

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

Martin Fuchs Volker Blanz
Hendrik Lensch Hans-Peter Seidel

MPI-I-2005-4-001

January 2005

FORSCHUNGSBERICHT RESEARCH REPORT

MAX-PLANCK-INSTITUT
FÜR
INFORMATIK

Stuhlsatzenhausweg 85 66123 Saarbrücken Germany

Authors' Addresses

Martin Fuchs, Volker Blanz, and Hans-Peter Seidel
Max-Planck-Institut für Informatik
Stuhlsatzenhausweg 85
66123 Saarbrücken
Germany
{mfuchs|blanz|hpseidel}@mpi-sb.mpg.de

Hendrik Lensch
Stanford University
253 Serra Mall
Gates Bldg 3B-364
Stanford, CA 94305
USA
lensch@stanford.edu

Keywords

Color, shading, shadowing, and texture; Reflectance

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

Martin Fuchs Volker Blanz Hendrik Lensch
Hans-Peter Seidel

January 6, 2005

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.

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 (Figure 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 non-linear 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 does not only simplify the measurement process, making the system applicable for a broader set of users, but it is also specially designed for non-rigid 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 straight-forward 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 to sample 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, also measurements of indoor scene with complex interreflections have been demonstrated [6, 7, 8]. Ramamoorthi and Hanrahan [9] and Nishino et al. [10] 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. 3D 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 who applied 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 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. 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 [15]. In the second part of their work, they perform a color space analysis to separate 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, Lee 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 illumi-

nation of the face is sufficient for recovering a surface description [21, 22]. The reconstruction is possible due to a underlying data base 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 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 non-metallic 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 of 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 Figure 2, we measure reflectances with a HMI point light source (which is behind a protective pane that looks overly bright in Figure 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 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 semi-automatic 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

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

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° in azimuth.

A common parameterization is achieved by automatically establishing correspondence $(h, \varphi) \mapsto (h_i, \varphi_i)$ 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:

$$\mathbf{S}_0 = (x_1, y_1, z_1, x_2, \dots, x_n, y_n, z_n)^T \quad (2)$$

$$\mathbf{T}_0 = (r_1, g_1, b_1, r_2, \dots, r_n, g_n, b_n)^T. \quad (3)$$

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

4.2 Fitting the Morphable Model to Images

The goal of the fitting process is to find coefficients a_i, b_i (Equation 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 re-computing 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 1000 iterations. The following rendering parameters are automatically optimized:

- 3D rotation (3 angles)
- 3D translation (3 dimensions)
- focal length of the camera (1 variable)
- angle of directed light (2 angles)
- intensity of directed light (3 color channels)
- intensity of ambient light (3 color channels)
- color contrast (1 variable)
- gain in each color channel (3 variables)
- offset in each color channel (3 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 towards 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 Figure 3.

4.3 Sampling Facial Reflectance

With the face geometry reconstructed, we can re-sample the input images and thus obtain, from each measurement, a texture map of observed radiance

in the common parameterization of facial surface.

Figure 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}

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

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

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

We linearly scale all reflectance values, so that $(1, 1, 1)^T$ 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.

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 (Figure 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

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

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

$$M : \mathbb{R}^n \rightarrow (S^2 \times S^2 \rightarrow \mathbb{R}^3) \quad (7)$$

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} :

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

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

$$\sigma := \frac{1}{l_z \cdot l_z \cdot v_z} \quad (9)$$

The first division by l_z simply cancels out with the division by $|\hat{\mathbf{l}} \cdot \mathbf{n}|$ in Equation 4. Because of that, camera noise, which is independent from 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 Non-linear Optimization

We perform a non-linear 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 few samples, while we are still far from the optimum, and few iterations on many samples, which reduces the overall fit time significantly. A result is shown in Figure 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 Figure 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 less 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 reliably be estimated, even from a small number of well-lit measurements.

The restriction to less measurements may induce artefacts 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 Equation 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 artefacts 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 re-define

$$\sigma := \frac{d_{\max}}{\min\{d, d_{\max}\} \cdot l_z \cdot l_z \cdot v_z} \quad (10)$$

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

6 Results

The geometric quality of our estimations produces plausible results, as shown in Figure 3. View 16 and 17 in Figure 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 Figure 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 view 8 and 9 in Figure 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 1000 x 1000 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 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 micro-geometry 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 Figure 8, a comparison of original images and a superimposed reconstruction for the same illumination can be seen in Figure 9. A rendering approximating real world lighting conditions is given in Figure 10.

