Enhanced Dynamic Reflectometry for Relightable Free-Viewpoint Video

Christian Theobalt Naveed Ahmed Hendrik Lensch Marcus A. Magnor Hans-Peter Seidel

MPI–I–2006–4–006 July 2006

Authors' Addresses

Christian Theobalt, Naveed Ahmed, Hendrik Lensch, Hans-Peter Seidel
Max-Planck-Institut für Informatik
Stuhlsatzenhausweg 85
66123 Saarbrücken

Marcus Magnor
TU Braunschweig
Computer Graphics Lab
Mühlenpfordtstr. 23
38106 Braunschweig, Germany

Acknowledgements

This project has been supported by EU 3DTV NoE project No 511568. The authors would also like to thank Jan Petersen, Art Tevs, Natascha Sauber, Gernot Ziegler, and Edilson de Aguiar for their assistance. We are also grateful to Vlastimil Havran for providing us his importance sampling code.

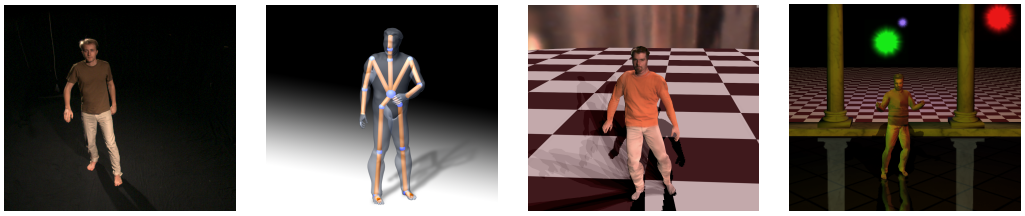
Abstract

Free-Viewpoint Video of Human Actors allows photo-realistic rendering of real-world people under novel viewing conditions. Dynamic Reflectometry extends the concept of free-view point video and allows rendering in addition under novel lighting conditions. In this work, we present an enhanced method for capturing human shape and motion as well as dynamic surface reflectance properties from a sparse set of input video streams. We augment our initial method for model-based relightable free-viewpoint video in several ways. Firstly, a single-skin mesh is introduced for the continuous appearance of the model. Moreover, an algorithm to detect and compensate lateral shifting of textiles in order to improve temporal texture registration is presented. Finally, a structured resampling approach is introduced which enables reliable estimation of spatially varying surface reflectance despite a static recording setup. The new algorithm ingredients along with the Relightable 3D Video framework enables us to realistically reproduce the appearance of animated virtual actors under different lighting conditions, as well as to interchange surface attributes among different people, e.g. for virtual dressing. Our contribution can be used to create 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

Contents



(a) (b) (c) (d)

Figure 1: Our method jointly estimates shape, motion and dynamic surface reflectance of a human actor from input video (a). A dynamic scene description based on a template body model (b) enables us to photo-realistically render moving actors from arbitrary viewpoint and under both captured real-world (c) and artificial illumination (d).

1 Introduction

Recent advances in graphics hardware and rendering algorithms enable the creation of images of unprecedented realism in real-time. In order to capitalize on these novel rendering possibilities, 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 [16]. Many of these techniques show how to interactively render photo-realistic views from real world-captured, dynamic scenes (see also Sect. 2). While the ability to realistically display dynamic events from novel viewpoints has by itself already a number of intriguing applications, the next step is to use objects that have been captured in the real world 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 [7, 26] as well as model-based [22, 18] 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 an approach that jointly captures shape, motion and time-varying surface reflectance of people. An overview of our method is given in Sect. 3. As input to our algorithm, we only require a handful of calibrated and synchronized video recordings, Sect. 4. The algorithm returns subject-adapted 3D geometry and animation parameters by employing a marker-less motion capture method. In a preprocessing step, multi-view texture consistency is enhanced and lateral shift of textiles detected, Sect. 5. From the video input, we estimate a dynamic reflectance description comprising of a spatially-varying BRDF and a dynamic surface normal field, Sect. 6. An optional structured resampling method enhances reflectance estimation greatly and reduces the bias introduced by a static recording setup, Sect. 7. 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, Sect. 8. We validate our algorithm, Sect. 9, and present results for several subjects wearing different clothes made of non-lambertian textiles, Sect. 10. The paper concludes in Sect. 11.

Contributions of our paper are:

- An algorithm that warps input video images in order to guarantee multi-view photo-consistency in conjunction with inexact object geometry,
- A method to detect and compensate lateral shifting of textiles,
- dynamic reflectometry from multi-view video, i.e.,
  - per-texel BRDF parameter estimation,
  - reconstruction of time-varying normal maps to capture small, variable detail of surface geometry (e.g., wrinkles in clothing),
- A structured resampling method for reliable estimation of spatially varying surface reflectance despite a static recording setup, and
- the integration of the acquisition, estimation, and rendering methods into a working system.

The following section highlights related work on free-viewpoint video and BRDF estimation.

2 Related Work

We capitalize on previous research in many areas, but primarily pick up ideas from the fields of free-viewpoint video and image-based reflectance estimation.

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 way 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 [24, 36] or polyhedral visual hull methods [23], as well as approaches performing point-based reconstruction [14]. The combination of stereo reconstruction with visual hull rendering leads to a more faithful reconstruction of surface concavities [20]. Stereo methods have also been applied to reconstruct and render dynamic scenes [42, 16], some of them employing active illumination [34]. Alternatively, a complete parameterized geometry model can be used to pursue a model-based approach towards free-viewpoint video [6]. On the other hand, light field rendering [19] is employed in the 3D TV system [25] 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.

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 consisting of only one material. 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. Often, the parameters of an analytic BRDF model are fit to the measured data [32, 18] or a data-driven model is used [26]. Zickler et al. [41] proposed a scattered data interpolation method to reconstruct a reflectance model. Reflectance measurements of scenes with more complex incident illumination can be derived by either a full-blown inverse global illumination approach [37, 12, 4] or by representing the incident light field as an environment map and solving for the direct illumination component only [38, 30, 27]. 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 [40]. Rushmeier et al. estimate diffuse albedo and normal map from photographs with varied incident light directions [31, 3]. A linear light source is employed by Gardner et al. [10] to estimate BRDF properties and surface normal. In [11, 13], 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 [15] a method to generate animatable and relightable face models from images taken with a special light stage is described. Wenger et al. [35] extend the light stage device such that it enables capturing of dynamic reflectance fields. Their results are impressive, however it is not possible to change the viewpoint in the scene. Einarsson et. al. [8] extends it further by using a large light stage, a trade-mill where the person walks on, and light field rendering for displays, human performances can be rendered from novel perspectives and relit. Unfortunately the method can only present single periodic motion, such as walking, and is only suitable for low frequency relighting. 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, that allows for arbitrary viewpoint changes as well as high-frequency relighting.

Carceroni and Kutulakos present a volumetric method for simultaneous motion and reflectance capture for non-rigid objects [5]. They have shown nice results for spatially confined 3D scenes where they used a coarse set of surfels as shape primitives.

In this paper, we propose a model-based approach that captures shape, motion parameters and dynamic surface reflectance of the whole human body at high accuracy. As input, we only expect a handful of synchronized video streams showing arbitrary human motion. We augment our initial research work [33] in several ways. Firstly, a single-skin mesh is introduced for the continuous appearance of the model. Moreover, algorithm to detect and compensate shifting of textiles in order to improve temporal texture registration is presented. Finally, a structured resampling approach is introduced which enables reliable estimation of spatially varying surface reflectance despite a static recording setup. Our reconstructed dynamic scene description enables us to render virtual people in real-time from arbitrary viewpoints and under arbitrary lighting conditions.

