

# The Impact of Specular Highlights on 3D-2D Face Recognition

Vincent Christlein<sup>1</sup>, Christian Riess<sup>1</sup>, Elli Angelopoulou<sup>1</sup>,  
Georgios Evangelopoulos<sup>2</sup> and Ioannis A. Kakadiaris<sup>2</sup>

<sup>1</sup> University of Erlangen-Nuremberg, Martensstr. 3, 91048 Erlangen, Germany;

<sup>2</sup> University of Houston, 4800 Calhoun Rd., Houston, Texas, USA

## ABSTRACT

One of the most popular form of biometrics is face recognition. Face recognition techniques typically assume that a face exhibits Lambertian reflectance. However, a face often exhibits prominent specularities, especially in outdoor environments. These specular highlights can compromise an identity authentication. In this work, we analyze the impact of such highlights on a 3D-2D face recognition system. First, we investigate three different specular removal methods as preprocessing steps for face recognition. Then, we explicitly model facial specularities within the face detection system with the Cook-Torrance reflectance model. In our experiments, specular removal increases the recognition rate on an outdoor face database by about 5% at a false alarm rate of  $10^{-3}$ . The integration of the Cook-Torrance model further improves these results, increasing the verification rate by 19% at a FAR of  $10^{-3}$ .

**Keywords:** Face Recognition, Specular Highlights, Specularity Removal, Reflectance Model

## 1. INTRODUCTION

The most prominent biometric identifiers are fingerprints, irises, hand geometries, faces and voice. Face images are easy to acquire, and can even be collected passively without the explicit cooperation of a person<sup>1</sup>. Several applications use face recognition techniques to secure doors, notebooks and mobile phones.

However, pose, facial expressions and illumination still constitute challenges for face recognition (FR) systems. Sinha *et al.*<sup>2</sup> reported that it is difficult even for humans to recognize a subject under varying incident illumination, shadows, different light colors and specularities. Thus, it is not surprising that computational methods also suffer from these difficulties. In particular, differing illumination conditions between a face gallery and the probe image makes reliable verification of a person challenging. Most methods model face reflection with the Lambertian reflectance model<sup>3-5</sup>, assuming that faces reflect light purely diffusely. This assumption holds well in indoor scenes, in particular when the subject is wearing makeup. However, for instance in outdoor scenes, strong specularities may occur, in conjunction with a considerable number of pixels with partially specular and diffuse reflection. As Li and Jain [6, p. 187] point out, specularities are one source of verification errors.

In this work, we investigate the impact of specular highlights on a 3D-2D FR system. Two approaches can be used to address the presence of specular reflection. First, we consider specular removal. Here, a diffuse-only image is created, which can be used in an existing diffuse-reflectance-based FR system. Alternatively, specular reflectance is directly incorporated in the FR system, using the well-known Cook-Torrance model<sup>7</sup>. To demonstrate the impact of specular highlights, we use the 3D-2D FR system proposed by Toderici *et al.*<sup>8</sup>. The evaluation is applied on a realistic face recognition database containing indoor and outdoor images. While specular removal slightly improves the verification rate, the largest performance gain is obtained from the direct incorporation of specular reflection within the FR system.

In Sec. 2, we briefly review the most closely related work. The 3D-2D FR system employed is presented in Sec. 3. In Sec. 4, we describe the investigated algorithms for specular removal, and the explicit specular modelling using the Cook-Torrance model. Quantitative and qualitative performance evaluations, and a discussion of the results are provided in Sec. 5. We conclude this work in Sec. 6.

---

Further author information: (Send correspondence to Ioannis A. Kakadiaris)  
Ioannis A. Kakadiaris: E-mail: ioannisk@uh.edu, Telephone: +1 (713) 743-1255

## 2. RELATED WORK

Ishiyama and Tsukada<sup>9</sup> proposed a specular removal method particularly designed for face recognition. The authors model the diffuse part of the facial reflectance with spherical harmonics. The residual obtained from subtracting this model from the input image yields the specular pixels. The estimated albedo of the gallery is obtained by dividing the residual with the diffuse approximation. For each gallery texture, first the pose of the probe image is adapted using landmark correspondences. Then, the illumination is fit by projecting the computed illumination subspace of the gallery according to the computed pose of the probe texture. The approach achieves better recognition results than without specular removal. Note that, this kind of specular removal is already integrated in the bidirectional relighting of the face recognition system we use<sup>8</sup> without explicit computation of the albedo. The method by Ishiyama and Tsukada also considers the specularities of the probe image by modeling it with a Phong specular component. In this work, we investigate a solution that relaxes the requirement for the Phong model by removing the specularities in advance. In this way, we allow a relighting process which only considers diffuse reflectance. In contrast to the work of Ishiyama and Tsukada, we also remove specularities from the probe facial textures. The probe textures contain specularities with a much higher probability than the gallery ones, which are usually captured in a controlled environment with low specularities.

Incorporating in FR a more sophisticated illumination model than the pure Lambertian or Phong model has been rarely addressed so far. Georghiadis<sup>10</sup> used the Torrance-Sparrow model<sup>11</sup> for photometric stereo and to recover the 3D shape of surfaces, and tested the model on face recognition. Evaluation on 10 subjects achieved about 15% better results than using a Lambertian model. In contrast to this work we employ the Cook-Torrance model (an extension of the Torrance-Sparrow model) and evaluate it on a more realistic outdoor face dataset.

## 3. A 3D-2D FACE RECOGNITION SYSTEM

In face recognition, a test image (the *probe*) is matched against a set of known images (the *gallery*). Different representations have been proposed for the probe and the gallery. For this work, we use a 3D-2D face recognition system. In such a system, the gallery data is captured in 3D, but the probe images for verification are captured as 2D images. This approach offers a compromise between 2D-2D systems which are cheaper, but also less accurate, and 3D-3D systems, where increased accuracy can be achieved at the expense of an high-priced, sophisticated capturing process: for 3D acquisition, more expensive, calibrated equipment is required on every authentication site. In comparison, 3D-2D face recognition systems require a subject to enroll once at a site with full 3D equipment, but authentication is very flexible, as 2D imaging systems are widely spread and straightforward to set up. In particular, we use the 3D-2D face recognition system by Toderici *et al.*<sup>8</sup>. This system has two stages: in the *enrollment phase*, the subject is entered into a 3D+2D gallery database. In the *verification*, a subject is verified by its 2D facial image.

