A Kaleidoscopic Approach to Surround Geometry and Reflectance Acquisition

Ivo Ihrke[1,2]   Ilya Reshetouski[1,2]   Alkhazur Manakov[1,2]   Art Tevs[2]

Michael Wand[1,2]   Hans-Peter Seidel[2]
[1]Saarland University   [2]MPI Informatik

Abstract

We describe a system for acquiring reflectance fields of objects without moving parts and without a massively parallel hardware setup. Our system consists of a set of planar mirrors which serve to multiply a single camera and a single projector into a multitude of virtual counterparts. Using this arrangement, we can acquire reflectance fields with an average angular sampling rate of about 120+ view/light pairs per surface point. The mirror system allows for freely programmable illumination with full directional coverage. We employ this setup to realize a 3D acquisition system that employs structured illumination to capture the unknown object geometry, in addition to dense reflectance sampling. On the software side, we combine state-of-the-art 3D reconstruction algorithms with a reflectance sharing technique based on non-negative matrix factorization in order to reconstruct a joint model of geometry and reflectance. We demonstrate for a number of test scenes that the kaleidoscopic approach can acquire complex reflectance properties faithfully. The main limitation is that the multiplexing approach limits the attainable spatial resolution, trading it off for improved directional coverage.

1. Introduction

The reflectance field as introduced by Debevec et al. [2] is an eight-dimensional function describing the response of a scene to incident illumination. It encompasses all potential views of an object, its light field [10], under arbitrary incident illumination, the incident light field [12]. So far, only subsets of the reflectance field have been acquired, usually for a static view point [2, 12, 3] or static illumination [10]. Capturing that encompasses both, view and illumination variation, is currently restricted to a narrow set of viewpoints and lighting positions [4]. If the object geometry and/or a model of the reflectance behavior of the surface are known, the dimensionality of the sampling problem reduces considerably [9]. Lensch et al. [9] selected a multi-lobe Lafortune [7] model as their reflectance model. Later work, referred to as *reflectance sharing*, mitigated the manual selection of a parametric BRDF model [20, 18].

Recent work has demonstrated the capability of estimating the object geometry in conjunction with its spatially varying BRDF [6, 15]. Holroyd et al. [6] use an optical setup that exploits Helmholtz reciprocity [19] while mechanically repositioning the apparatus around the object to cover the required view and illumination directions. Schwartz et al. [15], on the other hand, employ a massively hardware parallel setup consisting of 151 point-and-shoot cameras arranged in a hemispherical dome to sample the reflectance field densely.

In this paper, we investigate an alternative method for dense reflectance sampling. Our goal is to avoid the cost and complexity of parallel hardware setups, as well as the limitations of mechanically moving cameras and/or light sources. We therefore propose a multiplexing approach that folds full spherical acquisition and lighting into a single, fixed-device setup: We explore the use of kaleidoscopic mirror systems to densely sample the reflectance of an object while simultaneously acquiring data for surface reconstruction.

Mirror systems have previously been employed for generating virtual camera-projector systems. Closest to our work are the apparatus employed by Garg et al. [4], and Han and Perlin's work on BTF measurements by use of kaleidoscopes [5]. Garg et al. sample a spatially restricted portion of the reflectance field using two sets of coaxial camera-projector pairs multiplied by a set of 16 mirrors to generate 32 virtual projectors and cameras. Their setup avoids the complexity of mirror inter-reflection and they do not estimate the object geometry. Han and Perlin, on the other hand, exploit inter-reflections to generate a large number of virtual illumination and viewing directions using a single projector-camera pair. Their work is restricted to sample the reflectance of planar samples. Another setup exploiting mirrors for achieving virtual view distributions for surround structured light scanning is presented by Lanman et al. [8]. They use a two-mirror setup and a carefully placed orthographic projection unit to avoid the super-positioning of active light projection patterns. This allows for the estimation of object geometry. The authors, however, do not sample reflectance.

We propose a system that can identify the virtual view and projection directions in the camera and projector images, respectively. This enables the use of structured light scanning within arbitrary camera-projector-mirror systems. In our experiments, we achieve approximately 250×150 virtual camera-projector pairs, a number comparable to a dedicated hardware setup [15]. The virtual cameras and projectors cover most of a full sphere surrounding the object. The inherent directional multiplexing, however,
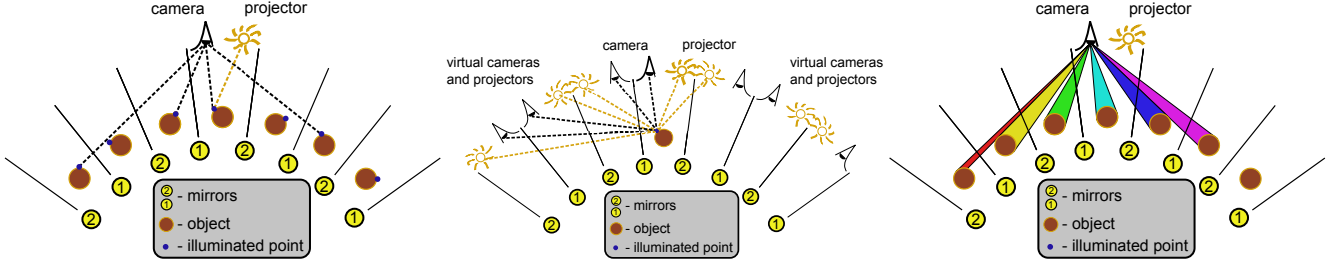
Figure 1. Illustration of imaging and projection inside a mirror system.

