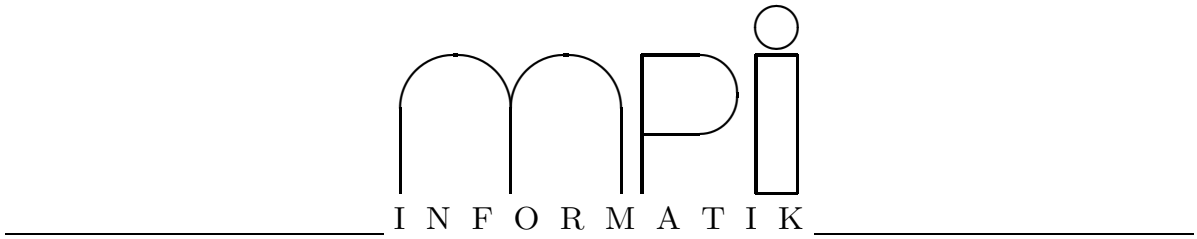


Joint Motion and Reflectance Capture for Creating Relightable 3D Videos

Christian Theobalt Naveed Ahmed
Edilson de Aguiar Gernot Ziegler
Hendrik Lensch Marcus A. Magnor
Hans-Peter Seidel

MPI–I–2005–4–004 April 2005

FORSCHUNGSBERICHT RESEARCH REPORT

MAX-PLANCK-INSTITUT
FÜR
INFORMATIK

Stuhlsatzenhausweg 85 66123 Saarbrücken Germany

Authors' Addresses

Christian Theobalt, Naveed Ahmed, Edilson de Aguiar, Gernot Ziegler, Marcus A. Magnor, Hans-Peter Seidel
Max-Planck-Institut für Informatik
Stuhlsatzenhausweg 85
66123 Saarbrücken
{theobalt,nahmed,edeaguia,gziegler,magnor,
hpseidel}@mpi-sb.mpg.de

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

Abstract

Passive optical motion capture is able to provide authentically animated, photo-realistically and view-dependently textured models of real people. To import real-world characters into virtual environments, however, also surface reflectance properties must be known. We describe a video-based modeling approach that captures human motion as well as reflectance characteristics from a handful of synchronized video recordings. The presented method is able to recover spatially varying reflectance properties of clothes by exploiting the time-varying orientation of each surface point with respect to camera and light direction. The resulting model description enables us to match animated subject appearance to different lighting conditions, as well as to interchange surface attributes among different people, e.g. for virtual dressing. Our contribution allows creating realistic 3D renditions of real-world people under arbitrary novel lighting conditions on standard graphics hardware.

Keywords

Refelctance Estimation, BRDF, Dynamic Reflectometry, Multi-View Video, Free-Viewpoint Video, Optical Human Motion Capture, Acquisition, Scene Digitization

Seeing People in Different Light: Joint Motion and Reflectance Capture

Abstract

Passive optical motion capture is able to provide authentically animated, photo-realistically and view-dependently textured models of real people. To import real-world characters into virtual environments, however, also surface reflectance properties must be known. We describe a video-based modeling approach that captures human motion as well as reflectance characteristics from a handful of synchronized video recordings. The presented method is able to recover spatially varying reflectance properties of clothes by exploiting the time-varying orientation of each surface point with respect to camera and light direction. The resulting model description enables us to match animated subject appearance to different lighting conditions, as well as to interchange surface attributes among different people, e.g. for virtual dressing. Our contribution allows creating realistic 3D renditions of real-world people under arbitrary novel lighting conditions on standard graphics hardware.

Categories and Subject Descriptors (according to ACM CCS): I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture, I.4.8 [Image Processing and Computer Vision]: Scene Analysis, I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism

1. Introduction

Rendering realistic images in real-time is solved, or at least that may be the impression after having sampled the latest generation of computer games. Higher and higher hardware polygon throughput, steadily increasing fill-rates, more and more programming flexibility in the rendering pipeline, and state-of-the-art rendering algorithms enable creating images of unprecedented realism at interactive frame rates. To convert ongoing advances in graphics hardware and software into attainable visual realism, however, ever more detailed and accurate scene descriptions must be available. The price to pay can be measured in working hours spent to create detailed geometry meshes, complex textures, convincing shaders, and authentic animations: Apparently, scene modeling is becoming a limiting factor in realistic rendering.

One alternative to avoid excessive modeling times consists of capturing suitable models directly from "the real thing". Image- and video-based rendering (IBR/VBR) approaches pursue this notion, aiming at automatically generating visually authentic computer models from real world-recorded objects and events [KRN97]. In recent years, a number of IBR/VBR researchers have demonstrated how to successfully recover approximate geometry [MBM01], animation sequences [CTMS03], and view-dependent textures [DTM96] from multi-image or multi-video data. These (and many more) techniques show how to interactively render photo-realistic views from real world-captured, dynamic scenes [MP04, ZKU*04].

While the ability to realistically display dynamic events from novel viewpoints has already a number of intriguing applications, the next step is to use such real world-captured objects for augmenting virtual scenes. To import a real-world object into surroundings different from the recording environment, however, its appearance must be adapted to the new illumination situation. To do so, the bi-directional reflectance distribution function (BRDF) must be known for all object surface points. Data-driven [DHT*00, MPBM03] as well as model-based [Mar98, LKG*03] methods have been proposed to recover and represent the BRDF of real-world materials. Unfortunately, these methods cannot be directly applied to dynamic objects exhibiting time-varying surface geometry and constantly changing local illumination.

We present a joint motion capture and BRDF reconstruction approach to acquire animation parameters and surface reflectance properties of people. As input to our algorithm we require only a handful of calibrated and synchronized video recordings. The algorithm automatically returns

subject-adapted 3D geometry, animation parameters, diffuse texture, per-texel BRDF model parameter values, as well as time-varying surface normals. PC graphics hardware-assisted rendering then allows us to photo-realistically visualize recorded people at interactive frame rates in changing lighting conditions and from arbitrary perspective. We present results for several subjects wearing different clothes made of non-lambertian fabrics.

The following Section highlights related work on free-viewpoint video and BRDF estimation. An overview of our approach is provided in Sect. 3. Input data acquisition issues are considered in Sect. 4, before model-based motion capture and multi-view texture generation is outlined in Sect. 5. The heart of the paper is presented in Sect. 6, where we describe in detail our BRDF estimation and time-varying normal map reconstruction algorithm. The employed real-time rendering technique on off-the-shelf graphics hardware is detailed in Sect. 7. Results for different people wearing different apparel follow in Sect. 8, before we conclude with Sect. 9.

2. Related Work

In our research we pick up and extend ideas that have been published in the fields of marker-less human motion capture, image-based reflectance estimation and free-viewpoint video. As the nuts and bolts of motion estimation are not the focus of this work, for the sake of brevity, we refer the interested reader to overview articles on human motion estimation [MG01]. In the following, we review relevant work from the other two fields.

