Reflectance from Images: A Model-Based Approach for Human Faces

Martin Fuchs, Volker Blanz, *Member*, *IEEE*, Hendrik Lensch, and Hans-Peter Seidel

Abstract—In this paper, we present an image-based framework that acquires the reflectance properties of a human face. A range scan of the face is not required. Based on a morphable face model, the system estimates the 3D shape and establishes point-to-point correspondence across images taken from different viewpoints and across different individuals' faces. This provides a common parameterization of all reconstructed surfaces that can be used to compare and transfer BRDF data between different faces. Shape estimation from images compensates deformations of the face during the measurement process, such as facial expressions. In the common parameterization, regions of homogeneous materials on the face surface can be defined a priori. We apply analytical BRDF models to express the reflectance properties of each region and we estimate their parameters in a least-squares fit from the image data. For each of the surface points, the diffuse component of the BRDF is locally refined, which provides high detail. We present results for multiple analytical BRDF models, rendered at novel orientations and lighting conditions.

Index Terms— Color, shading, shadowing, texture, reflectance.

1 Introduction

In movie productions and interactive applications featuring virtual characters, vivid renderings involve subtle effects of light interacting with skin and expressive scenes may need to be shot with harsh lighting from grazing angles. The realistic reproduction of human faces by computer graphics under such conditions requires sophisticated models of the visual properties of skin. The problem of obtaining these without the time-consuming effort of an artist has been addressed by a number of approaches which use measurements of the reflectance of human skin. Many approaches require a large number of measurements or dedicated equipment for geometry reconstruction, such as range scanners, or complex setups for automated control of the illumination conditions.

In this paper, we focus on a precise method which reduces the measurement equipment and simplifies the procedure by exploiting computer vision techniques. Ambiguities, noise, and sparseness of the data obtained with the simplified measurement paradigm are compensated for by a model-based approach that uses plausibility constraints to reduce the set of possible solutions in terms of the reflectance function and the 3D shape of the surface.

Our acquisition pipeline can be divided in three main steps (Fig. 1). In the first step, digital photographs of a face at different orientations and illumination directions are taken in a calibrated environment.

Then, we fit a morphable model of 3D faces to each image. The reconstruction of 3D shape from each image is essential for recovering the reflectance function from the image data, which can be interpreted as sets of reflectance samples. Moreover, the morphable model provides a common parameterization of the facial surface that establishes correspondence across images of the face taken from different viewpoints. This correspondence is necessary for combining information from multiple images.

In the third step, we apply a nonlinear optimization method in order to estimate parameters of a spatially varying BRDF model for the face surface. The estimate can be used to render the face in novel views and lighting conditions.

Our contribution is an entirely image-based method that does not rely on a range scan of the face that is investigated. The method not only simplifies the measurement process, making the system applicable for a broader set of users, but it is also specially designed for nonrigid objects that may slightly change during the measurement process. This is due to our adaptive, model-based registration scheme that establishes correspondence between different images of the same face. Moreover, the common parameterization of different persons' faces in our unified framework for face representation makes it straightforward to automatically transfer the reflection properties of structures, such as the lips, from one person's face on the corresponding geometry of the new face.

2 Related Work

The measurement of reflection properties requires sampling the 4D space of incident and outgoing directions. Image-based measurements are based on the idea that, in one image of a curved surface, many combinations of light and viewing directions can be observed simultaneously [1], [2], [3], [4], [5], which results in a much more efficient acquisition paradigm than classical gonioreflectometers. While, typically, a point light source is applied for BRDF measurements, measurements of indoor scene with complex interreflections have also been demonstrated [6], [7], [8]. Ramamoorthi and Hanrahan [9] and Nishino et al. [10]

- *M. Fuchs, V. Blanz, and H.-P. Seidel are with the Max-Planck-Institut für Informatik, Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany. E-mail: {mfuchs, blanz, hpseidel}@mpi-sb.mpg.de.*
- *H. Lensch is with Stanford University, 253 Serra Mall, Gates Bldg. 3B-364, Stanford, CA 94305. E-mail: lensch@stanford.edu.*

Manuscript received 6 Aug. 2004; revised 20 Dec. 2004; accepted 3 Jan. 2005; published online 10 Mar. 2005.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org, and reference IEEECS Log Number TVCG-0087-0804.

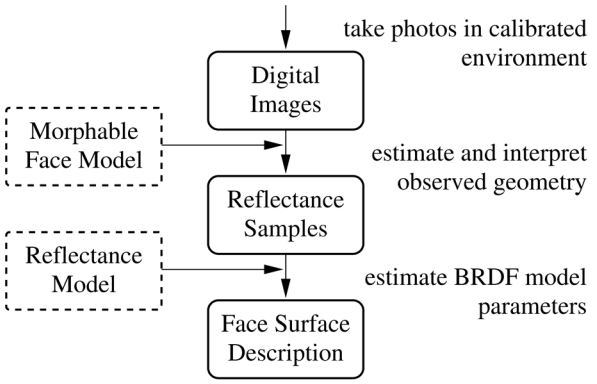

Fig. 1. Acquisition pipeline of our model-based approach.

even succeeded in reconstructing unconstrained incident illumination and reflection properties at the same time. Spatial variation in the surface reflection properties is only partially recovered by these methods, i.e., if at all, only the diffuse part is allowed to vary across the surface.

A very general approach for acquisition of reflectance properties was proposed by Lensch et al. [11]. The method is able to automatically cluster material properties and describe the spatial variations across the surface in terms of these clusters, including varying specular reflectivity. Three-dimensional shape recorded with range scanners is registered with the images using a silhouette matching algorithm. For human faces, the method would be difficult to apply since the silhouette is not informative enough to precisely determine 3D orientation.

