.

Of course, one can apply some color correction techniques to improve the quality. For example, Do et al. used sclera region of the eye to estimate illumination color and then apply skin detection [4]. Even a nonuniform illumination field is possible to correct if the light colors are given by the user [16]. However, the failure in white balancing may cause information loss, which is generally very difficult to correct properly.

9.3.2.1 Canonical Images and Colors

Even though an image is taken under canonical condition, it does not guarantee that the objects appear in the same colors under different canonical illumination. White, grays and black do appear at least in most of the cases very similarly under different light sources, but of course there are some limitations. It is not possible even in theory to perceive all RGB components for a gray object if the prevailing illumination does not have spectral output in components' spectral range. The ideal camera RGB responses for white in canonical case should have equal RGB values even under different light sources, given that the sources are not very extreme. Cameras do reproduce a white surface quite well over a range of light sources, but of course there is a physical limitation due to gain control, for example.

If a camera has linear response over a certain input signal range, then those grays falling the range will be reproduced in gray colors if the color signal from scene falls into the input range. The grays here refer to those objects whose spectra is constant

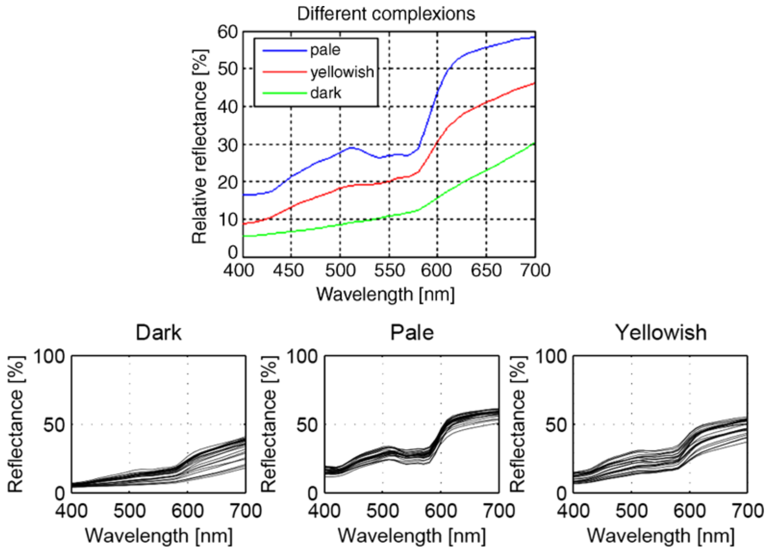

Fig. 9.4 Skin complexions of three skin groups: pale, yellowish, and dark. Their reflectances are smooth and similar, mainly separated by their levels. Measured skin reflectances are available, for example, in the Physics-based face database [31]

over the wavelength range, but the value of this constant is smaller than the maximum value ("white"). When the spectra is not constant over the range, the effect of illumination cannot be canceled out from the reproduced RGB values. Thus, the object colors will be affected to a different degree between different light sources. Therefore, a camera can reproduce only the achromatic colors similarly under different light sources assuming that the camera is white balanced to the prevailing light sources. This means that skin can have different colors under images taken with different conditions.

The reproduction differences can be demonstrated very easily. First, the objects need to be selected, and, in this case, three skin complexions (pale, yellowish and dark) are used. The spectral reflectances for the complexions are shown in Fig. 9.4. The reflectances are smooth and similar. They are separated mainly by their level, but not their shape [8, 17, 51], which suggests the similar reproduction in color. Due to this, skin spectra can be reconstructed at high quality using only three basis vectors [18, 38]. The similarity is due to the colorants (melanin, carotene, and hemoglobin) determining the reflectance [5]. As a natural object, skin has not uniform coloration.

Using (9.1), the RGB values for skin are calculated using the Sony camera's responses. The RGB values are then converted into NCC chromaticity. These theoretical skin chromaticities are displayed in Fig. 9.5. The canonical skin values are dissimilar under different illuminations even in an ideal case.

Cameras produce even bigger variations in skin colors: Fig. 9.6 shows skin chromaticities for a Sony camera taken under the same light sources as the ones used in simulation (Horizon 2300 K or light at sunset/sunrise), Incandescent A 2856 K,

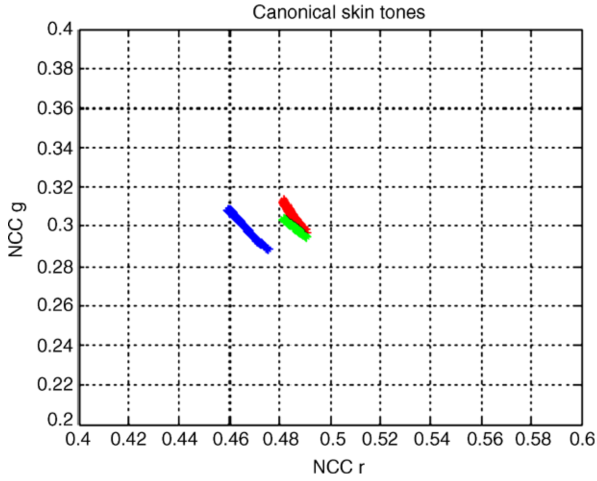

Fig. 9.5 Canonical skin tones were obtained by converting the theoretical skin RGBs to Normalized Color Coordinate (NCC) space

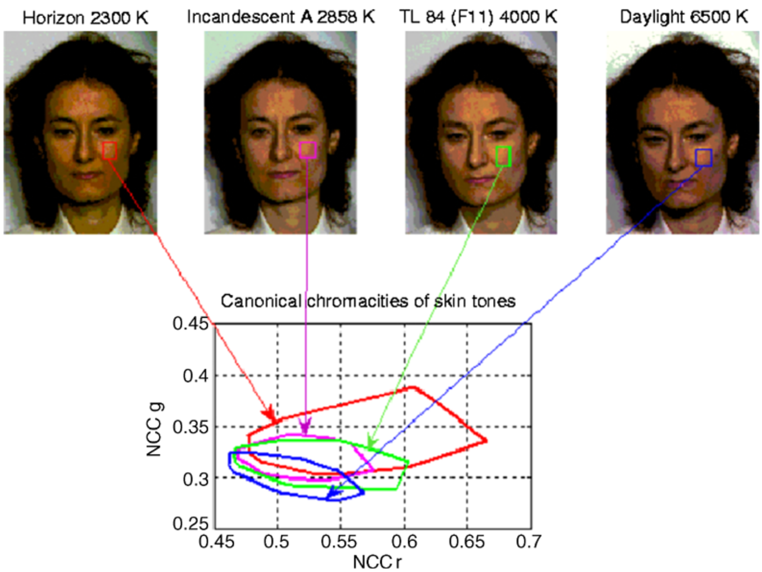

Fig. 9.6 The skin tone appearance difference can be clearly observed in the four images taken with the Sony camera (see Fig. 9.2). From the selected area marked with a box the RGB values were taken and converted to NCC color space. As shown in the *graph below the images*, the areas of canonical chromaticities more or less overlap

fluorescent lamp TL84, and daylight D65 6500 K. The overlap between loci is significant. Note that the locus obtained using Horizon light covers a bigger area than for other light sources. This might be due to unsuccessful white balancing.