2.1. Free-Viewpoint Video

Research in free-viewpoint video aims at developing methods for photo-realistic, real-time rendering of previously captured real-world scenes. The goal is to give the user the freedom to interactively navigate his or her viewpoint freely through the rendered scene. Early research that paved the trail for free-viewpoint video was presented in the field of image-based rendering (IBR). Shape-from-silhouette methods reconstruct geometry models of a scene from multi-view silhouette images or video streams. Examples are image-based [MBR*00, WLSG02] or polyhedral visual hull [MT02] methods, as well as approaches performing voxel-based [MTG97] or point-based [GWN*03] reconstruction. The combination of stereo reconstruction with visual hull rendering leads to a more faithful reconstruction of surface concavities [LSMS02]. Stereo methods have also been applied to reconstruct and render dynamic scenes [ZKU*04, KRN97]. Alternatively, a complete parameterized geometry model can be used to pursue a model-based approach towards free-viewpoint video [CTMS03]. On the other hand, light field rendering [LH96] is employed in the 3D TV system [MP04] to enable simultaneous scene acquisition and rendering in real-time.

IBR methods can visualize a recorded scene only for the same illumination conditions that it was captured in. For correct relighting, it is inevitable to recover complete surface reflectance characteristics.

2.2. Image-based Reflectance Estimation

The estimation of reflection properties from still images has been addressed in many different ways. Typically, a single point light source is used to illuminate an object of known 3D geometry. One common approach is to take HDR images of a curved object, yielding a different incident and outgoing directions per pixel and thus capturing a vast number of reflectance samples in parallel. Quite often the parameters of an analytic BRDF model are fit to the measured data [SWI97, LKG*03] or a data-driven model [MPBM03] is used. Reflectance measurements of scenes with more complex incident illumination can be derived by either a full-blown inverse global illumination approach [YDMH99, GHH01, BG01] or by representing the incident light field as an environment map and solving for the direct illumination component only [YM98, RH01, NSI01]. In our approach we will approximate the incident illumination by multiple point light sources and estimate BRDF model parameters taking only direct illumination into account.

Reflection properties together with measured photometric data can also be used to derive geometric information of the original object [ZTCS99]. Rushmeier et al. estimate diffuse albedo and normal map from photographs with varied incident light directions [RTG97, BMR01]. A linear light source is employed by Gardner et al. [GTHD03] to estimate BRDF properties and surface normal. In [Geo03, GCHS04], reflectance and shape of static scenes are simultaneously refined using a single light source in each photograph.

Instead of explicitly reconstructing a mathematical reflectance model it has also been tried to take an image-based approach to relighting. In [HWT*04] a method to generate animatable and relightable face models from images taken with a special light stage is described. Using deformable geometry, the face is rendered under novel illumination by reconstructing from a large database of images that show the face under different incident illumination directions, different viewing directions and with different expressions. For our 3D video scenario, we prefer a more compact scene description based on parametric BRDFs that can be reconstructed in a fairly simple acquisition facility.

Carceroni and Kutulakos present a volumetric method for simultaneous motion and reflectance capture for non-rigid objects [CK01]. In contrast to their work, we propose a model-based approach that captures shape, motion parameters and reflectance of the whole human body at high accuracy.

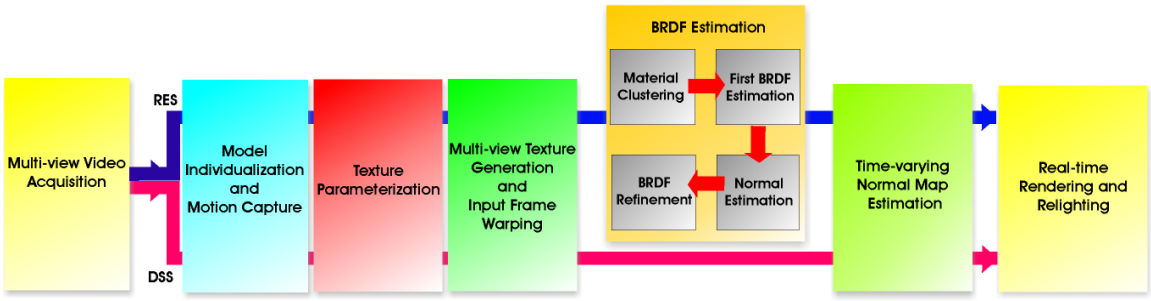


Figure 1: *Algorithmic workflow of our method.*

3. Overview

Fig. 1 illustrates the workflow between the components that make up our joint motion and reflectance capture approach. The input to our system consists of multi-view video sequences that are recorded using eight synchronized color video cameras (Sect. 4). The reflectance estimation sequence (RES) is used to estimate surface reflectance properties. Arbitrary human motion is captured in the dynamic scene sequences (DSS), and these sequences are later visualized and relit. In both types of sequences, the person wears identical clothes. The respective data paths for both input sequences are shown in Fig. 1. A generic body model is adapted to match the shape and proportions of the recorded person. Subsequently, human pose parameters are computed for all time frames by means of a silhouette-based marker-free motion capture approach (Sect. 5.1). To store all per-surface element data needed during reflectance estimation in texture space, we make use of a texture atlas as surface parameterization of the body model. Multi-view video (MVV) textures are generated by transforming each input video image into the texture domain (Sect. 5.2). To correct for photo-inconsistencies due to inexact body geometry, the input images can be warp-corrected prior to MVV texture generation (Sect. 5.3). From the RES video data, BRDF model parameter values are estimated for each surface element (texel) of the geometry model individually (Sect. 6.1). The recovered local reflectance properties then allow us to estimate the time-varying surface normal field in the DSS sequences (Sect. 6.2). The moving body model, its spatially-varying reflectance, and the time-varying normal field enable us to interactively render and instantaneously relight the DSS sequences from arbitrary viewpoint and illumination direction. (Sect. 7 and Sect. 8). Contributions of our paper are:

- An algorithm to warp-correct input video images in order to guarantee multi-view photo-consistency in conjunction with inexact object geometry,
- dynamic reflectometry, i.e., per-texel, per-time step BRDF parameter estimation from multi-view video footage,
- reconstruction of time-varying normal maps to capture small, variable detail of surface geometry (e.g., wrinkles in clothing), and
- the integration of the recording facilities, the motion capture method, the reflectance estimation aproach and the renderer into one working system.

4. Acquisition

As input to our system, we record multi-view video (MVV) sequences. A multi-view recording setup enables us to capture an area of approx. $4 \times 4 \times 3$ m with eight externally synchronized video cameras. The cameras are placed in a semi-circular arrangement around the center of the stage. Since we estimate both motion and reflectance properties, we have strict requirements concerning the spatial, temporal, and color resolution of our imaging devices. Only recently, suitable production-line video cameras have become available that meet our requirements. We employ Imperx^{TM}MDC-1004 cameras that feature a 1004x1004 CCD sensor with linear 12 bits-per-pixel resolution. The CCD features a Bayer mosaic to record red, green, and blue channel information. With two on-board processors, the cameras deliver 48 fps at full resolution. Compromising between high speed motion capture and data volume, the sequences used throughout the paper are recorded at 25 fps. The video data is captured in parallel by eight frame grabber cards to be streamed in real-time to a RAID system consisting of eight hard drives. The cameras are calibrated, and radial and tangential lens distortions are corrected up to second order. Color-consistency across cameras is ensured by applying a color-space transformation to each camera stream. This transformation has been derived from recordings of a reference color pattern.