Marschner et al. [12] were the first to apply image-based BRDF measurements to human skin. From a sparse set of input images, they fitted parameters of a reflectance model [13] to the data, recovering homogeneous reflection properties. With spatially uniform reflectance functions, local details of skin, which are important for the realistic appearance of faces, cannot be expressed. Marschner et al. later added a detailed albedo texture which was obtained from registered measurements captured from different viewpoints [14].

Debevec et al. [15] proposed a face relighting method that uses thousands of images recorded in a special setup, the light stage. Initially, the method involves no analytic function for skin reflectance, but, instead, exploits the superposition principle of light: Arbitrary lighting conditions can be simulated by computing a weighted sum of the recorded images of the face, each lit by a point light source at a different position. In the second part of their work, they perform a color space analysis to separate the contribution of specular from diffuse reflection and create a description of surface reflectance using the Torrance-Sparrow model [16].

In contrast, Paris et al. presented a method that obtains a rough approximation of a face's reflection properties by fitting the simple Phong reflectance model [17] to a single image of a face at frontal illumination [18]. All methods that were described so far require a priori input of the 3D geometry, which is usually acquired using a range scanner.

Face geometry can also be reconstructed from images; for instance, there is recent work by Moghaddam et al. [19], [20] which takes silhouette images as input and reconstructs a face as a linear combination in a model space.

When restricting the spatially varying component of the reflectance model only to the diffuse component of a Phong model, Blanz and Vetter have shown that a single image without knowledge about geometry and illumination of the face is sufficient for recovering a surface description [21], [22]. The reconstruction is possible due to an underlying database of previously recorded faces. Our work builds on this technique for acquiring the 3D shape.

Carceroni and Kutulakos [23] recovered shape, motion, and reflection properties from video streams with known directional lighting by exploiting the dense coherence in subsequent frames. They have demonstrated the application of their technique to human skin, but have not acquired a complete face.

Based on the Torrance-Sparrow model, Georghiades estimated both shape and reflectance in an exclusively image-based acquisition framework from 12 images taken from the same viewpoint [24]. Methods that only record from a single viewpoint, however, have access to only a few data at grazing viewing and illumination angles, which affects the reliability of the estimated BRDFs.

Image analysis of human faces is not restricted to measuring reflective properties. Tsumura et al. presented a color space method which separates the contribution of specular reflection and shading from the melanin and hemoglobin component of scattered light within the facial tissue [25].

There is recent work by Krishnaswamy and Baranoski [26] which addresses the choice of an appropriate model for human skin, including spectral phenomena and subsurface light transport. They provide an overview of the physical composition of human skin and contain references to in-detail work on anatomy. For our purposes, it is sufficient to observe that, on the interface between stratum corneum and air, specular reflection occurs, while light penetrating the surface is scattered diffusely. Skin being a nonmetallic material implies that the reflection is governed by Fresnel's Law. The Cook-Torrance model explicitly contains a Fresnel term to express that. Some other parametric models of reflectance can partly reproduce the angular dependency described by the Fresnel term implicitly. In this paper, we investigate the appropriateness of existing well-known models.

Subsurface light transport plays an important role for the appearance of human faces. It can be expressed using a BSSRDF. Jensen et al. have proposed a BRDF approximation for their subsurface light transport model [27]. Following this idea, we consider local reflectance models.

Except for the methods by Blanz-Vetter and Carceroni-Kutulakos, all methods for the reconstruction of individual face appearance assume the face to be static. If facial expressions or small deformations are involved, the resulting quality is degraded. In our approach, the geometry of the face is estimated from each individual input image while still maintaining a common parameterization. This offers the capability of measuring BRDFs using images from different viewpoints in different lighting directions, even with different facial expressions.

3 Measurement Setup

The measurements used in our system are digital photographs recorded in a calibrated environment. As shown in Fig. 2, we measure reflectances with an HMI point light

Fig. 2. The measurement setup. The camera (on the left) remains at the same position, while the orientation of the subject's face and the position of the light source (which is behind a protective pane) are varied. The steel spheres behind the subject are used for measuring the position of the light source relative to the camera.

source (which is behind a protective pane that looks overly bright in Fig. 2) in a photo lab with black walls, carpet, and ceiling [28]. The position of the camera is kept fixed, while the photographs show different orientations of the face and different positions of the light source. The position of the light relative to the camera is computed from the images using steel spheres as targets with the method of Lensch et al. [11]. Once per measurement session, we calibrate the digital camera with Bouguet's "Camera Calibration Toolbox" [29]. Our measurement loop involves the following steps:

- For each face, measure three metric distances on the face (such as distance between eye corners) for absolute scaling.
- Then, loop over a set of light conditions.
  - Record two pictures of the steel spheres to compute the light source position.
  - Loop over up to three poses of the subject.
    * Take a digital picture in point light conditions.
    * At the same pose, take an additional picture with diffuse light or take a set of pictures with varying exposure.
    * The combination of these pictures forms one measurement.

Taking several pictures per pose is helpful because it provides additional information about shadowed regions of the face, which is important for identifying features and for precise fitting of the model. This can be either accomplished by high dynamic range—allowing a more accurate placement of the feature points by the user in dark image areas—through the different pictures, or by one additional image with diffuse illumination, which, in addition, allows a more accurate geometry fit as the morphable face model has been shown to be more accurate in such lighting conditions.

In each case, the BRDF model will always be fitted against a single exposure image.

We have collected measurements from two subjects in varying pose and lighting situations. In total, we use 22 measurements for the female face and 21 for the male face.

4 Model-Based Shape Estimation and Registration

An essential step in our framework is the estimation of 3D shape from the image data, which is achieved by fitting a deformable model of human faces to each image. This approach has a number of advantages:

- No 3D scan of the face is required for reconstructing BRDF.
- The deformable model compensates deformations within the face, such as small facial expressions that occur during the measurement.
- Due to the global nature of the model, a consistent and realistic surface is estimated even if large portions of the face are shadowed in the images.
- The 3D surface is registered with each image in a semiautomatic way, using manually defined feature points.
- After fitting the model to images, corresponding pixels in different views of the face are automatically identified by computing a mapping from a common reference face to each input image. Combining information from different images for each surface point is essential for BRDF measurements.
- A common parameterization of different individual faces can be used for transferring reflectances between faces.