Specifically, during enrollment a 3D facial mesh and a 2D image is acquired. Landmarks are automatically computed or manually annotated on the raw mesh data. The 3D facial mesh is then preprocessed to overcome sensor-specific problems (i. e., fill holes and remove spikes). An Annotated Face Model (AFM) is aligned and fitted to the 3D facial mesh. It allows a uniform representation of 3D facial meshes and provides a global 2D parametrization in the form of geometry maps. The texture is lifted (i. e., a geometry map is created with the texture values assigned from the projection of the fitted AFM on the 2D image plane) the 3D mesh of the fitted AFM is projected to the 2D image plane of the image). Additionally, a visibility map, which denotes non-visible areas of the 2D texture, is computed. The fitted AFM, the lifted 2D texture and the visibility map are stored as meta data for a specific subject in the gallery.

During the recognition phase, a 2D facial probe image is compared to the meta data in the gallery. The process is outlined in Fig. 1. First, reference landmarks on the 2D face have to be annotated or detected. The following steps are computed for each subject of the gallery in order to compute a one-to-one similarity score. The AFM of the gallery is registered to the 2D facial image using 3D-2D landmark correspondences. A pose estimation step compensates mismatches of the head pose between the probe and a gallery image. The probe 2D texture is generated by the projection of the fitted AFM. Similarly to the enrollment phase, a visibility map is computed. Differences in the illumination conditions of the probe are normalized by an optimization-based

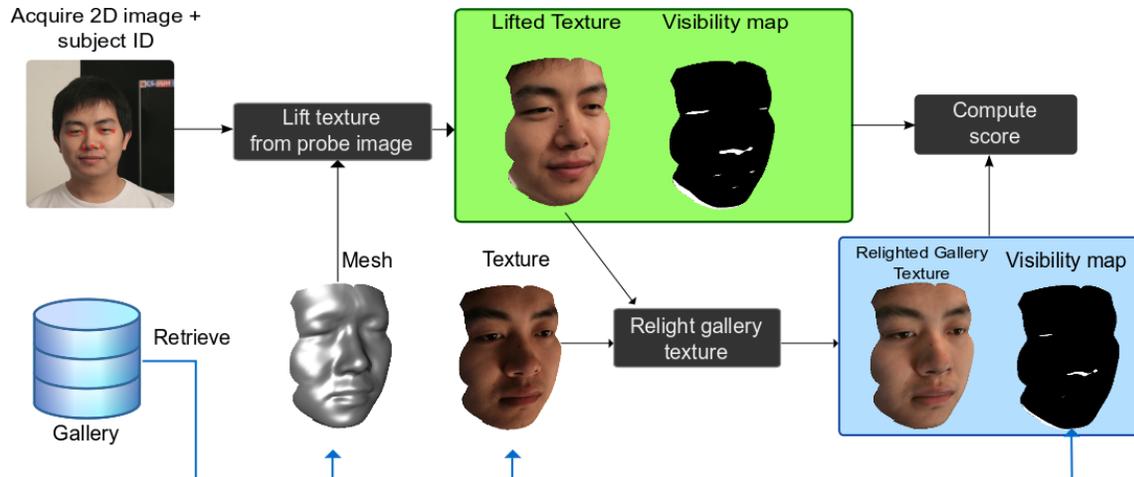


Figure 1: The face verification phase: From the fitted and pose aligned mesh, the texture is lifted and compared with the relit gallery texture before a score is computed between the textures (image courtesy of Toderici *et al.*<sup>8</sup>).

relighting of the gallery texture. Finally, a similarity score is computed between the lifted textures. From this score, the best match in the gallery is determined (identification), and a threshold-based ACCEPT/REJECT decision is made (verification).

#### 4. SPECULAR HIGHLIGHTS IN 3D-2D FR

Given an inhomogeneous opaque surface, *diffuse* or body reflection occurs from light rays which penetrate the body and then, leave the body again. *Specular* or interface reflection is caused by light rays which are directly reflected from the surface body. Many algorithms in computer vision (e. g., for stereo matching, segmentation, recognition and tracking) work under the assumption of perfect diffuse (Lambertian) reflection. However, in the real world, most materials exhibit a combination of diffuse and specular reflection and this can cause errors in the aforementioned algorithms (e. g., stereo mismatching). Thus, in this work, we aim to quantify the performance gain in face recognition when specular highlights are considered.

##### 4.1 Methods for Specularity Removal

The first investigated approach is to remove specularities prior to face detection. A variety of general purpose methods for specularity removal exist<sup>12</sup>. For application on FR, we require a specularity removal method that operates on a single image. Furthermore, the method must be fully automatic.

Specifically, we employ three different specularity removal methods<sup>13-15</sup>, based on the dichromatic reflection model<sup>16</sup>,

$$I(\lambda, \mathbf{x}) = w_d(\mathbf{x})S_d(\lambda, \mathbf{x})E(\lambda) + w_s(\mathbf{x})E(\lambda) , \quad (1)$$

where the reflected light  $I$  at a certain wavelength  $\lambda$  and surface location  $\mathbf{x} = \{x, y\}$  is a weighted linear combination of the diffuse reflectance function  $S_d(\lambda, \mathbf{x})$  and the spectral energy distribution function of the illumination  $E(\lambda)$ . The weights  $w_d(\mathbf{x})$  and  $w_s(\mathbf{x})$  denote geometric weighting factors for the corresponding diffuse and specular reflection. Note that this reflection model assumes uniform illumination color throughout the scene. Furthermore, the specularity removal methods utilize the *chromaticity space*. Let  $c = \{r, g, b\}$  a color channel, and  $I_c$  the channel  $c$  of an input image  $I$ . Then, the *image chromaticity*  $\sigma_c$  for a color channel  $c$  can be directly computed by

$$\sigma_c(\mathbf{x}) = \frac{I_c(\mathbf{x})}{I_r(\mathbf{x}) + I_g(\mathbf{x}) + I_b(\mathbf{x})} . \quad (2)$$

Similarly, the *diffuse chromaticity*  $\Lambda_c$  and *illumination chromaticity*  $\Gamma_c$  can be defined by setting  $w_s = 0$ , and  $w_d = 0$ , respectively. Hence, Eqn. 1 can be rewritten in terms of diffuse chromaticity and illumination chromaticity,

$$I_c(\mathbf{x}) = m_d(\mathbf{x})\Lambda_c(\mathbf{x}) + m_s(\mathbf{x})\Gamma_c, \quad (3)$$

where  $m_d$  and  $m_s$  are the weighting factors for the diffuse and specular reflection component, respectively. The goal is to decompose the image intensity  $I_c$  with a known illumination chromaticity  $\Gamma_c$  into the reflection components,  $m_d(\mathbf{x})\Lambda_c(\mathbf{x})$  and  $m_s(\mathbf{x})\Gamma_c$  for  $c = \{r, g, b\}$ <sup>13</sup>.