The lighting conditions in our studio are fully controllable. No exterior light can enter the recording area, and the influence of indirect illumination is minimized by covering up all the walls and the studio floor with black cloth and carpet. Two different lighting setups are used. Lighting setup 1 (LS1) illuminates the scene with only one K5600^{TM}Jokerbug 400 spot light. In lighting setup 2 (LS2), additional light sources on the ceiling are used in order to illuminate the set more evenly. In our simulations we approximate the contribution of the single spot light with one point light source and the illumination from the ceiling light with two additional point light sources, Fig.2. Light source

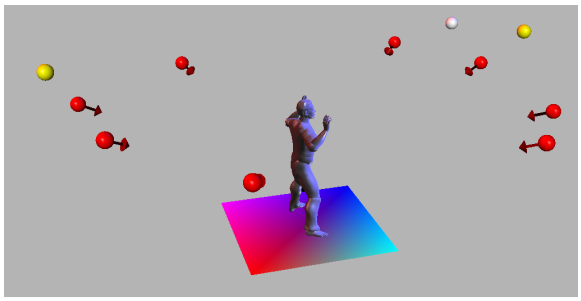


Figure 2: *Illustration of camera arrangement (red), lighting setup 1 (white) and lighting setup 2 (white+yellow).*

positions, intensities and color response of the cameras are calibrated off-line.

We successively record three MVV sequences for each person and each type of apparel. A short sequence of the scene background recorded with illumination setup LS2 later facilitates color-based background subtraction of the motion sequences. The second sequence, referred to as the reflectance estimation sequence (RES), serves as input to the BRDF estimation algorithm. While BRDF parameter value estimation works best if the scene is illuminated by only one light source (LS1), robust motion capture is practically impossible if large parts of the subject are in shadow. To resolve the conflict, the RES is acquired in single-shot mode. The person strikes an initialization pose and turns between shots by approximately 5° until having completed a full 360°-circle. At each orientation step, a set of eight images is captured for lighting setup 1, and a second set of images is recorded for setup 2. The first set is used for BRDF estimation, the second set for recovering body pose. Prior to reflectance estimation we fit our geometry model to each body pose in the RES (Sect. 5.1). For each point on the model's surface, the RES contains as many different appearance samples as there are images depicting the repective point. Over time, the surface element normal points in various directions, and we obtain a large number of reflection samples for our large-scale moving object. While surface normal orientation varies freely, our static camera and lighting setup allows for only a limited number of half vector directions $\vec{h} = \hat{l} \cdot \hat{v_j}/2$, i.e., angular separations between spot light $\hat{l}$ and camera directions $\hat{v_j}$. By placing the cameras non-symmetrically with respect to the spot light, we gather samples for up to eight different light-to-camera angles, which we found sufficient to robustly fit our isotropic BRDF models (Sect. 6).

Finally, the dynamic scene sequences (DSS) capture the motion sequences from which the actual relightable 3D videos are generated. The scene is now illuminated using lighting setup 2. From the DSS we also reconstruct a time-varying surface normal field (Sect. 6.2).

5. Motion Capture and Texture Generation

We use a model-based approach to represent the time-varying geometry in a 3D video. To estimate the geometry model and merge the video data from multiple viewpoints we perform model-based motion capture, precompute a static texture parameterization, and resample and align the input streams using a novel warp-correction technique.

5.1. Model-based Motion Capture

To simultaneously recover dynamic 3D surface geometry as well as animation parameters, we make use of the model-based, passive optical motion capture approach proposed in [CTMS03]. The method relies on a parameterized generic body model that is adapted to the proportions and pose of the recorded individual. The model consists of sixteen individual body segments. The surface of each segment is modeled as a closed triangle mesh. Each segment can be uniformly scaled and its surface geometry deformed. The segments are connected by 17 joints which define the hierarchical skeleton structure. 35 translation and joint rotation parameters then specify any body pose.

To match the rendered model to all video image silhouettes at one time step, model parameters are automatically adapted by a non-linear optimization scheme. As error measure, the overlap between rendered model and image silhouettes is evaluated by employing the pixel-wise XOR operation on graphics hardware and summing up the resulting set pixels. The same criterion is employed to customize the model such that it matches its real world counterpart in shape and proportions.

5.2. Texture Parameterization

As surface parameterization, each body segment is parameterized separately over a planar rectangular domain using patches of minimal distortion. The sixteen planar patch layouts are finally assembled into one texture atlas for the complete model. This way, we obtain a pose-independent bijective 3D-to-2D mapping between a surface element and a texel in the texture domain. All data related to surface elements (normals, light vectors, visibility etc.) can now be conveniently stored as textures. Throughout our experiments, we use 1024x1024-texel texture maps.

The graphics hardware is used to transform each video camera image into the texture domain. For each video time step, eight so-called multi-view video textures (MVV textures) are created.

5.3. Warp Correction

Although the body model initialization procedure yields a faithful representation of the person's true geometry, small

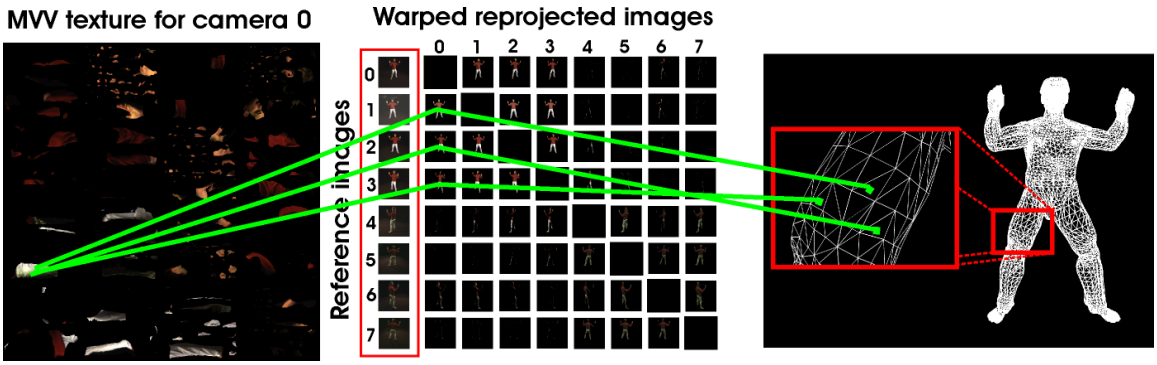


Figure 3: *MVV texture generation for camera 0: The color information for each surface point on the body model is not looked up in the original input video frame recorded from camera 0. Instead the texel color is taken from an image that has been obtained by reprojecting the model that has been textured with camera image 0 into the camera view that sees the surface point most head-on.*

inaccuracies between the real human and its digital counterpart are inevitable. Due to these geometry inaccuracies, pixels from different input views may get mapped to the same texel position in different MVV textures, even though they do not correspond to the same surface element of the true body geometry.

One common strategy to enhance model-to-texture consistency is to deform the geometry until an overall photo-consistency measure is maximized. Geometry deformation-based optimization, however, tends to give unstable results, e.g., due to sudden visibility changes. We take an alternative approach. Instead of moving surface elements to their correct locations in 3D, we move the image pixels within the 2D input image planes until they all become photo-consistent given the available geometry. The following example illustrates our modified MVV texture generation scheme, Fig.3:

Let's assume we want to assemble an MVV texture from the video image $I_y(t)$ seen by camera Y at time t. For texel K in the MVV texture, we find out which camera sees it best by searching for the minimal deviation between camera viewing vectors and the surface element normal. If the camera that sees the surface point best is Y, the texel color is taken from $I_Y(t)$. In case camera $X \neq Y$ sees the point best and it is not occluded, we regard the video image $I_x(t)$ as the reference image. The model at time t is projectively textured with $I_Y(t)$ and rendered into camera view X. The image of the reprojected textured model is warped such that it is optimally aligned with the reference image. The color of K is taken from the warped image. This way, all texel color values stem from the same physical camera image. The texel color, however, is always taken from a version of that camera image that has been brought into optimal registration with the camera view that sees the corresponding surface element most head-on.