As all of models employed express diffuse reflection using a Lambertian term, differences in non-specular 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 over-fitting and measurement noise. An example of this can be seen on the overly bright reflex in the eyebrow region, in the leftmost rendering in Figure 8.

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 model, which both are 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 piece-wise 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 were observed. Consequently, the distinction between diffuse and specular reflection is weakened, however, the overall result is improved (as can be seen in Figure 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 Figure 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 off-line 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] Yoichi Sato, Mark D. Wheeler, and Katsushi Ikeuchi. Object Shape and Reflectance Modeling from Observation. In *Proceedings of SIGGRAPH 97*, pages 379–388, August 1997.
- [2] R. Lu, J. Koenderink, and A. Kappers. Optical Properties (bidirectional reflectance distribution functions) of velvet. *Applied Optics*, 37(25):5974–5984, September 1998.
- [3] Yizhou Yu and Jitendra Malik. Recovering Photometric Properties of Architectural Scenes from Photographs. In *Proceedings of SIGGRAPH 98*, pages 207–218, July 1998.
- [4] Wojciech Matusik, Hanspeter Pfister, Matthew Brand, and Leonard McMillan. A data-driven reflectance model. *ACM SIGGRAPH*, 22(3):759–769, July 2003.
- [5] Wojciech Matusik, Hanspeter Pfister, Matthew Brand, and Leonard McMillan. Efficient isotropic brdf measurement. In *Eurographics Symposium on Rendering: 14th Eurographics Workshop on Rendering*, pages 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. In *Proc. SIGGRAPH*, pages 215–224, August 1999.
- [7] Simon Gibson, Toby Howard, and Roger Hubbard. Flexible image-based photometric reconstruction using virtual light sources. *Computer Graphics Forum*, 20(3), 2001. ISSN 1067-7055.
- [8] Samuel Boivin and André Gagalowicz. Image-based rendering of diffuse, specular and glossy surfaces from a single image. In *Proc. of SIGGRAPH 2001*, Computer Graphics Proceedings, Annual Conference Series, pages 107–116, August 2001.
- [9] Ravi Ramamoorthi and Pat Hanrahan. A signal-processing framework for inverse rendering. In *Proc. of SIGGRAPH 2001*, Computer Graphics Proceedings, Annual Conference Series, pages 117–128, August 2001. ISBN 1-58113-292-1.
- [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”. In *in Proc. of Eighth IEEE International Conference on Computer Vision ICCV '01*, pages 599–606, July 2001.
- [11] Hendrik P. A. Lensch, Jan Kautz, Michael Goesele, Wolfgang Heidrich, and Hans-Peter Seidel. Image-based reconstruction of spatial appearance and geometric detail. *ACM Transactions on Graphics*, 2:27, April 2003.
- [12] S. Marschner, S. Westin, E. Laforge, K. Torrance, and D. Greenberg. Image-based BRDF Measurement Including Human Skin. In *10th Eurographics Workshop on Rendering*, pages 131–144, June 1999.
- [13] Eric P. F. Laforge, Sing-Choong Foo, Kenneth E. Torrance, and Donald P. Greenberg. Non-linear approximation of reflectance functions. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 117–126. ACM Press/Addison-Wesley Publishing Co., 1997.
- [14] S. Marschner, B. Guenter, and S. Raghupathy. Modeling and Rendering for Realistic Facial Animation. *11th Eurographics Workshop on Rendering*, pages 231–242, June 2000. ISBN 3-211-83535-0.
- [15] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. Acquiring the reflectance field of a human face. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 145–156. ACM Press/Addison-Wesley Publishing Co., 2000.
- [16] K. Torrance and E. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of Optical Society of America*, 57(9):1105–1114, 1967.
- [17] Bui Tuong Phong. Illumination for Computer Generated Pictures. *Communications of the ACM*, 18(6):311–317, June 1975.
- [18] Sylvain Paris, Francois X. Sillion, and Long Quan. Lightweight face relighting. In *Proceedings of the 11th Pacific Conference on Computer Graphics and Applications*, page 41. IEEE Computer Society, 2003.
- [19] Baback Moghaddam, Jinho Lee, Hanspeter Pfister, and Raghu Machiraju. Model-based 3d face capture with shape-from-silhouettes. In *AMFG '03: Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, page 20. IEEE Computer Society, 2003.

