An Illumination invariant skin-color model for face detection

Ukil Yang, Minsung Kang, Kar-Ann Toh and Kwanghoon Sohn, Member, IEEE

Abstract— Face detection is an important step towards a fully automatic face recognition system. Among existing techniques in the literature, methods based on skin-color information have shown computational effectiveness as well as robustness in terms of rotation, scaling and partial occlusion. However, due to color variations resulted from illumination changes, many color-based techniques have yet to demonstrate a stable state of performance. In this paper, we present an illumination invariant color space model to address the color variation issue. The proposed method is evaluated both in terms of skin-color detection and face detection performances. Our empirical experiments evidenced both effectiveness and usefulness of the proposed method.

I. INTRODUCTION

Face detection techniques have been broadly studied, particularly as a pre-processing technique in many vision-based applications, such as content-based image retrieval, video conferencing, crowd surveillance and intelligent human–computer interfaces [1]. Particularly, face detection is an essential step in face recognition, with the purpose of localizing and extracting face regions from the background.

Various approaches to face detection are well listed in [1] and [2]. Many approaches utilize techniques such as principle component analysis, neural networks, machine learning, information theory, geometrical modeling, template matching, Hough transform, motion extraction and color analysis [3]. Among these techniques, methods based on skin-color information have shown computational effectiveness as well as robustness in terms of rotation, scaling and partial occlusion [2]. However, the color-based approaches are error prone under different illumination conditions [4].

To address the issue of color changes under illumination variations, this paper presents an illumination invariant color space, and proposes a method for skin-color. Then, we evaluate the face detection method using public face image database.

The remainder of the paper is organized as follows: Section II gives a brief description of an image acquisition model. An illumination invariant color space and the proposed skin-color detection method are described in Section III. Our experimental results and analyses are described in Section IV. Finally, section V provides a summary and some concluding remarks.

II. AN IMAGE ACQUISITION MODEL

As illustrated in Fig. 1, an image data is usually acquired by a camera under a certain illumination condition. The image acquisition process could be modeled based on a Lambertian surface given by Eq. (1) [5, 6].

\[
\begin{align*}
\rho_R &= \sigma \int E(\lambda)S(\lambda)Q_R(\lambda) \, d\lambda \\
\rho_G &= \sigma \int E(\lambda)S(\lambda)Q_G(\lambda) \, d\lambda \\
\rho_B &= \sigma \int E(\lambda)S(\lambda)Q_B(\lambda) \, d\lambda
\end{align*}
\]

where \(\lambda\) denotes the wavelength of light and \(\rho_R\), \(\rho_G\) and \(\rho_B\) denote the respective color values acquired by a RGB camera. The spectral sensitivities of the RGB sensors are denoted by \(Q_R(\lambda)\), \(Q_G(\lambda)\) and \(Q_B(\lambda)\), and the spectral power distribution of a light and surface reflectance of an object are denoted by \(E(\lambda)\) and \(S(\lambda)\), respectively. \(\sigma\) is a constant factor called the Lambertian shading term which represents the angle between the surface normal and the illumination direction.

Two assumptions are commonly adopted in order to derive useful illumination factors for the RGB color values in Eq. (1). According to [6, 7], a delta function has been assumed to simplify the RGB spectral sensitivity formulation of a camera. This is grounded on the argument that the RGB color values of a pixel \((\rho_R, \rho_G, \rho_B)\) are integral values of all wavelengths where the spectral sensitivities of the RGB camera sensors could be approximated using delta functions with an appropriate choice of scaling factors: \(q_R\), \(q_G\), and \(q_B\) (see Fig. 2). Based on this assumption, Eq. (1) can be simplified as shown in Eq. (2).

The second assumption is that the spectral power distribution of light \((E(\lambda))\) follows the spectral power distribution of a black body given by Planck's law [5] as shown in Eq. (3).
Fig. 1. Image acquisition process

Fig. 2. The example of the spectral sensitivities of a real camera: Measurement form (a) [7], approximation form (b)

\[
\begin{align*}
\rho_b &= \sigma E(\lambda_b)S(\lambda_b)q_b \\
\rho_g &= \sigma E(\lambda_g)S(\lambda_g)q_g \\
\rho_r &= \sigma E(\lambda_r)S(\lambda_r)q_r
\end{align*}
\]

\[E(\lambda) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/\lambda T} - 1} = \frac{c_1}{\lambda^2} \frac{1}{e^{hc/\lambda T} - 1}
\]

where \( T \) denotes the temperature of a black body. \( c, h \) and \( k \) are constants which denote the speed of light, 3.8 × 10^8 (m/s), Planck constant, 6.626 × 10^{-34} (Js) and Boltzmann constant, 1.380 × 10^{-23} (J/K), respectively. \( c_1 \) and \( c_2 \) are also constants being defined as:

\[c_1 = 2hc^2, \quad c_2 = \frac{hc}{k}
\]