For a correct separation the methods typically assume that the surface color is chromatic, i.e., the surface color is not a black, gray or white color ( $r \neq g$  or  $r \neq b$  or  $g \neq b$ ). Further assumptions are that each color region always includes at least one diffuse pixel and that the illumination chromaticity is known. Additionally, the output of the camera must be linear to the flux of the incident light, i.e. the camera gamma equals 1 and camera gain and camera noise are ignored.

In order to separate the reflection components the specular component must be pure white (i.e.,  $\Gamma_r = \Gamma_g = \Gamma_b$ ). Since in the real world a pure white illumination is very unlikely, the image is normalized with the illumination chromaticity  $\Gamma_c$ ,  $c = \{r, g, b\}$ <sup>13</sup>. The illumination chromaticity can be measured and compensated with a white reference reflectance target in the scene. Since in most cases of FR under uncontrolled conditions, such targets may not be available, we rely on a color constancy algorithm, the physics-based Illuminant Estimation by Voting algorithm by Riess *et al.*<sup>17</sup>, to estimate the illumination chromaticity. We correct this estimate by a fixed Fresnel term<sup>18</sup>, as skin reflectance slightly biases the output of the method.

#### 4.1.1 Method of Tan and Ikeuchi

The method of Tan and Ikeuchi (abbreviated to TAN) first computes a specular-free (or pseudo diffuse) image which has the same geometric profile as the diffuse image<sup>13</sup>. This is then used in a *specular-to-diffuse* mechanism which iteratively decreases the specular component of each pixel.

The *specular-to-diffuse* mechanism relies on the maximum image chromaticity, which is defined as:

$$\sigma_{\max}(\mathbf{x}) = \max(\sigma_r(\mathbf{x}), \sigma_g(\mathbf{x}), \sigma_b(\mathbf{x})) \quad (4)$$

where  $\sigma_{\max}(\mathbf{x})$  is a scalar value. The maximum diffuse chromaticity  $\Lambda_{\max}(\mathbf{x})$  and the maximum illumination chromaticity  $\Gamma_{\max}$  are similarly defined. Tan and Ikeuchi showed that maximum chromaticities of diffuse points in the image are always larger than those of the specular points. Furthermore, the projection of points from a single material form a curve in maximum chromaticity space<sup>13</sup> that can be described as:

$$I_{\max}(\mathbf{x}) = \frac{m_d(\mathbf{x})(\Lambda_{\max}(\mathbf{x}) - 1/3)\sigma_{\max}(\mathbf{x})}{\sigma_{\max} - 1/3}. \quad (5)$$

By setting  $\Lambda_{\max}(\mathbf{x})$  to a fixed scalar value ( $1/3 < \Lambda_{\max}(\mathbf{x}) \leq 1$ ), a specular-free image can be created by setting the value of each pixel regardless of its color. In our implementation, we chose  $\Lambda_{\max} = 0.6$ .

This pseudo diffuse image contains no specularities, but results in an image with a different color profile. However, as the geometric profile of the specular-free image is identical to the one of the input image, it can be used to separate the image diffuse and specular components. The logarithmic differentiations of two neighboring pixels are computed from both the normalized input image and the specular-free image. If, per pixel, both logarithmic differentiations are equal, the pixel is diffuse. This can be formulated as:

$$\Delta(\mathbf{x}) = \frac{d}{d\mathbf{x}} \log \left( \sum_{c \in \{r, g, b\}} I_c(\mathbf{x}) \right) - \frac{d}{d\mathbf{x}} \log \left( \sum_{c \in \{r, g, b\}} I_{\text{sf},c}(\mathbf{x}) \right), \quad (6)$$

where  $I_{\text{sf},c}$  denotes the specular free image at color channel  $c$ . If  $\Delta(\mathbf{x})$  is zero, the pixel is diffuse. Otherwise, it is either (partially) specular or a color discontinuity, caused by noise or a boundary between two surface colors.

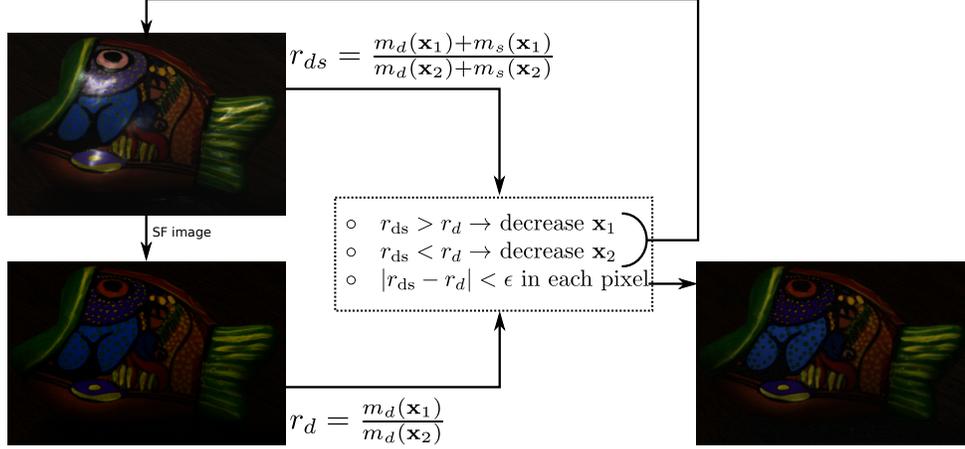


Figure 2: YOON: First, a specular-free image<sup>13</sup> (bottom left) of the input image (top left) is computed. From these images two ratios  $r_d$  and  $r_{ds}$  are computed for each two neighboring pixels upon which the input image is iteratively adjusted until convergence. Result: bottom right image.

Tan and Ikeuchi distinguish these cases via thresholding on the red and green chromaticity differences of two neighboring pixels (i. e., if they are both above 0.1 the current pixel is declared as boundary pixel). Furthermore, if the difference of the maximum chromaticity of two neighboring pixels is smaller than 0.01 the current pixel is marked as noise. Specular pixels are transferred to diffuse pixels with the *specular-to-diffuse* mechanism. Here, the maximum chromaticities of neighboring pixels are compared. The pixel with lower maximum chromaticity adopts the maximum chromaticity of its neighbor. The algorithm converges when no pixel is labeled as specular.

#### 4.1.2 Method of Yoon, Choi and Kweon

The method by Yoon *et al.*<sup>15</sup> (abbreviated to YOON) shares the basic idea of TAN, but is considerably simpler and faster. The method is outlined in Fig. 2. First, a specular-free image (Fig. 2, bottom left) is computed from the input image (Fig. 2, top left). An iterative framework based upon comparing local ratios of adjacent pixels in the input and the specular-free image is then used for generating a correct-color, specular-free version of the input image. This update is applied until convergence (Fig. 2, bottom right).