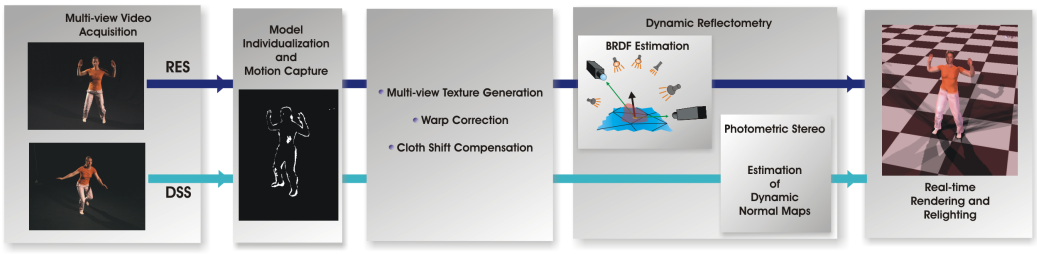


Figure 2: Algorithmic workflow of our method.

3 Overview

Fig. 2 illustrates the workflow between the components of our joint shape, 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 it is these sequences that 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. 2. 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 surface parameterization methods. 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). Furthermore, texture registration problems due to shifting of the apparel over the body surface are prevented by means of an automatic cloth shift detection procedure (Sect. 5.4). From the RES video data, BRDF model parameter values are estimated for each surface texture element (texel) of the geometry model individually (Sect. 6.1). An optional structured resampling method enhances reflectance estimation greatly and reduces the bias introduced by a static recording setup, Sect. 7. 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. 8, Sect. 9, and Sect. 10).

4 Acquisition

As input to our system, we record multi-view video (MVV) sequences. A multi-view recording setup enables us to capture a volume of approx. $4 \times 4 \times 3$ m with eight externally synchronized video cameras. Seven of the cameras are placed in an approximately circular arrangement around the center of the scene, the eighth camera records the set from an overhead position. Since we estimate both motion and reflectance properties, we have strict requirements concerning the spatial, temporal, and color resolution of our imaging devices. We employ Imperx$^{\text{TM}}$ MDC-1004 cameras that feature a 1004x1004 pixel CCD sensor with linear 12 bits-per-pixel dynamic range providing a sustained frame rate of 25 fps. The captured video data is streamed in real-time to eight parallel RAID systems. The cameras are calibrated, and inter-camera color-consistency is ensured by applying a color-space transformation to each camera stream.

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 floor with black cloth and carpet. We employ two K5600$^{\text{TM}}$ Jokerbug 800 spot lights to illuminate our scenes. They are placed in opposite corners of our stage, and they are oriented towards the center of the recording area, Fig. 3. During estimation, we approximate the contribution of each spot light by a single point light source. Light source positions, intensities and color response of the cameras are calibrated off-line. Our flexible approach can also handle more general setups comprising of spot lights and diffuse illumination from the ceiling.

We successively record two MVV sequences for each person and each type of apparel. The first sequence, referred to as the reflectance estimation sequence (RES), serves as input to the BRDF estimation algorithm. The human subject performs, on the very spot, a simple 360° rotation around his vertical body axis which only takes a few seconds. The actor keeps the upper body in a static initialization posture. This way, we can safely assume that, for the duration of the RES, the apparel does not shift over the body. Prior to reflectance estimation we fit our geometry model to each recorded body pose (Sect. 5.1). For each point on the model's surface, the RES contains (in the best case) as many different appearance samples as there are images depicting the respective point. Over time, the surface element's

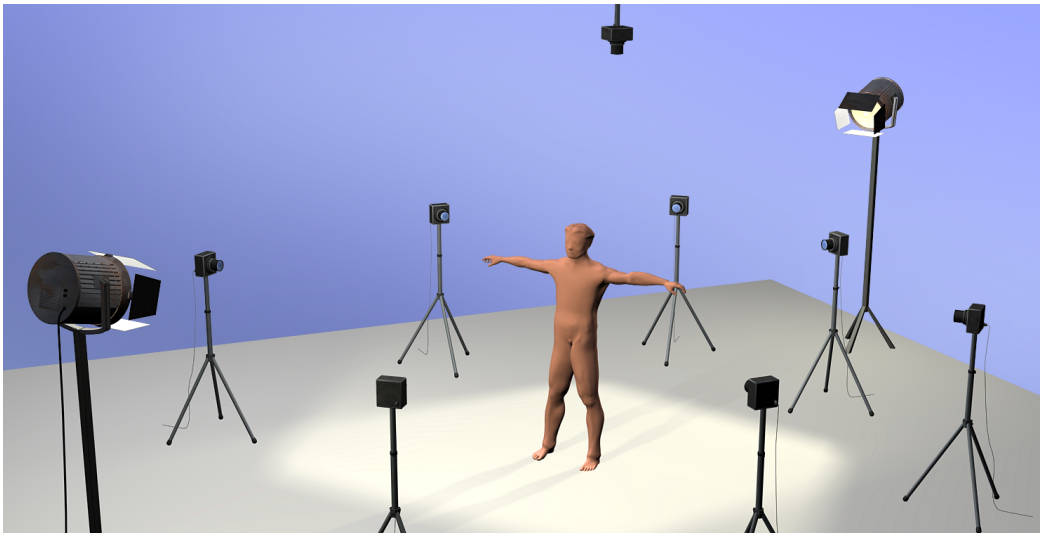


Figure 3: Illustration of camera and light source arrangement.

normal points in various directions with respect to light sources and imaging sensors. While surface normal orientation varies freely, our static camera and lighting setup allows for only a limited number of half vector directions $\hat{h_j} = \hat{l}+\hat{v_j}$, i.e., angular separations between spot light $\hat{l}$ and camera directions $\hat{v_j}$. By placing the eight cameras non-symmetrically with respect to the two spot lights, we gather samples for up to 16 different light-to-camera angles, which we found sufficient to robustly fit our isotropic BRDF models (Sect. 6).

The second type of input video footage, the dynamic scene sequences (DSS), capture the actor's performances from which the actual relightable 3D videos are to be generated. 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 of the actor. To estimate our model's shape and motion parameters, we perform maker-less motion capture. In order to merge the video data from multiple viewpoints, we compute static texture parameterizations, resample and align the input streams using a novel warp-correction technique, and temporally register the video streams by detecting shifting of the apparel.

5.1 Model-based Motion Capture

To simultaneously recover dynamic 3D surface geometry as well as animation parameters, we use an extended version of the model-based, passive optical motion capture approach proposed in [6]. A kinematic body model consisting of 16 separate segments represented by individual triangle meshes is used for automatic silhouette-based motion estimation and shape customization. In a semi-automatic procedure, the segmented model can be transformed into a new model with closed single-skin surface geometry, Fig. 1b. To do so, the triangle meshes of adjacent segments are zippered and surface skinning weights are assigned to correctly model motion-dependent geometry deformation. While we use the segmented version of the model for motion estimation and shape individualization, we use the single-skin model for reflectance estimation and later rendering.

5.2 Texture Parameterization

We jointly apply two different static surface parameterizations of the single-skin mesh over the planar domain based on [39]. To this end, the mesh is cut along a predefined path. Parameterization I leaves the mesh boundary free and minimizes surface distortion (the left image in Fig. 5 shows a texture in parameterization I). We use it during BRDF and normal estimation since it leads to a fairly uniform distribution of samples on the model's surface. Parameterization II forces the boundary of the mesh to lie on the border of a square (Fig. 4 shows two textures in parameterization II). This causes larger distortion, but the textures with fixed boundary are ideal for cloth shift detection (Sect. 5.4).

