The 12th International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST (2011)
M. Dellepiane, F. Niccolucci, S. Pena Serna, H. Rushmeier, and L. Van Gool (Editors)

Integrated High-Quality Acquisition of Geometry and Appearance for Cultural Heritage

C. Schwartz, M. Weinmann, R. Ruiters and R. Klein

University of Bonn, Germany

Abstract

Current research trends demonstrate that, for a wide range of applications in cultural heritage, 3D shape acquisition alone is not sufficient. To generate a digital replica of a real world object the digitized geometric models have to be complemented with information pertaining to optical properties of the object surface. We therefore propose an integrated system for acquiring both the 3D shape and reflectance properties necessary for obtaining a photo-realistic digital replica. The proposed method is suitable for the digitization of objects showing the complex reflectance behavior, for example specularities and meso-scale interreflections, often encountered in the field of cultural heritage. We demonstrate the performance of our system with four challenging examples. By using Bidirectional Texture Functions, our structured light based approach is able to achieve good geometric precision while preserving tiny details such as scratches and engravings.

Categories and Subject Descriptors (according to ACM CCS): Computer Graphics [I.3.3]: Picture/Image Generation—Digitizing and scanning; Computer Graphics [I.3.7]: Three-Dimensional Graphics and Realism—Color, shading, shadowing, and texture; Image Processing and Computer Vision [I.4.1]: Digitization and Image Capture—Reflectance;

1. Introduction

The generation of digital replicas of real world objects remains a challenging issue in the fields of computer vision and computer graphics. Although it is still current practice to describe objects via geometry, possibly in conjunction with texture mapped pictures, in recent years the research goal has shifted towards obtaining digital reproductions that allow to render images that cannot be distinguished from images of the real object captured with a real camera under the same viewing and illumination conditions. This requires not only a high-quality reconstruction of the geometry of the acquired object, but also of its optical material properties such as reflectance. This is especially important in the field of cultural heritage (CH).

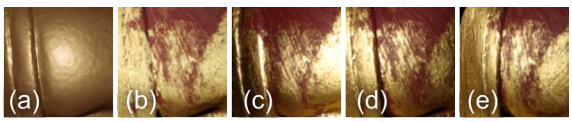


Figure 1: *Image-based representation of mesoscopic details. The 3D mesh (a) is missing several fine details. The texture (b) can reproduce some details. However, texture is limited to a single viewpoint and a single light-direction and hence details not seen in this particular combination will not be captured. In contrast, BTF-based representations allow variation in both light (c),(d) and view direction (d),(e). Note for example the changes in shading and highlights on the bumpy diagonal grooves in the gold leaf, that occur in (c), (d), (e).*

The appearance of an object, as it is seen by a human observer, depends on interactions of the incident light with the object, which are sorted into three categories [Mül09]: The macroscopic, the microscopic and the mesoscopic scale. On the *macroscopic* scale, 3D geometry is used to represent the shape of the object. Depending on the distance to the viewer its level of detail should be adjusted to the image resolution in such a way that the error along silhouettes remains strictly below the size of one pixel. At a *microscopic* scale, that is the scale where a local surface element can be assumed to be flat so that all incident light arriving at an opaque location is directly reflected without hitting the surface twice, the reflectance behavior is described by the *Bidirectional Reflectance Distribution Function* (BRDF). The BRDF describes the statistics of local surface reflectance, i.e. the probabilities for photons from an incoming direction

DOI: 10.2312/VAST/VAST11/025-032

ω_i to be reflected towards an outgoing direction ω_o, determining the view- and light-dependent intensity and shape of highlights. *Spatially varying BRDFs* (SVBRDFs) can be employed whenever these statistics exhibit a non-uniform behavior. In between these two extremes, the *mesoscopic* scale describes features that correspond to the size of a surface element which projects to about one pixel in image space. In most cases, these surface elements do not meet the requirements to be on the microscopic scale, because the they contain scratches, fibers or cracks.

Consider for example a scratch on the surface of the object that is slightly smaller than a pixel of the optical sensor used for acquiring reflectance. When lighting is changed, light transport inside the scratch, such as interreflection, scattering, self-shadowing or self-occlusion (masking), will change the appearance of the pixel drastically. In contrast to the microscopic case, where these changes can in general be described by smoothly varying lobe functions, approximated by the different BRDF models, the variations on the mesoscopic scale exhibit discontinuities depending on the local surface structure. Nevertheless, these variations on the mesoscopic scale under varying view- or light-directions are a visual clue for a human observer, who is still able to recognize these features even if they are below the pixel resolution in the image. To model these features [WHON97] coined the term *apparent BRDF* (ABRDF) to describe the reflectance behavior on the mesoscopic scale. However, due to the specific structure of the underlying surface elements, properties that hold for BRDFs, such as conservation of energy, reciprocity or smoothness do not hold for ABRDFs, which include effects such as masking, shadowing, interreflections or sub-surface scattering. This is especially important in the field of CH where each individual feature can still convey important information to the expert. Therefore, capturing this information so that it can be visualized afterwards is of great importance. On this point, one might argue that the best way would be to increase the resolution of the geometric reconstruction. However, the reconstruction of a sufficiently accurate 3D geometry is an ill-posed problem and would require impractical super-resolution approaches. Furthermore, storing such details in the geometry is comparatively expensive and would drastically increase complexity, and with it memory consumption and computational costs. Image-based approaches, in contrast, are capable of directly storing the ABRDF and thus handling both mesoscopic and microscopic appearance efficiently (see Figure 1).