Warped images are precomputed for all possible combinations of X and Y. In our case this corresponds to 56 warping computations for each time step (Fig. 3). To establish per-pixel correspondences, the warping operation itself is based on the optical flow [LK81] between the reference image and the image of the reprojected textured model, Fig.4. A regular 2D triangle mesh is superimposed on the reprojected model image, and per-vertex displacements are derived from the optical flow values in the immediate neighborhood of each vertex. The triangle mesh is deformed to adapt to these displacements by means of a thin-plate spline interpolation [Far99]. The new mesh configuration minimizes the thin-plate bending energy in a least-squares sense. Finally, the warped image is created on the GPU by rendering the textured deformed mesh into a floating point buffer.

The warping-based MVV texture assembly is an optional step that is activated if geometry inaccuracies are apparent. Optical flow is based on the assumption that all surfaces in the scene are diffuse. For reflectance estimation, though, we deliberately generate specular highlights in the images. Our experiments show that the method nonetheless produces good results since for the majority of garments the diffuse reflectance is predominant.

6. Dynamic Reflectometry

Our reflectance estimation approach consists of two steps. In the first step we determine BRDF parameter values per pixel from the reflectance estimation sequence. An iterative estimation process enables us to handle geometry inconsistencies between the real object and the much smoother human body model. In the second step we compute even time-varying normal maps per frame to capture surface detail such as wrinkles in clothing whose shape and extend depend on the current pose of the person. The underlying technique is similar to [LKG*03] which we have extended in order to cope with multiple light sources, time-varying data, and inter-frame consistency.

6.1. BRDF Estimation

We estimate a set of spatially-varying BRDFs for each person and each outfit from the respective reflectance estimation sequence (RES) explained in Sect. 4. The pose parameters for the RES have been determined beforehand. The goal is to estimate a separate parametric reflectance model for each surface element that is able to faithfully reproduce the appearance in each camera view and at each time step of the multi-video sequence. For each surface element, the BRDF representation consists of an individual diffuse color component that is specific to the surface point, and a set of specular parameters that are shared by all surface points belonging to the same material. Our framework is flexible enough to incorporate any parametric reflectance model. However, in the majority of our experiments we employ the parametric BRDF model proposed by Phong [Pho75]. We have also tested our method with the model proposed by Lafortune [LFTG97], using two specular lobes.

In general, our estimation of BRDF parameters and later the estimation of the time-varying normals is based on minimizing for each surface point $\vec{x}$ the error $E(\vec{x},\rho(\vec{x}))$ between the current model $\rho(\vec{x})$ and the measurements for this point from all cameras c at all time steps t:

$$E(\vec{x},\rho(\vec{x})) = \sum_{t}^{N}\sum_{c}^{8}\kappa_c(t)\Big(S_c(t) - [\sum_{j}^{J}\lambda_j(t)(f_r(\hat{l}(t),\hat{v}_c(t),\rho(\vec{x})) \cdot I_j(\hat{n}(t)\cdot\hat{l}(t)))]\Big)^2, \quad (1)$$

The term is evaluated separately in the red, green and blue color channel. $S_c(t)$ denotes the measured color samples at $\vec{x}$ from camera c, and I_j denotes the intensity of light source j. For BRDF estimation the number of light sources equals one (lighting setup 1). More light sources are used when the same energy functional is employed during time-varying normal estimation (Sect. 6.2). The viewing directions $\hat{v}_c(t)$ and light source directions $\hat{l}_j(t)$ are expressed in the point's local coordinate frame based on the surface normal $\hat{n}(t)$. Visibility of the surface point with respect to each camera is given by $\kappa_c(t)$ and with respect to the light sources by λ_j, both being either 0 or 1. f_r finally evaluates the BRDF. All information that is relevant for one texel thus can be grouped into an implicit data structure we called *dynamic texel* or *dyxel*:

$$Dyx(\vec{x},t) = [S_1(t),\ldots,S_8(t),\hat{v}_1(t),\ldots,\hat{v}_8(t), \hat{n}(t),\hat{l}(t),\kappa_1(t),\ldots,\kappa_8(t),\lambda_1(t),\ldots,\lambda_J(t)].$$

Using a non-linear optimization this formula in principle could be used to determine a full BRDF and the surface normal at the same time. However, we applied an iterative approach and carefully designed the reflectance estimation sequence to obtain a much more stable optimization. For example we use only a single light source during the RES. The subsequent steps of our iterative BRDF estimation scheme are *material clustering, first BRDF estimation, normal estimation* and *refined BRDF estimation* as depicted in Fig. 5.

Instead of determining the specular part of the BRDF per pixel we assume that there is only very little variation of the specular part within the same material, e.g. skin, hair or the different fabrics. By combining the measurements of multiple surface points exhibiting the same material we increase the number of samples and more importantly the variation in viewing and lighting directions in order to obtain a more faithful specular estimate. The *clustering* step determines to what material a surface element, i.e., each texel in the texture atlas, belongs. The number of materials is determined a priori. We employ a straightforward color-based clustering approach that considers the raw texel color values. The clustering output is a material texture map in which each texel is assigned a material label, Fig. 5.

During the *first BRDF estimation*, an optimal set of per-texel BRDF parameters is determined while the normals are taken from the default geometry. The estimation itself consists of a non-linear minimization of Eq. 1 in the BRDF parameters. For optimization, we make use of a Levenberg-Marquardt minimization scheme [PTVF02] in the same manner as [LKG*03]. Firtst, we find an optimal set of parameter values for each material cluster of texels. To quantify the estimation error per material cluster, we sum the error term in Eq.1 for all surface elements that belong to the cluster. Given the average BRDF for each material, we can render the model by applying only average specular reflectances. By subtracting this specular component from each sample, we generate new dyxels that contain purely diffuse reflectance samples. Using these purely diffuse sam-

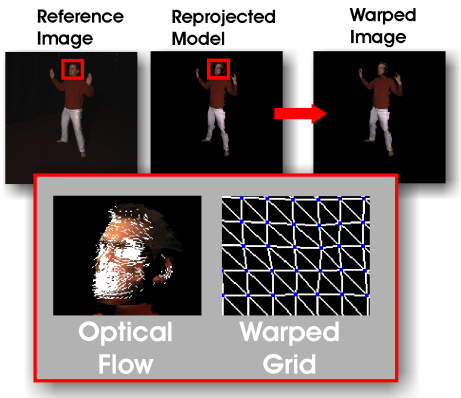


Figure 4: *Warp-correction via texture re-projection to achieve photo-consistency among all camera images. The inlay shows optical flow and warped mesh for the marked face region.*

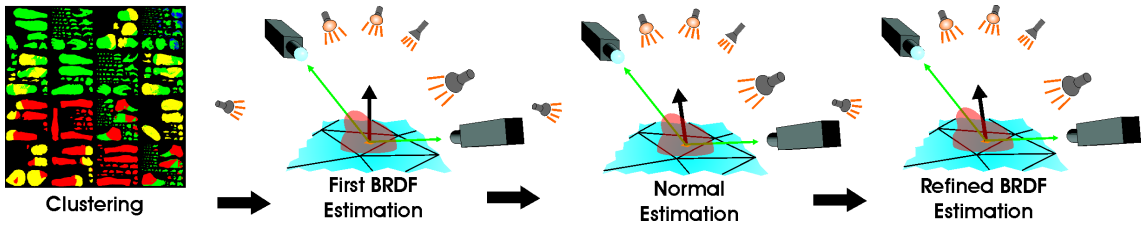


Figure 5: *Subsequent steps to derive per-texel BRDF.*

ples, an individual diffuse component is estimated for each surface element (texel) by minimizing Eq.1 over the diffuse color parameter. The output of the first BRDF estimation is then a set of spatially-varying BRDF parameters ρ_{first}.