Using our surface parameterizations, all data related to surface elements (e.g. image samples, normals, visibility information, light vectors) can be conveniently stored as textures. Throughout our experiments, we use 1024x1024-texel texture maps. Before any further processing commences, 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 for both of the two described texture parameterizations.

Figure 4: Detection of relative cloth shift between two subsequent combined textures C_t and C_{t+1} (in parameterization II) by means of optical flow. Areas where shifting textiles were found are marked in red on the 3D model.

5.3 Warp Correction

Although the body model initialization procedure yields a faithful representation of the person's true geometry, small 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, in particular due to nonlinear optimizations that are normally required. 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 (see Fig.5 and video):

Let's assume we want to assemble a MVV texture from the video image $I_j(t)$ seen by camera j 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 j, the texel color is taken from $I_j(t)$. In case camera $i \neq j$ sees the point best and it is not occluded, we regard the video image $I_i(t)$ as the reference image. The model at time t is projectively textured

with $I_j(t)$ and rendered into camera view i. 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 optimally aligned with the camera view that sees the corresponding surface element most head-on.

Warped images are precomputed for all 56 possible combinations of i and j at every time step (Fig. 5). To establish per-pixel correspondences, the warping operation itself is based on the optical flow [21] between the reference image and the image of the reprojected textured model. A regular 2D triangle mesh is superimposed on the reprojected model image, per-vertex displacements are derived from the optical flow values, and the mesh is deformed accordingly via thin-plate spline interpolation [9]. Finally, the warped reprojected image is created on the GPU.

Optical flow is based on the assumption that all surfaces in the scene are diffuse. For reflectance estimation, though, we deliberately generate specular highlights on surfaces in the scene. Experiments show that our method nonetheless produces good results since for the majority of garments diffuse reflectance is the predominant contribution.

5.4 Cloth Shift Detection

Our BRDF estimation procedure (Sect. 6.1) yields a fixed assignment of BRDF parameter values to individual body surface locations which it stores as static textures. During dynamic normal estimation (Sect. 6.2) and rendering (Sect. 8), however, this rigid attribution of reflectance properties may not hold since in a DSS, in contrast to the RES, it is likely that textiles undergo transversal shift over the body surface. It is thus necessary to detect cloth shifting and to take it into account for these two steps. We have developed the following procedure: For each time step of a DSS, and the last time step of the corresponding RES, all MVV textures in parameterization II are weightedly blended into combined textures. Each combined texture describes the complete surface appearance of the person in a single square image, Fig. 4. The combined texture of the single RES time step is regarded as the reference texture since it captures the surface appearance from which BRDF parameters are estimated. Cloth shift in 3D corresponds to image motion in the texture domain which we compute via optical flow (Sect. 5.3). For each time step of a DSS we need to compute the motion of each texel

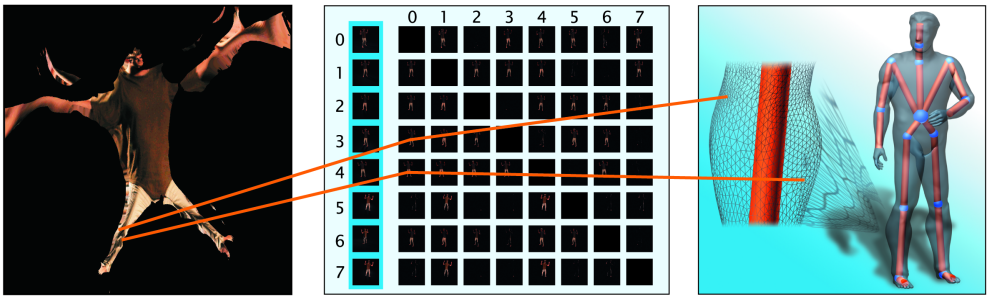


Figure 5: MVV texture generation (parameterization I) for camera 0: The color information for each surface point on the body model is not looked up in the original input video frame recorded from camera 0 (column with darker blue background). Instead the texel color is taken from an image that has been 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 (second column of images). Given 8 cameras, 56 warped reprojected images exist for one time step (image matrix in center).

with respect to its position in the BRDF parameter textures. Due to lighting changes this absolute shift-motion cannot be computed by direct comparison to the reference texture. Instead, we first compute the relative shift-motion between all subsequent combined textures C_t and C_{t+1}, Fig. 4. The motion between the reference texture and C_0 is also computed. The absolute shift-motion of a texel in C_t is the sum of all its shift motions that have been detected at previous time steps. We use the GPU to render a texture-space representation of absolute shift-motion of each texel at each time step. To this end, we use the mesh-based warping technique described in Sect. 5.3 to warp the color-coded BRDF texture coordinates in parameterization II according to the absolute shift-motion.

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 extent depend on the current pose of the person. The underlying technique

is similar to [18] 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. 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 [28]. We have also tested our method with the model proposed by Lafortune [17], 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 i at all time steps t:

$$
\begin{aligned}
&E(\vec{x}, \rho(\vec{x})) = \\
&\sum_{t}^{N} \sum_{i}^{8} \kappa_i(t) \Bigg(S_i(t) - [\sum_{j}^{J} \lambda_j(t) (f_r(\hat{l}(t), \hat{v}_i(t), \rho(\vec{x})) \\
&\qquad \cdot I_j(\hat{n}(t) \cdot \hat{l}(t)))] \Bigg)^2 .
\end{aligned} \tag{1}
$$

The term is evaluated separately in the red, green and blue color channel. $S_i(t)$ denotes the measured color samples at $\vec{x}$ from camera i, and I_j denotes the intensity of light source j. The viewing directions $\hat{v}_i(t)$ 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 $\kappa_i(t)$ and with respect to the light sources by $\lambda_j(t)$, 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 call *dynamic texel* or *dyxel*.

Using non-linear optimization, this formula can in principle be used to determine a full BRDF and the surface normal at the same time. However, we

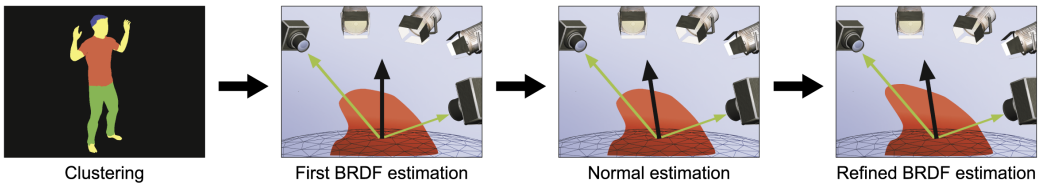


Figure 6: Subsequent steps to estimate per-texel BRDFs.

apply an iterative approach. We carefully designed the reflectance estimation sequence to obtain a much more stable optimization. The subsequent steps of our iterative BRDF estimation scheme are *material clustering, first BRDF estimation, normal estimation* and *refined BRDF estimation*, Fig. 6.

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 maps, belongs. The number of materials is preset a priori. We employ a straightforward color-based clustering approach that considers the raw texel color values.

During the *first BRDF estimation* pass, an optimal set of per-texel BRDF parameters is determined using the normals of the default geometry. The estimation itself consists of a non-linear minimization of (1) in the BRDF parameters. For optimization, we make use of a Levenberg-Marquardt minimization scheme [29] in the same manner as [18]. First, 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 (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 dynamic texels that contain purely diffuse reflectance samples. Using these purely diffuse samples, an individual diffuse component is estimated for each surface element (texel) by minimizing (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. During normal estimation we minimize the following extended version of the energy functional Eq. 1 in the local surface normal direction $\hat{n}(\vec{x})$ of each surface point $\vec{x}$:

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

The additional regularization term $\Delta(\hat{n}(\vec{x}))$ penalizes angular deviation from the default normal of the body model. This way, we also enforce convergence to a plausible local minimum of the non-convex energy functional. The coefficients 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. In the *second BRDF estimation* we capitalize on the refined surface normals and compute a more accurate set of BRDF parameters, ρ_{final}. As output, we generate one floating point texture for each BRDF model parameter.