In the following, we briefly summarize the concept of the morphable model. For details, see [21], [22].

4.1 Morphable Model of Faces

The morphable model [30], [21] is a vector space of 3D shapes and textures spanned by a set of examples, capturing the common properties and the main modes of variation found in an object class such as human faces. Shape and texture vectors are defined such that any linear combination,

$$\mathbf{S} = \sum_{i=1}^{m} a_i \mathbf{S}_i, \quad \mathbf{T} = \sum_{i=1}^{m} b_i \mathbf{T}_i, \tag{1}$$

of examples $\mathbf{S}_i$, $\mathbf{T}_i$ is a realistic face if $\mathbf{S}$ and $\mathbf{T}$ are within a few standard deviations from their averages.

The morphable model is constructed from 200 textured *Cyberware* (TM) laser scans, which are given in a cylindrical parameterization at a resolution of 0.615mm in height and 0.7^o in azimuth.

A common parameterization is achieved by automatically establishing correspondence $(h, \phi) \mapsto (h_i, \phi_i)$ from a reference scan to each individual face i [21], [22] such that corresponding points, such as the tip of the nose, are assigned the same parameters (h, ϕ) on all faces.

Concatenating the vertex coordinates of the n sampling points of the reference face, we define shape and texture vectors:

$$\mathbf{S_0} = (x_1, y_1, z_1, x_2, \ldots, x_n, y_n, z_n)^T, \tag{2}$$

$$\mathbf{T_0} = (r_1, g_1, b_1, r_2, \ldots, r_n, g_n, b_n)^T. \tag{3}$$

Then, we can use the common parameterization to define $\mathbf{S}_i$ and $\mathbf{T}_i$ accordingly.

Fig. 3. Some reconstructed face geometry examples, rendered into the original images (lower row), in comparison to the original images (upper row). These geometries are used as input for the BRDF estimation. Measurements 1 and 9 are quite close to the original face, 8 shows a small artifact on the ridge of the nose, 16 and 17 do not visually equal the subject's face texture, but are still satisfactory in terms of geometry.

4.2 Fitting the Morphable Model to Images

The goal of the fitting process is to find coefficients a_i, b_i (1) and rendering parameters such that the rendered image I_{model} is as similar as possible to the input image I_{input} in terms of image difference [21], [22]. Rendering involves recomputing the surface normals at each iteration, perspective projection, and Phong illumination. The Phong model serves as an initial step in a bootstrapping approach to the problem addressed in this paper, which is to estimate 3D shape and BRDF from the same image data. Since the fitting algorithm makes a conservative estimate on shape and texture, we do not expect significant errors caused by assuming Phong illumination. In addition to local shading, the algorithm computes cast shadows with a shadow map once every 1,000 iterations. The following rendering parameters are automatically optimized:

- 3D rotation (three angles),
- 3D translation (three dimensions),
- focal length of the camera (one variable),
- angle of directed light (two angles),
- intensity of directed light (three color channels),
- intensity of ambient light (three color channels),
- color contrast (one variable),
- gain in each color channel (three variables),
- offset in each color channel (three variables).

The optimization starts with default values defining a frontal view and frontal illumination. Fitting the model is achieved by a stochastic Newton algorithm [22] minimizing a cost function $E = E_I + E_F + E_P$. E_I is the sum of square differences in the three color channels, E_F is the sum of square differences of image-plane positions of a set of about 7-20 manually defined feature points, such as the corners of the eyes. This term, which contributes to E only in the first iterations, helps the system to converge by pulling some feature points toward the predefined positions. E_P is a regularization term that penalizes linear combinations that are unplausible in terms of the probability density estimated by a Principal Component Analysis of the database of 3D faces. E_P is the Mahalanobis distance of the solution from the average face (for details, see [22]).

Fitting the model to the images provides a correspondence mapping from texture coordinates of the morphable model to image coordinates in each image. In order to achieve precise correspondence, the slight facial deformations due to movements during the long delay between measurements are compensated by fitting the model to each image separately, starting from the average shape. This essentially makes each fit an independent estimate of shape, which introduces measurement errors in our procedure, even though the explicit and implicit constraints in the morphable model and the image data make sure that we still obtain realistic surface normals that are consistent across measurements.

In each surface point of the male model, we have measured the deviation of the surface normal of each reconstructed head from the average of all reconstructions: The mean angular deviation is 6.3 degrees.

While the fitting algorithm places model structures, such as the corner of the mouth, to the desired position in each image, we cannot expect that individual structures which are not captured by the morphable model, such as moles or freckles, are set in correspondence with the same vertex or texture coordinate of the 3D model across different images. We therefore extract the person's texture from a frontal view at frontal illumination with the method described in [21] and use this texture instead of the average texture as a starting value in each fitting process: As a consequence, misalignments of individual texture details will be penalized throughout the fit.

Even though the deformations between images are due to facial expressions, it was not necessary to use the extended vector space of facial expressions presented in [31]: It turned out that the small variations of expression found in the database of 200 close-to-neutral individual faces are sufficient. The quality of the fit is shown in Fig. 3.

4.3 Sampling Facial Reflectance

With the face geometry reconstructed, we can resample the input images and thus obtain, from each measurement, a texture map of observed radiance in the common parameterization of facial surface.