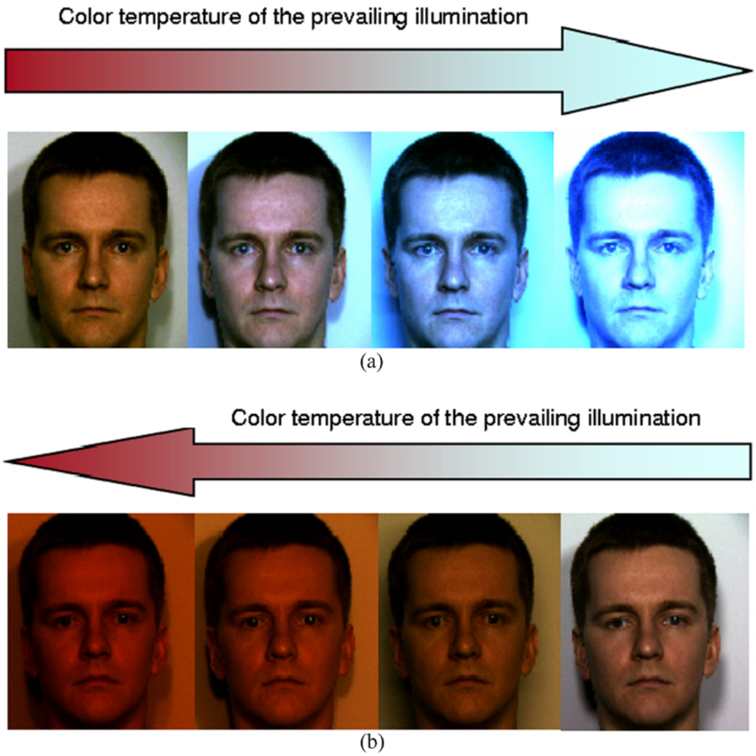

Fig. 9.7 The color appearance shift is apparent in (**a**) and (**b**). The color temperature of the light sources increases from left to right. The *arrow* indicates the change in the color of the light. The limited dynamic response range causes distortion in color: pixels can saturate to a maximum value (the *rightmost image at the upper row*) or be under-exposed to zero (the *leftmost image at the lower row*)

9.3.2.2 Non-canonical Images and Colors

If images are not taken under the illumination used in camera calibration, the colors are distorted even more. The distortion will appear as a shift in colors, as can be seen in Fig. 9.7 which displays images taken under four different light sources while the camera is calibrated to one of them. In the upper image series, the camera was calibrated to the light source Horizon (first image on the left) and after light source was changed to incandescent A, TL84 and daylight, respectively. In the lower image series, the camera was calibrated to daylight (first image on the right) and then images were taken under TL84, A and Horizon.

The skin color tends to shift in the direction of illumination color change. More reddish prevailing illumination causes color shift towards red, while more bluish one adds blue components. Of course, a light source with strong spikes in spectra can cause additional distortions for certain colors. Since cameras have limited dynamic response ranges, the colors can be distorted also due to saturation or under-exposure.

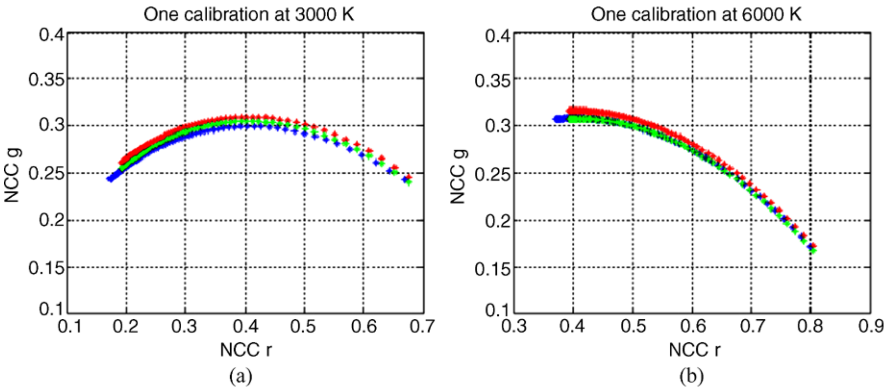

Fig. 9.8 The skin NCC chromaticities were simulated using the data of the Sony camera (see Fig. 9.2) and the skin reflectances from Fig. 9.4. **a** shows the possible skin chromaticities when the camera was calibrated to a Planckian of 3000 K and **b** when the calibration illumination was a Planckian of 6000 K. The chromaticity range depends on the calibration illumination and the possible color temperature range of prevailing illuminations

Manual or automatic brightness control in the camera can alleviate this problem, but manual operation tends to be tedious and automatic control might cause problems by itself.

Figure 9.8 shows simulated skin chromaticities using only one calibration. The chromaticity range obtained depends on the calibration light and the color temperature range of the prevailing illumination. The possible range of skin colors (locus [44]) is affected by the amount of calibrations. Figure 9.8 shows that different white balancing illuminants have dissimilar ranges of possible skin chromaticities and produce separate skin locus. When the loci of all different calibrations are gathered together, a bigger locus is obtained, as shown in Fig. 9.9. Of course, the illumination range as well as different camera settings affect the locus size.

9.4 Separating Sources of Skin Data

Many materials, like inks and dyes, are used to imitate the appearance of skin. Some studies have been already done to examine how well the imitation works and how the real skin can be separated from imitation.

The skin data can come from different sources like real faces, photos or print [37]. The source cannot often be determined from normal RGB data, so spectral data is needed. An interesting spectral data region is near infrared. Figure 9.10 shows near infrared spectra for real faces, facial skin from photos and facial skin from a print of three different skin complexions. The spectra from photos and prints, which are flat, are clearly different from that of real faces. Thus simple ratio between two channels can be used to separate real skin from other sources. The level difference in real spectra between different complexions start to diminish as a degree of wavelength. Skin complexion groups are separable in print spectra, but not in photo spectra.

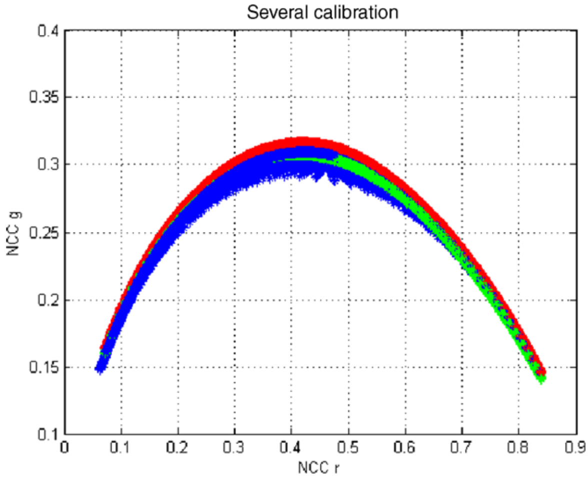

Fig. 9.9 The skin locus formed with all prevailing illumination/white balancing combinations

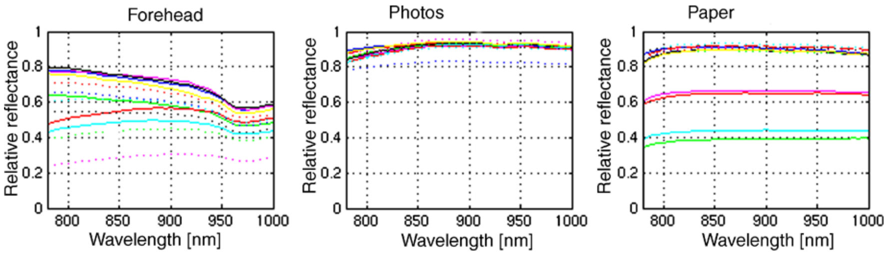

Fig. 9.10 Near infrared skin spectra from real faces (*left*), photos (*middle*) and paper (*right*)

The skin color appearance for mannequins is also sought after, but it clearly is different from real skin [1]. Kim et al. [24] have studied the differences between masked fake faces and real skin. They concluded that wavelengths of 685 nm and 850 nm can be used to discriminate them.