As seen from Eq. (3), the Planck’s law describes the spectral power distribution at all wavelengths from a black body at temperature \( T \). Although light could be radiated from sources other than a black body, the sun which constitutes the major light source of the earth is a black body where Planck’s Law is a good approximation for incandescent and daylight illuminants [5, 6]. Therefore, adopting Planck’s law to describe the spectral power distribution of a light \( E(\lambda) \) in Fig. 1 is reasonable for our applications.

Moreover, when a color image is acquired under an ordinary illumination condition, the Planck’s law can be further approximated by Eq. (5) as shown below. This is because the RGB camera sensors generally work within the visible range (380 nm ≤ λ ≤ 770 nm, see Fig. 2), and the color temperature of natural illumination is below 10,000K. Putting this second assumption into the first leads to a further simplification of Eq. (2) into Eq. (6).

\[E(\lambda) = \frac{c_1}{\lambda^2} \frac{1}{e^{hc/\lambda T} - 1} = \frac{c_1}{\lambda^2} e^{-hc/\lambda T}
\]

Consequently, when an object whose surface reflectance \( S(\lambda) \) is captured by a camera having a weighted delta spectral sensitivity (Eq. (2)) under an illumination condition \( \left( \frac{c_1}{\lambda^2} e^{-hc/\lambda T} \right) \), the RGB colors of the object can be represented as \( \rho_b, \rho_g, \rho_r \) in Eq. (6).

That is, according to the image acquisition model, the color values of an image are determined by non-illumination terms \( (S(\lambda) \) which contains the object characteristics and \( \lambda_b, \lambda_g, \lambda_r \), \( q_b, q_g, q_r \) and \( q_b \) for camera characteristics) and illumination terms \( (T \) is the color temperature of illumination source and \( \sigma \) is the angle between the surface normal of an object and the illumination direction).

III. PROPOSED METHOD

An image acquisition model has been presented in section II. This model will be investigated here in order to eliminate or reduce those illumination related terms such as ‘\( T \)’ and ‘\( \sigma \)’ in Eq. (6). The RGB values are first normalized using the geometric mean to get rid of the impact of the ‘\( T \)’ and ‘\( \sigma \)’ term. Then, as a pre-processing step to extract the ‘\( T \)’ term, a natural logarithm is taken for the normalized values where we obtain Eq. (7-1) [6]. As shown in Eq. (7-2), since the sum of these components is ‘0’, three components can be reduced to two components without any loss of color information. In this paper, we use \( \rho_b' \) and \( \rho_g' \) to represent the color information. Eq. (7-1) is next re-written as Eq. (8-1) in vector form for \( \rho_b' \) and \( \rho_g' \).

\[\rho_b' \triangleq \ln \left( \frac{\rho_b}{\sqrt{\rho_b \rho_g \rho_r}} \right), \quad \rho_g' \triangleq \ln \left( \frac{\rho_g}{\sqrt{\rho_b \rho_g \rho_r}} \right), \quad \rho_r' \triangleq \ln \left( \frac{\rho_r}{\sqrt{\rho_b \rho_g \rho_r}} \right)
\]

\[\rho_b' + \rho_g' + \rho_r' = \ln \left( \frac{\rho_b}{\sqrt{\rho_b \rho_g \rho_r}} \right) + \ln \left( \frac{\rho_g}{\sqrt{\rho_b \rho_g \rho_r}} \right) + \ln \left( \frac{\rho_r}{\sqrt{\rho_b \rho_g \rho_r}} \right)
\]

\[= \ln \left( \frac{\rho_b}{\sqrt{\rho_b \rho_g \rho_r}} \times \frac{\rho_g}{\sqrt{\rho_b \rho_g \rho_r}} \times \frac{\rho_r}{\sqrt{\rho_b \rho_g \rho_r}} \right) = 0
\]
\[ \tilde{\rho} = \tilde{S} + \frac{c_2}{T} \tilde{\lambda} \]  