- [20] Jinho Lee, Baback Moghaddam, Hanspeter Pfister, and Raghu Machiraju. Finding optimal views for 3d face shape modeling. In *IEEE International Conference on Automatic Face and Gesture Recognition (FG)*, May 2004.
- [21] V. Blanz and T. Vetter. A morphable model for the synthesis of 3D faces. In *Computer Graphics Proc. SIGGRAPH'99*, pages 187–194, Los Angeles, 1999.
- [22] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. *IEEE Trans. on Pattern Analysis and Machine Intell.*, 25(9):1063–1074, 2003.
- [23] Rodrigo L. Carceroni and Kiriakos N. Kutulakos. Multi-view scene capture by surfel sampling: From video streams to non-rigid 3d motion, shape and reflectance. *International Journal of Computer Vision*, 49(2-3):175–214, 2002.
- [24] Athinodoros S. Georghiades. Recovering 3-d shape and reflectance from a small number of photographs. In *Rendering Techniques 2003: 14th Eurographics Workshop on Rendering*, pages 230–240. Eurographics, June 2003.
- [25] Norimichi Tsumura, Nobutoshi Ojima, Kayoko Sato, Mitsuhiro Shiraishi, Hideto Shimizu, Hirohide Nabeshima, Syuuichi Akazaki, Kimihiko Hori, and Yoichi Miyake. Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. *ACM Trans. Graph.*, 22(3):770–779, 2003.
- [26] Aravind Krishnaswamy and Gladimir V. G. Baranoski. A biophysically-based spectral model of light interaction with human skin. *Eurographics*, 2004.
- [27] H. W. Jensen, S. Marschner, M. Levoy, and P. Hanrahan. A Practical Model for Subsurface Light Transport. In *Proceedings SIGGRAPH*, pages 511–518, August 2001.
- [28] Michael Goesele, Hendrik P. A. Lensch, Wolfgang Heidrich, and Hans-Peter Seidel. Building a photo studio for measurement purposes. In *Proceedings of Vision, Modeling, and Visualization (VMV-00)*, pages 231–238, 2000.
- [29] Jean-Yves Bouguet. Camera Calibration Toolbox for Matlab. See <http://www.vision.caltech.edu/bouguetj>.
- [30] T. Vetter and T. Poggio. Linear object classes and image synthesis from a single example image. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(7):733–742, 1997.

- [31] V. Blanz, C. Basso, T. Poggio, and T. Vetter. Reanimating faces in images and video. In P. Brunet and D. Fellner, editors, *Computer Graphics Forum, Vol. 22, No. 3 EUROGRAPHICS 2003*, pages 641–650, Granada, Spain, 2003.
- [32] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge Univ. Press, 2nd edition, 1994.
- [33] Paul Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. pages 189–198, July 1998.
- [34] Robert L. Cook and Kenneth E. Torrance. A reflectance model for computer graphics. *ACM Trans. Graph.*, 1(1):7–24, 1982.
- [35] G. Ward Larson. Measuring and Modeling Anisotropic Reflection. In *Proc. SIGGRAPH*, pages 265–272, July 1992.
- [36] P. Hanrahan and W. Krueger. Reflection from layered surfaces due to subsurface scattering. In *Proc. of SIGGRAPH 93*, Computer Graphics Proceedings, Annual Conference Series, pages 165–174, August 1993.
- [37] Jan Koenderink and Sylvia Pont. The secret of velvety skin. *Mach. Vision Appl.*, 14(4):260–268, 2003.
- [38] Michael Goesele, Hendrik P. A. Lensch, Jochen Lang, Christian Fuchs, and Hans-Peter Seidel. Disco: acquisition of translucent objects. *ACM Trans. Graph.*, 23(3):835–844, 2004.