The specular-free image is obtained by subtraction of the minimum color channel  $I_{\min}(\mathbf{x})$ :

$$I_{\text{sf},c}(\mathbf{x}) = I_c(\mathbf{x}) - I_{\min}(\mathbf{x}) = m_d(\mathbf{x})(\Lambda_c(\mathbf{x}) - \Lambda_{\min}(\mathbf{x})) \text{ , for } c = \{r, g, b\} \text{ ,} \quad (7)$$

where  $\Lambda_{\min} = \min\{\Lambda_r, \Lambda_g, \Lambda_b\}$ .  $I_{\text{sf},c}(\mathbf{x})$  denotes the specular-free image, similarly to TAN, in color channel  $c$  at pixel  $\mathbf{x}$ ,  $im$  denotes the normalized input image. This resulting image is independent of the specular component and has the same geometric profile as the input. Based on  $I_{\text{sf},c}(\mathbf{x})$ , the specularities in the input image can be iteratively removed by comparing the local ratios  $r_d$  of the specular-free image with  $r_{ds}$  of the input image (see Fig. 2). The two ratios are equalized by decreasing one of the two neighboring pixels in the input image in every step. If  $r_{ds}$  is larger than  $r_d$ , the pixel  $\mathbf{x}_1$  is decreased by:

$$m = n \left( \sum_{c \in \{r, g, b\}} I_{d,c}^n(\mathbf{x}_1) - r_d \sum_{c \in \{r, g, b\}} I_{d,c}^n(\mathbf{x}_2) \right) \text{ ,} \quad (8)$$

where  $I_d^n$  denotes the normalized input image at iteration  $n$ . Otherwise, the pixel  $\mathbf{x}_2$  is decreased by:

$$m = n \left( \sum_{c \in \{r, g, b\}} I_{d,c}^n(\mathbf{x}_2) - \frac{\sum_{c \in \{r, g, b\}} I_{d,c}^n(\mathbf{x}_1)}{r_d} \right) \text{ .} \quad (9)$$

In our implementation, we use  $n = 1/3$  and apply the update step by testing two neighboring pixels first in the  $x$ - and then in the  $y$ -direction.

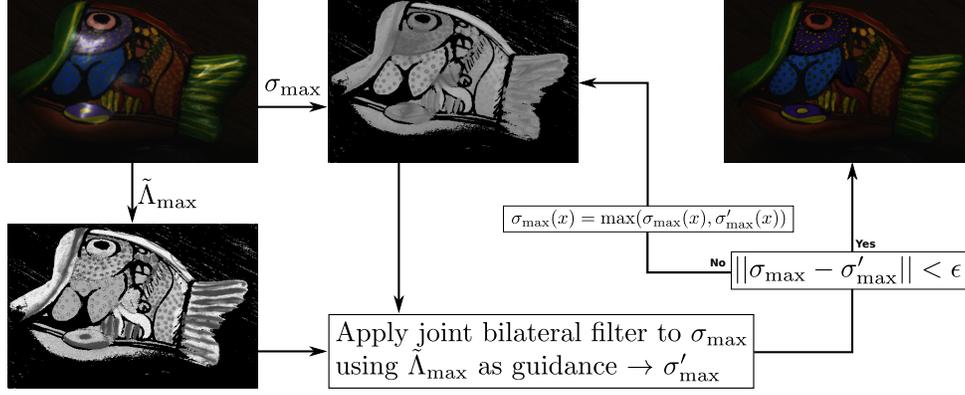


Figure 3: Method YANG: From the input image<sup>13</sup> (top left) compute  $\sigma_{\max}$  (top middle) and an estimate of  $\Lambda_{\max}$  (bottom left), then apply bilateral filtering and  $\sigma_{\max}$  update until convergence. Result: top right image.

### 4.1.3 Method of Yang, Wang and Ahuja

The method of Yang *et al.*<sup>14</sup> (denoted as YANG) is also based on TAN. In contrast to TAN, no specular-free image is computed. Instead, the problem of specular removal is reduced to finding the maximum diffuse chromaticity  $\Lambda_{\max}(\mathbf{x})$  in each pixel. The authors exploit that in local patches, the maximum image chromaticity  $\sigma_{\max}$  and the maximum diffuse chromaticities often do not change significantly. Using a low-pass filter, the maximum diffuse chromaticity values can be propagated to the specular pixels. For this task, Yang *et al.* propose to use iterative bilateral filtering towards the maximum chromaticity  $\sigma_{\max}(\mathbf{x})$ . The bilateral filter is guided by the maximum diffuse chromaticity to avoid blending different surface materials. In contrast to TAN, the method does not suffer from artifacts of non-converged pixels that were induced by discontinuities in surface colors. Additionally, fast bilateral filtering makes this method about 200 times faster than TAN.

The algorithm is outlined in Fig. 3. First,  $\sigma_{\max}$  (Fig. 3, top middle) and an estimate of  $\Lambda_{\max}$  (Fig. 3, bottom left) are computed. Then,  $\sigma_{\max}$  is iteratively smoothed using  $\tilde{\Lambda}_{\max}$  as guidance, until convergence (Fig. 3, top right).  $\Lambda_{\max}$  is unknown and needs to be approximated through  $\tilde{\Lambda}_c(\mathbf{x}) \approx \Lambda_c(\mathbf{x})$ , which depends only on the material, as

$$\tilde{\Lambda}_c(\mathbf{x}) = \frac{\sigma_c(\mathbf{x}) - \sigma_{\min}(\mathbf{x})}{1 - 3\sigma_{\min}} \quad \text{for } c = \{r, g, b\}, \quad (10)$$

where  $\sigma_{\min} = \min\{\sigma_r, \sigma_g, \sigma_b\}$  is the minimum image chromaticity,  $\tilde{\Lambda}_{\max}$  denotes the maximum chromaticity channel of  $\tilde{\Lambda}_c(x)$  for  $c = \{r, g, b\}$ . The smoothing process is repeated until the update for  $\sigma_{\max}$  is smaller than  $\epsilon = 0.003$  (Fig. 3, top right image). Note that, typically, the method converges within three iterations.