results in a rather low spatial resolution both for projection and for imaging. Calibration efforts as compared to a hardware parallel system are similarly complex but on different levels: As an example, radiometric calibration needs to estimate the parameters of the mirror system instead of individual response curves for each hardware device. A similar argument holds for the geometric calibration.

2. Operation Principle

Our system is build around imaging and projection within systems of planar mirrors. It exploits the multitude of reflections and inter-reflections to compress a full surround field of view into the aperture of a single camera and/or projector. The basis for our design is the recent analysis of planar mirror systems by Reshetouski et al. [14]. Due to Helmholtz reciprocity, their analysis applies to cameras as well as projectors with one difference: If we want to avoid computational recovery of raw reflectance data, as e.g. in [16], we have to avoid the superposition of light due to different mirror reflections of the light source.

2.1. Point Scanning in Planar Mirror Systems

Consider a two-dimensional mirror system as shown in Fig. 1. On the left, we illustrate how a projector is illuminating a single point on the surface of an object. By unfolding of the system [14], i.e. by mirroring the world instead of the ray, we can visualize the mirror world with a number of virtual objects. The illuminated point can be considered to be present on all copies of the object. The camera can then observe those illuminated points that are not occluded. The four points visible to the camera are e.g. generated by the reflection sequences (from left to right) $\{1, 2, 1\}$, $\{1\}$, $\{\}$, and $\{2, 1\}$. These numbers indicate the mirrors that are being traversed by the ray before it hits the object.

Alternatively, we can consider the system as consisting of virtual cameras and projectors observing the real object as shown in the middle of Fig. 1. In our example, the virtual cameras observe the illuminated point via the reflection sequences (from left to right) $\{1, 2, 1\}$, $\{2, 1\}$, $\{1\}$, and $\{\}$. As expected, these sequences are the same as in the previous case. This alternative interpretation immediately shows that a pair of reflectance values is sampled simultaneously: For one projector illumination direction, we obtain reflectance samples from four viewing directions. Furthermore, since multiple viewing rays are available that observe the same object point, we can triangulate the point and obtain part of the object geometry. Finally, the same object point can be illuminated by different

virtual projectors, one at a time. This results in $N_k \times M_k$ reflectance samples per object point $\mathbf{x}_k$, where N_k is the number of unoccluded views and M_k the number of unoccluded illumination directions. In our example, this yields 20 reflectance samples for the illuminated point.

Of course, point sampling an object is inefficient. It would require the acquisition of $\mathcal{O}(K \times M)$ images for sampling K surface points, and $M = \max_k M_k$. Key to our method is the use of structured light in the present context. The main challenge is that we have to avoid illuminating a surface point from more than one direction in a single image of the scan since this would lead to the superposition of light and thus to integral reflectance samples.

2.2. Structured Light in Planar Mirror Systems

As shown by Reshetouski et al. [14], a camera image of a planar mirror system can be decomposed into a number of sets of image coordinates for which a certain virtual camera view is valid, Fig. 1 (right). We call these image regions camera chambers, or short, C_c. A virtual camera in a virtual multi-view system is then given by this set of image coordinates in conjunction with the calibration parameters that define the projection operation $\pi_c : \mathbb{R}^3 \mapsto \mathbb{R}^2$ of the virtual camera $V_c^n := \{C_c^n, \pi_c\}$. Applying the Helmholtz reciprocity principle, we describe virtual projectors, similar to virtual cameras, as a set of coordinates of the original projector (not shown). A virtual projector is given by a projector chamber C_p, i.e. a set of projector coordinates for which the virtual projector is valid, in conjunction with its calibration parameters that define the projection operation $\pi_p : \mathbb{R}^3 \mapsto \mathbb{R}^2$ for the virtual projector $V_p^m := \{C_p^m, \pi_p\}$.

By illuminating only a single projector chamber at a time, we are able to decompose reflectance acquisition into non-overlapping regions of illumination.

The camera and projector chambers enable the decomposition of imagery, obtained from projecting light into a planar mirror system, into different light/direction pairs. Each combination of camera and projector chambers gives rise to a unique viewing/lighting direction pair. We are thus able to sample the reflectance field of an object using $N \times M$ (sub-)images. In practice, we achieve an average of about 120 unoccluded view/light directions per surface point. Without exact surface geometry which serves as an interpolation guide, parallax effects, both for view and lighting interpolation become apparent. We therefore opt for structured light projection which enables geometry reconstruction on top of reflectance sampling.