The default normals of the human body model cannot represent subtle details in surface geometry, such as wrinkles in clothing. In a *normal estimation* step, we make use of the first set of estimated BRDF parameters ρ_{first} in order to reconstruct a refined normal field via photometric stereo. In order to make this reconstruction tractable, we implicitly assume that the local normal directions do not change while the person is rotating in place. We found that normal estimation robustness is improved if the error function (Eq. 1) is extended into

$$E_{normal}(\vec{x},\rho(\vec{x})) = \alpha E(\vec{x},\rho(\vec{x})) + \beta\Delta(\hat{n})^{\gamma}. \qquad (2)$$

The additional term $\Delta(\hat{n})$ penalizes angular deviation from the default normal of the body model. The terms α and β are weighting factors summing to one, the exponent γ controls the penalty impact. Appropriate values are found through experiments. Normal estimation robustness is further improved if only those color samples in a dyxel are used that come from the two best camera views. For each texel, the modified error function is now minimized by varying the local normal direction $\hat{n}$.

The refined normal field is used for a *second BRDF estimation*. The same computations as for the first BRDF estimation, ρ_{first}, are performed, but now with the more accurate normal field. By this means, we obtain the final set of per-texel BRDF parameters ρ_{final}.

The results are stored in parameter texture maps. For the Phong model, we obtain one texture map containing the per-texel diffuse component, and two texture maps that store the per-material specular colors and exponents. In case of the Lafortune model, the number of specular parameter maps depends on the number of specular lobes.

6.2. Time-varying Normal Map Estimation

The BRDF reconstructed in the previous step enables us to relight any dynamic scene in which the person wears the same apparel as in the respective RES. To generate a visually compelling rendition, however, we found that we need not only accurate reflectance, but also a representation of the small surface geometry details that appear and disappear while a person is moving. We are able to capture these geometry details by estimating a time-varying surface normal field for each DSS via photometric stereo.

Motion parameters for the DSS are found by means of our silhouette-based tracking approach (Sect. 5). The video frames show the scene illuminated by lighting setup 2. During the estimation, we approximate the incident illumination with three point light sources,

The Time-varying normal direction is estimated for each surface point individually. We assume that the transverse motion of the cloth on the body is negligible, and, in consequence, that over time an MVV texel always corresponds to the exact same cloth surface point. The estimation procedure is a non-linear minimization of the regularized energy function, Eq. 2, in the normal direction. During optimization, the BRDF parameters for each surface element are taken from the parameter textures estimated from the corresponding RES.

In order to robustly perform photometric stereo and to minimize the influence of measurement noise, a sufficient number of samples has to be collected for each surface point. To serve this purpose, we assume that changes in local normal direction within a short window in time can be neglected. This way, all samples for a surface point that are taken from a chunk of subsequent time steps in the input footage can be applied to infer a single normal direction. The input sequence of length N is therefore split into C subsequent chunks of odd length d, the last chunk being allowed a different length. Typically, the chunk size is $d = 5$ time steps. For every point $\vec{x}$ on the body surface we fit an optimal normal $\hat{n}$ to each chunk of video individually. After the time-varying normals have been estimated at this coarse scale, the normal directions between subsequent chunks are interpolated via spherical linear interpolation, Fig.6.

This way, a normal field is generated that represents a compromise between smoothness in the temporal domain and local normal accuracy. It faithfully models subtle details in surface structure, and it exhibits no normal discontinuities at chunk boundaries that would appear as flickering in the final renditions of the 3D video. The results we obtain with this approach confirm that it is permissible to assume that during a sufficiently small time period the local normal direction does not change dramatically. A comparison to the

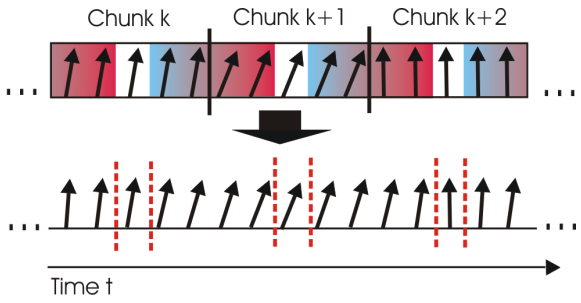


Figure 6: *2D illustration of robust time-varying normal map estimation. Top: The sequence subdivided into short chunks. For each chunk, one best-matching normal is derived per texel which is assigned to the chunk's center time step (white). Intermediate time steps are interpolated (Bottom).*

video footage shows that we are able to capture even the subtle wrinkles that are due to limb bending, Fig. 10a,b,c.

7. Rendering

The outcome of our approach is a relightable dynamic object description that consists of the animated geometry and the material properties.

The geometry is comprised of the 3D body model mesh, the underlying skeleton and the joints' motion parameters.

The material properties consist of the time-independent BRDF textures and the dynamic normal maps. The number of BRDF data parameters depends on the employed reflectance model. In the case of Phong we store a floating point diffuse component, a specular component, and a specular exponent for all color channels in each texel. The normal maps are represented as vectors in the tangent space of the triangles, where (0,0,1) represents an unaltered triangle normal (see Fig. 10a).

For the rendering we extend the human animation system in [CTMS03]. After reading the customized human character model and preparing the static BRDF textures, real-time rendering can commence. For each time step, we now read the pose parameters from the stream and apply the respective rigid transformations to the body model. The player also loads the fitting normal map. The final outlook is now determined by the shader programs, which use similar techniques as in [Fer04] to perform per-fragment lighting computations with BRDF textures and normal maps. On a 3.0 GHz Pentium 4 and an Nvidia GeForce^{TM} 6800 graphics board, we achieve 25 fps sustained rendering frame rate at 1024×1024-pixel resolution while illuminating the scene with three moving light sources.

To better demonstrate relighting effects while articulated body motion is performed in the scene, we have decided to illuminate the 3D video with point lights, so that the viewer can see the light source positions and corresponding shadows on the floor. Since we use high-level Cg shaders [MGAK03], our system can be switched to different parametric reflectance models with low effort. We currently can demonstrate Phong and Lafortune model implementations. A comparison between renderings with a two-lobe Lafortune model and the Phong model is shown in Fig. 7. The results shown in Figs. 8, 9 and 10 have been generated using Phong reflectance.

Figs.10a show that wrinkles in the apparel are faithfully identified and represented in the normal maps. Under varying illumination, the wrinkles are realistically rendered, Figs.10b. Figs.10c show rendered images of the pants at three consecutive time steps, illustrating the dynamic nature of the normal maps employed by the renderer. Small rendering artifacts are noticeable that are due to texture resampling.