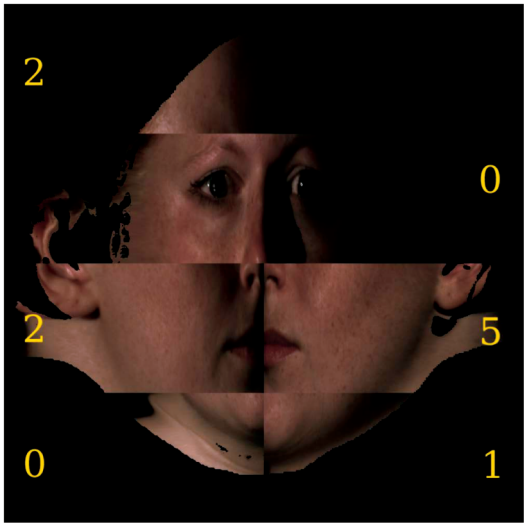

Fig. 4. Measured radiance on the face surface in the common surface parameterization for faces in a combination of different measurements. While the correspondence between the chin region of measurement 0 and 1 is suboptimal, good correspondence of sharp face features such as lips and nose edges is established between measurements 0, 2, and 5.

Fig. 4 shows several of these maps combined in one image, illustrating the correspondence between different measurements. In the same parameterization, we can store vectors $\mathbf{l}$ and $\mathbf{v}$ pointing to the light and to the viewer, along with surface normals $\mathbf{n}$ and point coordinates $\mathbf{p}$ in some global coordinate system. $\mathbf{l}$ and $\mathbf{v}$ are then transformed into an orthonormal local coordinate system defined by the normal $\mathbf{n}$ and the tangents $\mathbf{t}$ and $\mathbf{b}$, and normalized in length to be elements of the unit sphere, resulting in according vectors $\hat{\mathbf{l}}$ and $\hat{\mathbf{v}}$.

With the reflectances $\mathbf{r}$ computed from the observed radiances $\mathbf{R}$

$$\mathbf{r} := \frac{\mathbf{R}}{|\hat{\mathbf{l}} \cdot \mathbf{n}| \cdot \|\mathbf{l}\|^2}, \tag{4}$$

each point in the discretized face parameterization can be considered as a reflectance sample

$$S \ni s = ((\hat{\mathbf{l}}, \hat{\mathbf{v}}), \mathbf{r}, \sigma) \in \left(\left(S^2 \times S^2\right) \times \mathbb{R}^3 \times \mathbb{R}_0^+\right). \tag{5}$$

We linearly scale all reflectance values so that $(1,1,1)^{\mathrm{T}}$ is the reflectance of a diffuse white calibration target. σ will be used to express the confidence into the sample and it will be defined below in Section 5.1.

Some of the reflectance samples s in the set S of measured data cannot contribute to a meaningful measurement and are, therefore, discarded: These are samples in cast shadow, which can be identified using the shadow buffer technique. Also, samples which were seen or lit from a flat angle (l_z or v_z are low) should be discarded as the uncertainty in the normal estimation causes the largest error there. We use a morphological erosion filter in order to remove additional samples along the margin of shadows and occluding contours.

Fig. 5. In the parameterization defined by the reference texture (left), manually marked regions of different surface properties (right) apply to all faces automatically.

4.4 A Priori Clustering

The surface of a human face has regions with different types of tissue, such as skin, lips, eyes, and hair, and the reflectance properties of these regions differ considerably. Performing a separate data analysis on these regions helps to achieve high-quality models of reflectance within each region. A BRDF measurement technique for general objects would have to assign these clusters in an automated way [11]. Focusing on human faces, however, we can exploit the common parameterization of our face representation that defines correspondence from face to face and manually define regions of different reflectance properties in terms of this parameterization (Fig. 5), which can then be applied to all faces in an automated way.

5 Estimating BRDF Model Parameters

Generalization from the samples of the BRDF that we observed in the different face regions to novel viewing and lighting conditions requires interpolation and extrapolation of the BRDF. Due to measurement noise in the data within each cluster, we have to solve statistical regression problems, estimating a BRDF

$$f : S^2 \times S^2 \to \mathbb{R}^3, \quad \left(\hat{\mathbf{l}}, \hat{\mathbf{v}}\right) \mapsto \mathbf{r} \tag{6}$$

which maps the incoming light and viewing direction to the reflectance $\mathbf{r}$ in the three color channels, red, green, and blue. Estimating f from the sparse data is heavily underdetermined, so we have to assume that the solution f is an element of a general BRDF model, formally given as

$$M : \mathbb{R}^n \to \left(S^2 \times S^2 \to \mathbb{R}^3\right). \tag{7}$$

M maps a parameter vector $\mathbf{a} \in \mathbb{R}^n$ to a BRDF f. Thus, the regression problem is transformed into finding an optimal parameter vector $\mathbf{a}$.

5.1 Error Functional

In order to define a criterion for the optimal solution, we use a quadratic error functional which expresses the distance between the measurements in the sample set S and the BRDF given by the model M for the parameters $\mathbf{a}$:

$$E(S, M(\mathbf{a})) = \frac{1}{2} \sum_{((\hat{\mathbf{l}},\hat{\mathbf{v}}),\mathbf{r},\sigma) \in S} \frac{\left(\mathbf{r} - M(\mathbf{a})(\hat{\mathbf{l}}, \hat{\mathbf{v}})\right)^2}{\sigma^2}. \tag{8}$$

Each sample is assigned a confidence value, σ, which weights the cost function. As we do not know standard

Fig. 6. Result of fitting the Cook-Torrance model against the samples in a priori clusters.

Fig. 7. Result of fitting the Cook-Torrance model against the samples in a priori clusters after estimating spatially varying diffuse components.

deviations of measurement errors yet, we use σ to heuristically control the cost function. For each sample $((\hat{\mathbf{l}}, \hat{\mathbf{v}}), \mathbf{r}, \sigma)$, let

$$\sigma := \frac{1}{l_z \cdot l_z \cdot v_z}. \tag{9}$$

The first division by l_z simply cancels out with the division by $|\hat{\mathbf{l}} \cdot \mathbf{n}|$ in (4). Because of that, camera noise, which is independent of the observed surface normal and the incident light direction, has the same effect on all samples and is treated evenly by the cost function.