In summary, the final goal of object acquisition has to be capturing the appearance of the object on all three scales; allowing the artifacts to be considered from arbitrary viewpoints under arbitrary illumination conditions in different virtual environments on their highest possible resolution, only limited by the resolution of the capturing device. Although much effort has gone into this problem, full appearance capture on all scales is still a challenging problem with no satisfactory solution. Currently, the standard approach is to capture geometry and reflectance in two separate steps. While there are techniques that provide good results either for geometry or reflectance behavior, combining them is often a challenging task, as it usually requires a difficult registration of the captured data. Therefore, an integrated setup that is capable of capturing high-quality geometry as well as view- and light-dependent reflectance properties of arbitrary 3D objects is desirable.

In this work, we present such an integrated, automatable, high-quality acquisition system for capturing both a highly detailed 3D geometry of the artifact and its mesoscopic and microscopic reflectance properties. Our contributions are:

- fully automatable, integrated capture of high-quality geometry and appearance for the creation of digital replicas
- highly accurate self-calibration of the acquisition device
- consequent exploitation of *high dynamic range* (HDR) data throughout the whole pipeline, allowing the system to deal with optically complicated objects
- substantial improvements in the accuracy and resolution of the digitized objects over prior work
- matrix factorization based hole filling approach for imputing missing data in sparsely sampled regions.

2. Previous Work

Most existing approaches for automatically generating digital 3D models only focus on capturing geometry without explicitly considering the photo-realistic depiction of the observed object, which in many cases heavily depends on the object's materials and reflection behavior. An in-depth survey of approaches for 3D object acquisition is given in [IKL*08]. Another survey on sophisticated techniques for appearance acquisition and representation, without considering the issue of geometry capture, is given in [WLL*09]. Similarly, in CH, there are also some attempts which focus on acquiring reflectance behavior alone, without capturing the geometry of an object. In 2001, [MGW01] proposed *Polynomial Texture Maps* (PTMs), which capture the reflectance of an object under a fixed viewpoint for varying lighting-directions. Further studies such as [EMM10, PCC*10] focus on the enhancement of surface details via fixed viewpoint *Reflectance Transformation Imaging*, which e.g. results in a more accurate perception of cuneiform inscriptions. In [HCD01], a light-dome setup similar to ours but without multiple viewpoints is proposed for capturing the *reflectance field* of an object. However, these approaches have the severe drawback, that objects can only be inspected from a fixed viewpoint and cannot be combined in a 3D scene, such as a virtual recreation of a historical site.

To facilitate a rendering from novel viewpoints, more recent work [GWS*09] suggests the use of multiple single-viewpoint PTMs or similar image-based reflectance information, such as *Hemispherical Harmonics* (HSH). Interpolation between views is then performed by using optical flow. However, using the optical flow basically means the implicit use of 3D-structure with doubtful quality and considerable regularization artifacts at object boundaries. In [MMC*08] it is hinted that it would be desirable to instead

use geometry information obtained from structured light for the viewpoint interpolation. In this paper, we also propose the use of structured light geometry but prefer *Bidirectional Texture Functions* (BTFs) [DNVGK97] for the reflectance information. These are well-suited to representing mesoscopic as well as microscopic effects, but also provide several advantages over PTMs in terms of compression, high-quality real-time visualization and interaction methods. For a more detailed discussion we refer to Section 6.

There have been previous approaches [FKIS02, MBK05] that also acquired BTFs on objects. Unfortunately, both methods rely on a shape-from-silhouette approach based on the reconstruction of a visual hull, which has severe limitations. The major drawback to using such a method is that most concave regions of an object cannot be reconstructed, even when using an infinite number of views (cf. Figure 4). This is problematic, as macroscopic occlusion, parallax and shadowing effects that were not captured by the visual hull's geometry are then shifted into the BTF, which is not well-suited for representing this kind of information. In [FKIS02], additional geometry was obtained with a laser-range-scanner to compensate for this weakness. However, the registration of the geometry poses a serious problem. In contrast, our approach will overcome this drawback by acquiring geometry and surface reflection behavior simultaneously. Recent investigation in using high-resolution photographs with 3D geometry to provide a consistent texture for the surface (e.g. [DCC*10]) may also be considered an attempt to acquire the surface reflectance. However, mesoscopic and microscopic information is only expressed by a single color, which (except for perfect lambertian objects) is not enough for a faithful photo-realistic reproduction.