(8-1)

where, \( \tilde{\rho} \), \( \tilde{\lambda} \) and \( \tilde{S} \) are as follow;

\[
\tilde{\rho} = \begin{bmatrix} \rho'_R \\ \rho'_G \\ \rho'_B \end{bmatrix}, \quad \tilde{\lambda} = \begin{bmatrix} \lambda'_R \\ \lambda'_G \\ \lambda'_B \end{bmatrix}, \quad \tilde{S} = \begin{bmatrix} S'_R \\ S'_G \\ S'_B \end{bmatrix}
\]

\[
\left[ \begin{array}{c} \frac{\lambda'_G}{\lambda'_R} \cdot S(G)q_R \\ \frac{\lambda'_R}{\lambda'_G} \cdot S(R)q_G \end{array} \right] \left[ \begin{array}{c} \frac{\lambda'_G}{\lambda'_R} \cdot S(G)q_R \\ \frac{\lambda'_R}{\lambda'_G} \cdot S(R)q_G \end{array} \right] = \left[ \begin{array}{c} \lambda'_G \cdot S(G)q_R \\ \lambda'_R \cdot S(R)q_G \end{array} \right]
\]

As shown in Eq. (8-1), \( \tilde{\rho} \) is a color vector defined in the new color space represented by ‘geometric mean normalization’ and ‘taking logarithm’. (We just call it ‘Normal-Log color space’.) Fig. 3 is graphical representation of Eq. (8-1). As shown in Fig. 3, the essential color information of an object (\( \tilde{S} \)) is distorted by illumination factor \( \frac{c_2}{T} \tilde{\lambda} \). As a result, \( \tilde{S} \) is represented as \( \tilde{\rho} \) which is the color information of an object in an image. Thus, due to color variations resulted from illumination changes, the color information of an object become \( \tilde{\rho} \) in an image, even though the essential color information of an object is \( \tilde{S} \).

Fig. 3. Graphical representation of Eq. (8-1)

If we take an image under different illumination conditions for the same object using the same camera, how is the essential color information of an object (\( \tilde{S} \)) represented in an image? According to Eq. (8-1), the variation of illumination only affects the \( \frac{c_2}{T} \tilde{\lambda} \) term. Moreover, because \( \tilde{\lambda} \) is defined by using the parameters of Dirac delta camera model (see Eq. (8-2)), \( \tilde{\lambda} \) is not affected by the variation of illumination. Therefore, illumination only affects the scale of the \( \frac{c_2}{T} \tilde{\lambda} \) term, and if the scale factor of the \( \frac{c_2}{T} \tilde{\lambda} \) term \( \left( \frac{c_2}{T} \right) \) could be removed, the illumination invariance can be guaranteed.

The procedure to remove the \( \frac{c_2}{T} \) term is described as follows. In Eq. (8-1), \( \tilde{\rho} \) is calculated from the RGB values of an acquired image where \( c_2 \) is a constant and \( \tilde{S} \), \( \tilde{\lambda} \), \( T \) are unknown variables, respectively. In order to remove the \( \frac{c_2}{T} \) term, we apply an inner product with an ortho-normal vector of \( \tilde{\lambda} \), which is also an unknown variable and denoted as \( \tilde{\lambda}^\perp \) \( (\tilde{\lambda} \cdot \tilde{\lambda}^\perp = 0 \) and \( \| \tilde{\lambda}^\perp \| = 1 \), to Eq. (8-1) resulting in Eq. (9).

\[
\tilde{\rho} \cdot \tilde{\lambda}^\perp = \left( \tilde{S} + \frac{c_2}{T} \tilde{\lambda} \right) \cdot \tilde{\lambda}^\perp \\
\Rightarrow \tilde{\rho} \cdot \tilde{\lambda}^\perp = \tilde{S} \cdot \tilde{\lambda}^\perp \quad \text{since} \quad \tilde{\lambda} \cdot \tilde{\lambda}^\perp = 0
\]

Eq. (9) could be graphically represented in Fig. 4. Thus, changes of the illumination condition when an image is acquired lead to changes of an object color in an image, and these changes lie on a line which is parallel to \( \tilde{\lambda} \). Moreover, the essential color information of an object (\( \tilde{S} \)) also lies on the line. Therefore, when the essential color information of an object (\( \tilde{S} \)) and the color information of the object represented in an image (\( \tilde{\rho} \)) are projected on \( \tilde{\lambda}^\perp \), the two projected values are always the same irrespective of the illumination condition. Namely, what Normal-Log color space projected on \( \tilde{\lambda}^\perp \) is an illumination invariant color space.