9.5 Modeling Skin Colors

Skin color model is a description of possible skin tones. To create such a model, one has to first select the color space in which the model is formed, then the mathematical model to describe the possible skin colors, and finally, the data upon which the model is defined. The performance of the model depends on all these factors and is a trade-off between generality of the model and accuracy for a certain image.

Skin detection methods have been compared in several studies using different data [22, 35, 46]. The studies disagree, which might be because the optimality of the model depends on its purpose, data, material and modeling parameters.

9.5.1 Behavior of Skin Complexions at Different Color Spaces Under Varying Illumination

Color space in which skin data is processed, has also an effect on detection. Not all color spaces are equal: they can map RGB values differently, which can be used to separate certain colors. Even a mixture of color spaces can be used like in [47], at least for canonical or nearly canonical images.

As mentioned earlier, a color space conversion does not remove chromaticity shifts due to illumination or effects caused by noise. In fact, noise can be detrimental for low RGB values or near thresholds. The brightness control or lack of it can have a strong effect on the possible skin chromaticities. If there is no automatic brightness or gain controller, it is possible for one channel to have low values or even underclipping. Therefore, the skin colors have been studied under varying illuminations [34].

RGB coordinates are device-oriented, but they can be converted into human vision oriented spaces like XYZ or CIE Lab. A correct conversion requires an illumination-dependent transform matrix, including also the effect of device characteristics. Of course, there exist general transforms matrices. None of the matrix transforms reduce the effect of changing light since it has already affected RGBs.

The more device oriented color spaces can be classified, based on the conversion method, into two groups: those using linear transforms from RGB and those obtained via non-linear transforms. For example, linear transform based color spaces are: I1I2I3, YES, YIQ, YUV, YCrCb (Rec. 601–5 and 709). Among the nonlinear transforms are: NCC rgb, modified rgb, natural logarithm ln-chromaticity, P1P2, l1l2l3, ratios between channels (G/R, B/R, and B/G), HSV, HSL, modified ab, TLS and Yuv.

Overlap between different skin complexions vary in color spaces. In [34], the overlaps between two complexions (pale and yellowish) were compared in different color spaces and across different cameras: the overlaps between them were reasonably high in all color spaces (ranging from 50–75 percent) when using different canonical images. When using both canonical and uncanonical images, the overlap still increased due to the fact that more colors fall into the region. However, when comparing skin data from different cameras, the overlaps between skin RGBs were smaller and dependent on the cameras used in comparison. Therefore, one can argue that color spaces and cameras used do have an effect on skin detection and thus for face recognition.

9.5.2 Color Spaces for Skin

Several color spaces have been suggested for general skin color modeling, but thus far, none of them has been shown to be superior to the other. The list of comparison studies for color spaces can be found, for example, in [33] or [22]. However, it seems that those spaces in which intensity is not separated so clearly from chromaticity are similar to RGB. The separation can be evaluated using linear or linearized RGB data: RGB is transformed into color space using substitution $R \to cR$, $G \to cG$, and $B \to cB$, in which c describes a uniform change in the intensity levels. If the factor c does not cancel out for chromaticity descriptors, the separation is incomplete.

Normalized color coordinates (NCC) are quite often used in modeling, and they separates the intensity and chromaticity. To avoid the intensity changes, only the chromaticity coordinates are used. In [46], different color spaces are compared in terms of efficiency and transferability of the model. The performance of NCC and CIE xy was superior to several other skin color models. It was also shown in [55] that NCC has a good discrimination power. More details of color spaces for skin detection can be found in [33] or [22].

A color can be uniquely defined in by its intensity and two chromaticity coordinates since $r + g + b = 1$. The chromaticity coordinates for NCC color space are defined as

$$r = \frac{R}{R + G + B}, \tag{9.3}$$

$$g = \frac{G}{R + G + B}. \tag{9.4}$$

The intensity is canceled from chromaticity coordinates since they are calculated by dividing the descriptor value of the channel by the sum of all descriptor values (intensity) at that pixel.

The modeling can be done using only the chromaticity coordinates to reduce the effect of illumination intensity changes, which are common in videos and images. Some models do include intensity (like in [14]), but more data is needed to construct the model and computational costs are increased due to a third component.

9.5.3 Skin Color Model and Illumination

Section 9.3 showed that illumination affects skin color both in canonical and uncanonical images. What is more, this dependency is camera-specific: the camera sensors and internal image preprocessing of the camera affect the color production and thus on the end results (see Fig. 9.11). Therefore, creating a universal model is difficult.

Many face detection algorithms assume that the images are taken under canonical or near canonical conditions. For many data sets, this is true. An example of this kind of image data set is a set of personal photos.

When the illumination varies, the previous approaches have a high risk of failure. Of course, the images can be subjected to color correction or color constancy algorithm, but sometimes this can lead even more serious color distortions [35].

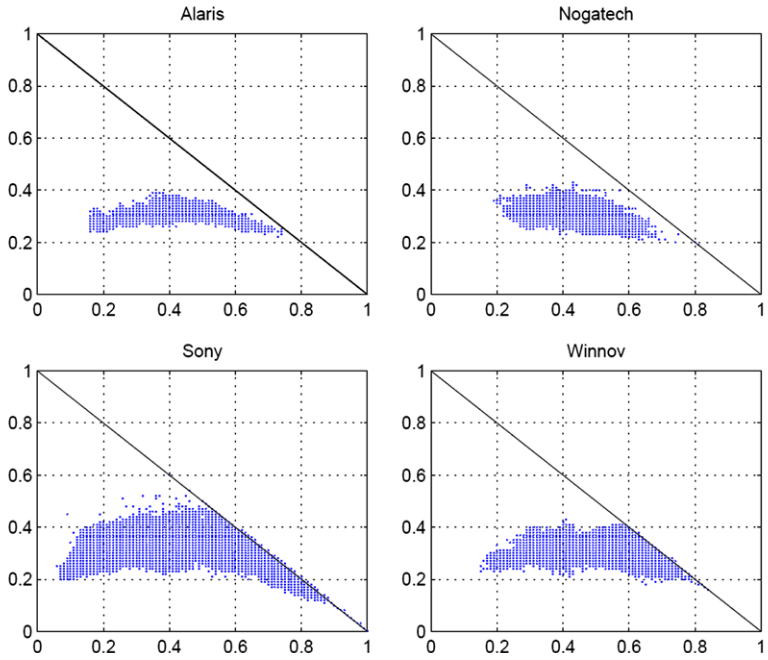

Fig. 9.11 The camera and its properties determine the skin locus, as indicated by the loci of four cameras. However, some regions are common to all, most notable the region of skin tones

Color correction based approach has been suggested, for example, by Hsu et al. [15]: the colors in image are corrected so that the skin would appear in skin tones and after this segment the image using skin color model. The color correction is based on a pixel with a high brightness value which are assumed to belong to a white object. These pixels are used to calculate correction coefficient which are applied to the image. This approach can fail for many reasons like data loss due to saturation, or if a pixel with high brightness belongs to a nonwhite object. The latter case is demonstrated in Fig. 9.12.