Other existing approaches for the acquisition of geometry and reflectance advocate the use of SVBRDFs instead of BTFs [HLZ10, LKG*03, WMP*06]. While [LKG*03] acquires the geometry with a separate scanner, thereby introducing registration problems, [HLZ10] also uses structured light within the same measurement system to acquire the geometry, although this approach still requires the registration of the structured light measurements against each other and cannot find a global solution as proposed in this work. [WMP*06] follow a different approach by employing per-view photometric stereo to obtain geometry and achieve very good results with a specialized reflectance model for their distinct application of measuring appearance of human faces, combining SVBRDFs and additional sub-surface scattering measurements. However, fitting analytical BRDFs imposes strong assumptions about the materials that will distort the appearance of the digitized object (see Section 1). Additionally, in contrast to SVBRDFs, the BTF has the advantage that it inherently captures complex light transport in the mesoscale structure, due to its image-based nature. This eliminates the need to reproduce these light paths via a costly simulation, which would pose a problem when considering real-time rendering. An alternative approach would be to combine the setup in [HLZ10] with the capture of

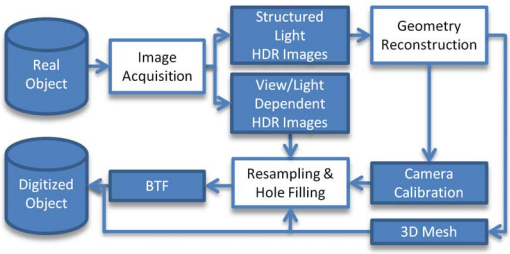


Figure 2: *Overview of our proposed processing pipeline for Cultural Heritage objects*

BTFs. However, for a reasonably dense sampling of the BTF a much larger number of measurements is required, leading to extremely high acquisition times. Also the individual measurements still need to be registered.

The method we build upon is the approach proposed in [MBK05]. The proposed setup with the multiple camera array and the use of BTFs is similar to our setting. As argued above, we use structured light to obtain a better 3D geometry. We found the structured light based super-resolution method proposed in [WSRK11] to be most suitable, as it allows high-quality reconstructions of objects even with complicated reflectance behavior (cf. Figure 4b) and can be easily combined with the setup proposed by [MBK05].

3. Overview

In this section, we will give a brief overview of the digitization pipeline that will create a digital model from a measurement of a real CH artifact (cf. Figure 2). The details for every step can be found in the consecutive sections.

From an application point of view, our pipeline requires the user to place the object that is to be digitized in our multi-view, multi-light setup (Section 4) and run the automatic image acquisition. The following processing steps can be fully automated and hence require no further user interaction.

We split the HDR images acquired in the first step into a set of geometry measurements, and reflectance data. Geometry measurements consist of images of the object captured under a structured light illumination and reflectance data consists of images of the object captured from different viewpoints under varying light positions. From the geometry measurement, we create a 3D mesh, i.e. our final geometry, by using the super-resolution structured light approach proposed by [WSRK11] (Section 5). Using the reconstructed 3D mesh and an improved camera calibration from the geometry step, we can project the view- and light-dependent HDR images onto a texture-atlas of the parameterized geometry (see Section 6). Furthermore, a resampling and hole filling approach is applied for gathering the reflectance samples for every point on the surface and creating a BTF representation. As a result we get a 3D mesh and the corresponding BTF that can now be rendered from arbitrary views under novel lighting conditions. Even real-time viewing is possible.

4. The Acquisition Setup

For capturing the reflectance information, ideally, we would need to observe the object from every possible view- and under every possible illumination-direction. This would require a setup with cameras and light-sources densely distributed on a full sphere.

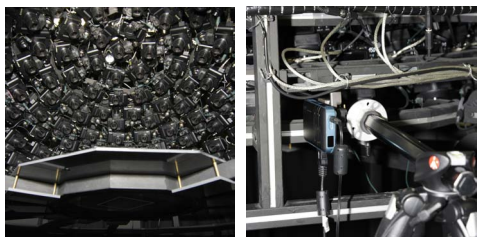
To reduce the complexity we approximate this setting with a hemispherical gantry (see Figure 3a) mounting $N = 151$ uniformly distributed consumer cameras whose flashes also serve as light-sources, first proposed in [MBK05]. In contrast to their setup, we used *Canon PowerShot G9* cameras with a considerably higher spatial resolution of 12 Megapixels and always take HDR sequences, i.e. increase the radiometric resolution. Furthermore, we propose the installation of L additional projectors on the gantry to project structured light patterns (see Section 5). In our experiments, we simulated this setup by placing one *Acer C20 Pico LED-Projector*, mounted on a tripod (see Figure 3b), at five to eight different positions: five 60° azimuthal steps on the height of the lowest camera ring, which is placed at a declination of 15°, and two or three positions from about 70° declination. This particular projector has a resolution of 848×480 pixels. To fully utilize this resolution, we project $P = 38$ different patterns.