Fig. 4. Graphical representation of Eq. (9)

To use an illumination invariant color space for skin-color detection, we have to know \( \tilde{\lambda}^\perp \). However, there is no information about \( \tilde{\lambda}^\perp \), but \( \tilde{\lambda} \cdot \tilde{\lambda}^\perp = 0 \) and \( \| \tilde{\lambda}^\perp \| = 1 \). \( \tilde{\lambda} \) is defined by using the parameters of Dirac delta camera model (see Eq. (8-2)). Since \( \tilde{\lambda}^\perp \) is an ortho-normal vector of \( \tilde{\lambda} \), \( \tilde{\lambda}^\perp \) could be estimated from the color characteristics of a camera. However, it is difficult to measure the color characteristics of a camera without certain equipment (such as ‘camSPECS’ made by image engineering). Moreover, in general, camera companies do not reveal the color characteristics of their cameras. For these reasons, we estimated \( \tilde{\lambda}^\perp \) by using a training dataset.
To estimate $\lambda^i$, we use Eq. (9) and a training dataset which is acquired under some established conditions. As shown in Fig. 5, we collected a training dataset under various illumination conditions for each object using the same camera. Under this condition, the variance of $\bar{\lambda} \cdot \lambda^i$ in Eq. (9) should be ideally ‘0’, because $\bar{\lambda}$ and $\lambda^i$ are static for all $\lambda$ in the training dataset. Eventually, $\lambda^i$ can be estimated from the training dataset by minimizing the variance of $\bar{\lambda} \cdot \lambda^i$ given by Eq. (10) below. Eq. (10) is easily solved based on principle components analysis (PCA) [8, 9]. $\lambda^i_{\text{opt}}$ is eigen-vector corresponding to the smallest eigen-value of PCA by using a training dataset. Fig. 6 shows some examples of a training dataset.

$$\lambda^i_{\text{opt}} = \arg\min_{\lambda^i} \left\{ \text{var} \left[ \bar{\lambda} \cdot \lambda^i \right] \right\}$$  \hspace{1cm} (10)

The methodology for illumination invariant skin-color detection is summarized in Fig. 7. The method is divided into ‘Training phase’ and ‘Testing phase’. In the training phase, the skin-color database is acquired under various illumination conditions for a skin-color object with the same camera in order to define the ‘$\lambda_{\text{opt}}^i$’ of the camera. Based on Eq. (10), the $\lambda_{\text{opt}}^i$ of skin-color is calculated using the acquired database. Based on the $\lambda_{\text{opt}}^i$, the $\bar{\lambda} \cdot \lambda_{\text{opt}}^i$ values of the skin-color database for training is calculated. By using the above values, the illumination invariant skin-color model is finally generated.

In the testing phase, the same camera which was used for image acquisition in the training database should be used to acquire a new input image for accurate results, and the skin-color pixels are extracted from a new input image based on the results of the training phase (the skin-color model) and a classification tool (such as a Neural Network, a Support Vector Machine, a Bayesian classifier, and etc.).
We manually cropped those skin-color patches from an Expanded Skintone colors in the Digital ColorChecker® SG as shown in Fig. 8-(c). Some cropped samples are shown in Fig. 9. There are fourteen mutually different skin-color patches in Digital ColorChecker® SG, and each patch shows the color variations according to color temperature and intensity variations in illumination. As shown in Fig. 9, even if an object color was not changed, the represented color of the object in an image was observed to be varying due to color temperature and intensity variations under illumination.

To confirm the practicality of the illumination invariant color space derived the image acquisition model based on Lambertian surface, we calculate the \((\rho_B', \rho_R')\) values of the cropped patches based on Eq. (7), and then plot their values under \(\rho_B'\) and \(\rho_R'\) axes. As shown in Fig. 10, the \((\rho_B', \rho_R')\) values are seen to vary according to illumination variation, and the distribution of their values show a certain direction. Moreover, these directions appear to be similar for all acquired patches. That is, it is confirmed that the directions are not affected by illumination, and Eq. (8-1) is the practical model for the color distortion by an illumination. In other words, the proposed method for illumination invariant skin-color detection which is derived from Eq. (8-1) is adoptable in real applications.