For a more general skin model, one should use the knowledge of illumination changes, calibration and camera settings like in the skin locus-based approach [43]. The drawback of this model is that it is not so specific as canonical models—more color tones are included. Thus, more nonskin objects will be considered skin candidates. Since color itself is rarely enough to determine whether the target is skin or not, the face candidates are in case subjected for further processing.

9.5.4 Mathematical Models for Skin Color

The model for skin color can be either a mathematically defined area in color space or a statistical approach in which a probability to belong skin is attached to color

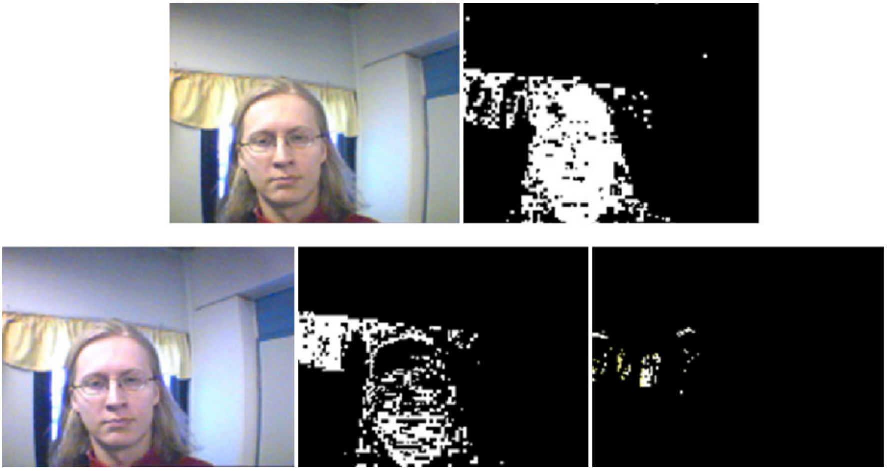

Fig. 9.12 The *upper row* displays the color segmentation results using Hsu et al. model [15] without the color correction part. The *lower row* shows the segmentation with their color correction method. The color correction fails because the yellow curtains have the highest brightness values and is assumed to be a white object

tones. The model may be fixed or adaptive, and in the latter case, the update depends whether it is applied on single images or video frames. A more detailed review can be found, for example, in [33] or [22].

The area based approach uses a spatial constraint in the color space to define possible skin areas. The shape of the constraint can be simple thresholds like in [3] or a more complex shaped function like in [15]. Generally no thresholding is done, since the colors that fall inside the area are considered skin. These models often assume that skin has or can be corrected to have skin tone appearance. An exception is the skin locus in which the illumination changes are included in the model.

It is possible to adapt the model even for single images (e.g., [3, 26, 45]) although the successfulness depends on the validity of assumptions behind the adaptation criteria. The adaptation schema generally use a general skin model obtained from a representative image set and after that fine-tune into an image specific model. For example, in Cho et al. [3], the fine-tuning phase assumes that the skin color histogram is unimodal and skin color occurs mainly on real skin areas. This approach can fail if the image has dominant skin-colored, nonfacial object or the histogram is not unimodal.

The challenge of the probability-based approach is to be able to reliably find the probability distribution of skin colors. This requires collecting a representative data set of images for forming the model. An example of a statistical model is the one presented by Jones and Rehg [21]. They calculate the histogram and Gaussian models using over 1 billion labeled pixels. Many other statistical models like SOM or neural networks has been suggested and a review of them can be found, for example, in [33] or [22]. In addition to the statistical model, one has to determine the threshold limit for separating the skin from nonskin. It is difficult to automatically

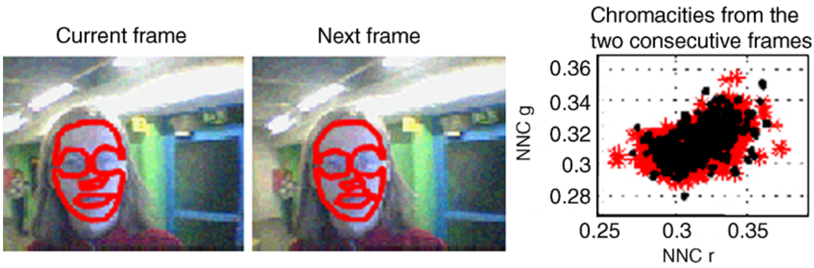

Fig. 9.13 Two consecutive frames are taken from a video sequence (the *first* and *second image from the left*). The facial skin areas of the frames are manually extracted and their skin RGB values are then converted to the NCC chromaticity space. The chromaticities from these two frames are marked with different colors in the right image. As can be observed from the rightmost image, the chromaticities overlap significantly

find the threshold value because the probability model found may not be valid for all images.

9.5.4.1 Video Sequences

The processing of video sequences is similar to that of single, independent images. Thus, the skin detection presented earlier can also be used for videos. The fixed skin color models are suitable for videos in which changes in illumination are minimal. Generally, this is not the case and the skin color models need to be updated. The model adaptation relies often on the dependencies between consecutive frames, which is true for many videos: The consecutive frames often exhibit sequential dependency. This can be observed in Fig. 9.13: the overlap between the chromaticities from two consecutive frames is significant.

If the illumination changes between images are slow (no abrupt, drastic object color changes) or the person moves in a nonuniform illumination field slowly enough, the skin color model can adapt to the color changes. This required some constraint for selecting the pixels used in the model update. Three different adaptive schemes have been suggested: two of them use spatial constraints [39, 57] (see Fig. 9.14) and one skin locus [35]. The basic idea is the same: to use some constraint to select the pixels for model updating. The spatial constraints use different ideas to select candidate pixels from a located face: the method of Raja et al. [39] updates the skin color model using pixels inside the localized face area. The pixels are selected from an area which is 1/3 of the localization area and 1/3 from the localization boundaries. Yoo and Oh [57] argued that the localization should resemble the shape of the object (face) and they used all pixels inside the elliptical face localization. The skin locus can be used in two ways: either the whole locus or partial locus is used to select skin colored pixels from the localized face and its near surroundings.