The second division by l_z and v_z addresses noise caused by registration errors. In regions where the camera observed the face surface from an almost normal angle, a small delocalization of the estimated surface against the measured surface causes few problems as the normals do not vary much there. Under grazing angle conditions, however, normals vary more strongly for changing coordinates in the image; accordingly, a small delocalization has a stronger effect there and such samples need to be treated as less reliable. An analogous argument for the incident light direction motivates the division by l_z.

Thus, we effectively obtain an error functional similar to the one used by Lafortune et al. [13].

5.2 Hierarchical Nonlinear Optimization

We perform a nonlinear optimization using the Levenberg-Marquardt algorithm as described by Press et al. [32]. For each material cluster, we start with a conservative initial value and fit the BRDF model parameters against a small, randomly chosen subset of the samples of the cluster. In subsequent fitting iterations, we increase the size of the subset until all available data are considered. This algorithm spends many iterations on a few samples, while we are still far from the optimum, and a few iterations on many samples, which reduces the overall fit time significantly. A result is shown in Fig. 6.

The performance of the first Levenberg-Marquardt fitting step depends on the initial value; it needs more iterations if chosen far from the optimum. As the following steps start with the result of the previous steps, their runtime depends less on the initial value.

Tests which we performed with an isotropic Lafortune model on synthetic data indicate that BRDF model parameters which express diffuse light scattering can almost always be accurately recovered; direction dependent scattering is estimated reliably in cases where the kind of scattering is suggested by the initial parameters (e.g., one forward pointing lobe for forward scattering).

5.3 Refining Spatial Detail

By dividing the face into regions with separate BRDFs, we are able to estimate the full parameter vector, including specular properties. However, this results in a BRDF that is spatially uniform within the surface region, which looks unrealistic (see Fig. 6). In order to reproduce the spatial variation caused by surface details on the skin, we fit the BRDF again for each point of the discretized face surface, this time only considering samples for this precise point.

As we have much fewer samples now, we cannot estimate the full parameter set, so we only consider spatial variation of the parameters that model the diffuse reflection. As the diffuse reflection can be observed from most viewing and light conditions and it contributes, for most BRDF models, linearly to the total reflectance, it can be reliably estimated, even from a small number of well-lit measurements.

The restriction to fewer measurements may induce artifacts at borders of different measurements in the texture space because small errors in the estimation of the distance between face and light source cause the overall brightness of the diffuse texture to vary, which may show up as visible discontinuity at measurement borders. At the face contour and at the shadow line, these are not noticeable as our cost function from (8) assigns low confidence to these samples. Along the border of cast shadows and to regions where samples are filtered out because their brightness is outside a confidence range for the observed cluster, this mechanism alone is too weak and small artifacts remain which may result in visible discontinuities.

We therefore define, for each sample, the Euclidean distance in texture space d from this sample to the closest position without sample in the same measurement. With that, we can redefine

$$\sigma := \frac{d_{\max}}{\min\{d, d_{\max}\} \cdot l_z \cdot l_z \cdot v_z}. \tag{10}$$

The results in this paper have been obtained for $d_{\max} = 10$. An example is shown in Fig. 7.

