Color Constancy in 3D-2D Face Recognition

Manuel Meyer\textsuperscript{1}, Christian Riess\textsuperscript{1}, Elli Angelopoulou\textsuperscript{1}, Georgios Evangelopoulos\textsuperscript{2} and Ioannis A. Kakadiaris\textsuperscript{2}

\textsuperscript{1} University of Erlangen-Nuremberg, Martensstr. 3, 91048 Erlangen, Germany; \textsuperscript{2} University of Houston, 4800 Calhoun Rd., Houston, Texas, USA

ABSTRACT

Face is one of the most popular biometric modalities. However, up to now, color is rarely actively used in face recognition. Yet, it is well-known that when a person recognizes a face, color cues can become as important as shape, especially when combined with the ability of people to identify the color of objects independent of illuminant color variations. In this paper, we examine the feasibility and effect of explicitly embedding illuminant color information in face recognition systems. We empirically examine the theoretical maximum gain of including known illuminant color to a 3D-2D face recognition system. We also investigate the impact of using computational color constancy methods for estimating the illuminant color, which is then incorporated into the face recognition framework. Our experiments show that under close-to-ideal illumination estimates, one can improve face recognition rates by 16%. When the illuminant color is algorithmically estimated, the improvement is approximately 5%. These results suggest that color constancy has a positive impact on face recognition, but the accuracy of the illuminant color estimate has a considerable effect on its benefits.

Keywords: Face Recognition, Color Constancy, Illumination Normalization, Illumination Invariance

1. INTRODUCTION

The human face is an important identifier for biometric analysis.\textsuperscript{1} It does not require a cooperative subject, like for instance fingerprint or iris authentication, and is socially well accepted.\textsuperscript{1} At the same time, the human face is one of the most descriptive and easiest-to-access inputs to biometric systems. Already today, notebooks, door-locks, or even smartphones offer face-recognition based authentication. Despite the impressive progress on face recognition (FR) systems, variations in pose, expression, and illumination still affect the reliability of such systems.

A significant factor in face recognition by humans is pigmentation. As Sinha \textit{et al.}\textsuperscript{2} demonstrated, for humans the intricacies of face reflectance (i.e., albedo, hue, translucency, shadows, specularities, illumination direction and their spatial variation) are at least as important as shape information when recognizing a face. Automated face recognition (FR) systems have tried to address the influence of these factors, but have mostly focused on albedo and illumination direction.\textsuperscript{3-5}

An important aspect that, to our knowledge, has not been sufficiently addressed in FR systems is the ability of humans to perceive the color of objects independently of the color of the incident illumination, a property also known as color constancy. Given that pigmentation plays a considerable role in face recognition, one would prefer to use a pigmentation descriptor that is independent of the color of the illuminant. Though color constancy has been extensively studied,\textsuperscript{6} its impact on FR systems has not been rigorously evaluated.

In this paper, we investigate the impact that the illumination color has on the performance of a FR system. We examine how the recognition accuracy is affected when information about the illuminant color is explicitly incorporated in face recognition. For that purpose we chose to use the FR system proposed by Toderici \textit{et al.},\textsuperscript{7} since it is a fully-automated 3D-2D FR system, which allows, through its relighting process, the inclusion of illuminant color information. We created a small database of facial images under different illumination colors. Our evaluation showed that optimal information about the illuminant color has a considerable impact

Further author information: (Send correspondence to Ioannis A. Kakadiaris)
Ioannis A. Kakadiaris: E-mail: ioannisk@uh.edu, Telephone: +1 (713) 743-1255
on FR accuracy. We also drew from the work on computational color constancy and directly estimated the illuminant color in the face images, using a state-of-the-art physics-based method.\textsuperscript{8} Our experiments show that the automatically estimated illuminant color has a positive effect on FR accuracy, especially at lower false positive rates. However, the largest improvement in recognition rates is achieved when approximately ideal information about the illuminant color is available. Hence, our analysis suggests that the effects of illuminant color on a FR system should be explicitly addressed. The inclusion of color constancy methods for automatic color correction in FR systems is feasible and has a positive impact. There is, however, a need for color constancy methods explicitly developed for skin reflectance (see also Sec. 2) in order to achieve the full potential of incorporating illuminant color in FR.

In summary, our contributions are:

- the creation of a small database for evaluating the influence of color constancy in FR,
- the explicit incorporation of illuminant color in a 3D-2D FR system,
- the affirmation that integrating a known illuminant greatly improves FR, and
- the insight that it is important to estimate the illuminant color in at least the probe image.