In order to reliably capture objects that exhibit complicated reflectance behavior, such as highlights or interreflections, we employ HDR data for the reconstruction of geometry and reflectance. Thus, we need to acquire several exposure steps S for every captured HDR image (depending on the dynamic range of the object's material).

Our proposed setup allows the acquisition of extremely large amounts of data, e.g. 206,115 images for the Buddha dataset ($N \times N \times S = 151 \times 151 \times 3 = 68{,}403$ images for reflectance and $N \times L \times P \times S = 151 \times 8 \times 38 \times 3 = 137{,}712$ images for geometry), in about 3.7 hours (2.5 hours for the reflectance and 1.2 hours for the geometry). This is due to the massive parallelization of taking 151 pictures simultaneously. Although our angular-sampling is much denser (151×151), these timings are well within the region of other related work ([HLZ10]: 4.9h for 6×7, [LKG*03]: 2.5h for 15×20).

Please note that due to the parallelization, the object in question is exposed to a total of only $N \times S$ flashes, even though $N \times L \times S$ pictures are recorded. We estimated the harmfulness of our acquisition technique to organic materials by comparing the energy in the spectral bands of our employed flashes with an off-the-shelf 100W tungsten-halogen lamp. Our measurements indicate that one acquisition, using $151 \times 3 = 453$ flash triggerings, emits roughly as much UV light (integral over $287 - 400\,nm$ wavelength) as a few *seconds* of exposure to the lamp. Even when taking the energy over the complete visible spectrum into account, the situation is is similar.

Calibration. For the realistic depiction of the digitized object, a precise calibration of the measurement setup is of high importance. A fine-grained 3D reconstruction of the geometry of an object requires an accurate geometric camera calibration, i.e. intrinsic and extrinsic parameters of the camera projection. Such a precise calibration is obtained using the method proposed in [WSRK11] (see Section 5).



(a) cameras ***(b)*** projector tripod

Figure 3: *The measurement setup: (a) 151 cameras mounted on a hemisphere. (b) placement of the projector.*

In our case, however, we also need to consider the radiometric calibration, as we want to be able to combine the information gathered from 70,000 or more individual pictures of the reflectance measurement, each with slightly different color-characteristics, due to the CCD-sensors and flashes that vary from image to image.

The response curves of the sensors of the cameras are calibrated beforehand, since their color characteristics remain reasonably stable, whereas the radiometric attributes of the flashes have to be determined for every flash triggering. For that purpose we installed four *SphereOptics Zenith Ultra* reflectance standards with known BRDF on the sample-holder. The standards exhibit different albedos ($3\%, 10\%, 30\%$ and 97%) to allow the calibration of flashes with different intensities and are captured by the camera at 90° declination simultaneously with the object. It should be taken into account that the artifact casts a shadow onto the sample-holder. Therefore, we recommend placing redundant reflectors on all sides of the sample-holder.

5. Geometry reconstruction

Within our acquisition system, the reconstruction of the overall shape of the observed objects is performed using a structured light approach, since structured light has proven to produce accurate 3D models [Bla04]. Techniques based on structured light project a sequence of patterns onto the object, assigning every point on the surface a unique code. Recognizing these codes in one or more cameras provides correspondences, which allow the reconstruction of the surface points through triangulation. This has the additional advantage, that the geometry and the reflectance measurement are both undertaken using the same sensors, i.e. the camera-hemisphere, avoiding registration of the measurements.

To increase geometric details and obtain a very precise camera-calibration, we employ the structured light super-resolution framework proposed by [WSRK11]. This allows a robust acquisition of surfaces exhibiting complex reflectance behavior. We combine patterns from multiple projectors to generate larger codewords and thus obtain very dense and

(a) NextEngine *(b)* WSRK 2011 *(c)* MBK 2005

Figure 4: *Quality of different geometry reconstruction methods. We use [WSRK11].*

accurate correspondences. These correspondences are then used in a *Bundle Adjustment* (BA) approach that computes an optimized 3D point cloud and camera calibration. Finally, a triangle mesh is generated using Poisson surface reconstruction. Even though the Poisson reconstruction will produce a slightly smoothed surface, fine surface details will still be captured by the BTF representation (cf. Figure 1).

The framework is based on Gray codes and thus requires $\log_2(n)$ patterns for distinguishing n locations per projector and orientation (vertical or horizontal). Following the proposal of Weinmann et al., we use HDR imaging and inverse patterns to achieve a reliable classification of whether a pixel is illuminated or not, and hence, enable a robust decoding, which avoids misclassification. Thus, with our projector (see Section 4) a total number of $P = (\lceil \log_2 848 \rceil + \lceil \log_2 480 \rceil) \cdot 2 = 38$ different patterns is used.

As mentioned in Section 4 the geometric calibration of our measurement setup is obtained from the BA step. For computational complexity and accuracy reasons, only a subset of the given correspondences that we identified with high confidence is used.