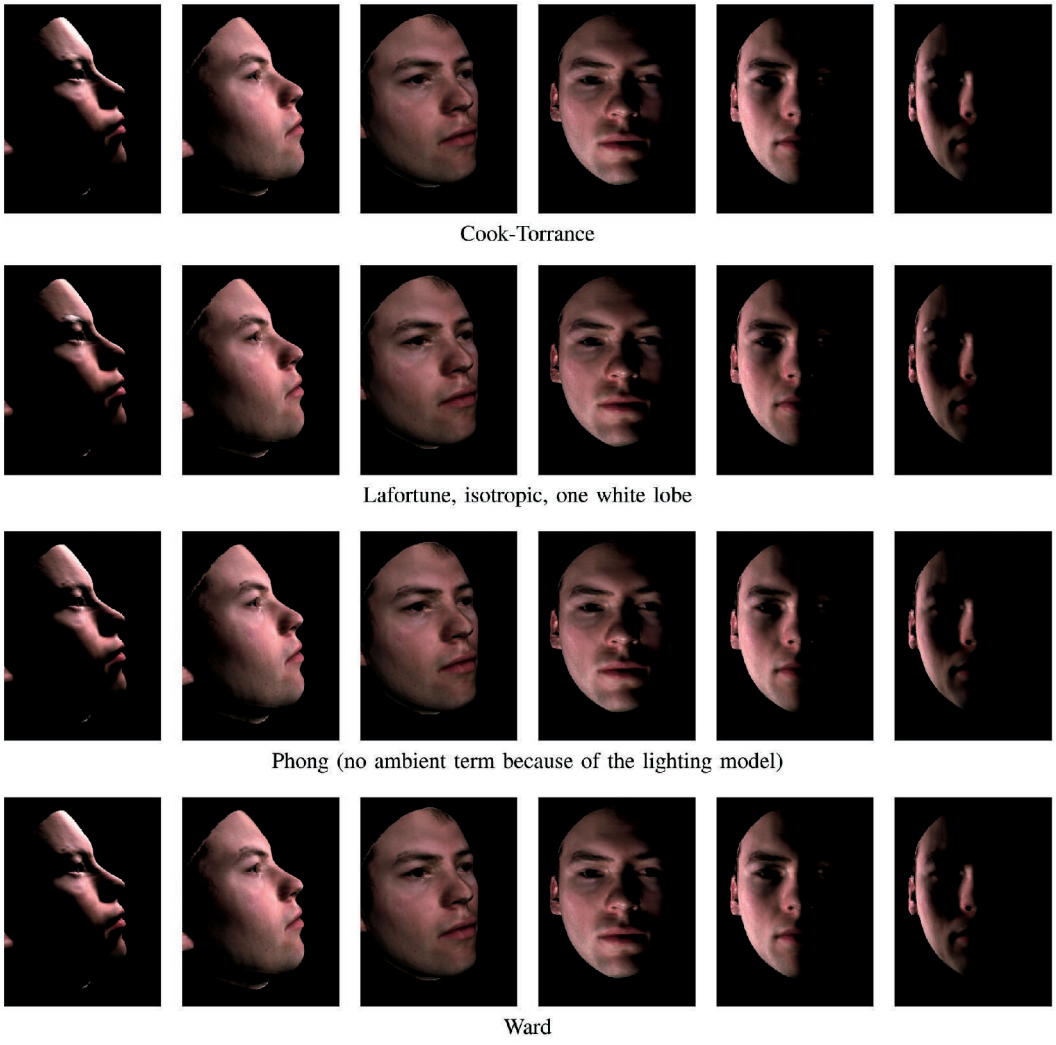

Fig. 8. Renderings of novel pose and varying point lighting, in BRDFs obtained from various models.

6 Results

The geometric quality of our estimations produces plausible results, as shown in Fig. 3. View 16 and 17 in Fig. 3 demonstrate the necessity of an additional image with the same pose, either with diffuse lighting or with extended dynamic range. The geometry reconstruction would fail if only sparse image information is available due to extreme lighting conditions. With the additional image, the correct 3D shape is reconstructed and the correspondence between different views is accurately established (see Fig. 4).

Since the 3D shape is adapted to each individual input image, deformations on the face geometry can be easily compensated for. We observed such deformations over time (compare views 8 and 9 in Fig. 3) and as reactions to the light intensity, e.g., when the light moves close to the face, humans tend to partially close the eyelids.

For the resolution of the measured reflectance maps, we choose $1,000 \times 1,000$ in h and φ. With respect to the quality of the obtained reflectance samples, we observed that samples at grazing light and viewing angles were subject to more measurement noise than perpendicularly lit and observed samples. This is accounted for by assigning a lower confidence value which increased the stability of our fitting process.

6.1 Comparison of BRDF Models

We have performed the estimation of model parameters for several well-established lighting models. Specifically, these were:

- the Phong Model [17], which is probably the most widely used of all available models due to its simplicity and early availability in hardware,
- the Cook-Torrance [34] model, which models specular reflectance according to explicit microgeometry assumptions,
- the Ward model [35], which is designed for fitting against experimental data, while still providing physical parameters, and
- the Lafortune model [13], which models a superset of the effects modeled by the Phong model, among them off-specular reflections, while still being computationally simple.

In all cases, we used isotropic variants of the models. Resulting renderings are shown in Fig. 8, a comparison of original images and a superimposed reconstruction for the same illumination can be seen in Fig. 9. A rendering approximating real world lighting conditions is given in Fig. 10.

Fig. 9. Renderings of original measurement and superimposed reconstruction. The same gamma curve has been applied to both.

As all of the models employed express diffuse reflection using a Lambertian term, differences in nonspecular areas are hardly visible. For the specular terms, differences can be more easily observed. The results indicate that our system found realistic parameter settings for all models that were tested and for most of the material clusters.

Upon visual inspection of the results, we tend to prefer the Cook-Torrance or, alternatively, the Ward model for rendering. The Phong model lacks generality with respect to specular effects and, accordingly, fails to reproduce brighter specularities under grazing angles. The Lafortune model is superior in this respect; however, it does not inherently restrict the form of the model to plausible solutions and is, therefore, more easily affected by overfitting and measurement noise. An example of this can be seen in the overly bright reflex in the eyebrow region in the leftmost rendering in Fig. 8.

Fig. 10. Rendering in an approximation of a real-world situation. Environment by Paul Debevec [33].

The original Lafortune model contains a linear combination of several lobes for complex scattering effects and one lobe alone already allows a much more expressive BRDF than the Phong model. Possibly in conjunction with measurement error, the stability of the fit of several lobes was an issue in our experiments; when fitting the parameters for several lobes, the resulting shapes of the additional lobes depended heavily on the choice of initial parameters and varied from person to person.

Higher quality in the reproduction of specular properties is observed for the Ward and Cook-Torrance models, which are both expressive. Both models inherently produce plausible specular lobes.

The fits of all models show one general aspect of our system which might motivate further improvement: While our approach so far restricts the specular variation to a piecewise constant inside the clusters, specular properties may vary continuously over the face. Thus, some areas in the face have, in reality, a stronger specular highlight (such as the nose) or a weaker one (such as in the beard shadow of the male model). This is to some extent compensated for in the diffuse refinement step by assigning higher or lower Lambertian diffuse reflectance to areas where highlights

Fig. 11. Rendered at novel pose and lighting, the images on the left and right show shapes and spatially varying Cook-Torrance BRDFs of two individuals. The images in the center demonstrate that BRDFs can be transferred between shapes.

were observed. Consequently, the distinction between diffuse and specular reflection is weakened, however, the overall result is improved (as can be seen in Fig. 8).

6.2 Transfer of BRDF on Faces

In the common face surface parameterization, the locations of features do not depend on the individual. Therefore, the individual surface description is intrinsically transferable among different faces. We demonstrate this by exchanging the surface description of two significantly different models, as seen in Fig. 11. While the geometry of the male and the female model are distinct both in size and in local features, an exchange is still possible, creating paradox effects of a male texture on a female face and vice versa. One possible application of this could lead to offline make-up if combined with measured illumination.