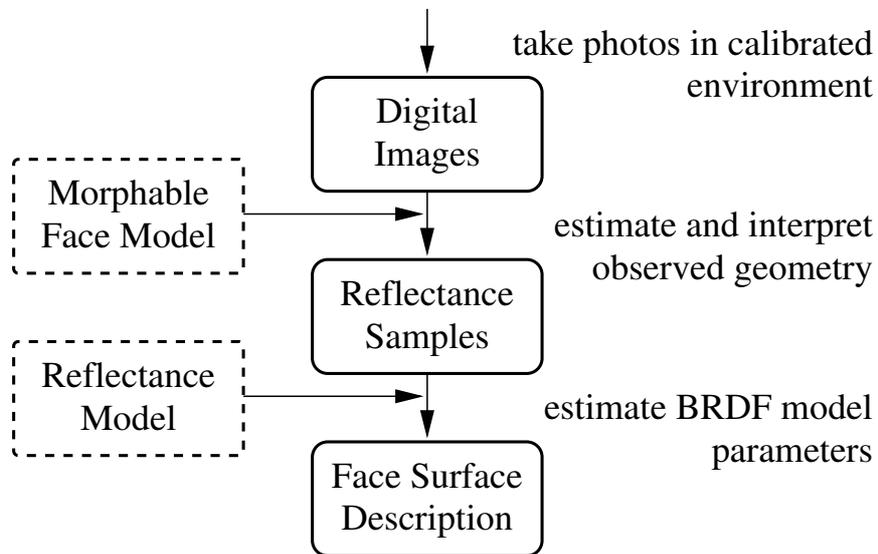


Figure 1: Acquisition pipeline of our model-based approach.



Figure 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.



Figure 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. Measurement 1 and 9 are quite close to the original face, 8 shows a small artefact 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.

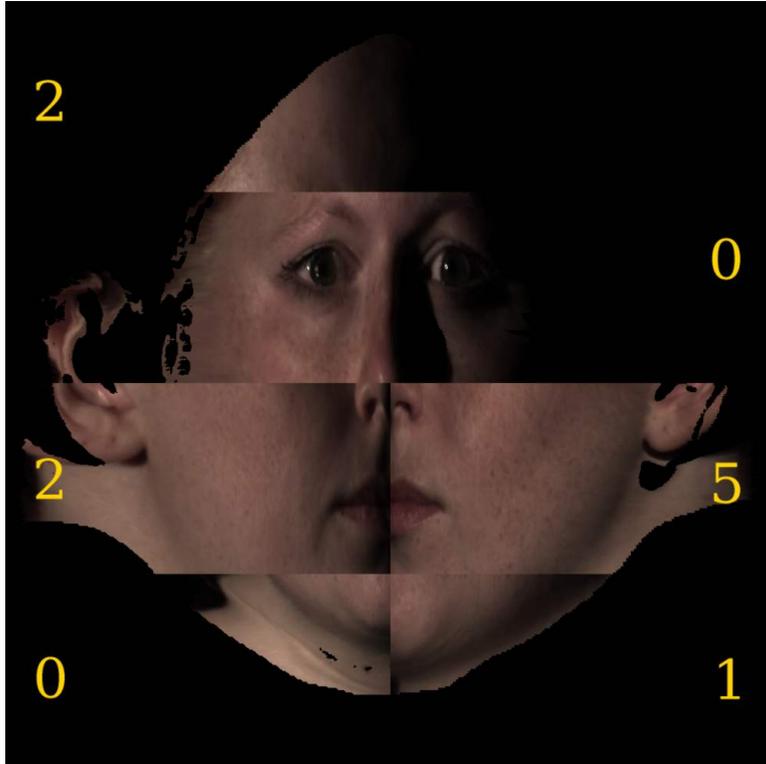


Figure 4: Measured radiance on the face surface in the common surface parameterization for faces, in a combination of different measurements. While the correspondence between the chin region of measurement 0 and 1 is sub-optimal, good correspondence of sharp face features such as lips and nose edges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