## 4.2 Incorporating a Specular Reflection Model

We also investigate the direct incorporation of specularities for illumination normalization in FR. Most face recognition algorithms assume Lambertian reflectance<sup>19,20</sup>. To model also specular reflectance, we rely on the Cook-Torrance Bidirectional Reflectance Distribution Function (BRDF). Its specular BRDF  $f_s$  is defined as

$$f_s(\mathbf{l}, \mathbf{v}) = \frac{DGF}{4\pi \overline{\cos \theta_i} \cos \theta_o}, \quad (11)$$

where  $\overline{\cos \theta} = \max(0, \cos \theta)$  denotes a clamped cosine,  $\mathbf{v}$  the viewing direction and  $\mathbf{l}$  the light direction.  $\theta_i$  and  $\theta_o$  denote the angle between  $\mathbf{l}$  and the surface normal, and the angle between  $\mathbf{v}$  and the surface normal, respectively. The distribution of the microfacets is denoted by  $D$  while  $G$  is a geometrical attenuation factor that takes into

account the partial interception of incident and reflected light. The Fresnel term  $F$  defines the amount of and the spectral composition of the specularly reflected light.

The Cook-Torrance model describes the proportion of outgoing radiance at the viewing direction  $\mathbf{v}$  to the incoming radiance of the light source from direction  $\mathbf{l}$ . For  $D$ , we chose to use the Beckmann distribution, and for  $G$  we take the geometry factor proposed by Torrance and Sparrow<sup>11</sup>. The Fresnel term  $F$  is computed using Schlick’s approximation<sup>21</sup>, for which  $F(0^\circ)$  can be computed from the refractive indices<sup>18</sup>. Since the surface material is mainly skin, we use the values of the epidermis layer<sup>18</sup>, i. e., 1.433 for the red, 1.448 for the green and 1.449 for the blue color channel.

The rendering equation for relighting, in which the BRDF is integrated, defines for multiple point light sources the outgoing radiance  $L_o$  as

$$L_o = \mathbf{c}_a \otimes L_A + \sum_{k=1}^n f(\mathbf{l}_k, \mathbf{v}) \otimes d_k \mathbf{e}_{L_k} \overline{\cos} \theta_{i_k} , \quad (12)$$

where  $\otimes$  denotes element-wise vector multiplication,  $\mathbf{c}_a \otimes L_A$  describes ambient illumination with the ambient light  $L_A$  and the ambient surface color  $\mathbf{c}_a$ . Furthermore,  $k$  denotes the number of light sources,  $f(\mathbf{l}_k, \mathbf{v})$  the actual BRDF,  $\mathbf{e}_{L_k}$  the illuminant color of light  $L_k$ ,  $d_k$  is an attenuation factor, and  $\overline{\cos} \theta$  models foreshortening. We assume  $L_A$  to differ from the illuminant color by only a constant factor  $k_a$ . For multiple light sources, we set  $L_A = k_a \mathbf{e}_a$  as the sum of all illumination colors (i. e.,  $\mathbf{e}_a = \sum_{k=1}^n \mathbf{e}_k$ ). Furthermore, we set  $\mathbf{c}_a$  equal to the diffuse color  $\mathbf{c}_d$ .

Incorporating the diffuse BRDF ( $1/\pi$ ) and the specular BRDF from Eqn. 11 in the rendering Eqn. 12 yields

$$L_o = \mathbf{c}_d \otimes k_a \mathbf{e}_a + \sum_{k=1}^n \left[ \left( \frac{\mathbf{c}_d}{\pi} + k_s \frac{D_B(\theta_h, m_k) G_{TS}(\mathbf{l}, \mathbf{v}) F_S(\alpha_h, F(0^\circ))}{4\pi \overline{\cos} \theta_i \overline{\cos} \theta_o} \right) \otimes d_k \mathbf{e}_{L_k} \overline{\cos} \theta_{i_k} \right] , \quad (13)$$

where  $D_B$ ,  $G_{TS}$  and  $F_S$  denote our implementation choices for  $D$ ,  $G$  and  $F$ , and  $k_s$  is a normalization factor for the Beckmann distribution<sup>22</sup>. The angle  $\theta_h$  is the angle between the surface normal and  $\mathbf{h}$ , where  $\mathbf{h}$  is the halfway vector between  $\mathbf{l}$  and  $\mathbf{v}$ .  $m$  denotes the root-mean-square slope of the microfacets and  $\alpha_h$  the angle between  $\mathbf{h}$  and the viewing direction  $\mathbf{v}$ .

Most of these factors can be precomputed. The unknown parameters to be estimated in Eqn. 13 are:  $k_a$  and  $k_s$ , the per light parameters  $m_k$  for the Beckmann distribution, and the per-light attenuation factors  $d_k$ . The unknown albedo  $\mathbf{c}_d$  depends only on the diffuse reflectance component. Within the framework of Toderici *et al.*<sup>8</sup>, it is implicitly computed within the relighting step. Three lights were used as proposed by Toderici *et al.*

Note that modelling skin with a BRDF is only approximately correct. Skin consists of multiple layers. Incident photons entering the skin at one point may scatter below the surface of the skin and exit at a different point [6, p.171]. This behavior is known as subsurface scattering, and should ideally be modeled with a Bidirectional Surface Scattering Reflectance Distribution Function (BSSRDF). However, we restricted this study to employing a BRDF function.

## 5. EVALUATION

The specularly removal methods were evaluated within two variants of the face recognition framework. In the first scenario, specularly removal was applied *before relighting* (i. e., before the light condition of the probe face texture is applied to the gallery face texture). In the second scenario, *relighting is omitted* (i. e., after specularly removal, gallery and probe images are directly passed to score computation). The advantage of the latter approach is the reduced computational load, since relighting is the computationally most demanding step of the recognition pipeline.



Figure 4: Results of specular removal on a facial texture containing a strong specular component. TAN shows clear artifacts. YANG and YOON perform better. The albedo is best preserved by YANG.

### 5.1 Database

The evaluation was performed on the University of Houston Database 30 (UHDB30)<sup>23</sup>. The face images were captured with a Flea2 camera. The dataset consists of nine subjects, 22 outdoor and seven indoor sessions. We chose to use one indoor session as gallery and the remaining images as probe, thus resulting in  $9 \times 252$  (Gallery  $\times$  Probes) comparisons. This makes the comparison more challenging, because the light conditions differ considerably between indoor and outdoor images. Additionally, outdoor images contain typically stronger specular highlights than indoor images.

### 5.2 Integration of Specularity Removal Methods

Several parameters have to be set for the specular removal methods. YANG typically converges after 3 iterations. We set the number of iterations to 10, to make sure that convergence is always reached. For YOON, we set the maximum difference of the ratios  $r_d$  and  $r_{ds}$ , to 0.1. Additionally, we set the maximum number of iterations to 500. To normalize the illumination color, we used Illuminant Estimation by Voting (IEbV)<sup>17</sup>, with 100 samples per image. The remaining thresholds were set according to the values stated in the respective papers. Each specular removal method is applied on both, gallery and probe face textures.