7 Conclusion

We have presented a novel acquisition paradigm of reflection properties of human faces. The method is entirely image-based and recovers 3D geometry and spatially varying BRDFs from a sparse set of digital photographs by estimating model parameters.

The morphable face model that is used for the geometry reconstruction accounts for deformations of the face and establishes a common parameterization for different input images and even between different individuals. Based on this novel parameterization, one can define surface regions with distinct material properties for which separate BRDFs are obtained. Within each region, spatial variation of the diffuse component is captured to account for details of natural skin. The framework produces realistic renderings of human faces, even for harsh lighting conditions.

The common parameterization of faces and the general, simple acquisition pipeline is suitable for a wide range of future developments, e.g., the investigation of effects such as subsurface light transport [36] and scattering from velvety hair covering facial skin [37]. Subsurface scattering, which is a topic of recent research [38], is of special interest here.

The rich variety of illumination effects on human skin and the fact that observers are sensitive to subtle details of facial appearance make human faces a relevant and challenging test bed for modeling reflection properties.

References

[1] Y. Sato, M.D. Wheeler, and K. Ikeuchi, "Object Shape and Reflectance Modeling from Observation," *Proc. SIGGRAPH '97,* pp. 379-388, Aug. 1997.

[2] R. Lu, J. Koenderink, and A. Kappers, "Optical Properties (Bidirectional Reflectance Distribution Functions) of Velvet," *Applied Optics,* vol. 37, no. 25, pp. 5974-5984, Sept. 1998.

[3] Y. Yu and J. Malik, "Recovering Photometric Properties of Architectural Scenes from Photographs," *Proc. SIGGRAPH '98,* pp. 207-218, July 1998.

[4] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "A Data-Driven Reflectance Model," *ACM SIGGRAPH,* vol. 22, no. 3, pp. 759-769, July 2003.

[5] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "Efficient Isotropic BRDF Measurement," *Proc. Eurographics Symp. Rendering: 14th Eurographics Workshop Rendering,* pp. 241-248, June 2003.

[6] Y. Yu, P. Debevec, J. Malik, and T. Hawkins, "Inverse Global Illumination: Recovering Reflectance Models of Real Scenes from Photographs," *Proc. SIGGRAPH,* pp. 215-224, Aug. 1999.

[7] S. Gibson, T. Howard, and R. Hubbold, "Flexible Image-Based Photometric Reconstruction Using Virtual Light Sources," *Computer Graphics Forum,* vol. 20, no. 3, 2001.

[8] S. Boivin and A. Gagalowicz, "Image-Based Rendering of Diffuse, Specular and Glossy Surfaces from a Single Image," *Proc. SIGGRAPH 2001, Computer Graphics Proc., Ann. Conf. Series,* pp. 107-116, Aug. 2001.

[9] R. Ramamoorthi and P. Hanrahan, "A Signal-Processing Framework for Inverse Rendering," *Proc. SIGGRAPH 2001, Computer Graphics Proc., Ann. Conf. Series,* pp. 117-128, Aug. 2001.

[10] K. Nishino, Z. Zhang, and K. Ikeuchi, "Determining Reflectance Parameters and Illumination Distribution from a Sparse Set of Images for View-Dependent Image Synthesis," *Proc. Eighth IEEE Intl Conf. Computer Vision (ICCV '01),* pp. 599-606, July 2001.

[11] H.P.A. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H.-P. Seidel, "Image-Based Reconstruction of Spatial Appearance and Geometric Detail," *ACM Trans. Graphics,* vol. 27, no. 2, Apr. 2003.

[12] S. Marschner, S. Westin, E. Lafortune, K. Torrance, and D. Greenberg, "Image-Based BRDF Measurement Including Human Skin," *Proc. 10th Eurographics Workshop Rendering,* pp. 131-144, June 1999.

[13] E.P.F. Lafortune, S.-C. Foo, K.E. Torrance, and D.P. Greenberg, "Non-Linear Approximation of Reflectance Functions," *Proc. 24th Ann. Conf. Computer Graphics and Interactive Techniques,* pp. 117-126, 199.7

[14] S. Marschner, B. Guenter, and S. Raghupathy, "Modeling and Rendering for Realistic Facial Animation," *Proc. 11th Eurographics Workshop Rendering,* pp. 231-242, June 2000.

[15] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the Reflectance Field of a Human Face," *Proc. 27th Ann. Conf. Computer Graphics and Interactive Techniques,* pp. 145-156, 2000.

[16] K. Torrance and E. Sparrow, "Theory for Off-Specular Reflection from Roughened Surfaces," *J. Optical Soc. Am.,* vol. 57, no. 9, pp. 1105-1114, 1967.

[17] B.T. Phong, "Illumination for Computer Generated Pictures," *Comm. ACM,* vol. 18, no. 6, pp. 311-317, June 1975.