In Sec. 2, we briefly review the most closely related work. Our illuminant-adaptive FR framework is presented in Sec. 3. We describe our dataset in Sec. 4. The evaluation on ground truth illumination, on estimated illumination, and the impact of both on face recognition is presented in Sec. 5. A summary and brief remarks are provided in Sec. 6.

### 2. PREVIOUS WORK

A typical approach for including skin color information into a 3D-2D face recognition system, is to directly compute the skin albedo (i.e., the illumination-invariant reflection coefficient) from the 2D input. For instance, Smith and Hancock\textsuperscript{9} and Lee et al.\textsuperscript{10} do so by assuming Lambertian reflection. Biswas et al.\textsuperscript{5} propose a method that assumes shadow-free images and does not handle specular reflection. Toderici et al.\textsuperscript{7} demonstrate that, compared to albedo estimation, bidirectional relighting results in improved FR rates. However, they assume a white illuminant.

In real-world scenarios, this is not often the case. The deviation from white light has a negative impact on many imaging applications. Therefore, a number of computer vision methods, known collectively as color constancy algorithms, aim to create illumination-invariant representations of a scene. For a recent survey on existing techniques, see Gijsenij et al.\textsuperscript{6} In general, color constancy itself is an underconstrained problem. Statistical methods circumvent this by learning a set of commonly used illuminants.\textsuperscript{11–13} Hence, they rely heavily on the quality of the training data. Unusual illuminants are typical failure cases for these methods. Alternatively, physics-based approaches introduce additional constraints like the presence of purely diffuse or specular surfaces,\textsuperscript{12,14} whose reliable identification is a challenging task in itself.

A couple of computational color constancy methods which exploit the presence of skin have been developed. Störring et al.\textsuperscript{15} introduced a methodology that can estimate the illuminant color from skin regions. However, they assume known camera parameters and Planckian illuminants. They also require an initial estimate of the illuminant color. Thus, though this work was pioneering in exploiting skin reflectance, it has limited applicability in FR systems. Very recently, Bianco and Schettini\textsuperscript{16} developed a statistical method that exploits the presence of skin clusters in color space. They learn the expected gamut of skin in HSV space under canonical illumination conditions. The illuminant color is then the mapping between the gamut of detected skin pixels and the skin canonical gamut. However, as a gamut-based method, their approach has difficulty handling illuminant colors that deviate from the expected canonical light set.\textsuperscript{16}

Despite the large body of work on color constancy, there has been very limited work in face recognition on skin pigmentation under varying illumination colors. Kanan et al.\textsuperscript{17} studied the impact of color constancy algorithms on object recognition and face recognition in particular. However, they only examined FR under
different illuminant directions, so the impact of varying illuminant color on face recognition has not yet been analyzed.

The main goal of our work is to quantify the impact of illumination color on face recognition. In order to do that, we measure the performance of a FR system:

1. under the assumption of a white illuminant
2. when incorporating near-perfect illuminant color information, as obtained from a ground truth white Lambertian sample present in the scene,
3. in a more realistic setup, which uses a color constancy (CC) algorithm to estimate the illuminant color.

For the algorithmic estimation of the illuminant color we employed the physics-based method by Riess et al. It performs comparably to other state-of-the-art CC methods but: (i) it requires no training; (ii) it requires no specularity segmentation; (iii) it operates on small image regions; and (iv) it can recover unusual illumination conditions. While preparing this paper, Bianco and Schettini published a method for illuminant color estimation on faces, which is not yet incorporated in this study. The FR system used in our study is the 3D-2D system by Toderici et al., which uses bidirectional relighting and, thus, allows for the incorporation of the illuminant color information in both/either the probe and/or the gallery image.

3. FRAMEWORK OF ILLUMINANT-COLOR AWARE FR

Our FR framework uses jointly 3D and 2D captures for the gallery. The probe data is a 2D image. An overview of an authentication scenario is shown in Fig. 1. Typically, methods where both the gallery and probe are 3D datasets offer the highest recognition rates. Yet in real-world setups, it is much easier to generate 3D gallery data than 3D probes, since the latter should be captured on demand at any time and place. This necessitates a flexible, easy-to-deploy probe acquisition system, like a color camera. Hence, the FR system we are using is a compromise between a very accurate but cumbersome 3D-3D pipeline, and a flexible but more error-prone 2D-2D system.