![Fig. 10. Distribution of \((\rho_B', \rho_R')\) values of each Skintone patch according to illumination variation (red points) and estimation results of the \(\hat{\lambda}_{opt}\) (line)

To evaluate the performance of the proposed method for skin-color detection in terms of the illumination invariant property, we performed the skin-color detection test with the Face Recognition Grand Challenge (FRGC) database [10] (Fall 2002 session, 247 people, 1976 pieces). First, as shown in Fig. 11, we manually segment the skin regions, and define the segmentation results as truth maps for performance evaluation regarding skin-color detection. In training phase, we estimated \(\hat{\lambda}_{opt}\) based on the minimum variance estimator described in Eq. (10) by using the segment results (3 pieces per person acquired in different illumination conditions). Based on the estimation results \((\hat{\lambda}_{opt}\)), we next constructed the skin-color model. The RGB color values of skin regions are converted into values in illumination invariant color space using Eq. (7-1) and \(\rho \cdot \hat{\lambda}_{opt}^L\). Then, the skin-color model is constructed based on the distributions of the obtain values. In the testing phase, we extracted the skin-color pixels from the test images (1235 pieces, not include training dataset) by the Bayesian rule based on the skin-color model used in the training phase with the assumption that non-skin-color is uniformly distributed.

![Truth map](image)

Fig. 11. Examples of the FRGC database (up) and segmentation results of skin regions (down)

Table 1 shows the detection results of the skin-color pixels. As a test result, we calculated the False Acceptance Rate (FAR) and the False Rejection Rate (FRR) based on the truth map as shown in Fig. 11.

<table>
<thead>
<tr>
<th>Color space</th>
<th>Classification</th>
<th>FRR</th>
<th>FAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed method</td>
<td>RGB</td>
<td>1.05%</td>
<td>5.32%</td>
</tr>
<tr>
<td>Ruiz-del-Solar’s</td>
<td>RGB</td>
<td>15.27%</td>
<td>27.60%</td>
</tr>
<tr>
<td>method [11]</td>
<td>YCbCr</td>
<td>8.72%</td>
<td>15.33%</td>
</tr>
<tr>
<td>Garcia’s method</td>
<td>YCbCr</td>
<td>3.38%</td>
<td>4.72%</td>
</tr>
<tr>
<td>[12]</td>
<td>Bayesian rule</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigal’s method</td>
<td>I, RG, BY</td>
<td>5.83%</td>
<td>7.82%</td>
</tr>
<tr>
<td>[13]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fleck’s method</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[14]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As shown in Table 1, the FAR and FRR based on the proposed method are a lot better than those of other methods. Particularly, FRR is distinctively low. From these results, we confirm that the proposed method for skin-color detection is robust to variations of illumination conditions.

As a real application, we applied the proposed method for illumination invariant skin-color detection. The experiment on face detection was performed using the University College Dublin Colour Face Image (UCDC) database [15] and face detection library (FDLIB) [16]. Fig. 12 shows some example results of face detection. In Fig. 12-(a), FDLIB was used for face detection. Here, the FRR is 6/55 (10.91%), and the FAR is 8/57 (14.04%). Fig. 12-(b) shows the result when the proposed method was applied to pre-process the image before using FDLIB. Here, the FRR is 6/55 (10.91%), and the FAR is 0/49 (0%). From these result, we see that when the proposed method is applied for pre-processing prior to face detection, the FAR is distinctively decreased while the FRR
being unaffected. When we performed the experiments on face detection with all the images in UCDC database (94 images, 321 faces), the FRR is seen to decrease from 37/321 (11.53%) to 31/321 (9.66%), the FAR is seen to decrease from 64/348 (18.39%) to 2/292 (0.68%) with the application of our pre-processing method. These results show that when the proposed method for illumination invariant skin-color detection is applied as a pre-processing step prior to face detection, the FAR is significantly decreased with the FRR being slightly decreased (or unchanged).

(a)

(b)

Fig. 12. Result of face detection:
(a) FDLIB only, (b) Skin-color detection + FDLIB

V. CONCLUSION

The color attributes of an image provide essential information in many vision based tasks. Although a general color-based digital image processing technique may show effectiveness in terms of implementation and processing time, it suffers from color variations caused by irregular illumination changes. To overcome this problem in skin-color detection, we have presented an illumination invariant color space derived from an image acquisition process based on the Lambertian surface. Based on this result, we have proposed a new methodology for illumination invariant skin-color detection.

By removing those terms related to illumination, we sought an illumination invariance model. Since the formulation involved an illumination model and a camera model, both the illumination characteristics and the camera characteristics have been utilized. This established an illumination invariant color space and a new methodology for skin-color detection.

To evaluate the performance of the proposed method, we performed skin-color extraction using public face image database acquired under varying illumination conditions. Our experiments verified that the proposed method was robust to skin-color detection under those tested variation of illumination conditions.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea(NRF) through the Biometrics Engineering Research Center(BERC) at Yonsei University. (No. R112002105070030(2010))

REFERENCES