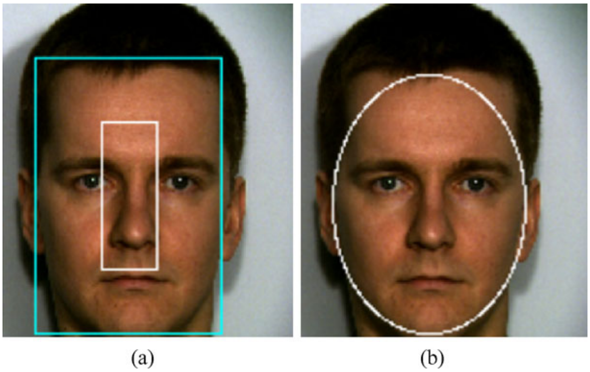

Fig. 9.14 Spatial constraints suggested for adaptive skin color modeling: the *left image* shows the method suggested by Raja et al. [39]. The *outer box* indicates the localized face while the pixels inside the *inner box* are used for model updating. The *image on the right* shows elliptical constraint by Yoo and Oh [57]

There are many possible methods for updating the skin color model, but perhaps a common method is the moving average, as presented in (9.5):

$$\check{M} = \frac{(1-\alpha) * M_t + \alpha * M_{t-1}}{\max((1-\alpha) * M_t + \alpha * M_{t-1})}, \tag{9.5}$$

where $\check{M}$ is a new, refreshed model, M is the model, t is the frame number and α is a weighting factor. Quite often, the weighting factor is set to 0.5 to get equal emphasis on the skin color model of current and previous frames. The moving average method provides a smooth transition between models from different frames. It also reduces the effect of noise, which can change pixel color without any variation in external factors and thus be detrimental to the models.

However, the spatial constraint models have been shown to be very sensitive to localization errors, therefore, they can easily adapt to nonskin objects [35]. The failure due to these constraints can happen even under a fairly moderate illumination change. In Fig. 9.15, Raja et al.'s method has failed while tracking a face on a video sequence and the skin color model is adapted to nonskin colored target, as shown in this image.

The constraint suggested by Raja et al. easily fails under a nonuniform illumination field change, as demonstrated in Fig. 9.16. The model is updated using the pixel inside the localization and therefore, it can adapt only to global illumination changes, but not to the nonuniform illumination field variation.

The correct localization of face is not so sensitive for a skin locus based approach since the nonskin colored pixels can be filtered out. Large skin colored objects connected to the face are problematic and cues other than color are needed to solve this.

9.6 Color Cue for Face Detection

As mentioned above, color is a useful cue for face detection as it can greatly reduce the search area by selecting only the skin-like regions. However, it is obvious that

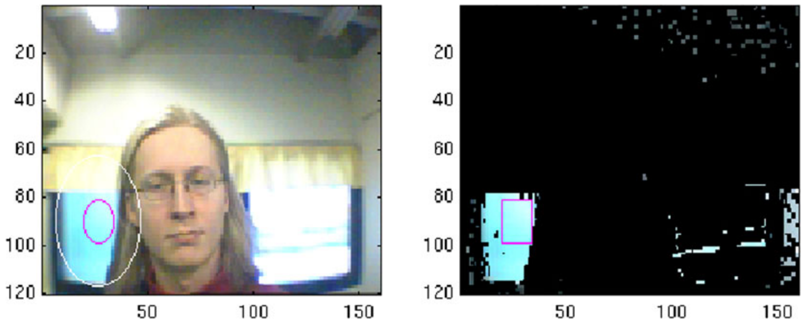

Fig. 9.15 The face tracking based on Raja et al.'s method failed and adapted to a nonfacial target. The *left image* displays the "localized face". The *right image* shows the pixels selected by the current skin color model. The *red box* shows the pixels used for refreshing the model

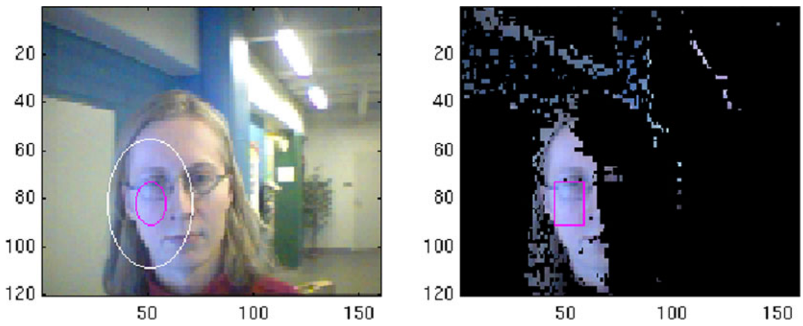

Fig. 9.16 The constraint suggested by Raja et al.'s selects a nonpresentative set of skin pixels

the use of skin color only is not enough to distinguish between faces and other objects with a skin-like appearance (such as hands, wood, etc.). Therefore, other procedures are needed to verify whether the selected regions are (or contain) faces or not. Depending on the robustness of the skin model and changes in the illumination conditions, one can notice two cases:

- Case #1: The initial skin color detection step produces consistently reliable results. The skin color model is valid for the illumination conditions, the camera and its settings. The skin color model can be designed either for stable, controlled illumination (typical case) or for variable illumination (skin locus). In such cases, it is generally enough to consider each connected resultant component from the skin detection as a face candidate. Then, one can verify the "faceness" of the candidate by simple and fast heuristics.
- Case #2: The initial skin color detection step produces unsatisfactory results or even fails. In this case, the skin color model does not correspond to the prevailing illumination, used camera or settings of the camera. One can hope that the results would indicate the locations of the faces, but their size estimation is too unreliable. Therefore, a different method for face detection (either an appearance-based

or feature-based one) should be used when searching for the faces in and around the detected skin regions.

In both cases, the use of color accelerates the detection process. In the following, we review some methods based on color information for detecting faces. Most of the color-based face detectors start by determining the skin pixels which are then grouped using connected component analysis. Then, for each connected component, the best fit ellipse is computed using geometric moments, for example. The skin components which verify some shape and size constraints are selected as face candidates. Finally, features (such as eyes and mouth) are searched for inside each face candidate based on the observation that holes inside the face candidate are due to these features being different from skin color. Therefore, most of the color-based face detection methods mainly differ in the selection of the color space and the design of the skin model. In this context, as seen in Sect. 9.5, many methods for skin modeling in different color spaces have been proposed. For comparison studies, refer to [35, 46] and [34].

Among the works using color for face detection is Hsu et al.'s system which consists of two major modules: (1) face localization for finding face candidates, and (2) facial feature detection for verifying detected face candidates [15]. For finding the face candidates, the skin tone pixels are labeled using an elliptical skin model in the YC_bC_r color space, after applying a lighting compensation technique. The detected skin tone pixels are iteratively segmented using local color variance into connected components which are then grouped into face candidates. Then, the facial feature detection module constructs eye, mouth and face boundary maps to verify the face candidates. Good detection results have been reported on several test images. However, no comparative study has been made thus far.

In [7], Garcia and Tziritas presented another approach for detecting faces in color images. First, color clustering and filtering using approximations of the YC_bC_r and HSV skin color subspaces are applied to the original image, providing quantized skin color regions. Then a merging stage is iteratively performed on the set of homogeneous skin color regions in the color quantized image, in order to provide a set of face candidates. Finally, constraints related to shape and size of faces are applied, and face intensity texture is analyzed by performing a wavelet packet decomposition on each face area candidate in order to detect human faces. The authors have reported a detection rate of 94.23% and a false dismissal rate of 5.76% on a data set of 100 images containing 104 faces. Though the method can handle nonconstrained scene conditions, such as the presence of a complex background and uncontrolled illumination, its main drawback lies on that fact that it is computationally expensive due to its complicated segmentation algorithm and time-consuming wavelet packet analysis.