### 5.3 Qualitative Evaluation

**Specularity Removal and Purely Diffuse Relighting:** Figure 4 depicts a facial texture image of UHDB30 which was captured outdoors and thus contains strong specularities. Noticeable in all results are the dark pixels in the middle of the nose which are saturation artifacts. As the specular part is directly subtracted from the input image, the output diffuse component contains black-pixel artifacts. This comes from saturation of specularities, which can either be averaged out<sup>24</sup>, masked out<sup>25</sup> or avoided by careful control of the capturing setup. Fortunately, such overexposed pixels occur relatively rarely in UHDB30. Another issue is the sparkles in the facial hair regions. Here, TAN produces large gray areas.

To alleviate the artifacts at the beard and forehead, we blurred the original image in Fig. 4a with a Gaussian kernel of  $3 \times 3$  pixels size with a standard deviation of 0.8 (Fig. 5a). For TAN, the artifacts increased. However, YOON preserved the facial color better. The only noticeable artifact in Fig. 5c is a slight albedo edge in the right half of the face. In contrast, YANG did not benefit from the Gaussian blur preprocessing. Artifacts at the forehead disappeared, but new artifacts around the eyebrows were introduced. Additionally, blurring did not achieve a complete removal of the saturated specularities at the nose bridge.

As a general remark, we observed that TAN is weak at preserving edges in the image. For specular removal, YOON performs best. In comparison, YANG fails to remove some specularities at the forehead, nose and lips. However, the diffuse component generated by YOON is too dark, while YANG much better preserves the face albedo.



Figure 5: Results of specular removal on a facial texture containing a strong specular component which has been blurred prior to removal. TAN has strong artifacts while YOON has the most significant improvement, in comparison to the non-blurred version.

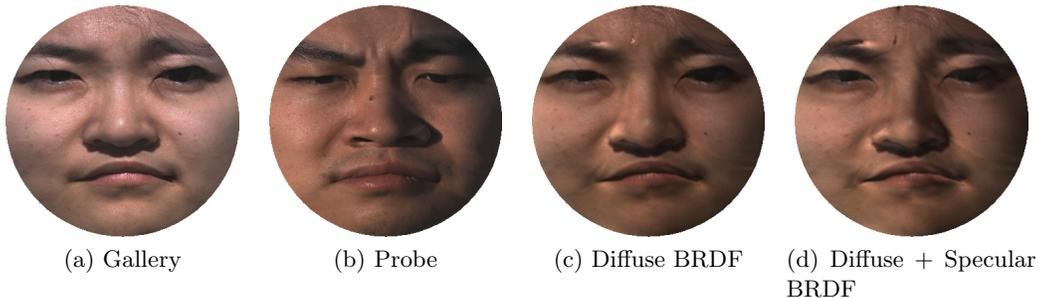


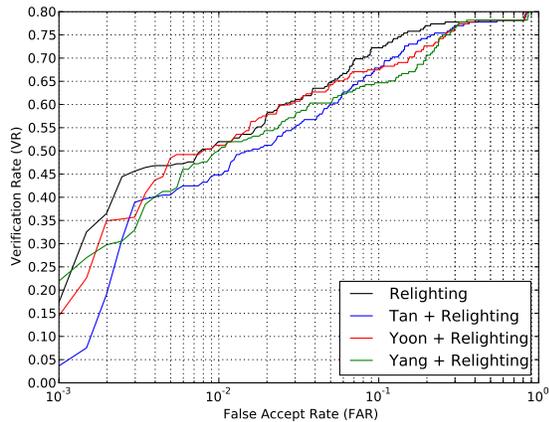
Figure 6: Qualitative results of applying different BRDFs. (a) Gallery (source) image to be relit with the illumination condition of probe/target (b), (c) using a purely diffuse BRDF, (d) using the Cook-Torrance BRDF. The diffuse and specular BRDF covers shadowed areas better (e.g., the right side of the nose bone) and the specular highlight at the nose tip are better transferred from the probe image. All images are in the texture geometry space of the 3D face model.

**Incorporation of Specularities using the Cook-Torrance Reflectance Model** An example result can be seen in Fig. 6, where the gallery texture (6a) is relit to the probe’s lighting condition (6b). Compared to the result using a purely Lambertian model (6c), the use of a combined diffuse/specular BRDF looks much more plausible (e.g., the shadow at the right side of the nose bone is much darker). However, the difference in the specular part is difficult to see in the image.

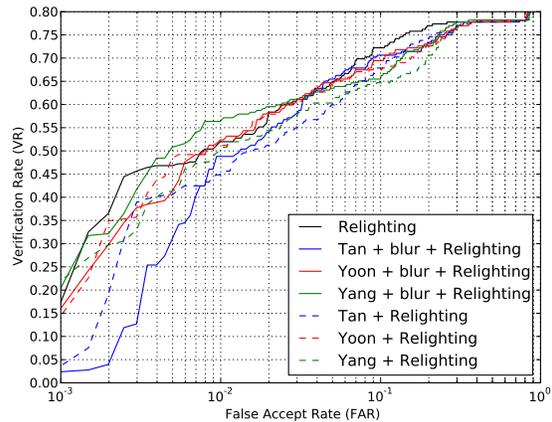
#### 5.4 Quantitative Evaluation

We present quantitative results for face recognition using receiver operating characteristic (ROC) curves. A ROC curve describes the 1:1 verification rate (i.e., it answers the question whether an individual is the claimed person or not). The verification rate (or true acceptance rate, respectively) is plotted as a function of the false acceptance rate (FAR). Each point on the curve represents a decision threshold to obtain the respective FAR. Note that the x-axis denoting the FAR is in logarithmic scale. The closer the ROC curve is to the upper left corner, the higher the overall accuracy of the test. As the ROC curve is sufficient to show the performance<sup>26</sup>, we omit the also often reported cumulative match characteristic (CMC) curves. Although the full ROC curves are presented, the performance at very low false acceptance rates (e.g.,  $10^{-3}$ ) is often an important metric for assessing the system performance. Also note that the plotted scores are normalized using Z-score normalization.