6.2 Estimation of Time-varying Normals

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 time-varying normal direction is estimated for each surface point individually. The estimation procedure is a non-linear minimization of the regularized energy function (2) in the normal direction.

Our cloth shift detection method (Sect. 5.4) has stored information about potential temporal displacement of textiles as a stream of warped BRDF texture coordinates. We account for this by performing BRDF parameter lookups using these temporally varying texture coordinates. 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

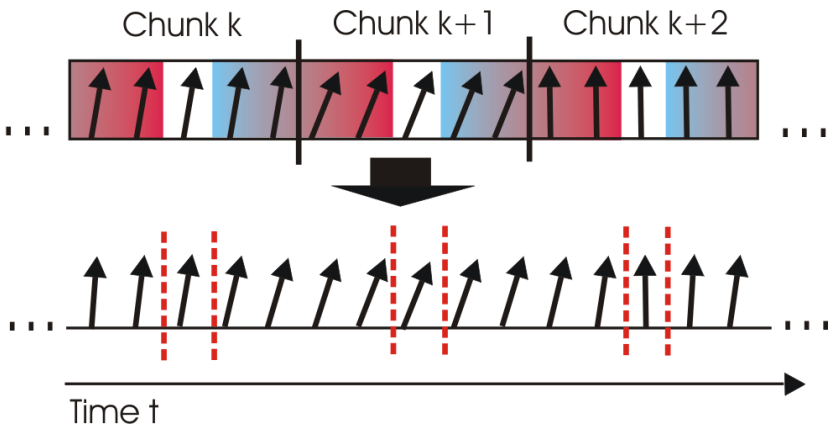


Figure 7: 2D illustration of robust time-varying normal map estimation. Top: The sequence subdivided into short chunks. For each chunk, one best-matching normal is derived per texel which is assigned to the chunk's center time step (white). Intermediate time steps are interpolated (bottom).

achieve this purpose, we assume that changes in local normal direction within a short window in time can be neglected. The input sequence is split into subsequent chunks of frames (typically of length 5). 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.7. For some sequences, we are able to correct undersampling problems by adaptively enforcing consistency with the template model. To this end, in those surface areas exhibiting a high fitting error we re-run the normal estimation with a higher influence of the penalty term in Eq. 2. This way, we obtain a normal field that is smooth in the temporal domain and faithfully models subtle dynamic details in surface geometry.

7 Exploiting Spatial Coherence for Structured Resampling

Our acquisition setup comprising of only 8 cameras and 2 light sources is comparably simple and inexpensive. However, the fixed relative arrangement of cameras and light sources can induce a bias in the reflectance estimation. There are two primary reason for the bias:

- The scene is captured under a limited number of half-vector directions.

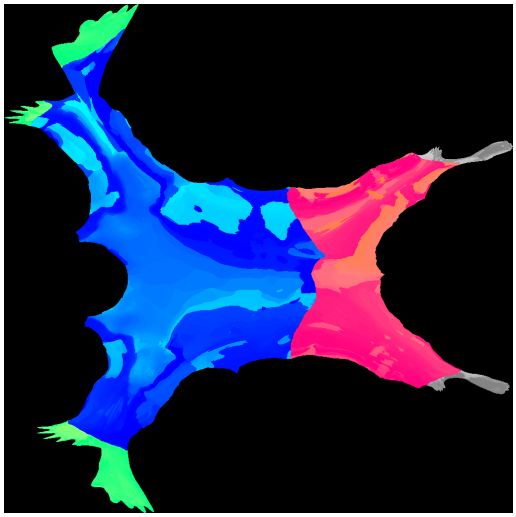


Figure 8: Patches of same material clustered according to average normal direction.

- only "slice" of samples is taken for the 4D sample space due to the single in place rotation in the RES.

Both of these factors possibly lead to BRDF estimate that may not generalize well to lighting conditions that are very different to the acquisition conditions. We can reduce the bias by applying a structured resampling method that exploits spatial coherence in the BRDF to generate more reliable reflectance estimation. In order to achieve this purpose, we employ for reflectance estimation of a single surface point not only samples of the particular point, but also samples from other surface locations of similar material that are seen under different viewing and lighting directions. It is the main challenge to resample the reflectance data in such a way that the spatial variation in the surface is not compromised, while at the same time, the estimated BRDF models remain general.

We select additional samples that come from surface locations with different orientation than the texel currently estimated and that cover a region in sample space that has not been captured by the original texel. We put this idea into practice in the following way:

In a first step, the surface is clustered into patches of similar normal directions and same material, Fig. 8. For each texel, the reflectance properties are estimated as follows. The normal direction $\hat{n}$ of the texel defines the reference normal direction. The list L of patches of the same material is selected in the order of increasing deviation of the average patch normal direction $\hat{n}$. Now,

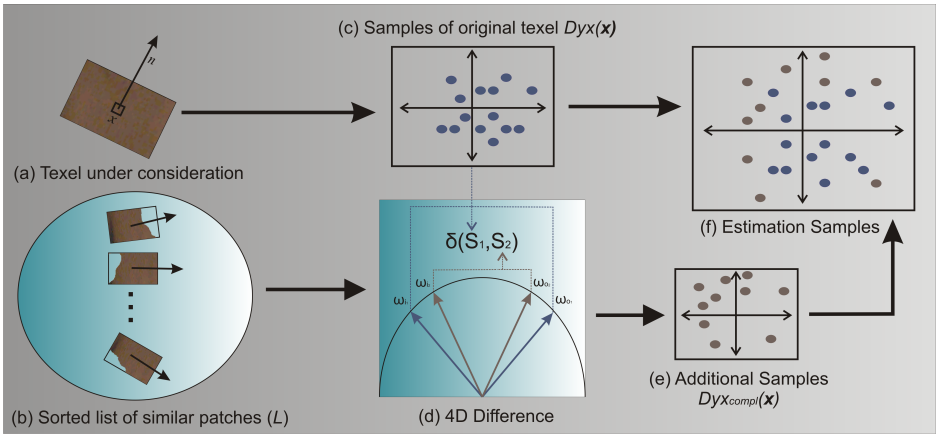


Figure 9: Weighted selection of samples.

n_p many $P_0, ..., P_{n\_p}$ patches are drawn from L. From each patch, a texel is selected randomly, resulting in a set of texels, $T = \vec{X}_{P_0}, ..., \vec{X}_{P_{n\_p}}$. The set of texels T has been selected in a way that maximizes the number of different surface orientations. From the reflectance samples associated with texels in T, we now select a subset that maximizes the coverage of 4D sample space (correspond to the surface from $\vec{X}$ alone).