Sobottka and Pitas presented a method for face localization and facial feature extraction using shape and color [42]. First, color segmentation in HSV space is performed to locate skin-like regions. After facial feature extraction, connected component analysis and best fit ellipse calculation, a set of face candidates are obtained. To verify the "faceness" of each candidate, a set of eleven lowest-order-geometric

Fig. 9.17 Examples of face detection results using the color-based face detector in [10]

moments is computed and used as inputs to a neural network. The authors reported a detection rate of 85% on a test set of 100 images.

In [11], Haiyuan et al. presented a different approach for detecting faces in color images. Instead of searching for facial features to verify the face candidates, the authors modeled the face pattern as a composition of a skin part and a hair part. They made two fuzzy models to describe the skin color and hair color in CIE XYZ color space. The two models are used to extract the skin color regions and the hair color regions which are compared with the prebuilt head-shape models by using a fuzzy theory based pattern-matching method to detect the faces.

In [10], Hadid et al. presented an efficient color-based face detector, using the skin locus model to extract skin-like region candidates, and then performing the selection by simple yet efficient refining stages. After ellipse fitting and orientation normalization, a set of criteria (face symmetry, presence of some facial features, variance of pixel intensities and connected component arrangement) are evaluated to keep only facial regions. The refining stages are organized in a cascade to achieve high accuracy and to keep the system fast. The system was able to detect faces and deal with different conditions (size, orientation, illumination and complex background). Figure 9.17 shows some detection examples performed by the system under different conditions.

Several other approaches using color information for detecting and tracking faces and facial features in still images and video sequences have been proposed [13, 54].

Fig. 9.18 Examples of face detection results using the color-based face detector in [9]

It appears that most of the methods have not been tested under practical illumination changes (usually only mild changes are considered), which makes them belonging to the first category (Case #1) described above.

More recently, to detect faces in natural and unconstrained environments, Hadid and Pietikänen [9] proposed an approach which considers the fact that color is a very powerful and useful cue for face detection, but unfortunately, it may also produce unsatisfactory results or even fail. The proposed approach consists of first preprocessing the images to find the potential skin regions, avoiding thus scanning the whole image when searching for faces, and then performing an exhaustive search in and around the detected skin regions. The exhaustive search is performed using a two-stage SVM based approach, exploiting the discrimination power of the Local Binary Patterns (LBP) features. The obtained results are interesting in the sense that the proposed approach inherits the speed from the color-based methods and the efficiency from the gray scale-based ones. Some detection results are shown in Fig. 9.18.

One problem of color-based face detectors lies in the fact that they are generally camera specific. Most of the methods have reported their results on specific and limited data sets and this fact does not facilitate performing a comparative analysis between the methods. Among the attempts to define a standard protocol and a common database for testing color-based face detector is the work of Sharma and Reilly [41].

Currently, most methods for face detection rely only on gray scale information even when color images are available. Generally these methods scan the images at all possible locations and scales and then classify the sub-windows either as face or nonface, yielding in more robust but also computationally more expensive processing methods, especially with large-sized images. Among robust approaches based only on gray scale information is Viola and Jones's approach [49]. The approach uses Haar-like features and AdaBoost as a fast training algorithm. AdaBoost is used to select the most prominent features among a large number of extracted features and construct a strong classifier from boosting a set of weak classifiers. Such systems generally run in real-time for small-sized images (e.g., 240 × 320 pixels), but tend to be slow for larger images. Including other cues such color or motion information may thus be very useful for speeding-up the detection process.

9.7 Color Cue for Face Recognition

The role of color information in the recognition of nonface objects has been the subject of much debate. However, there has only been a small amount of work which examines its contribution to face recognition. Most of the work has only focused on the luminance structure of the face, thus ignoring color cues, due to several reasons.

The first reason lies in the lack of evidence from human perception studies about the role of color in face recognition. Indeed, a notable study in this regard was done in [23], in which the authors found that the observers were able to quite normally process even those faces that had been subjected to hue-reversals. Color seemed to contribute no significant recognition advantages beyond the luminance information. In another piece of work [56], it is explained that the possible reason for a lack of observed color contribution in these studies is the availability of strong shape cues which make the contribution of color not very evident. The authors then investigated the role of color by designing experiments in which the shape cues were progressively degraded. They concluded that the luminance structure of the face is undoubtedly of great significance for recognition, but that color cues are not entirely discarded by the face recognition process. They suggested that color does play a role under degraded conditions by facilitating low-level facial image analysis such as better estimations of the boundaries, shape and sizes of facial features [56].

A second possible reason for a lack of work on color-based face recognition relates to the difficulties of associating illumination with white balancing of cameras. Indeed, as discussed in Sect. 9.3, illumination is still a challenging problem in automatic face recognition, therefore, there is no need to further complicate the task.

A third possible reason for ignoring color cues in the development of automatic recognition systems is the lack of color image databases[1] available for the testing of the proposed algorithms, in addition to the unwillingness to develop methods which cannot be used with the already existing monochrome databases and applications.

However, the few attempts to use color in automatic face recognition includes the work conducted by Torres et al. [48] who extended the eigenface approach to color by computing the principal components from each color component independently in three different color spaces (RGB, YUV and HSV). The final classification is achieved using a weighted sum of the Mahalanobis distances computed for each color component. In their experiments using one small database (59 images), the authors noticed performance improvements for the recognition rates when using YUV (88.14%) and HSV (88.14%) color spaces, while a RGB color space provided the

[1] Note that recently some color image databases have finally been collected (e.g., the color FERET database and the FRGC version 2 database).

same results (84.75%) when using R, G or B separately and exactly the same results as using the luminance Y only. Therefore, they concluded that color is important for face recognition. However, the experiments are very limited, as only one small face database is used and the simple eigenface approach is tested.

In another piece of work that deals with color for face recognition [20], it has been argued that a performance enhancement could be obtained if a suitable conversion from color images to a monochromatic form would be adopted. The authors derived a transformation from color to gray-scale images using three different methods (PCA, linear regression and genetic algorithms). They compared their results with those obtained after converting the color images to a monochromatic form by using a simple transformation $I = \frac{R+G+B}{3}$, and they noticed a performance enhancement of 4% to 14% using a database of 280 images. However, the database considered in the experiments is rather small, thus, one should test the generalization performance of the proposed transformation on a larger set of images from different sources.

In [40], Rajapakse et al. considered an approach based on Nonnegative Matrix Factorization (NMF) and compared the face recognition results using color and gray scale images. On a test set of 100 face images, the authors have claimed a performance enhancement when using also color information for recognition.

In [19], Jones has attempted to extend the Gabor-based approach for face recognition to color images by defining the concept of quaternions (four component hypercomplex numbers). On a relatively limited set of experiments, the author has reported a performance enhancement on the order of 3% to 17% when using the proposed quaternion Gabor-based approach instead of the conventional monochromatic Gabor-based method.