3.1 3D-2D Face Recognition Framework

We briefly describe the enrollment and recognition pipeline. The basic steps of our system follow the 3D-2D system of Toderici et al. (see Fig. 1). During the enrollment of each gallery subject, a annotated face model (AFM) is registered to the subject’s 3D mesh. The AFM is then deformed to smoothly fit the original
mesh. Next, a connection between this adapted 3D mesh and the 2D gallery data is established. Based on localized landmarks in 3D and 2D, the face image is registered to the AFM. A texture map is computed through mapping 2D image pixels to 3D points in the face model. An additional visibility map handles occlusions. During authentication, the 2D probe image is mapped to each gallery subject’s deformed AFM. The process is basically the same as for the gallery enrollment. The probe’s image data is registered to the gallery subject’s mesh, a texture and visibility map are generated. Using a bidirectional relighting method, the illumination conditions of the gallery image are adapted to the probe’s conditions. Finally, a similarity score is computed between the probe texture and a single gallery subject’s relit texture. This process is repeated for all subjects in the gallery.

3.2 Relighting and Illumination Color

Our particular interest lies in the bidirectional relighting. Its purpose is to adapt a subject’s texture in the gallery to the lighting conditions of the probe. It has been shown that relighting outperforms albedo estimation, which makes it our preferred framework for the adaptation to changing lighting conditions. The main idea is to consider a texture $T$ evaluated at point $\vec{x}$ as

$$T^g(\vec{x}) = \tilde{i}_s(\vec{x}) + (\tilde{i}_d(\vec{x}) + \tilde{i}_a(\vec{x})) \cdot \vec{a}(\vec{x})$$

(1)

where $\tilde{i}_s(\vec{x})$ and $\tilde{i}_d(\vec{x})$ are the specular and diffuse components of a scene’s reflectance at $\vec{x}$, $\tilde{i}_a(\vec{x})$ denotes the ambient illumination at $\vec{x}$ and $\vec{a}(\vec{x})$ denotes the albedo at $\vec{x}$. These reflectance components are defined as

$$\tilde{i}_s(\vec{x}) = k_s(\vec{x}) \vec{c}(\vec{x}) (\vec{r}(\vec{x}), \vec{v}(\vec{x}))^n$$

(2)

$$\tilde{i}_d(\vec{x}) = \vec{c}(\vec{x}) \cos \theta$$

and

(3)

$$\tilde{i}_a(\vec{x}) = \vec{c}(\vec{x})$$

(4)

In these equations, $k_s(\vec{x})$ is the specular reflection coefficient and $\vec{c}(\vec{x})$ is the spectral distribution of the light source at $\vec{x}$ (i.e., the illuminant color). In Eqn. 2, $\vec{r}(\vec{x})$ is the vector of the reflected light direction, $\vec{v}(\vec{x})$ is the viewing direction vector, and $n$ is a scalar. The diffuse term in Eqn. 3 is dependent on the angle $\theta$, spanned by the surface normal and the light direction vector. Toderici et al. proposed to adjust the illumination conditions of two images without explicitly estimating the albedo, by minimizing the residual error

$$\epsilon = \sum_{\vec{x}} \| R(\vec{x}) \|_2$$

(5)

Here, $R$ denotes the error (or residual) between the relit representation and the probe:

$$R(\vec{x}) = T^p(\vec{x}) - \tilde{T}_s^p(\vec{x}) - (\tilde{T}_d^p(\vec{x}) + \tilde{T}_a^p(\vec{x})) \cdot \frac{T^g(\vec{x}) - \tilde{T}_s^g(\vec{x})}{\tilde{T}_d^g(\vec{x}) + \tilde{T}_a^g(\vec{x})}$$

(6)

The term $T^g(\vec{x})$ represents the gallery texture, and $\tilde{T}_s^g(\vec{x})$, $\tilde{T}_d^g(\vec{x})$, and $\tilde{T}_a^g(\vec{x})$ denote this texture’s reflectance components. The probe texture, is denoted by $T^p(\vec{x})$ with its parameters $\tilde{T}_s^p(\vec{x})$, $\tilde{T}_d^p(\vec{x})$, and $\tilde{T}_a^p(\vec{x})$. Note that multiplication and division are computed pointwise (i.e. per color channel separately). Toderici et al. do not optimize the illuminant color $\vec{c}$. Instead, $\vec{c}$ is assumed to be perfectly white in the RGB color space, (1.0, 1.0, 1.0). Though a common assumption in FR algorithms, it is a poor approximation of the variable illumination in many real-world situations.