6. Reflectance Acquisition

Since we want to be able to achieve accurate, photo-realistic results under novel viewpoints and illuminations, we need to capture and represent the reflectance behavior at every point under varying view- and light-directions. We use the BTF [DNVGK97], which represents spatial-, view- and light-variance in an image-based approach, avoiding the problems of explicitly modeling the mesoscopic and microscopic material properties. A BTF $\rho(x, y, \phi_i, \theta_i, \phi_o, \theta_o)$ is a 6D approximation of the 8D reflectance field, depending on the point on the surface $(x, y) \in S \subset \mathbb{R}^2$, the view-direction $(\phi_o, \theta_o) \in \Omega_l$ and the light-direction $(\phi_i, \theta_i) \in \Omega_l$ on a hemisphere Ω_l in a local tangent coordinate system of (x, y). For a more comprehensive overview, we refer to [MMS*04, WLL*09, FH09].

We choose a surface parameterized material representation over 4D image-space interpolation methods, for three reasons. Firstly, there is the aspect of data compression. When considering high frequency surface details, such as shadows or highlights, PTMs show a very poor reproduction quality (cf. Figure 6). HSHs perform slightly better, but it is a well known fact that a large number of coefficients are needed

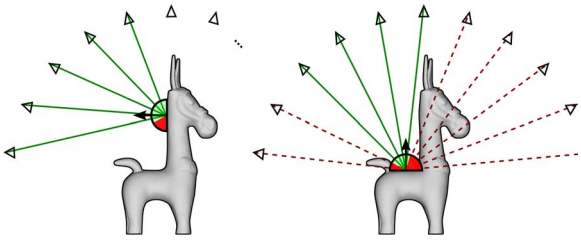


Figure 5: *Sketch illustrating the interpolation and hole filling problem. Left: missing data due to hemispherical setup. Right: missing data due to occlusion.*

for reproducing sharp features. Additionally, BTFs can utilize the high redundancy between view-directions, as parallax and occlusion is not an issue.

Secondly, for 3D surfaces the problem of data interpolation is a serious issue. Depending on the geometry of the object, there may be significant holes in the view- and light-sampling (cf. Figure 5). For rendering of novel view- and light-directions, simple real time capable interpolation schemes between single views will introduce artifacts. In contrast to this, in our approach, the task of interpolation and hole filling is performed as a preprocessing step, making more sophisticated techniques applicable.

Finally, if the surface of the virtual object should be deformed, e.g. for simulating waving cloths or flags or for illustrating interactions with an object, correct rendering is easily possible using a BTF representation.

Parameterization and Projection. Since the function is defined over a surface S and depends on the orientation of the local coordinate system, we need a consistent parameterization of the reconstructed geometry. Finding a good parameterization for a mesh is a field of research on its own. For our purposes, a parameterization should impose a low stretch and be reasonably fast for large meshes. We use ABF++ [SLMB05] for creating a parameterization and a texture-atlas.

Resampling and Hole Filling. To obtain a BTF, we first project the captured HDR images onto the parameterized surface. However, the images were taken under the *global* hemispherical camera setup Ω_g, while our BTFs are parameterized over a *local* hemisphere Ω_l. For efficient compression and rendering, the data has to be regularly sampled in the local coordinate system. Hence, we have to perform a resampling of the irregular samples obtained from the projections. Further problems arise from the fact that some of the directions in Ω_l may have been occluded by the object's geometry. As a result, the sampling over Ω_l contains holes, i.e. there was no meaningful reflectance data captured for these directions (cf. Figure 5). For this purpose, we apply a hole filling approach. We resample the available irregular samples into the target sampling Ω_l using *Radial Basis Functions* (RBF) (Shepard's method [She68] with a Gaussian weight-function). For every entry in the source sampling, a confidence, depending on visibility and steepness of the view- and light-directions, was computed.

(a) Texture *(b)* PTM *(c)* BTF *(d)* Ground truth

Figure 6: *Comparison between HDR-texture, PTM, compressed BTF and ground truth. (a) and (c) are tonemapped versions of pathtraced HDR renderings, (d) is a tonemapped HDR photography and (b) was created from tonemapped HDR images. While the spatial resolution is similarly good, reflectance properties are best preserved using the BTF. However, slight differences between BTF and ground truth occur in regions with large contributions of indirect illumination. These are discussed in Section 7.*

The confidence is also interpolated by the RBF. We draw a subset of RBF interpolated BRDFs with a high overall confidence. Using these representatives we construct a basis using *Non-negative Matrix Factorization* (NMF) [LS00] and project all RBF-interpolants into this basis using non-negative least squares optimization, considering the interpolated confidence as weights. To avoid artifacts in very sparsely sampled regions, we employ a data-driven regularization prior by punishing deviations from the mean reflectance, whereat the strength of the prior is chosen in inverse relation to the confidence of the data. The final BRDFs are then computed by blending the RBF-interpolation and the reconstructed NMF-projection according to confidence.