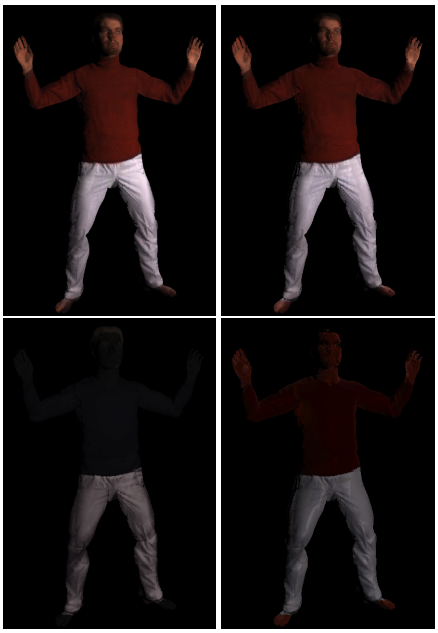


Figure 7: *Top row: person rendered with Phong (left) and Lafortune (right) model while being illuminated by one light source. Bottom row: Only specular component rendered for Phong (left) and Lafortune (right) model under the same lighting conditions.*

8. Results and Discussion

For validation, we have five different sequences of a male and a female subject available. Each sequence is between 50 and 250 frames long. Unfortunately, ground truth BRDF data and normal maps are not at our disposal. Thus, we assess the estimation accuracy in both cases by means of visual comparison to the actual video footage. We found that our method is capable of nicely reproducing the appearance of the actor in the video frames.

Our BRDF estimation approach captures surface reflectance characteristics of different materials simultaneously, as seen in the renderings of Figs.10d,e. The animated male and female models are accurately relit for illumination conditions very different from the recording setup. The approach reliably discriminates between diffuse and specular reflectance. The realistically reproduced specular reflection of the trousers of the male model is shown in the accompanying video.

Once we have estimated the BRDF for one type of clothing, we can also use the surface appearance description to change the apparel of a person even for motion sequences in which the person was originally dressed differently. Fig.8 depicts an example of dynamic reclothing.

The entire estimation process including motion capture and reflectance estimation takes approximately three minutes per time step. Optional input frame warping takes around 10 seconds for one pair of reference image and reprojected image. We assess the multi-view warping quality by comparing the image differences between reference views and reprojected model views before and after the warp. Typically, we achieve an average reduction in absolute image difference in the range of 6% over a whole sequence. The local registration improvements in single image pairs lead to a global improvement in multi-view texture-to-model consistency. In Fig. 9 the texture registration improvement due to the warp-correction step is demonstrated. However, in some rare cases local deteriorations in the final texture can be observed despite an improvement on the global level. The decision if the warp-correction is applied is thus left to the user.

Our method is subject to a couple of limitations: First, our method is based on the assumption that interreflections on the body surface can be neglected. In the RES, interreflections potentially play a role between the wrinkles in clothing. To prevent this effect from degrading the estimation accuracy, we have taken care to minimze the number of wrinkles in the RES.

Another limitation of our approach is that visual quality deteriorates if the fabric shifts substantially across the body. Furthermore, we cannot account for loose apparel whose surface can deviate almost arbitrarily from the body model.

For some body poses, rendering artifacts due to undersampling may occur. Especially the lower side of the arms sometimes can not be seen by any of the cameras and thus the true normal directions cannot be inferred. Additional appropriately positioned imaging sensors would solve this problem.

Finally, we intend to employ a single-skin surface model instead of our current segmented one in the future. With the current body representation, occlusions of parts of the surface geometry in the RES complicate the reflectance and normal estimation processes. If a surface point on the model is never seen by any camera, we cannot reconstruct its reflectance. In that case, we interpolate missing parts in the BRDF textures from neighboring regions. However, discontinuities in the texture when frequently occluded surface patches suddenly appear may still be visible. Alternatively, recording the person in more than one body pose can solve that problem already during acquisition. Moreover, if the face geometry of the template model is too different from the shape of the real actor's face, blurring artifacts occur in the final rendering. One possibility to solve this would be to precede the reflectance estimation with a face model reconstruction from high-resolution images of the head. We'd like to emphasize that all limitations inflicted by our specific body geometry are not principal limitations of our method.

Our results demonstrate that we have developed an effective novel method for simultaneous capture of dynamic scene geometry, per-texel BRDFs and time-varying normal maps from multi-view video. The acquired scene description enables realistic real-time rendition of relightable 3D videos.

9. Conclusions

Our video-based modeling approach jointly captures motion and surface reflectance of a person. From eight synchronized multi-video streams, we recover all information necessary to photo-realistically render a recorded person from arbitrary viewpoint and in arbitrary illumination. The ability to perform convincing relighting enables us to implant real-world, animated people into virtual surroundings. The abstract description of people appearance in terms of geometry, animation and surface reflectance further allows us to separate surface appearance from geometry. This way, we can interchange surface attributes among different people, e.g., for re-dressing one person with another person's clothes. Moreover, we employ a compact data format for our scene description that can be acquired with only a handful of imaging sensors.

Joint motion and reflectance capture can not only be applied to humans but to any dynamic object whose motion can be described by a kinematic chain and for which a suitably parameterized geometry model is available. For BRDF parameter recovery, the proposed algorithm currently assumes that the subject is illuminated by one point light source. While this setup has been chosen to maximize observed reflection variations, the approach can be extended towards more general illumination configurations captured, e.g., via HDR environment maps. To overcome the fixed relationship between light and camera direction, alternatively, a number of spotlights may be applied that are switched on and off during acquisition to illuminate the person sequentially from different directions.

References

[BG01] BOIVIN S., GAGALOWICZ A.: Image-based rendering of diffuse, specular and glossy surfaces from a single image. In *Proc. of ACM SIGGRAPH 2001* (2001), pp. 107–116. 2

[BMR01] BERNARDINI F., MARTIN I. M., RUSHMEIER H.: High-quality texture reconstruction from multiple scans. *IEEE TVCG 7*, 4 (2001), 318–332. 2

[CK01] CARCERONI R. L., KUTULAKOS K. N.: Multi-view scene capture by surfel sampling: From video streams to non-rigid 3D motion shape & reflectance. In *ICCV* (2001), pp. 60–67. 2

[CTMS03] CARRANZA J., THEOBALT C., MAGNOR M., SEIDEL H.-P.: Free-viewpoint video of human actors. In *Proc. of SIGGRAPH'03* (2003), pp. 569–577. 1, 2, 4, 8

[DHT*00] DEBEVEC P., HAWKINS T., TCHOU C., DUIKER H.-P., SAROKIN W., SAGAR M.: Acquiring the reflectance field of a human face. *Proc. of ACM SIGGRAPH'00* (2000), 145–156. 1

[DTM96] DEBEVEC P., TAYLOR C., MALIK J.: Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach. *Proc. of ACM SIGGRAPH'96* (1996), 11–20. 1

[Far99] FARIN G.: *Curves and Surfaces for CAGD: A Practical Guide*. Morgan Kaufmann, 1999. 5

[Fer04] FERNANDO R.: *GPU Gems: Programming Techniques, Tips and Tricks for Real-time Graphics*. Nvidia, 2004. 8

[GCHS04] GOLDMAN D. B., CURLESS B., HERTZMANN A., SEITZ S.: *Shape and Spatially-Varying BRDFs From Photometric Stereo*. Tech. rep., University of Washington, 2004. 2

[Geo03] GEORGHIADES A. S.: Recovering 3-d shape and reflectance from a small number of photographs. In *Eurographics Symposium on Rendering* (2003), pp. 230–240. 2