**Specularity Removal and Purely Diffuse Relighting:** Figure 7a shows the verification rate with relighting. For a low FAR of  $10^{-3}$ , YANG is the best performing method. For higher false acceptance rates, it is favorable to omit specularity removal. In a second experiment (Fig. 7b), we added a Gaussian blur filter as a preprocessing

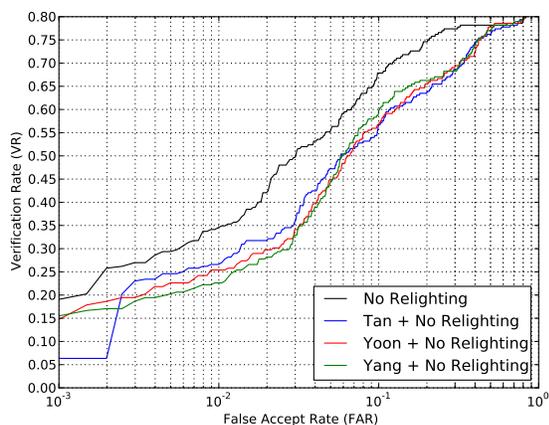


(a) Unmodified input

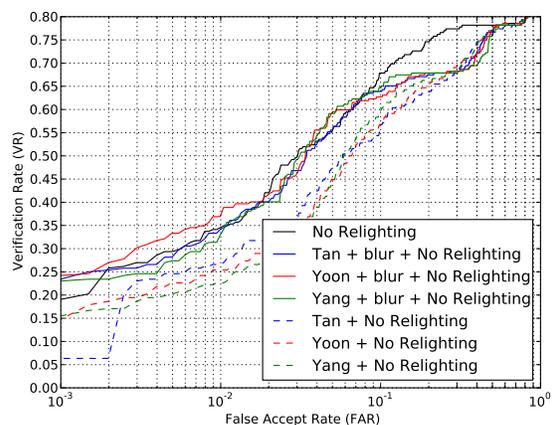


(b) Gaussian blurred input

Figure 7: Specularity removal on UHDB30: (a) with unmodified input images (b) with Gaussian blurred input images. For comparison, the non-blurred results are also shown as dashed lines in (b).



(a) Unmodified input



(b) Gaussian blurred input

Figure 8: Specularity removal on UHDB30 without relighting: (a) with unmodified input images (b) with Gaussian blurred input images. For comparison, the non-blurred results are also shown as dashed lines in (b).

step to the specular removal methods. The non-blurred results of Fig. 7a are shown as dashed lines, while the blurred versions are denoted with solid lines. Applying the blurring step negatively affects the specular removal method TAN compared to its non-blurred version, while YANG and YOON achieve slightly better results.

Figure 8a shows the results from the experiment without relighting. If the relighting step is omitted, all specular removal approaches perform worse than the baseline result. Interestingly, Fig. 8b indicates that the performance of each method improves from Gaussian smoothing. With the Gaussian blur filter, all methods improve by about 5% over the raw (no relighting) curve at  $10^{-3}$  FAR.

Consequently, specular removal can improve the face recognition rate on outdoor, natural illumination conditions. An important preprocessing step is a slight blurring of the images to attenuate saturation artifacts and to reduce the noise in the image. This preprocessing better supports the specular-to-diffuse process for neighborhood-based specular removal methods.

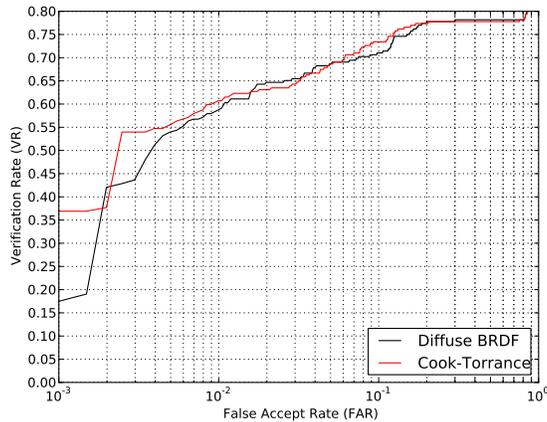


Figure 9: Results of incorporating the Cook-Torrance reflection model into 3D-2D FR.

**Incorporation of Specularities using the Cook-Torrance Reflectance Model:** Figure 9 indicates that the Cook-Torrance model has a significant impact on the face recognition rate. Note that the score of the diffuse BRDF (black curve) is slightly different in comparison to Fig. 7 since we had to choose different bounding parameters in the optimization process for the Cook-Torrance model, and thus, for a more convenient comparison, we chose to use the same optimization parameters for the relighting process as proposed by Toderici *et al.* The verification rate significantly increased, particularly for lower FARs. At a FAR of  $10^{-3}$ , we obtain an improvement of about 19%. Thus, we consider this approach superior to specularity removal and relighting of purely diffuse pixels.

Nevertheless, simplifications of the Cook-Torrance BRDF should be tested as well. The Beckman distribution for the microfacets' distribution is computationally very expensive. Therefore, in computer graphics, the distribution function is often modeled similarly to the Blinn-Phong specular term. Kemelmen and Szirmay-Kalos<sup>22</sup> in turn give an approximation of the geometry factor  $G$ , which makes its computation numerically more stable and more efficient. These simplifications, and the proposed Cook-Torrance BRDF should also be tested against the Phong model which Toderici *et al.*<sup>8</sup> proposed to use during relighting.

## 5.5 Discussion

We showed that using the diffuse component of a specularity removal method can improve the verification rate in FR under uncontrolled illumination. One important preprocessing step is to apply Gaussian smoothing to the image, to attenuate artifacts from saturated pixels and to remove image noise. Even better results can be obtained if the specular component is explicitly incorporated in the actual illuminant normalization algorithm. To achieve that, we incorporated the physics-based Cook-Torrance model which simulates macroscopic reflectance by microscopic facets.

For a more comprehensive understanding of specularities in face recognition, both approaches should be further investigated. In future work, we plan to evaluate both techniques on larger databases and improve robustness, e.g., by applying the method of Lehmann and Palm to ignore unsaturated and too bright pixels in advance<sup>25</sup>. Another extension would be to further evaluate different existing specularity removal methods regarding their performance and robustness. Mallick *et al.* proposed a very promising method which seems to handle finely textured images, like faces, very well<sup>27</sup>.

Furthermore, the use of the Cook-Torrance BRDF has to be tested on a larger database to fully assess its impact. This would require methods with an overall lower computational complexity. Currently, as a workaround, the relighting process operates on images that are downsampled by a factor of 50%. On full-resolution images, the Cook-Torrance BRDF could achieve even higher verification rates.

## 6. CONCLUSION

We investigated the impact of specular reflection on FR on a challenging benchmark of outdoor images, where specular highlights occur more often than in controlled indoor environments. Our results indicate that specularities can have a considerable impact on face recognition performance. First, we applied methods for specular removal to create purely diffuse images. Face recognition on these images (instead of the original images) increased by about 5% at a FAR of  $10^{-3}$ . Alternatively, we directly incorporated the Cook-Torrance model for specular reflectance in the relighting step of a 3D-2D face recognition pipeline. This resulted in a significant improvement by 19% over the baseline at a FAR of  $10^{-3}$ . However, one important issue is computational complexity. For future work, computationally more efficient ways of modeling the specular component have to be researched.