Such an assumption can be considered a significant limitation. However, if an illuminant color is known, it can be directly incorporated into Eqn. 6 as $\vec{c}$. Thus, the algorithm can better adjust to different light conditions in the gallery and the probe scenes. The incorporation is also flexible. Consider, for example, the case where the illumination color is only known for one image, either the gallery or the probe. This information can be utilized in the corresponding illumination components in Eqn. 5. For the remaining parameters, the white light assumption can still be used. All other components in Eqns. 2, 3, and 4 are optimized by minimizing Eqn. 6. Bidirectional relighting can include the supplementary illuminant-color information when available.
Knowledge about the illuminant color is obtained in two ways in our analysis. Ground truth about the illuminant color is available via a white Lambertian sample present in the scene (see Fig. 2). However, in non-laboratory conditions, ground truth information is typically not available. Thus, we also compute the color of the incident illuminant using the method of Riess et al.\textsuperscript{8} It estimates the chromaticity (i.e., normalized RGB) of a globally dominant illumination in the face region. We used this physics-based illuminant estimation algorithm, because it involves few parameters, is computationally efficient and is, in general, well suited for controlled captures. Note that this illuminant estimator is not skin-specific. One interesting approach for skin-specific illuminant estimation is the recent work by Bianco and Schettini.\textsuperscript{16} However, the underlying approach is based on the Gamut mapping algorithm, whose performance is highly dependent on its training data. For this work, we preferred an estimator that can be used out-of-the-box, without any parameter configuration.

4. DATABASE

To the best of our knowledge, there is no publicly available FR database explicitly designed for evaluating the impact of skin pigmentation in combination with varying illumination color. Therefore, we created a new 3D-2D database which: (i) contains probe images under various colors of incident illumination and (ii) includes ground truth for the illuminant color. The database is composed of data from five subjects. The probe images were taken under nine different illuminant colors. They were combinations of a ceiling light and a spot light. The spot light was used either as-is, or in combination with a red, green, or blue foil. The gallery images were only taken under the no-foil spot light. It was placed at two distinct locations in order to also incorporate some variation in the illumination direction. In total we have 45 probe and 10 gallery 3D-2D datasets. For reference, we placed a white Lambertian target in the background (see Fig. 2). It is standard practice to use the color of these target pixels as an approximation of the true color of the incident light. We refer to this reflectance target as the Lambertian surface (LS).

The images were acquired with a Flea2 camera. Andreopoulos and Tsotsos\textsuperscript{19} recently reported that camera settings can affect the outcome of computer vision algorithms. Thus, when we collected the images, we tried to limit the camera’s influence as much as possible. White balancing and other post-processing were disabled. All camera settings were kept fixed throughout the image acquisition process. Figure 2 shows images of a subject under the nine illumination setups; L1: fluorescent ceiling light, L2: frontal spotlight, L3: L1 and L2 combined, L4: L2 with red covering foil, L5: L1 and L4 combined, L6: L2 with green covering foil, L7: L1 and L6 combined, L8: L2 with blue covering foil, L9: L1 and L8 combined.

5. EVALUATION

We evaluated seven variants of incorporating the illuminant color in the FR system as shown in Table 1. WHITE is the baseline FR setup, without any illuminant-color correction. In all other cases, the illuminant color is included in either the probe, or the gallery, or both. The illuminant color is either obtained from the LS, or is computationally estimated.

5.1 Varying-Illuminant-Color Database

The ROC curves for verification performance are shown in Fig. 3(a). The x-axis denotes the false acceptance rate. The y-axis denotes the percentage of correctly recognized subjects. The baseline experiment WHITE is plotted in red. Its area under the curve (AUC) is 0.605. The corresponding AUCs are shown in Table 1. The
best performance, with an AUC of 0.705, is obtained under ideal conditions, i.e., using the LS values in both the gallery and the probe images (shown in Fig. 3(a) in dark blue). Thus, we improve the AUC by about 16%. In general, including information about the illuminant color on both the probe and the gallery results in improved FR performance for an increase in AUC of approx. 4%, 7% and 5% for P ICE G LS, P LS G ICE and P+G ICE, respectively. Estimating the illuminant color on only the gallery or only the probe produced mixed results.