In order to properly balance the contributions of the samples of $\vec{X}$ and the samples drawn from T, we employ the following selection mechanism. A weighting function $\delta(S_1, S_2)$ is employed that measures the difference of two samples $S_1 = (\hat{\omega}_{i_1}, \hat{\omega}_{o_1})$ and $S_2 = (\hat{\omega}_{i_2}, \hat{\omega}_{o_2})$ in the 4D parameter space as follows:

$$\delta(S_1, S_2) = \Delta(\hat{\omega}_{i_1}, \hat{\omega}_{i_2}) + \Delta(\hat{\omega}_{o_1}, \hat{\omega}_{o_2}) \tag{3}$$

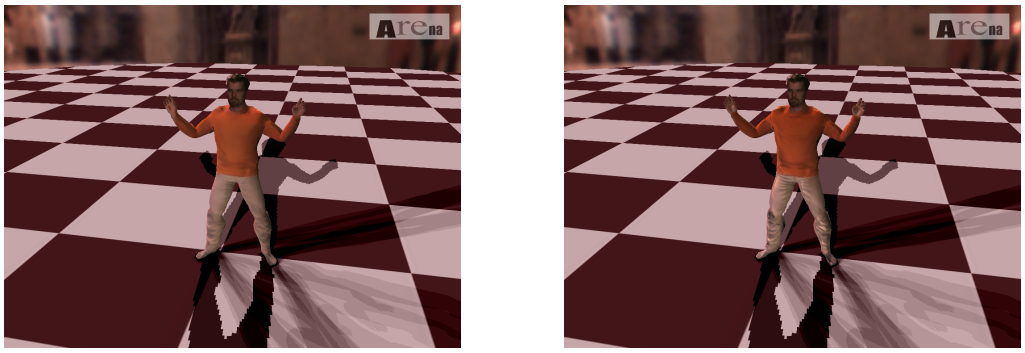
where Δ denotes the angular difference between two vectors. We use the function δ to weight each sample from T by the difference to the closest sample (according to δ) in $S(\vec{X})$.

The final set of samples used in reconstructing the reflectance of $\vec{X}$ comprises of the original samples $S(\vec{X})$ and the samples from T with the highest weights. Typically, the original samples accord for the 60% of the total samples and the additional samples for 40% respectively, Fig. 9.

With the structured resampling method, we obtain the diffuse map which preserves much of the spatial variation. Besides that it also results in better renditions for illumination conditions which differ starkly from the acquisi-

tion environment. To quantify the improvement we compare the input image with rendering generated only using the diffuse map obtained from the previous approach and the structured resampling method. With structured resampling we obtain a PSNR improvement of around 2 dB.

Fig. 10 shows a side by side comparison between the results obtained without the structured resampling and the results obtained with structured resampling. Once can see that with the exploitation of spatial coherence more surface detail is preserved under lighting conditions strongly different from acquisition setup.



(a) (b)

Figure 10: Comparison of renditions without sample redistribution (a), with structured sampling (b). One can see that in (b) subtle surface details are better reproduced.

8 Rendering

Our approach outputs a relightable dynamic object description that consists of the animated geometry and the reconstructed material properties. The material properties consist of the time-independent BRDF textures, the dynamic normal maps, and a sequence of warped BRDF texture coordinates modeling cloth shifting.

During real-time playback, our renderer *Arena* displays the single-skin body model in the smoothly interpolated sequence of captured body poses. The final appearance of the body surface is determined on the GPU by BRDF shader programs implemented in Nvidia's Cg^{TM} language. We can currently

demonstrate Phong and Lafortune shader implementations. Due to limited camera resolution, changing facial expressions cannot faithfully be captured by our dynamic reflectance representation. We therefore do not compute full BRDF lighting in the face but only simulate diffuse surface appearance. To this end, we use the fragment colors from the weightedly blended input video streams as diffuse reflectance components. Our real-time renderer can illuminate the dynamic scene from multiple point or directional light sources with activated shadow mapping. Relighting from static environment maps that are approximated by discrete directional light sources via importance sampling is also feasible. On a 3.0 GHz Pentium 4 we achieve 19 fps sustained rendering frame rate if a model with 90000 triangles and four point lights with activated shadow mapping are applied. Using an environment map for relighting, we can still render the scene at 6 fps if 16 approximating lights are employed.

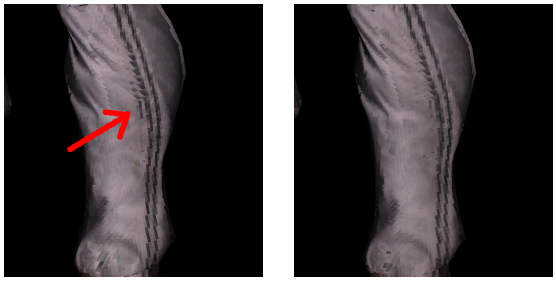
9 Validation and Discussion

We have validated our approach by visual inspection and quantitative evaluation. In total, we have processed 15 different input sequences using Phong and Lafortune BRDFs. They cover 3 different human subjects, 6 different types of apparel, and comprise 150 to 350 frames each. For numerical verifications, we restrict ourselves to Phong sequences.

Warp correction (Sect. 5.3) leads to an average reduction in absolute image difference between reference and reprojected image in the range of 6%. These pair-wise registration enhancements result in improvements in multi-view texture-to-model consistency, see Fig. 11.

Cloth shift compensation also contributes to more realistic renditions, especially since the positions of seams of pieces of apparel are more reliably captured, see Fig. 12. Although both of these methods lead to visual improvements in locally confined surface areas, they have a very noticeable positive influence on the overall visual quality.

We have also measured the average peak-signal-to-noise-ratio with respect to one input video stream not used for reconstruction. By employing warp correction, we obtain a PSNR improvement of 0.2 dB. Cloth shift compensation typically accounts for an additional gain of 0.1-0.2 dB. This numerically confirms the achieved reconstruction improvements. However, the global quality measure does not do full justice to the perceivable local visual improvements



(a) (b)

Figure 11: (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 artifacts have been significantly reduced due to better multi-view consistency. Block artifacts are due to limited texture resolution.

that we achieve.

For our captured test persons, we neither have ground truth BRDF parameters nor ground truth normal maps at our disposition. We therefore assess the quality of the reconstructed dynamic reflectance models by visual comparison to the input video footage. Our approach captures dynamic scene geometry reliably and jointly estimates dynamic surface reflectance of different materials, Fig. 15. Diffuse and specular reflectances are faithfully discriminated. The realistically reproduced specular reflection of the trousers of a male test subject is shown in the accompanying video that can be downloaded from [1]. Subtle details like time-varying wrinkles in clothing are faithfully captured in normal maps, Fig. 15a, and realistically relit under varying illumination conditions, Fig. 15b.

To get as-good-as-possible reflectance data for one RES, we generated a high-quality surface mesh using a laser range scanner and rotated the person on a turntable. We compared BRDF parameters and normal maps computed with our model to the ones obtained with the scanned model. The per-material BRDF parameters found with our approximate geometry only deviate minimally, see Tab. 1. Our surface normals nicely capture surface details, most importantly wrinkles and seams of pieces of apparel. In some places our normal field even exceeds the resolution of the scan (see Fig. 13).