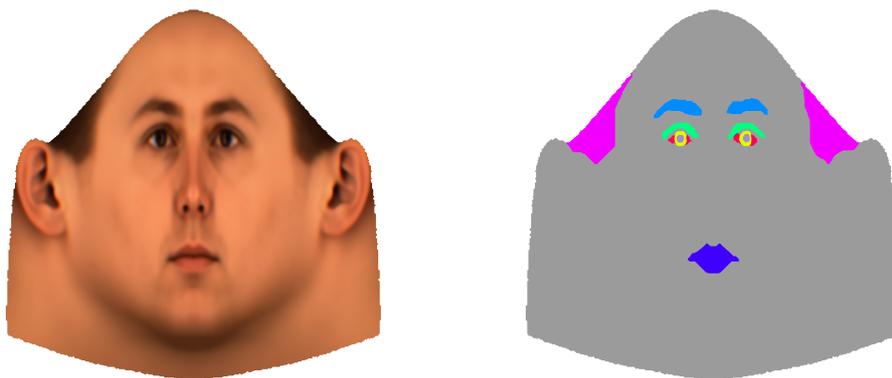


Figure 5: In the parameterization defined by the reference texture (left), manually marked regions of different surface properties (right) apply to all faces automatically.

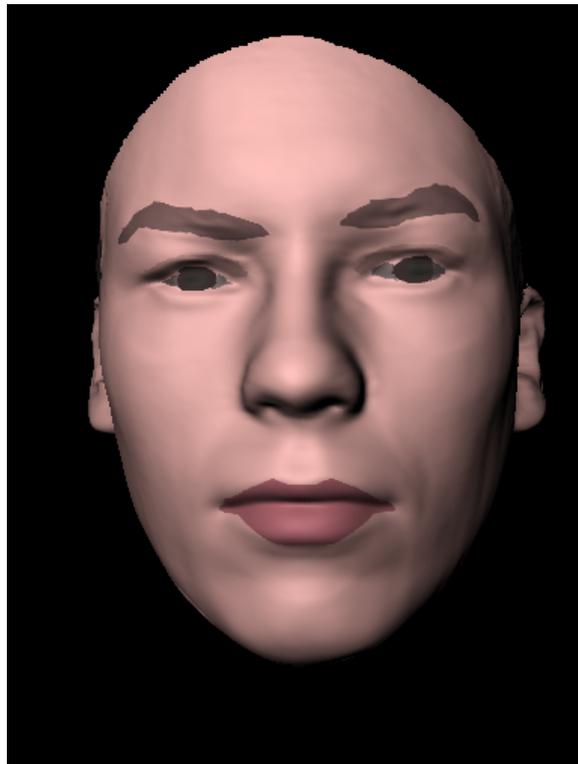


Figure 6: Result of fitting the Cook-Torrance model against the samples in a-priori clusters.



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



Cook-Torrance



Lafortune, isotropic, one white lobe



Phong (no ambient term because of the lighting model)



Ward

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



Original



Cook-Torrance



Lafortune, isotropic, one white lobe



Phong (no ambient term because of lighting model)



Ward

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



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

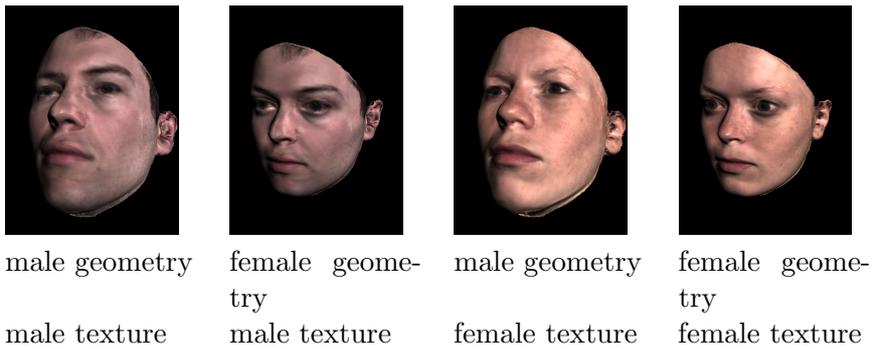


Figure 11: Rendered at novel pose and lighting, the images on the left and right show shapes and spatially varying Cook-Torrance BRDFs of two individuals. The images in the center demonstrate that BRDFs can be transferred between shapes.