[GHH01] GIBSON S., HOWARD T., HUBBOLD R.: Flexible image-based photometric reconstruction using virtual light sources. *Computer Graphics Forum 20*, 3 (2001). 2

[GTHD03] GARDNER A., TCHOU C., HAWKINS T., DEBEVEC P.: Linear light source reflectometry. *ACM Trans. Graphics. 22*, 3 (2003), 749–758. 2

[GWN*03] GROSS M. H., WÜRMLIN S., NÄF M., LAMBORAY E., SPAGNO C. P., KUNZ A. M., KOLLER-MEIER E., SVOBODA T., GOOL L. J. V., LANG S., STREHLKE K., MOERE A. V., STAADT O. G.: blue-c: a spatially immersive display and 3d video portal for telepresence. *ACM Trans. Graph. 22*, 3 (2003), 819–827. 2

[HWT*04] HAWKINS T., WENGER A., TCHOU C., GARDNER A., GÖRANSSON F., DEBEVEC P.: Animatable facial reflectance fields. In *Proc. of Eurographics Symposium on Rendering* (2004), pp. 309–319. 2

[KRN97] KANADE T., RANDER P., NARAYANAN P. J.: Virtualized reality: Constructing virtual worlds from real scenes. *IEEE MultiMedia 4*, 1 (1997), 34–47. 1, 2

[LFTG97] LAFORTUNE E., FOO S., TORRANCE K., GREENBERG D.: Non-Linear Approximation of Reflectance Functions. In *SIGGRAPH* (August 1997), pp. 117–126. 6

[LH96] LEVOY M., HANRAHAN P.: Light field rendering. In *in Proc. of ACM SIGGRAPH'96* (1996), pp. 31–42. 2

[LK81] LUCAS B., KANADE T.: An iterative image registration technique with an application to stereo vision. In *Proc. DARPA IU Workshop* (1981), pp. 121–130. 5

[LKG*03] LENSCH H. P. A., KAUTZ J., GOESELE M., HEIDRICH W., SEIDEL H.-P.: Image-based reconstruction of spatial appearance and geometric detail. *ACM Transactions on Graphics 22*, 2 (2003), 27. 1, 2, 5, 6

[LSMS02] LI M., SCHIRMACHER H., MAGNOR M., SEIDEL H.-P.: Combining stereo and visual hull information for on-line reconstruction and rendering of dynamic scenes. In *Proc. of IEEE MMSP* (2002), pp. 9–12. 2

[Mar98] MARSCHNER S.: *Inverse Rendering for Computer Graphics*. PhD thesis, Cornell University, 1998. 1

[MBM01] MATUSIK W., BUEHLER C., MCMILLAN L.: Polyhedral visual hulls for real-time rendering. In *Proceedings of 12th Eurographics Workshop on Rendering* (2001), pp. 116–126. 1

[MBR*00] MATUSIK W., BUEHLER C., RASKAR R., GORTLER S., MCMILLAN L.: Image-based visual hulls. In *Proceedings of ACM SIGGRAPH 00* (2000), pp. 369–374. 2

[MG01] MOESLUND T. B., GRANUM E.: A survey of computer vision-based human motion capture. *CVIU 81*, 3 (2001), 231–268. 2

[MGAK03] MARK W. R., GLANVILLE R. S., AKELEY K., KILGARD M. J.: Cg: a system for programming graphics hardware in a c-like language. *ACM Trans. Graph. 22*, 3 (2003), 896–907. 8

[MP04] MATUSIK W., PFISTER H.: 3d tv: a scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes. *ACM Trans. Graph. 23*, 3 (2004), 814–824. 1, 2

[MPBM03] MATUSIK W., PFISTER H., BRAND M., MCMILLAN L.: A data-driven reflectance model. *ACM Trans. Graph. (Proc. SIGGRAPH'03) 22*, 3 (2003), 759–769. 1, 2

[MT02] MATSUYAMA T., TAKAI T.: Generation, visualization, and editing of 3D video. In *Proc. of 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT'02)* (2002), p. 234ff. 2

[MTG97] MOEZZI S., TAI L.-C., GERARD P.: Virtual view generation for 3D digital video. *IEEE MultiMedia 4*, 1 (1997), 18–26. 2

[NSI01] NISHINO K., SATO Y., IKEUCHI K.: "eigen-texture method: Appearance compression and synthesis based on a 3d model". *IEEE Trans. PAMI 23*, 11 (nov 2001), 1257–1265. 2

[Pho75] PHONG B.-T.: Illumination for computer generated pictures. *Communications of the ACM* (1975), 311–317. 6

[PTVF02] PRESS W. H., TEUKOLSKY S. A., VETTERLING W. T., FLANNERY B. P.: *Numerical recipes in C++*. Cambridge University Press, 2002. 6

[RH01] RAMAMOORTHI R., HANRAHAN P.: A signal-processing framework for inverse rendering. In *Proceedings of SIGGRAPH 2001* (2001), ACM Press, pp. 117–128. 2

[RTG97] RUSHMEIER H., TAUBIN G., GUÉZIEC A.: Applying Shape from Lighting Variation to Bump Map Capture. In *Eurographics Workshop on Rendering* (June 1997), pp. 35–44. 2

[SWI97] SATO Y., WHEELER M. D., IKEUCHI K.: Object Shape and Reflectance Modeling from Observation. In *Proc. of SIGGRAPH'97* (1997), pp. 379–388. 2

[WLSG02] WUERMLIN S., LAMBORAY E., STAADT O., GROSS M.: 3d video recorder. In *Proc. of IEEE Pacific Graphics* (2002), pp. 325–334. 2

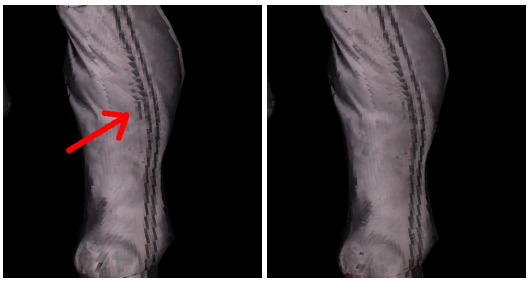
[YDMH99] YU Y., DEBEVEC P., MALIK J., HAWKINS T.: Inverse global illumination: Recovering reflectance models of real scenes from photographs. In *Proc. of SIGGRAPH'99* (August 1999), pp. 215–224. 2

[YM98] YU Y., MALIK J.: Recovering Photometric Properties of Architectural Scenes from Photographs. In *Proceedings of SIGGRAPH'98* (1998), pp. 207–218. 2

[ZKU*04] ZITNICK C. L., KANG S. B., UYTTENDAELE M., WINDER S., SZELISKI R.: High-quality video view interpolation using a layered representation. *ACM Trans. Graph. 23*, 3 (2004), 600–608. 1, 2

[ZTCS99] ZHANG R., TSAI P.-S., CRYER J., SHAH M.: Shape from Shading: A Survey. *IEEE Trans. PAMI 21*, 8 (1999), 690–706. 2

Figure 8: *Dynamic reclothing.*



(a) (b)

Figure 9: *Magnification of the lower leg of the rendered person. (a) Result without warp-correction prior to reflectance estimation - ghosting due to misalignments along the stripes of the trousers are visible. (b) Result with warp-correction - ghosting artifatcs have been significantly reduced due to better multi-view consistency. Block artifacts are due to limited texture resolution.*