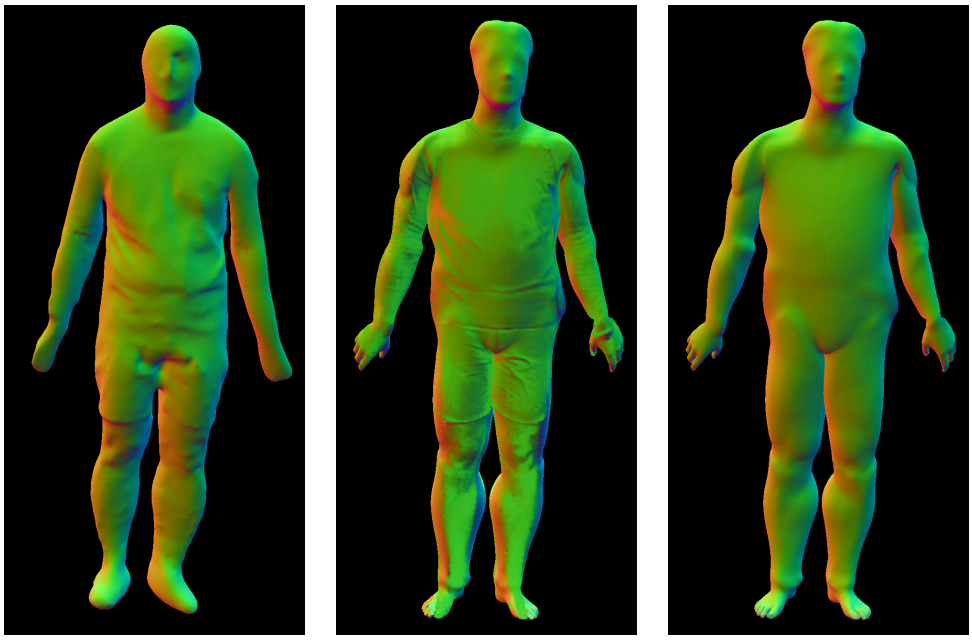
To verify the rendering quality under novel lighting conditions, we have recorded one of our male test subjects under two different calibrated lighting

(a) (b) (c)

Figure 12: Zoom-in on renditions of the t-shirt's waist seam. Block artifacts are due to limited texture resolution. (a) shows the model textured with the blended input camera views as reference. (b) shows the appearance of the rendered seam if cloth shift detection has not been used during reconstruction. Parts of the trousers are erroneously displayed in orange color. In contrast, (c) shows the correctly rendered seam if textile motion has been considered during the estimation. Although the shift detection is not pixel-accurate it still leads to a more realistic reproduction of the true appearance of the actor's apparel.

setups, henceforth termed LC A and LC B. In each of the lighting setups, just one of the spot lights has been employed. The positions of the light sources in LC A and LC B are (angularly) approximately 45° apart with respect to the center of the scene. We reconstructed the BRDF of the test subject using lighting setup LC A. Subsequently, we compared the ground truth images of the person illuminated by setup LC B, Fig. 14a, to the renderings of the person under LC B, Fig. 14b. This comparison illustrates that our method realistically reproduces the true appearance of even subtle surface details under incident illumination that is different from the lighting conditions used during reconstruction.

Our approach is subject to a couple of limitations: The single-skin surface model is generated in an interactive procedure. However, this is not a principal limitation of our method since model generation could be fully automated



(a) (b) (c)

Figure 13: Color-coded surface normals of a laser-scan of a person in shorts (a), the normal field estimated with our method (b), and the surface normals of our the template model (c). The 3 coordinates of the surface normal vectors in world space are encoded in the three RGB color channels. One has to note that our laser-scanner is not well-suited for full-body scanning, and thus exact geometry of the head, the hands, and the feet could not be reconstructed. Furthermore, due to resolution limitations, geometry details were smoothed out and subtle wrinkles were not captured. Even though, one can see that our estimated normal field (b) nicely reproduces the wrinkles that have been captured in the scan. These geometry features are not present in the smooth template geometry (c). One can even see that our image-based approach doesn't smooth out details like the seam of the t-shirt and the seam of the shorts. Instead it captures them accurately.

by a method similar to [2]. Furthermore, although we can handle normal every-day apparel, we can not account for loose apparel whose surface can deviate almost arbitrarily from the body model. Although cloth shift compensation and warp correction lead to visual improvements in the majority of cases, isolated local deteriorations are still possible. Cloth shift detection, for example, sometimes erroneously classifies evolving wrinkles as shifting of apparel. We thus leave the decision if either of the two methods are used

Table 1: Comparison of the estimated average diffuse and specular Phong BRDF parameters that have been estimated using the scanned and the template mesh. Values are given for three of the surface materials.

Material	Scanned	Template
Shirt (Diffuse)	0.352	0.336
Shorts (Diffuse)	0.773	0.743
Skin (Diffuse)	0.448	0.445
Shirt (Specular)	0.012	0.013
Shorts (Specular)	0.014	0.010
Skin (Specular)	0.015	0.012

to the user. Moreover, we currently neglect interreflections on the body. In the RES, they potentially play a role between the wrinkles in clothing. To prevent this effect from degrading the estimation accuracy, we have taken care to minimize the number of wrinkles in the RES. Sometimes, we observe small rendering artefacts due to undersampling (e.g. on the underneath of the arms). However, we have verified that the application of a RES sequence showing several rotation motions with different body postures almost completely solves this problem. If even in this extended RES a pixel is never seen by any of the cameras, we fill in reflectance properties from neighboring regions in texture space. Finally, our current face rendering technique is a compromise that produces visually compelling but not physically correct results. Capturing the reflectance of the face separately with a more closely spaced camera setup would permit full BRDF relighting.

Despite these limitations, our method is an effective combination of algorithmic tools that allows for the creation of realistic relightable dynamic scene descriptions.

10 Results

In the previous section, we have given evidence of the performance of individual algorithmic components of our joint motion and reflectance estimation scheme. In this section, we show that the interplay of these algorithmic ingredients enables the authentic rendition of moving real-world people in real-time. Our dynamic scene description allows us to photo-realistically

(a) (b)

Figure 14: Rendering under novel calibrated lighting condition: (a) shows the ground-truth (recorded) appearance of the actor from one input view under calibrated lighting condition LC B. (b) shows the rendered model under LC B from the same viewpoint. For BRDF reconstruction, a different calibrated lighting condition LC A was used. The comparison shows that we are able to reproduce the true appearance of the actor even under novel lighting conditions.

render human actors under both artificial, Fig. 1d, and real world illumination that has been captured in high-dynamic range environment maps, Fig. 1c. We can also implant actors into virtual environments as they are commonly used in computer games, such as a little pavilion, Fig. 15c, or a pool house, Fig. 15d. Here, our dynamic surface reflectance description even enables the realistic display of the person's reflection in the mirroring floor and the water surface. Relightable 3D videos can be displayed from arbitrary viewpoints and even real-world lighting environments can be rotated interactively, see Fig. 15e,f,g,h,i and the accompanying video (that can be downloaded from [1]).

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 (see video).

The motion estimation process takes around 3 s per time step. Reflectance

estimation at most takes 1 hour for the RES. Normal estimation takes approximately 50 s per time step, and it can be parallelized to bring the computation time down.

Optional input frame warping takes around 10 s for one pair of reference image and reprojected image. Cloth shift compensation accounts for an additional 35 s of computation time for one time step of video.

Even with the relighting option, our final data format is more compact than the traditional 3D video format [6] which comprises of 8 raw video streams for projective texturing and the moving model. As an example, for a 330 frames sequence, the data size of the complete relightable video (moving geometry+BRDF+normal maps) is 528 MB, whereas the total size of the raw input frames is 1452 MB.

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.

11 Conclusions

Our video-based modeling approach jointly captures shape, 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's 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 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 shape, motion and reflectance capture cannot 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 a small number of spot lights. Our approach, however, can be extended to use more general lighting configurations, such as the light stage device [35].