Below you find a list of the most recent technical reports of the Max-Planck-Institut für Informatik. They are available by anonymous ftp from [ftp.mpi-sb.mpg.de](ftp://ftp.mpi-sb.mpg.de) under the directory `pub/papers/reports`. Most of the reports are also accessible via WWW using the URL <http://www.mpi-sb.mpg.de>. If you have any questions concerning ftp or WWW access, please contact reports@mpi-sb.mpg.de. Paper copies (which are not necessarily free of charge) can be ordered either by regular mail or by e-mail at the address below.

Max-Planck-Institut für Informatik
Library
attn. Anja Becker
Stuhlsatzenhausweg 85
66123 Saarbrücken
GERMANY
e-mail: library@mpi-sb.mpg.de

MPI-I-2004-NWG3-001	M. Magnor	Axisymmetric Reconstruction and 3D Visualization of Bipolar Planetary Nebulae
MPI-I-2004-NWG1-001	B. Blanchet	Automatic Proof of Strong Secrecy for Security Protocols
MPI-I-2004-5-001	S. Siersdorfer, S. Sizov, G. Weikum	Goal-oriented Methods and Meta Methods for Document Classification and their Parameter Tuning
MPI-I-2004-4-005	I.P. Ivrisimtzis, W.-. Jeong, S. Lee, Y.a. Lee, H.-. Seidel	Neural Meshes: Surface Reconstruction with a Learning Algorithm
MPI-I-2004-4-004	R. Zayer, C. Rssl, H. Seidel	r-Adaptive Parameterization of Surfaces
MPI-I-2004-4-003	Y. Ohtake, A. Belyaev, H. Seidel	3D Scattered Data Interpolation and Approximation with Multilevel Compactly Supported RBFs
MPI-I-2004-4-002	Y. Ohtake, A. Belyaev, H. Seidel	Quadric-Based Mesh Reconstruction from Scattered Data
MPI-I-2004-4-001	J. Haber, C. Schmitt, M. Koster, H. Seidel	Modeling Hair using a Wisp Hair Model
MPI-I-2004-2-002	P. Maier	Intuitionistic LTL and a New Characterization of Safety and Liveness
MPI-I-2004-2-001	H.d. Nivelle, Y. Kazakov	Resolution Decision Procedures for the Guarded Fragment with Transitive Guards
MPI-I-2004-1-007	S. Wagner	?
MPI-I-2004-1-006	L.S. Chandran, N. Sivadasan	On the Hadwiger's Conjecture for Graph Products
MPI-I-2004-1-005	S. Schmitt, L. Fousse	A comparison of polynomial evaluation schemes
MPI-I-2004-1-004	N. Sivadasan, P. Sanders, M. Skutella	Online Scheduling with Bounded Migration
MPI-I-2004-1-003	I. Katriel	On Algorithms for Online Topological Ordering and Sorting
MPI-I-2004-1-002	P. Sanders, S. Pettie	A Simpler Linear Time $2/3 - \epsilon$ Approximation for Maximum Weight Matching
MPI-I-2004-1-001	N. Beldiceanu, I. Katriel, S. Thiel	Filtering algorithms for the Same and UsedBy constraints
MPI-I-2003-NWG2-002	F. Eisenbrand	Fast integer programming in fixed dimension
MPI-I-2003-NWG2-001	L.S. Chandran, C.R. Subramanian	Girth and Treewidth
MPI-I-2003-4-009	N. Zakaria	FaceSketch: An Interface for Sketching and Coloring Cartoon Faces
MPI-I-2003-4-008	C. Roessl, I. Ivrisimtzis, H. Seidel	Tree-based triangle mesh connectivity encoding