Very recently, color face recognition has been revisited by many researchers, with an aim to discover the efficient use of color for boosting the face recognition performance. For instance, inspired by the psychophysical studies indicating that color does play a role in recognizing faces under degraded conditions, Choi et al. [58] carried out extensive experiments and studied the effect of color information on the recognition of low-resolution face images (e.g., less than 20×20 pixels). By comparing the performance of grayscale and color features, the results showed that color information can significantly improve the recognition performance.

Yang et al. [55] compared the discriminative power of several color spaces for face recognition and found out that different color spaces display different discriminating power. Experiments on a large scale face recognition grand challenge (FRGC) problem also revealed that the RGB and XYZ color spaces are weaker than the I1I2I3, YUV, YIQ color spaces for face recognition. The authors proposed then color space normalization techniques for enhancing the discriminative power of different color spaces.

For color based face verification, Chan et al. [2] proposed a discriminative descriptor encoding the color information of the face images. The descriptor is formed by projecting the local face image acquired by multispectral LBP operators, into LDA space. The overall similarity score is obtained by fusing local similarity scores of the regional descriptors. The method has been tested on the XM2VTS and FRGC 2.0 databases with very promising results.

Liu and his colleagues extensively investigated the problem of color face recognition and reported very good results on FRGC database (Version 2 Experiment 4) [27–30, 52, 53]. For instance, in [27], the authors first derived new (uncorrelated, independent and discriminating) color spaces from the RGB color space by means of linear transformations. Then, vectors are formed in these color spaces by concatenating their component images to form augmented pattern vectors, whose dimensionality is reduced by PCA. Finally, an enhanced Fisher model (EFM) is used for recognition. The obtained results are better than those of methods using grayscale or RGB color images. In [29], the authors considered a hybrid color space by combining the R component image of the RGB color space and the chromatic components I and Q of the YIQ color space. Experiments on the Face Recognition Grand Challenge (FRGC) version 2 Experiment 4 showed the hybrid color space significantly improves face recognition performance due to the complementary characteristics of its component images. Since most of the experiments conducted by Liu and his team were mainly using the FRGC database, it is of interest to see how well the proposed methods generalize to other databases and settings.

9.8 Conclusions

Color is a useful cue in facial image analysis. Its use for skin segmentation and face detection is probably the most obvious, while its contribution to face recognition is not very clear. The first important issues when planning the use of color in facial image analysis are the selection of a color space and the design of a skin model. Several approaches have been proposed for these purposes, but unfortunately, there is no optimal choice. The choice made depends on the requirement of the application and also on the environment (illumination conditions, camera calibration, etc.).

Once a skin model has been defined, the contribution of color to face detection, not surprisingly, plays an important role in pre-processing the images and in the selection of the skin-like areas. Then, other refining stages can also be launched in order to find faces among skin-like regions. Color-based face detectors could be significantly much faster than other detectors which are based solely on gray-scale information, especially with large-sized images.

In relation to the contribution of color to face recognition, the issue is still under debate and among the open questions are: is color information useful for face recognition at all? If yes, how the three different spectral channels of face images should be combined to take advantages of the color information? What is the optimal color space which provides the highest discriminative power, etc.? The current results suggest that color cue has not yet shown its full potential and need further investigation. Therefore, it perhaps makes sense for current automatic face recognition systems not to rely on color for recognition because its contribution is not well established yet.

Acknowledgements The financial support of the Academy of Finland is gratefully acknowledged. J.B. Martinkauppi thanks Field NIRce which was sponsored by Botnia Atlantica, a project of the European Regional Development Fund.

References

1. Angelopoulou, E., Molana, R., Daniilidis, K.: Multispectral skin color modeling. In: Proc. IEEE Computer Society's Computer Vision and Pattern Recognition, pp. 635–642, December 2001
2. Chan, C., Kittler, J.V., Messer, K.: Multi-scale local binary pattern histograms for face recognition. In: ICB07, pp. 809–818 (2007)
3. Cho, K., Jang, J., Hong, K.: Adaptive skin-color filter. Pattern Recognit. **34**(5) (2001)
4. Do, H.C., You, J., Chien, S.: Skin color detection through estimation and conversion of illuminant color using sclera region of eye under varying illumination. In: Proc. 18th International Conference on Pattern Recognition, pp. 327–330, August 2006
5. Edwards, E.A., Duntley, S.: The pigments and color of living human skin. Am. J. Anat. **65**(1), 1–33 (1939)
6. Funt, B., Barnard, K., Martin, L.: Is machine colour constancy good enough. In: Proceedings of 5th European Conference on Computer Vision, pp. 445–459, June 1998
7. Garcia, C., Tziritas, G.: Face detection using quantized skin color regions merging and wavelet packet analysis. IEEE Trans. Multimed. **1**(3), 264–277 (1999)
8. Graf, H.P., Chen, T., Petajan, E., Cosatto, E.: Locating faces and facial parts. In: Proceedings of 1st International Workshop Automatic Face and Gesture Recognition, pp. 41–46, May 1995
9. Hadid, A., Pietikäinen, M.: A hybrid approach to face detection under unconstrained environments. In: Proc. 18th International Conference on Pattern Recognition (ICPR), vol. 1, p. 4, Hong Kong (2006)
10. Hadid, A., Pietikäinen, M., Martinkauppi, B.: Color-based face detection using skin locus model and hierarchical filtering. In: 16th International Conference on Pattern Recognition, pp. 196–200, Quebec, August 2002
11. Haiyuan, W., Qian, C., Yachida, M.: Face detection from color images using a fuzzy pattern matching method. IEEE Trans. Pattern Anal. Mach. Intell. **21**(6), 557–563 (1999)
12. Harwood, L.A.: A chrominance demodulator ic with dynamic flesh correction. IEEE Trans. Consum. Electron. **CE-22**, 111–117 (1976)
13. Hjelmas, E., Low, B.K.: Face detection: A survey. Comput. Vis. Image Underst. **83**(3), 236–274 (2001)
14. Hsu, R.L.: Face detection and modeling for recognition. PhD thesis, Michigan State University (2002)
15. Hsu, R.-L., Abdel-Mottaleb, M., Jain, A.K.: Face detection in color images. IEEE Trans. Pattern Anal. Mach. Intell. **24**(5) (2002)
16. Hsu, E., Mertens, T., Paris, S., Avidan, S., Durand, F.: Light mixture estimation for spatially varying white balance. ACM Trans. Graph. (TOG) **27**(3) (2008)
17. Hunke, M., Waibel, A.: Face locating and tracking for human-computer interaction. In: Proceedings of 1994 Conference Record of the Twenty-Eighth Asilomar Conference on Signals, Systems and Computers, pp. 1277–1281 October 1994
18. Imai, F.H., Tsumura, N., Haneishi, H., Miyake, Y.: Principal component analysis of skin color and its application to colorimetric reproduction on CRT display and hardcopy. J. Imaging Sci. Technol. **40**(5) (1996)
19. Jones, C.F.: Color face recognition using quaternionic Gabor wavelets. PhD thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia (2005)
20. Jones, C.F., Abbott, A.L.: Optimization of color conversion for face recognition. EURASIP J. Appl. Signal Process. **4**, 522–529 (2004)
21. Jones, M., Rehg, J.: Statistical color models with application to skin detection. Int. J. Comput. Vis. **46**(1) (2002)
22. Kakumanu, P., Makrogiannis, S., Bourbakis, N.: A survey of skin-color modeling and detection methods. Pattern Recognit. **40**(3) (2007)
23. Kemp, R., Pike, G., White, P., Musselman, A.: Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues. Perception **25**(1), 37–52 (1996)