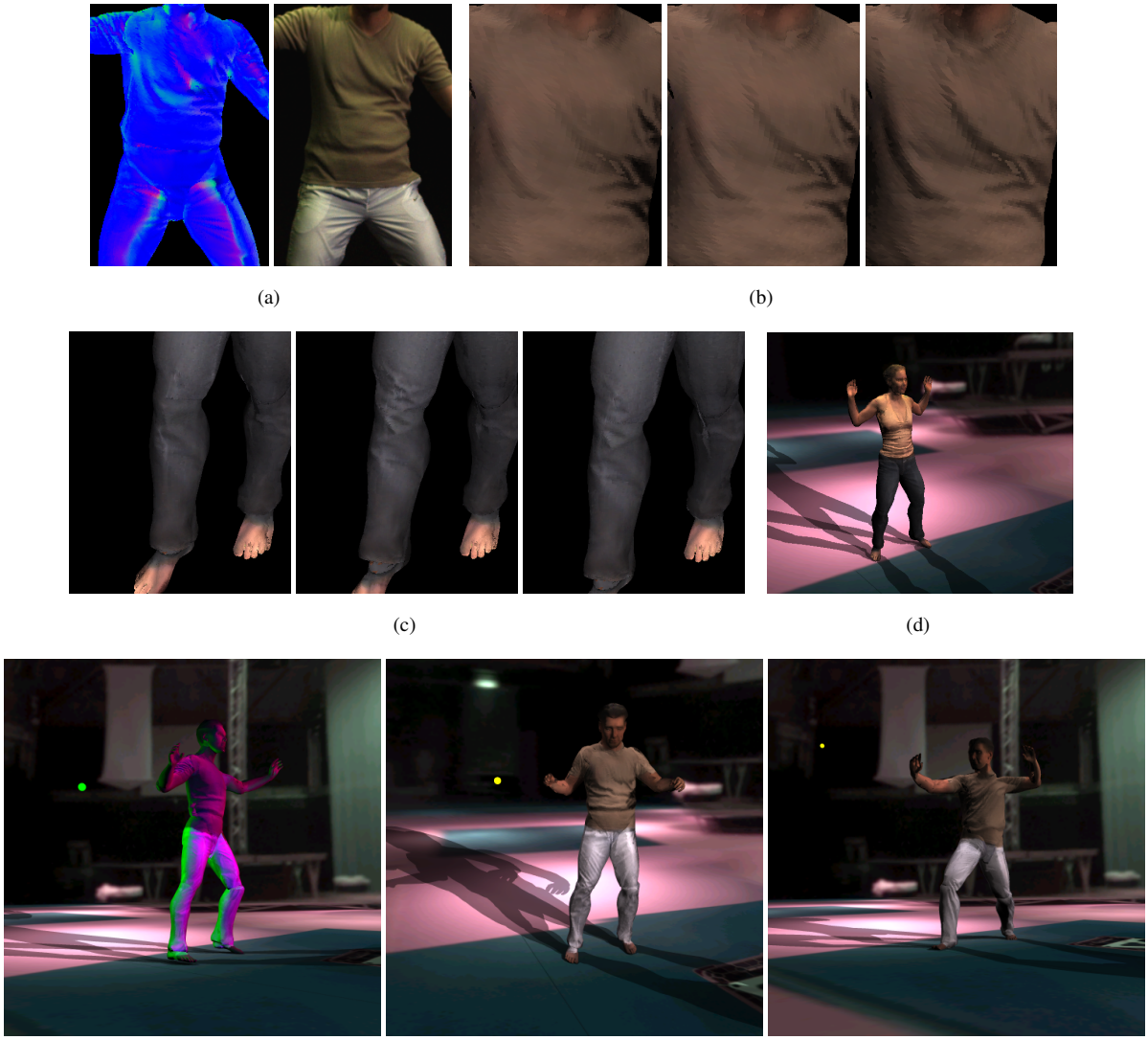
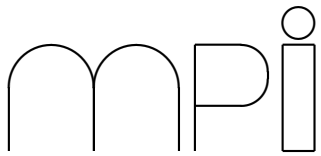


Figure 10: *(a) Color-coded normal map in local coordinates (left) and corresponding input video frame (right). The default normal in the tangent frame is the vector (0,0,1) which translates into a purely blue pixel in the local normal map. Normals deviating from the default one, e.g. due to wrinkles, appear in a different color. (b) Wrinkles on T-shirt rendered under different illumination conditions. (c) Rendered time-varying wrinkles in pants. (d) Single pose relit with different light positions. (e) Person rendered from different viewpoints and illuminations (colored dots: light source positions, colors are light source colors).*



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` 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-2005-5-002	S. Siersdorfer, G. Weikum	Automated Retraining Methods for Document Classification and their Parameter Tuning
MPI-I-2005-4-006	C. Fuchs, M. Goesele, T. Chen, H. Seidel	An Emperical Model for Heterogeneous Translucent Objects
MPI-I-2005-4-005	G. Krawczyk, M. Goesele, H. Seidel	Photometric Calibration of High Dynamic Range Cameras
MPI-I-2005-4-004	C. Theobalt, N. Ahmed, E. De Aguiar, G. Ziegler, H. Lensch, M.A.,. Magnor, H. Seidel	Joint Motion and Reflectance Capture for Creating Relightable 3D Videos
MPI-I-2005-4-003	T. Langer, A.G. Belyaev, H. Seidel	Analysis and Design of Discrete Normals and Curvatures
MPI-I-2005-4-002	O. Schall, A. Belyaev, H. Seidel	Sparse Meshing of Uncertain and Noisy Surface Scattered Data
MPI-I-2005-4-001	M. Fuchs, V. Blanz, H. Lensch, H. Seidel	Reflectance from Images: A Model-Based Approach for Human Faces
MPI-I-2005-2-004	Y. Kazakov	A Framework of Refutational Theorem Proving for Saturation-Based Decision Procedures
MPI-I-2005-2-003	H.d. Nivelle	Using Resolution as a Decision Procedure
MPI-I-2005-2-002	P. Maier, W. Charatonik, L. Georgieva	Bounded Model Checking of Pointer Programs
MPI-I-2005-2-001	J. Hoffmann, C. Gomes, B. Selman	Bottleneck Behavior in CNF Formulas
MPI-I-2005-1-007	I. Katriel, M. Kutz	A Faster Algorithm for Computing a Longest Common Increasing Subsequence
MPI-I-2005-1-003	S. Baswana, K. Telikepalli	Improved Algorithms for All-Pairs Approximate Shortest Paths in Weighted Graphs
MPI-I-2005-1-002	I. Katriel, M. Kutz, M. Skutella	Reachability Substitutes for Planar Digraphs
MPI-I-2005-1-001	D. Michail	Rank-Maximal through Maximum Weight Matchings
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-2004-4-006	K. Dmitriev, V. Havran, H. Seidel	Faster Ray Tracing with SIMD Shaft Culling
MPI-I-2004-4-005	I.P. Ivrissimtzis, 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-2004-2-007	S. Wagner	Summaries for While Programs with Recursion
MPI-I-2004-2-002	P. Maier	Intuitionistic LTL and a New Characterization of Safety and Liveness
MPI-I-2004-2-001	H. de Nivelle, Y. Kazakov	Resolution Decision Procedures for the Guarded Fragment with Transitive Guards
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. Ivrissimtzis, H. Seidel	Tree-based triangle mesh connectivity encoding
MPI-I-2003-4-007	I. Ivrissimtzis, 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. de 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 -