MPI-I-2003-4-007	I. Ivriissimtzis, W. Jeong, H. Seidel	Neural Meshes: Statistical Learning Methods in Surface Reconstruction
MPI-I-2003-4-006	C. Roessl, F. Zeilfelder, G. Nrnberger, H. Seidel	Visualization of Volume Data with Quadratic Super Splines
MPI-I-2003-4-005	T. Hangelbroek, G. Nrnberger, C. Roessl, H.S. Seidel, F. Zeilfelder	The Dimension of C^1 Splines of Arbitrary Degree on a Tetrahedral Partition
MPI-I-2003-4-004	P. Bekaert, P. Slusallek, R. Cools, V. Havran, H. Seidel	A custom designed density estimation method for light transport
MPI-I-2003-4-003	R. Zayer, C. Roessl, H. Seidel	Convex Boundary Angle Based Flattening
MPI-I-2003-4-002	C. Theobalt, M. Li, M. Magnor, H. Seidel	A Flexible and Versatile Studio for Synchronized Multi-view Video Recording
MPI-I-2003-4-001	M. Tarini, H.P.A. Lensch, M. Goesele, H. Seidel	3D Acquisition of Mirroring Objects
MPI-I-2003-2-004	A. Podelski, A. Rybalchenko	Software Model Checking of Liveness Properties via Transition Invariants
MPI-I-2003-2-003	Y. Kazakov, H. Nivelle	Subsumption of concepts in $DL \mathcal{FL}_0$ for (cyclic) terminologies with respect to descriptive semantics is PSPACE-complete
MPI-I-2003-2-002	M. Jaeger	A Representation Theorem and Applications to Measure Selection and Noninformative Priors
MPI-I-2003-2-001	P. Maier	Compositional Circular Assume-Guarantee Rules Cannot Be Sound And Complete
MPI-I-2003-1-018	G. Schaefer	A Note on the Smoothed Complexity of the Single-Source Shortest Path Problem
MPI-I-2003-1-017	G. Schfer, S. Leonardi	Cross-Monotonic Cost Sharing Methods for Connected Facility Location Games
MPI-I-2003-1-016	G. Schfer, N. Sivadasan	Topology Matters: Smoothed Competitive Analysis of Metrical Task Systems
MPI-I-2003-1-015	A. Kovcs	Sum-Multicoloring on Paths
MPI-I-2003-1-014	G. Schfer, L. Becchetti, S. Leonardi, A. Marchetti-Spaccamela, T. Vredeveld	Average Case and Smoothed Competitive Analysis of the Multi-Level Feedback Algorithm
MPI-I-2003-1-013	I. Katriel, S. Thiel	Fast Bound Consistency for the Global Cardinality Constraint
MPI-I-2003-1-012		- not published -
MPI-I-2003-1-011	P. Krysta, A. Czumaj, B. Voeking	Selfish Traffic Allocation for Server Farms
MPI-I-2003-1-010	H. Tamaki	A linear time heuristic for the branch-decomposition of planar graphs
MPI-I-2003-1-009	B. Csaba	On the Bollobás – Eldridge conjecture for bipartite graphs
MPI-I-2003-1-008	P. Sanders	Polynomial Time Algorithms for Network Information Flow
MPI-I-2003-1-007	H. Tamaki	Alternating cycles contribution: a strategy of tour-merging for the traveling salesman problem
MPI-I-2003-1-006	M. Dietzfelbinger, H. Tamaki	On the probability of rendezvous in graphs
MPI-I-2003-1-005	M. Dietzfelbinger, P. Woelfel	Almost Random Graphs with Simple Hash Functions
MPI-I-2003-1-004	E. Althaus, T. Polzin, S.V. Daneshmand	Improving Linear Programming Approaches for the Steiner Tree Problem
MPI-I-2003-1-003	R. Beier, B. Vcking	Random Knapsack in Expected Polynomial Time
MPI-I-2003-1-002	P. Krysta, P. Sanders, B. Vcking	Scheduling and Traffic Allocation for Tasks with Bounded Splittability
MPI-I-2003-1-001	P. Sanders, R. Dementiev	Asynchronous Parallel Disk Sorting
MPI-I-2002-4-002	F. Drago, W. Martens, K. Myszkowski, H. Seidel	Perceptual Evaluation of Tone Mapping Operators with Regard to Similarity and Preference

MPI-I-2002-4-001	M. Goesele, J. Kautz, J. Lang, H.P.A. Lensch, H. Seidel	Tutorial Notes ACM SM 02 A Framework for the Acquisition, Processing and Interactive Display of High Quality 3D Models
MPI-I-2002-2-008	W. Charatonik, J. Talbot	Atomic Set Constraints with Projection
MPI-I-2002-2-007	W. Charatonik, H. Ganzinger	Symposium on the Effectiveness of Logic in Computer Science in Honour of Moshe Vardi
MPI-I-2002-1-008	P. Sanders, J.L. Triff	The Factor Algorithm for All-to-all Communication on Clusters of SMP Nodes