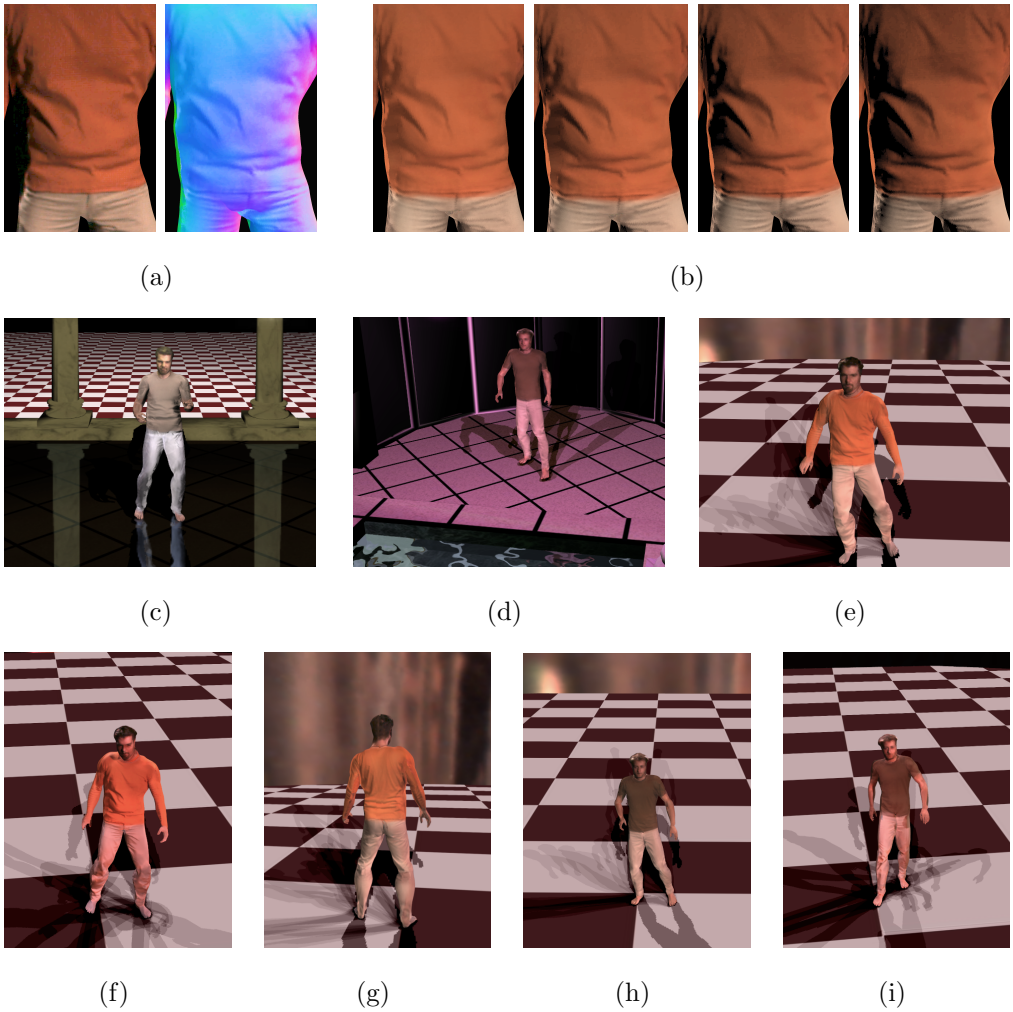


Figure 15: (a) Input video frame (left) and color-coded rendition of estimated normal map (right). In the color-coded rendition the three world-space coordinates of each surface normal have been used as RGB color components. Note that wrinkles are faithfully captured in the normal field. – (b) Under varying incident illumination conditions the wrinkles are authentically relit. – (c) Actor dancing in a pavillion while being illuminated from white spot lights. Note the physically-correct reflection in the mirroring floor. – (d) In a scene that could be taken from a future computer game the actor moves in a pool house lit by purple light sources. – Humans also appear photo-realistically under captured real-world illumination, such as the St Peter's Basilica environment map (e),(g),(h) and the Grace Cathedral environment map (f),(i) provided by Paul Debevec. Both viewpoint and environment lighting can be freely rotated in real-time.

Bibliography

[1] http://www.mpi-inf.mpg.de/∼theobalt/TVCG2006.avi .

[2] D. Anguelov, P. Srinivasan, D. Koller, S. Thrun, J. Rogers, and J. Davis. SCAPE - shape completion and animation of people. In *ACM Trans. on Graphics (Proc. of SIGGRAPH'05)*, volume 24(3), pages 408–416, 2005.

[3] F. Bernardini, I. M. Martin, and H. Rushmeier. High-quality texture reconstruction from multiple scans. *IEEE TVCG*, 7(4):318–332, 2001.

[4] S. Boivin and A. Gagalowicz. Image-based rendering of diffuse, specular and glossy surfaces from a single image. In *Proc. of ACM SIGGRAPH 2001*, pages 107–116, 2001.

[5] R. L. Carceroni and K. N. Kutulakos. Multi-view scene capture by surfel sampling: From video streams to non-rigid 3D motion shape & reflectance. In *ICCV*, pages 60–67, 2001.

[6] J. Carranza, C. Theobalt, M. Magnor, and H.-P. Seidel. Free-viewpoint video of human actors. In *Proc. of SIGGRAPH'03*, pages 569–577, 2003.

[7] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. *Proc. of ACM SIGGRAPH'00*, pages 145–156, 2000.

[8] P. Einarsson, C.-F. Chabert, A. Jones, W.-C. Ma, B. Lamond, im Hawkins, M. Bolas, S. Sylwan, and P. Debevec. Relighting human locomotion with flowed reflectance fields. In *Rendering Techniques*, pages 183–194, 2006.

[9] G. Farin. *Curves and Surfaces for CAGD: A Practical Guide*. Morgan Kaufmann, 1999.

[10] A. Gardner, C. Tchou, T. Hawkins, and P. Debevec. Linear light source reflectometry. *ACM Trans. Graphics. (Proc. of SIGGRAPH'03)*, 22(3):749–758, 2003.

[11] A. S. Georghiades. Recovering 3-d shape and reflectance from a small number of photographs. In *Eurographics Symposium on Rendering*, pages 230–240, 2003.

[12] S. Gibson, T. Howard, and R. Hubbold. Flexible image-based photometric reconstruction using virtual light sources. *Computer Graphics Forum*, 20(3), 2001.

[13] D. Goldman, B. Curless, A. Hertzmann, and S. Seitz. Shape and spatially-varying brdfs from photometric stereo. In *Proc. of ICCV*, pages 341–448, 2004.

[14] M. H. Gross, S. Würmlin, M. Näf, E. Lamboray, C. P. Spagno, A. M. Kunz, E. Koller-Meier, T. Svoboda, L. J. V. Gool, S. Lang, K. Strehlke, A. V. Moere, and O. G. Staadt. blue-c: a spatially immersive display and 3d video portal for telepresence. *ACM Trans. Graph. (Proc. of SIGGRAPH'03)*, 22(3):819–827, 2003.

[15] T. Hawkins, A. Wenger, C. Tchou, A. Gardner, F. Göransson, and P. Debevec. Animatable facial reflectance fields. In *Proc. of Eurographics Symposium on Rendering*, pages 309–319, 2004.

[16] T. Kanade, P. Rander, and P. J. Narayanan. Virtualized reality: Constructing virtual worlds from real scenes. *IEEE MultiMedia*, 4(1):34–47, 1997.

[17] E. Lafortune, S. Foo, K. Torrance, and D. Greenberg. Non-Linear Approximation of Reflectance Functions, August 1997.

[18] H. P. A. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H.-P. Seidel. Image-based reconstruction of spatial appearance and geometric detail. *ACM Transactions on Graphics*, 22(2):27, 2003.

[19] M. Levoy and P. Hanrahan. Light field rendering. In *in Proc. of ACM SIGGRAPH'96*, pages 31–42, 1996.