24. Kim, Y.S., Na, J., Yoon, S., Yi, J.: Masked fake face detection using radiance measurements. JOSA-A **26**(4) (2009)
25. Lee, E., Ha, Y.: Automatic flesh tone reappearance for color enhancement in TV. IEEE Trans. Consum. Electron. **43**(4), 1153–1159 (1997)
26. Li, B., Xue, X., Fan, J.: A robust incremental learning framework for accurate skin region segmentation in color images. Pattern Recognit. **40**(12) (2007)
27. Liu, C.: Learning the uncorrelated, independent, and discriminating color spaces for face recognition. IEEE Trans. Inf. Forensics Secur. **3**(2), 213–222 (2008)
28. Liu, Z., Liu, C.: Fusion of the complementary discrete cosine features in the yiq color space for face recognition. Comput. Vis. Image Underst. **111**(3), 249–262 (2008)
29. Liu, Z., Liu, C.: A hybrid color and frequency features method for face recognition. IEEE Trans. Image Process. **17**(10), 1975–1980 (2008)
30. Liu, Z., Liu, C.: Robust face recognition using color information. In: ICB, pp. 122–131 (2009)
31. Marszalec, E., Martinkauppi, B., Soriano, M., Pietikäinen, M.: A physics-based face database for color research. J. Electron. Imaging **9**(1), 32–38 (2000)
32. Martinkauppi, B., Finlayson, G.: Designing a simple 3-channel camera for skin detection. In: Proc. the 12th Color Imaging Conference: Color Science and Engineering: Systems, Technologies, and Applications, pp. 151–156, November 2004
33. Martinkauppi, B., Pietikäinen, M.: Facial skin color modeling. In: Li, S.Z., Jain, A.K. (eds.) Handbook of Face Recognition, pp. 109–131. Springer, Berlin (2005)
34. Martinkauppi, B., Soriano, M., Laaksonen, M.: Behavior of skin color under varying illumination seen by different cameras in different color spaces. In: Machine Vision in Industrial Inspection IX. Proc. SPIE, vol. 4301, pp. 102–113 (2001)
35. Martinkauppi, B., Soriano, M., Pietikäinen, M.: Comparison of skin color detection and tracking methods under varying illumination. J. Electron. Imaging **14**(4) (2005)
36. Martinkauppi, B., Hadid, A., Pietikäinen, M.: Color cue in facial image analysis. In: Lukac, R., Plataniotis, K. (eds.) Color Image Processing: Methods and Applications, pp. 285–308. CRC Press, Boca Raton (2006)
37. Martinkauppi, B., Lehtonen, J., Parkkinen, J.: Near-infrared images of skin. In: Proc. 4th European Conference on Colour in Graphics, Imaging, and Vision, 10th International Symposium on Multispectral Colour Science, pp. 508–511, June 2008
38. Nakai, H., Manabe, Y., Inokuchi, S.: Simulation and analysis of spectral distribution of human skin. In: Proc. 14th International Conference on Pattern Recognition, pp. 1065–1067 (1998)
39. Raja, Y., McKenna, S., Gong, G.: Tracking and segmenting people in varying lighting conditions using colour. In: Proceedings of IEEE 3rd International Conference on Automatic Face and Gesture Recognition, pp. 228–233, April 1998
40. Rajapakse, M., Tan, J., Rajapakse, J.: Color channel encoding with NMF for face recognition. In: IEEE Conference on Image Processing, vol. 3, pp. 2007–2010 (2004)
41. Sharma, P., Reilly, R.: A colour face image database for benchmarking of automatic face detection algorithms. In: EC-VIP-MC 2003 4th EURASIP Conference focused on Video/Image Processing and Multimedia Communications, pp. 423–428 (2003)
42. Sobottka, K., Pitas, I.: Face localization and facial feature extraction based on shape and color information. In: IEEE Conference on Image Processing, vol. 3, pp. 483–486 (1996)
43. Soriano, M., Martinkauppi, B., Huovinen, S., Laaksonen, M.: Adaptive skin color modeling using the skin locus for selecting training pixels. Pattern Recognit. **36**(3), 681–690 (2003)
44. Störring, M., Andersen, H.J., Granum, E.: Physics-based modelling of human skin colour under mixed illuminants. J. Robot. Auton. Syst. **35**(3–4), 131–142 (2001)
45. Sun, H.: Skin detection for single images using dynamic skin color modeling. Pattern Recognit. **43**(4) (2010)
46. Terrillon, J.C., Shirazi, M., Fukamachi, H., Akamatsu, S.: Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human face in color images. In: IEEE Int. Conf. on Automatic Face and Gesture Recognition, pp. 54–61 (2000)
47. Tomaschitz, J.A., Facon, J.: Skin detection applied to multi-racial images. In: Proc. 16th International Conference on Systems, Signals and Image Processing IWSSIP, pp. 1–3, June 2009

48. Torres, L., Reutter, J., Lorente, L.: The importance of the color information in face recognition. In: IEEE Conference on Image Processing, vol. 3, pp. 627–631 (1999)
49. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proc. Conf. Computer Vision and Pattern Recognition, pp. 511–518 (2001)
50. Wyszecki, G., Stiles, W.S. (eds.): Color Science Concepts and Methods, Quantitative Data and Formulae, 2nd edn. Wiley, New York (2000)
51. Yang, M.H., Ahuja, N.: Detecting human faces in color images. In: Proceedings of International Conference on Image Processing, pp. 127–130 (1998)
52. Yang, J., Liu, C.: A general discriminant model for color face recognition. In: ICCV, pp. 1–6 (2007)
53. Yang, J., Liu, C.: Color image discriminant models and algorithms for face recognition. IEEE Trans. Neural Netw. **19**(12), 2088–2098 (2008)
54. Yang, M.-H., Kriegman, D.J., Ahuja, N.: Detecting faces in images: a survey. IEEE Trans. Pattern Anal. Mach. Intell. **24**, 34–58 (2002)
55. Yang, J., Liu, C., Zhang, L.: Color space normalization: Enhancing the discriminating power of color spaces for face recognition. Pattern Recognit. **43**(4), 1454–1466 (2010)
56. Yip, A.W., Sinha, P.: Contribution of color to face recognition. Perception **31**(8), 995–1003 (2002)
57. Yoo, T., Oh, I.: A fast algorithm for tracking human faces based on chromaticity histograms. Pattern Recognit. Lett. **20**(10) (1999)
58. Young, C.J., Man, R.Y., Plataniotis, K.N.: Color face recognition for degraded face images. IEEE Trans. Syst. Man Cybern., Part B, Cybern. **39**(5), 1217–1230 (2009)