Figure 3(b) depicts the identification performance via CMC curves. The baseline variant, White, is plotted in red. It is consistently outperformed by the incorporation of the LS information in both the probe and the gallery images (dark blue line). Using the automatically extracted illuminant color estimate on the probe, and the Lambertian surface on the gallery leads to competitive results (plotted in black). However, for ranks 1 and 3, these results are worse than the baseline. Thus, although accurate illuminant color information improves the performance of a FR system, estimation errors can counter this advantage. Note that applying illuminant color information only to the probe does not improve over the baseline, whether ideal or estimated (yellow and green lines, resp.).

### 5.2 Standard Database

Some of the illuminant colors used in our database can be considered extreme. Thus, we also evaluated the impact of addressing the illuminant color in more typical FR setups for a larger testing sample size. We used a second dataset of eleven subjects, with limited (frontal) pose variation. For these eleven subjects we had six

![Figure 3](image_url)  
(a) Verification performance  
(b) Identification performance

Figure 3. FR performance in our 5-subject varying-illuminant-color database, explicitly designed for testing the impact of illuminant color (see Tab. 1 for the AUCs).
indoor sessions under varying lighting direction and intensity and more typical illuminant color variants. We also had 23 outdoor captures, in eight of which an additional light source was used. As gallery data we used one of the indoor sessions (taken under distant light source) for nine out of the eleven subjects. The remaining sessions were used as probe data, resulting in an evaluation comprised of $9 \times 252 = 2268$ comparisons. Note that ground truth illuminant information was not available, hence we evaluated only the cases White and P+G ICE. Two example images are shown in Fig. 1 (right).

Figure 4 shows the results on this set. The verification for very low false acceptance rates is somewhat higher for P+G ICE. Incorporating the illuminant color results in an AUC of 0.932 compared to 0.919, which is the AUC under the white light assumption. Thus, we observe only a marginal improvement of 1.4%. Note, however, that White performs better in identification.

5.3 Discussion

Our evaluation indicates that we should follow the human vision paradigm and not ignore the effects of the illuminant color in FR systems. Accurate illuminant color estimates can improve recognition rates by 16%, especially under strong illumination colors. In general, the recognition results are better when color constancy is applied on both the gallery and the probe images. However, our analysis also showed that the impact of incorporating the illuminant color is very much dependent on the accuracy of the illuminant estimate. The current, general-purpose color constancy methods are not sufficiently accurate for having a systematically positive impact on FR. However, the recent work by Bianco and Schettini\textsuperscript{16} (i.e., the development of a skin-specific method for illuminant color estimation), is a step in the right direction.

In summary, we conclude that the incorporation of color constancy in face recognition has considerable potential, as our ground truth experiments show. Operating on illuminant color estimates (rather than the ground truth) still yields performance gains. A further improvement can be expected when face- or FR-specific color constancy algorithms are developed and employed.

6. CONCLUSIONS

Biometric authentication via face recognition is descriptive, non-intrusive and socially well accepted. Skin pigmentation is an important cue in face recognition by humans. This is further supported by the ability of people to detect the color of objects independent of variations in the illumination color. We quantified the extend to which the combination of these two aspects have a comparable impact in FR systems. For that purpose, we generated a ground-truth database of facial images under varying illumination colors. Our analysis showed that under approximately correct illuminant color, the recognition rates considerably improve by about 16%. When a
color constancy algorithm is used to estimate the illuminant color, the improvement in FR is highly dependent on the accuracy of the estimate. When using a general-purpose illuminant color estimator on probe or gallery data, we still achieve an improvement, but by a smaller margin. Specifically, we obtained an improvement of 4% when estimating the illuminant color on the probe data and using a Lambertian surface on the gallery data, and an improvement of 7% in the opposite scenario. On a larger indoor/outdoor dataset, we achieved an improvement of 1.4%. As a perspective for future work, we consider the development and incorporation of FR-specific color constancy essential for shrinking the gap between the empirical optimum and the actual performance gain.

ACKNOWLEDGMENTS

Elli Angelopoulou gratefully acknowledges funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German National Science Foundation (DFG) in the framework of the excellence initiative. Part of this work was performed when Manuel Meyer was an intern at the Computational Biomedicine Lab, University of Houston. This work was supported in part by the UH Hugh Roy and Lillie Cranz Cullen Endowment Fund.

REFERENCES