[20] M. Li, H. Schirmacher, M. Magnor, and H.-P. Seidel. Combining stereo and visual hull information for on-line reconstruction and rendering of dynamic scenes. In *Proc. of IEEE Multimedia and Signal Processing*, pages 9–12, 2002.

[21] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. DARPA IU Workshop*, pages 121–130, 1981.

[22] S. Marschner. *Inverse Rendering for Computer Graphics*. PhD thesis, Cornell University, 1998.

[23] T. Matsuyama and T. Takai. Generation, visualization, and editing of 3D video. In *Proc. of 1st International Symposium on 3D Data Processing Visualization and Transmission (3DPVT'02)*, page 234ff, 2002.

[24] W. Matusik, C. Buehler, R. Raskar, S. Gortler, and L. McMillan. Image-based visual hulls. In *Proceedings of ACM SIGGRAPH 00*, pages 369–374, 2000.

[25] W. Matusik and H. Pfister. 3d tv: a scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes. *ACM Trans. Graph. (Proc. of SIGGRAPH'04)*, 23(3):814–824, 2004.

[26] W. Matusik, H. Pfister, M. Brand, and L. McMillan. A data-driven reflectance model. *ACM Trans. Graph. (Proc. SIGGRAPH'03)*, 22(3):759–769, 2003.

[27] K. Nishino, Y. Sato, and K. Ikeuchi. "eigen-texture method: Appearance compression and synthesis based on a 3d model". *IEEE Trans. PAMI*, 23(11):1257–1265, nov 2001.

[28] B.-T. Phong. Illumnation for computer generated pictures. *Communications of the ACM*, pages 311–317, 1975.

[29] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical recipes in C++*. Cambridge University Press, 2002.

[30] R. Ramamoorthi and P. Hanrahan. A signal-processing framework for inverse rendering. In *Proceedings of SIGGRAPH 2001*, pages 117–128. ACM Press, 2001.

[31] H. Rushmeier, G. Taubin, and A. Guéziec. Applying Shape from Lighting Variation to Bump Map Capture. In *Eurographics Workshop on Rendering*, pages 35–44, June 1997.

[32] Y. Sato, M. D. Wheeler, and K. Ikeuchi. Object Shape and Reflectance Modeling from Observation. In *Proc. of SIGGRAPH'97*, pages 379–388, 1997.

[33] C. Theobalt, N. Ahmed, E. de Aguiar, G. Ziegler, H. Lensch, M. A. Magnor, and H.-P. Seidel. Joint motion and reflectance capture for creating

relightable 3D videos. Technical Report MPI-I-2005-4-004, Max-Planck-Institut für Informatik, April 2005.

[34] M. Waschbüsch, S. Würmlin, D. Cotting, F. Sadlo, and M. Gross. Scalable 3D video of dynamic scenes. In *Proc. of Pacific Graphics*, pages 629–638, 2005.

[35] A. Wenger, A. Gardner, C. Tchou, J. Unger, T. Hawkins, and P. Debevec. Performance relighting and reflectance transformation with time-multiplexed illumination. In *ACM TOG (Proc. of SIGGRAPH'05)*, volume 24(3), pages 756–764, 2005.

[36] S. Würmlin, E. Lamboray, O. Staadt, and M. Gross. 3d video recorder. In *Proc. of IEEE Pacific Graphics*, pages 325–334, 2002.

[37] Y. Yu, P. Debevec, J. Malik, and T. Hawkins. Inverse global illumination: Recovering reflectance models of real scenes from photographs. In *Proc. of ACM SIGGRAPH'99*, pages 215–224, August 1999.

[38] Y. Yu and J. Malik. Recovering Photometric Properties of Architectural Scenes from Photographs. In *Proceedings of ACM SIGGRAPH'98*, pages 207–218, 1998.

[39] R. Zayer, C. Rössl, and H.-P. Seidel. Discrete tensorial quasi-harmonic maps. In *Proc. of Shape Modeling International*, pages 276–285. IEEE, 2005.

[40] R. Zhang, P.-S. Tsai, J. Cryer, and M. Shah. Shape from Shading: A Survey. *IEEE Trans. PAMI*, 21(8):690–706, 1999.

[41] T. Zickler, S. Enrique, R. Ramamoorthi, and P. N. Belhumeur. Reflectance sharing: Image-based rendering from a sparse set of images. In *Proc. of Eurographics Symposium on Rendering*, pages 253–264, 2005.

[42] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski. High-quality video view interpolation using a layered representation. *ACM TOC (Proc. SIGGRAPH'04)*, 23(3):600–608, 2004.

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-2006-5-006	G. Kasnec, F.M. Suchanek, G. Weikum	Yago - A Core of Semantic Knowledge
MPI-I-2006-5-005	R. Angelova, S. Siersdorfer	A Neighborhood-Based Approach for Clustering of Linked Document Collections
MPI-I-2006-5-004	F. Suchanek, G. Ifrim, G. Weikum	Combining Linguistic and Statistical Analysis to Extract Relations from Web Documents
MPI-I-2006-5-003	V. Scholz, M. Magnor	Garment Texture Editing in Monocular Video Sequences based on Color-Coded Printing Patterns
MPI-I-2006-5-002	H. Bast, D. Majumdar, R. Schenkel, M. Theobald, G. Weikum	IO-Top-k: Index-access Optimized Top-k Query Processing
MPI-I-2006-5-001	M. Bender, S. Michel, G. Weikum, P. Triantafilou	Overlap-Aware Global df Estimation in Distributed Information Retrieval Systems
MPI-I-2006-4-010	A. Belyaev, T. Langer, H. Seidel	Mean Value Coordinates for Arbitrary Spherical Polygons and Polyhedra in $\mathbb{R}^3$
MPI-I-2006-4-009	J. Gall, J. Potthoff, B. Rosenhahn, C. Schnoerr, H. Seidel	Interacting and Annealing Particle Filters: Mathematics and a Recipe for Applications
MPI-I-2006-4-008	I. Albrecht, M. Kipp, M. Neff, H. Seidel	Gesture Modeling and Animation by Imitation
MPI-I-2006-4-007	O. Schall, A. Belyaev, H. Seidel	Feature-preserving Non-local Denoising of Static and Time-varying Range Data
MPI-I-2006-4-005	A. Belyaev, H. Seidel, S. Yoshizawa	Skeleton-driven Laplacian Mesh Deformations
MPI-I-2006-4-004	V. Havran, R. Herzog, H. Seidel	On Fast Construction of Spatial Hierarchies for Ray Tracing
MPI-I-2006-4-003	E. de Aguiar, R. Zayer, C. Theobalt, M. Magnor, H. Seidel	A Framework for Natural Animation of Digitized Models
MPI-I-2006-4-002	G. Ziegler, A. Tevs, C. Theobalt, H. Seidel	GPU Point List Generation through Histogram Pyramids
MPI-I-2006-2-001	T. Wies, V. Kuncak, K. Zee, A. Podelski, M. Rinard	On Verifying Complex Properties using Symbolic Shape Analysis
MPI-I-2006-1-007	H. Bast, I. Weber, C.W. Mortensen	Output-Sensitive Autocompletion Search

MPI-I-2006-1-006	M. Kerber	Division-Free Computation of Subresultants Using Bezout Matrices
MPI-I-2006-1-005	A. Eigenwillig, L. Kettner, N. Wolpert	Snap Rounding of Bzier Curves
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-2005-2-003	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-2005-1-008	C. Gotsman, K. Kaligosi, K. Mehlhorn, D. Michail, E. Pyrga	Cycle Bases of Graphs and Sampled Manifolds
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-2004-4-001	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