[18] S. Paris, F.X. Sillion, and L. Quan, "Lightweight Face Relighting," *Proc. 11th Pacific Conf. Computer Graphics and Applications,* p. 41, 2003.
[19] B. Moghaddam, J. Lee, H. Pfister, and R. Machiraju, "Model-Based 3D Face Capture with Shape-from-Silhouettes," *AMFG '03: Proc. IEEE Int'l Workshop Analysis and Modeling of Faces and Gestures,* p. 20, 2003.
[20] J. Lee, B. Moghaddam, H. Pfister, and R. Machiraju, "Finding Optimal Views for 3D Face Shape Modeling," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition (FG),* May 2004.
[21] V. Blanz and T. Vetter, "A Morphable Model for the Synthesis of 3D Faces," *Computer Graphics Proc. SIGGRAPH '99,* pp. 187-194, 1999.
[22] V. Blanz and T. Vetter, "Face Recognition Based on Fitting a 3D Morphable Model," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 25, no. 9, pp. 1063-1074, 2003.
[23] R.L. Carceroni and K.N. Kutulakos, "Multi-View Scene Capture by Surfel Sampling: From Video Streams to Non-Rigid 3D Motion, Shape and Reflectance," *Int'l J. Computer Vision,* vol. 49, nos. 2-3, pp. 175-214, 2002.
[24] A.S. Georghiades, "Recovering 3-D Shape and Reflectance from a Small Number of Photographs," *Rendering Techniques 2003: Proc. 14th Eurographics Workshop Rendering,* pp. 230-240, June 2003.
[25] N. Tsumura, N. Ojima, K. Sato, M. Shiraishi, H. Shimizu, H. Nabeshima, S. Akazaki, K. Hori, and Y. Miyake, "Image-Based Skin Color and Texture Analysis/Synthesis by Extracting Hemoglobin and Melanin Information in the Skin," *ACM Trans. Graphics,* vol. 22, no. 3, pp. 770-779, 2003.
[26] A. Krishnaswamy and G.V.G. Baranoski, "A Biophysically-Based Spectral Model of Light Interaction with Human Skin," *Proc. Eurographics,* 2004.
[27] H.W. Jensen, S. Marschner, M. Levoy, and P. Hanrahan, "A Practical Model for Subsurface Light Transport," *Proc. SIGGRAPH,* pp. 511-518, Aug. 2001.
[28] M. Goesele, H.P.A. Lensch, W. Heidrich, and H.-P. Seidel, "Building a Photo Studio for Measurement Purposes," *Proc. Vision, Modeling, and Visualization (VMV-00),* pp. 231-238, 2000.
[29] J.-Y. Bouguet, "Camera Calibration Toolbox for Matlab," http://www.vision.caltech.edu/bouguetj, year?
[30] T. Vetter and T. Poggio, "Linear Object Classes and Image Synthesis from a Single Example Image," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 19, no. 7, pp. 733-742, July 1997.
[31] V. Blanz, C. Basso, T. Poggio, and T. Vetter, "Reanimating Faces in Images and Video," *Computer Graphics Forum,* P. Brunet and D. Fellner, eds., vol. 22, no. 3, pp. 641-650, 2003.
[32] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing,* second ed. Cambridge Univ. Press, 1994.
[33] P. Debevec, "Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-Based Graphics with Global Illumination and High Dynamic Range Photography," *journal?* pp. 189-198, July 1998.
[34] R.L. Cook and K.E. Torrance, "A Reflectance Model for Computer Graphics," *ACM Trans. Graphics,* vol. 1, no. 1, pp. 7-24, 1982.
[35] G. Ward Larson, "Measuring and Modeling Anisotropic Reflection," *Proc. SIGGRAPH,* pp. 265-272, July 1992.
[36] P. Hanrahan and W. Krueger, "Reflection from Layered Surfaces Due to Subsurface Scattering," *Proc. SIGGRAPH 93, Computer Graphics Proc., Ann. Conf. Series,* pp. 165-174, Aug. 1993.
[37] J. Koenderink and S. Pont, "The Secret of Velvety Skin," *Machine Vision Applications,* vol. 14, no. 4, pp. 260-268, 2003.
[38] M. Goesele, H.P.A. Lensch, J. Lang, C. Fuchs, and H.-P. Seidel, "Disco: Acquisition of Translucent Objects," *ACM Trans. Graphics,* vol. 23, no. 3, pp. 835-844, 2004.

Martin Fuchs studied computer science at the University of Saarland, Germany, receiving the diploma-degree in 2004. Since then, he has begun his PhD studies at the Max-Planck-Institute for Computer Science in Saarbrücken, Germany. His work focuses on image-based measurement techniques for relightable material representations in the context of human faces.

Volker Blanz received the diploma-degree from the University of Tübingen, Germany, in 1995. He then worked on a project on multiclass support vector machines at AT&T Bell Labs in Holmdel, New Jersey. He received the PhD degree in physics from the University of Tübingen in 2000 for his thesis on reconstructing 3D shape from images, written at the Max-Planck-Institute for Biological Cybernetics, Tübingen. He was a visiting researcher at the Center for Biological and Computational Learning at the Massachusetts Institute of Technology and a research assistant at the University of Freiburg. In 2003, he joined the Max-Planck-Institute for Computer Science, Saarbrücken, Germany. Since 2004, he has led an independent research group "Learning-Based Modeling of Objects" at the Max-Planck Center for Visual Computing and Communication (Saarbrücken and Stanford University). His research interests are in the fields of face recognition, machine learning, facial modeling, and animation. He is a member of the IEEE.

Hendrik Lensch studied computer science at the University of Erlangen and the Royal Institute of Technology (KTH) in Stockholm, receiving the diploma in 1999. He worked as a PhD student and research associate in the computer graphics group at the Max-Planck-Institut für Informatik in Saarbrücken, Germany. In 2003, he received the doctorate degree from Saarland University. He is a visiting assistant professor with Marc Levoy at Stanford University's Computer Graphics Lab in California, where he leads the research group "General Appearance Acquisition" of the Max Planck Center for Visual Computing and Communication (Saarbrücken/Stanford). His research centers on techniques for acquiring computer graphics models of real-world objects.

Hans-Peter Seidel is the scientific director and chair of the computer graphics group at the Max-Planck-Institut (MPI) Informatik and a professor of computer science at the University of Saarbrücken, Germany. The Saarbrücken computer graphics group was established in 1999 and currently consists of about 35 researchers. He has published some 200 technical papers in the field and has lectured widely on these topics. He has received grants from a wide range of organizations, including the German National Science Foundation (DFG), the European Community (EU), NATO, and the German-Israel Foundation (GIF). In 2003, he was awarded the "Leibniz Preis," the most prestigious German research award, from the German Research Foundation (DFG). He is the first computer graphics researcher to receive this award. In 2004, he was selected as the founding chair of the Eurographics Awards Programme.