Compression. Finally, the BTF dataset has to be compressed to enable efficient storage, transfer and rendering. A large number of different techniques for BTF compression have been proposed [FH09]. We use a *Decorrelated Full-Matrix-Factorization* (DFMF) compression technique, since it offers a good quality at high compression ratios [Mül09]. Furthermore, the technique is also suited for real-time rendering on the GPU. However, since it is based on the minimization of the L^2-norm, the high dynamic range of specular objects poses a difficult challenge. Like Matusik et al. [Mat03], we addressed the problem by first computing the logarithm of the intensity values and computing the factorization for the dynamic-compressed data. Since we use the YUV color model to decorrelate the color from intensity, we apply the logarithm only to the intensity channel Y and use normalized color channels by computing U/Y and V/Y.

7. Results

We tested our proposed acquisition system using four challenging examples (cf. Figure 7):
First, there is a specular metal Donkey ($10 \times 4 \times 17\,cm$). Due to the structured light reconstruction, we were able to reach a high geometric accuracy (see Figure 4). Nonetheless, the surface of the object is covered with many small scratches and cracks with patina, which are not captured in the geometry but can be identified by their difference in reflectance.
The most challenging example is a Buddha idol ($6.7 \times 11.8 \times 13.3\,cm$) composed of painted wood and partially covered by gold leaf. At many spots on the Buddha the gold leaf has started to crumble and drop off, showing the underlying red paint and thus exhibiting extremely varying reflectance behavior. Furthermore, the gold leaf itself shows strong structure and cracks that are successfully captured by our proposed method (see also Figures 1 and 8).
Our third test case is a Minotaur figurine ($4.5 \times 4.5 \times 10\,cm$), which shows a rich composition of different materials. These materials are comprised of diffuse and specular components. More interestingly, the geometry has several concave parts which could not be acquired with the previous approach by [MBK05], based on a visual hull reconstruction.
Finally, we also captured a Samurai statuette ($6.5 \times 7.5 \times 22.5\,cm$), made of dark terracotta. This object exhibits many fine geometric features, which were successfully captured by our system. Although at first glance the statue appears to be lambertian, it shows notable specular highlights under gazing angles, which would have been missed when using a simple texture representation.
All models were compressed using DFMF with a very precautious quality setting. Using this data we are able to off-line render novel images of arbitrary scenes in very high quality (see Figure 7) or even provide interactive real-time viewing on the GPU, still depicting the objects photo-realistically. With this representation, even web-based dissemination is easily possible with moderate bandwidth requirements [SRWK11]. Datasets presented in this paper and an interactive web-based viewer can be found on `http://btf.cs.uni-bonn.de`.

Figure 7: *Multiple objects from different measurements, composed in one virtual scene, rendered with full simulated light transport (pathtracing). All objects are to scale.*

Discussion. We have demonstrated the capabilities of our approach on a variety of challenging examples and shown that it is capable of faithfully reproducing the appearance of a CH artifact. Furthermore, we now discuss the computational costs, taking the definition of the measurement setup

into consideration. The main computational bottleneck can be identified in the resampling and compression steps that are executed to create the BTF representation. For the resampling of a 4.2 Megapixel BTF the processing time was about 25 hours, using an Intel Xeon processor with 2.33 GHz and 8 cores. Using a parallelizable FMF compression [RRK09] on the same computer takes about eight hours. However, the resampling as well as the compression can be performed in parallel on different machines, rendering the processing extremely scalable. The biggest issue is the massive amount of data which makes the use of out-of-core algorithms mandatory. However, we expect this to become more manageable in the future, with new technologies such as fast Solid State Disks becoming affordable and available with high storage capacities. Due to the vast parallelization capabilities, one could also consider using cloud services.

Regarding the measurement definition, the total achievable physical accuracy in terms of geometry and BTF resolution depends on the size of the object, since we utilize the zoom-lenses to cover different sized artifacts. Using similar cameras and projectors, [WSRK11] were able to achieve an RMSE of 23.3μm on a test object with ground truth geometry. Considering the Buddha with dimensions of roughly 6.7 $cm \times 11.8\,cm \times 13.3\,cm$ and a surface area of about 287 cm^2, the resolution of our cameras would allow for a 96 μm sampling of the surface. Probably, this limitation of camera resolution could also be relaxed by a super-resolution approach. However, even with the 96 μm accuracy, we need a texture-atlas of at least 3.1 Megapixel resolution. Unfortunately, due to the low-distortion parameterization, which leads to a partially empty atlas, we would need about 6.2 Megapixels resolution for the texture-map to exploit the full potential of the measurement setup. The objects tested within the scope of this paper were acquired using 4.2 Megapixel texture-maps (1 pixel $\approx 117\,\mu m$), resulting in uncompressed datasets of 534.4 GB. After DFMF compression with the generous amount of 100 components, the datasets are 1.59 GB in size. To facilitate mass storing and transmission, the data size can be further reduced, e.g. to 780 MB using a consecutive LZMA compression. Note that, for the purpose of rendering, the data needs to be unpacked in memory to the DFMF representation. Figure 8 demonstrates the quality of the compressed 4.2 Megapixel Buddha dataset.