## 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 Vincent Christlein 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

- [1] Bowyer, K. W., “Face Recognition Technology: Security versus Privacy,” *IEEE Technology and Society Magazine* **23**, 9–20 (Spring 2004). 1
- [2] Sinha, P., Balas, B., Ostrovsky, Y., and Russell, R., “Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About,” *Proceedings of the IEEE* **94**, 1948–1962 (Nov. 2006). 1
- [3] Zhou, Z., Ganesh, A., Wright, J., Tsai, S.-F., and Ma, Y., “Nearest-Subspace Patch Matching for Face Recognition under Varying Pose and Illumination,” in [*Proc. IEEE International Conference on Automatic Face & Gesture Recognition*], 1–8, IEEE (Sept. 2008). 1
- [4] Wang, Y., Zhang, L., Liu, Z., Hua, G., Wen, Z., Zhang, Z., and Samaras, D., “Face Relighting from a Single Image under Arbitrary Unknown Lighting Conditions,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**, 1968–1984 (Nov. 2009). 1
- [5] Biswas, S., Aggarwal, G., and Chellappa, R., “Robust Estimation of Albedo for Illumination-Invariant Matching and Shape Recovery,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**, 884–899 (May 2009). 1
- [6] Li, S. Z. and Jain, A. K., [*Handbook of Face Recognition*], Springer Publishing Company, Incorporated, 2nd ed. (Aug. 2011). 1, 7
- [7] Cook, R. L. and Torrance, K. E., “A Reflectance Model for Computer Graphics,” *ACM Transactions on Graphics* **1**, 7–24 (Jan. 1982). 1
- [8] Toderici, G., Passalis, G., Zafeiriou, S., Tzimiropoulos, G., Petrou, M., Theoharis, T., and Kakadiaris, I. A., “Bidirectional relighting for 3D-aided 2D face recognition,” in [*Proc. IEEE Computer Vision and Pattern Recognition*], 2721–2728, IEEE (June 2010). 1, 2, 3, 7, 11
- [9] Ishiyama, R. and Tsukada, M., “Specularity Removal for Enhancing Face Recognition,” in [*Proc. IEEE International Conference on Image Processing (ICIP)*], 57–60, IEEE (Nov. 2009). 2
- [10] Georghiades, A. S., “Incorporating the Torrance and Sparrow Model of Reflectance in Uncalibrated Photometric Stereo,” in [*Proc. IEEE International Conference on Computer Vision*], 816–823 vol.2, IEEE (Oct. 2003). 2
- [11] Torrance, K. E. and Sparrow, E. M., “Theory for Off-Specular Reflection From Roughened Surfaces,” *Journal of Optical Society of America* **57**, 1105–1112 (Sept. 1967). 2, 7
- [12] Artusi, A., Banterle, F., and Chetverikov, D., “A Survey of Specularity Removal Methods,” *Computer Graphics Forum* **30**, 2208–2230 (Aug. 2011). 3
- [13] Tan, R. T. and Ikeuchi, K., “Separating Reflection Components of Textured Surfaces using a Single Image,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**, 178–193 (Feb. 2005). 3, 4, 5, 6

- [14] Yang, Q., Wang, S., and Ahuja, N., “Real-Time Specular Highlight Removal Using Bilateral Filtering,” in [*Proc. European Conference on Computer Vision*], Daniilidis, K., Maragos, P., and Paragios, N., eds., *Lecture Notes in Computer Science* **6314**, 87–100, Springer, Berlin / Heidelberg (Sept. 2010). 3, 6
- [15] Yoon, K.-J., Choi, Y., and Kweon, I. S., “Fast Separation of Reflection Components using a Specularity-Invariant Image Representation,” in [*Proc. IEEE International Conference on Image Processing*], 973–976, IEEE (Oct. 2006). 3, 5
- [16] Shafer, S. A., “Using Color to Separate Reflection Components,” *Color Research & Application* **10**(4), 210–218 (1985). 3
- [17] Riess, C., Eibenberger, E., and Angelopoulou, E., “Illuminant Color Estimation for Real-World Mixed-Illuminant Scenes,” in [*Proc. IEEE Color and Photometry in Computer Vision Workshop*], 782–789 (Nov. 2011). 4, 8
- [18] Eibenberger, E. and Angelopoulou, E., “Beyond the Neutral Interface Reflection Assumption in Illuminant Color Estimation,” in [*Proc. IEEE International Conference on Image Processing*], 4689–4692, IEEE (Sept. 2010). 4, 7
- [19] Biswas, S., Aggarwal, G., and Chellappa, R., “Robust Estimation of Albedo for Illumination-Invariant Matching and Shape Recovery,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**, 884–899 (May 2009). 6
- [20] Zhang, L. and Samaras, D., “Face Recognition from a Single Training Image under Arbitrary Unknown Lighting using Spherical Harmonics,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**, 351–363 (Mar. 2006). 6
- [21] Schlick, C., “An Inexpensive BRDF Model for Physically-based Rendering,” *Computer Graphics Forum* **13**(3), 233–246 (1994). 7
- [22] Kelemen, C. and Szirmay-Kalos, L., “A Microfacet Based Coupled Specular-Matte BRDF Model with Importance Sampling,” in [*Eurographics 2001 short presentation*], 25–34 (Sept. 2001). 7, 11
- [23] University of Houston, “UHDB30 Face Database.” <http://cbl.uh.edu/URxD/datasets> (2013). 8
- [24] Ragheb, H. and Hancock, E. R., “A Probabilistic Framework for Specular Shape-from-Shading,” *Pattern Recognition* **36**, 407–427 (Feb. 2003). 8
- [25] Lehmann, T. M. and Palm, C., “Color Line Search for Illuminant Estimation in Real-World Scenes,” *Journal Optical Society of America* **18**, 2679–2691 (Nov. 2001). 8, 11
- [26] Bolle, R. M., Connell, J. H., Pankanti, S., Ratha, N. K., and Senior, A. W., “The Relation between the ROC Curve and the CMC,” in [*Proc. IEEE Workshop on Automatic Identification Advanced Technologies*], 15–20, IEEE (Oct. 2005). 9
- [27] Mallick, S., Zickler, T., Belhumeur, P., and Kriegman, D., “Specularity Removal in Images and Videos: A PDE Approach,” in [*Proc. European Conference on Computer Vision*], Leonardis, A., Bischof, H., and Pinz, A., eds., *Lecture Notes in Computer Science* **3951**, 550–563, Springer Berlin / Heidelberg (May 2006). 11