Limitations We have to point out that the class of objects that can be represented with the chosen BTF approach has its limits. Transparent objects, perfect mirroring surfaces and surfaces exhibiting strong subsurface scattering effects cannot be faithfully represented this way. Similar constraints also apply for the structured light reconstruction approach. However, our proposed system is robust enough to deal with objects that exhibit specularity (Buddha, Donkey). The geometry of optically complicated materials like the marble of the Minotaur can be reconstructed and convincing renderings under distant illumination can be generated from the BTF data, although subsurface scattering effects were not captured explicitly.

Figure 8: *The Buddha with the compressed BTF rendered under point light illumination. The insets show detail enlargements to demonstrate how well surface and material features are preserved.*

When considering arbitrary global illumination, our proposed approach will introduce some systematic errors, due to the fact that macroscale interreflections are still captured within the image-based BTF representation. Such interreflections are also simulated by the global illumination algorithm, resulting in a duplicate integration of these light paths and thus, a too bright appearance (see Figure 6). Using a setup with projectors as light sources, such interreflections can be eliminated at capture time (see [HLZ10]). However, this problem could also be handled in a post-processing step by removing the geometry induced interreflections from the BTF as proposed for heightfield geometry in [RK09], which is a task for future work. Another limitation with respect to real-time rendering is the fact that with the here utilized compression approach, larger scenes consisting of multiple objects will not fit into the memory of todays graphics hardware. This could be tackled by converting the data into a more compact representation for such purposes.

Finally, the high-quality results presented in this work come at the cost of relying on a rather complex and expensive capturing setup, which might hinder the widespread use of this system. However, as there is a certain demand for high-quality virtual replicas, building or renting such a setup could be worthwhile.

8. Conclusions and Future Work

In this paper, we presented an improved acquisition setup and processing pipeline capable of capturing a highly detailed geometry along with the extensive view- and light-dependent reflectance information. By capturing appearance at different scales and representing it as BTF we facilitate the creation of digital replicas of real world objects that can be viewed from arbitrary directions and illuminated by arbitrary illumination patterns. The proposed acquisition setup is completely automatable and requires no manual steps or user guidance. We furthermore tested our approach on a set of objects which exhibit different challenging characteristics in terms of geometry or reflectance behavior, which we managed to obtain faithfully. The consequent usage of HDR data throughout all the steps of our pipeline, although introducing new issues that needed to be addressed, allows for a robust high-resolution reconstruction of even specular ob-

jects. Our presented results show a considerable amount of detail that was not achieved before using a reflectance capturing approach. Furthermore, the results demonstrate that, after acquisition and processing, BTFs can be used with similar ease as conventional textures or SVBRDF models, while allowing a much more precise digital reproduction of real world objects.

In this work, we only captured directions on one hemisphere. For a faithful reproduction, it is however necessary to have reflectance data for the whole sphere. The acquisition of this data would easily be possible with our setup by turning the object and taking multiple measurements. We did not explore this option in the scope of this work, but instead relied on the hole filling technique to fill in the missing data. However, we consider capturing this data in the future, since even when rendering the object from viewpoints that lie on the captured hemisphere, reflectance information for the missing angles are needed for global illumination computations and thus for a faithful rendering of the object. Still, there will always be a demand for hole filling and interpolation. Therefore, we would like to extend our approach and also consider spatial neighborhoods on the surface of the object to further improve the visual quality. Moreover, we see the removal of macroscale interreflection in the BTF as a relevant direction for future work.

Acknowledgements. We would like to thank Martin Rump for his invaluable help with the BTF resampling, compression and rendering and Michael Zschippig for taking the measurements. The research leading to these results was partially funded by the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 231809; 2008-2012 and by the German Science Foundation (DFG) under research grant KL 1142/4-1.

References

[Bla04] BLAIS F.: Review of 20 years of range sensor development. *Jour. of Electronic Imaging 13*, 1 (2004), 231–243. 4

[DCC*10] DELLEPIANE M., CALLIERI M., CORSINI M., CIGNONI P., SCOPIGNO R.: Improved color acquisition and mapping on 3D models via flash-based photography. *ACM Jour. on Comp. and CH 2*, 4 (2010), 1–20. 3

[DNVGK97] DANA K. J., NAYAR S. K., VAN GINNEKEN B., KOENDERINK J. J.: Reflectance and texture of real-world surfaces. In *CVPR* (1997), pp. 151–157. 3, 5

[EMM10] EARL G., MARTINEZ K., MALZBENDER T.: Archaeological applications of polynomial texture mapping: analysis, conservation and representation. *Jour. of Arch. Sci. 37* (2010). 2

[FH09] FILIP J., HAINDL M.: Bidirectional texture function modeling: A state of the art survey. *PAMI 31* (2009), 1921–1940. 5, 6

[FKIS02] FURUKAWA R., KAWASAKI H., IKEUCHI K., SAKAUCHI M.: Appearance based object modeling using texture database: acquisition, compression and rendering. In *EGRW* (2002), pp. 257–266. 3

[GWS*09] GUNAWARDANE P., WANG O., SCHER S., RICKARDS I., DAVIS J., MALZBENDER T.: Optimized image sampling for view and light interpolation. In *VAST* (2009), pp. 93–100. 2

[HCD01] HAWKINS T., COHEN J., DEBEVEC P.: A photometric approach to digitizing cultural artifacts. In *VAST* (2001), pp. 333–342. 2

[HLZ10] HOLROYD M., LAWRENCE J., ZICKLER T.: A coaxial optical scanner for synchronous acquisition of 3d geometry and surface reflectance. In *SIGGRAPH* (2010), pp. 99:1–99:12. 3, 4, 7

[IKL*08] IHRKE I., KUTULAKOS K. N., LENSCH H. P. A., MAGNOR M., HEIDRICH W.: State of the art in transparent and specular object reconstruction. In *Eurographics STAR* (2008), pp. 87–108. 2

[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 ToG 22* (2003), 234–257. 3, 4

[LS00] LEE D. D., SEUNG H. S.: Algorithms for non-negative matrix factorization. In *NIPS* (2000), pp. 556–562. 6

[Mat03] MATUSIK W.: *A data-driven reflectance model*. PhD thesis, Massachusetts Institute of Technology, 2003. 6

[MBK05] MÜLLER G., BENDELS G. H., KLEIN R.: Rapid synchronous acquisition of geometry and BTF for cultural heritage artefacts. In *VAST* (2005), pp. 13–20. 3, 4, 6

[MGW01] MALZBENDER T., GELB D., WOLTERS H.: Polynomial texture maps. In *SIGGRAPH* (2001). 2

[MMC*08] MUDGE M., MALZBENDER T., CHALMERS A., SCOPIGNO R., DAVIS J., WANG O., GUNAWARDANE P., ASHLEY M., DOERR M., PROENCA A., ET AL.: Image-based empirical information acquisition, scientific reliability, and long-term digital preservation for the natural sciences and cultural heritage. *Preservation 30* (2008). 2

[MMS*04] MÜLLER G., MESETH J., SATTLER M., SARLETTE R., KLEIN R.: Acquisition, synthesis and rendering of bidirectional texture functions. In *EG STAR* (2004), pp. 69–94. 5

[Mül09] MÜLLER G.: *Data-Driven Methods for Compression and Editing of Spatially Varying Appearance*. PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, 2009. 1, 6

[PCC*10] PALMA G., CORSINI M., CIGNONI P., SCOPIGNO R., MUDGE M.: Dynamic shading enhancement for reflectance transformation imaging. *J. Comput. Cult. Herit. 3* (2010), 6:1–6:20. 2

[RK09] RUITERS R., KLEIN R.: Heightfield and spatially varying brdf reconstruction for materials with interreflections. *Computer Graphics Forum 28*, 2 (2009), 513–522. 7

[RRK09] RUITERS R., RUMP M., KLEIN R.: Parallelized matrix factorization for fast btf compression. In *EGPGV* (2009), pp. 25–32. 7

[She68] SHEPARD D.: A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 23rd ACM national conference* (1968), pp. 517–524. 5

[SLMB05] SHEFFER A., LÉVY B., MOGILNITSKY M., BOGOMYAKOV A.: Abf++: fast and robust angle based flattening. *ACM Trans. Graph. 24* (2005), 311–330. 5

[SRWK11] SCHWARTZ C., RUITERS R., WEINMANN M., KLEIN R.: Webgl-based streaming and presentation framework for bidirectional texture functions. In *VAST* (2011). 6

[WHON97] WONG T.-T., HENG P.-A., OR S.-H., NG W.-Y.: Image-based rendering with controllable illumination. In *EGWR* (1997), pp. 13–22. 2

[WLL*09] WEYRICH T., LAWRENCE J., LENSCH H. P. A., RUSINKIEWICZ S., ZICKLER T.: Principles of appearance acquisition and representation. *Foundations and Trends in Computer Graphics and Vision 4* (2009), 75–191. 2, 5

[WMP*06] WEYRICH T., MATUSIK W., PFISTER H., BICKEL B., DONNER C., TU C., MCANDLESS J., LEE J., NGAN A., JENSEN H. W., GROSS M.: Analysis of human faces using a measurement-based skin reflectance model. In *SIGGRAPH* (2006), pp. 1013–1024. 3

[WSRK11] WEINMANN M., SCHWARTZ C., RUITERS R., KLEIN R.: A multi-camera, multi-projector super-resolution framework for structured light. In *3DIMPVT* (2011). 3, 4, 5, 7