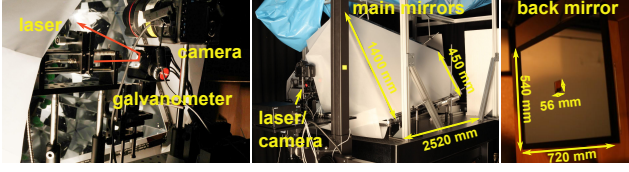Before we describe the actual implementation of our

Figure 2. Images of our acquisition system: *Left:* the laser/camera system, the orange line indicates a laser ray that can be steered inside the mirror system. *Middle:* an outside view of our system with dimensions. *Right:* the back mirror is attached to the back entrance of the system once an object is mounted inside.
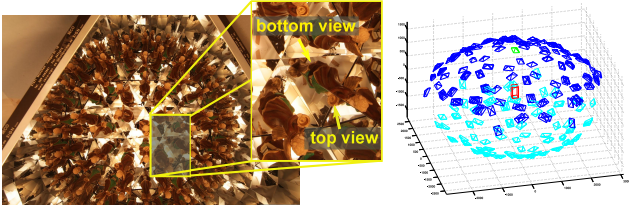

Figure 3. *Left:* Image of an object inside our acquisition system. The inset shows that both, views from the top and views from below are generated by our system. Side views from different directions and at different elevations are visible towards the boundary of the image. *Right:* Virtual camera distribution in our setup. The red box is the bounding box of the object, blue cones are virtual cameras in the upper hemisphere, cyan ones are located in the lower hemisphere. The camera marked in green is the real camera location. Units are given in mm.

system, we first discuss the hardware design that was used for our experiments.

3. Hardware Design

Our hardware setup consists of four planar mirrors, one digital SLR camera, and one RGB laser projector with a double axis mirror galvanometer for control. Next we discuss our design choices regarding these components.

Mirror System: We employ a kaleidoscopic mirror system consisting of four mirrors, three of which are arranged in the classical configuration of a truncated three-sided pyramid [5]. This type of system can only generate hemispherical view distributions which is sufficient to perform exhaustive reflectance sampling if the sample is flat. However, non-flat geometry requires sampling of the full sphere of viewpoints in order to obtain dense coverage of both, surface geometry and reflectance. To enable spherical sampling, we add one additional mirror at the back entrance of the system which generates reflections that also show the object from below, see Fig. 3. The three main mirrors are optical quality first surface foil mirrors stretched on an aluminum frame. The mirror on the back side is a glass mirror with a hole cut by a water jet. This hole enables an object holder assembly to protrude inside the system. Images of our system are shown in Fig. 2.

Projector: For our projector we use a system of coaxial red, green, and blue lasers with approximately 5mW output power per channel. The choice of using a laser system is motivated by the much larger depth-of-field of a laser beam as compared a standard digital projector: Since our system folds the rays into the mirror system, distances
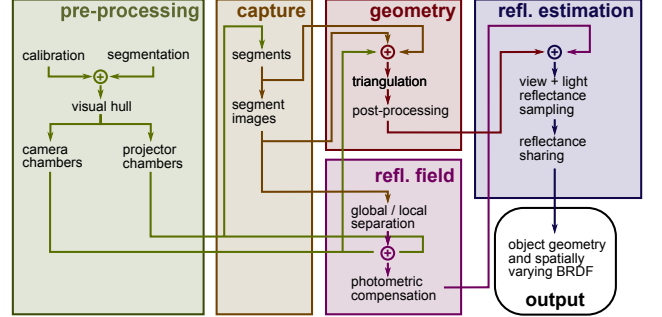

Figure 4. Overview of our reconstruction pipeline. The individual steps are covered in detail in Sect. 5.

between the apparent closest object (no reflection) and the farthest virtual objects are quite significant (about 4m). Since we need to be able to control the scan in both, x- and y-directions we are using a two-mirror galvanometer scanner. In order to ensure uniform illumination within one scan line, we use a constant angular scan speed of the projector in one axis. Careful adjustment of the system geometry allows for planar sheets to be scanned in this (horizontal) coordinate [11]. Reflectance data is inherently HDR. Exposure bracketing cannot be used to perform HDR imaging in case of laser scanning since the illumination is non-stationary. We therefore choose to generate the different exposures by scanning the same scan line multiple times. We take four exposures spaced by a factor of 4 between consecutive pairs of images. To cover the large field of view inside the mirror system, we employ a wide angle two-axis mirror galvanometer (Thorlabs GVS012) with an opening angle of 80 degrees in both dimensions. The projector is mounted close to the entrance of the mirror system to make full use of the aperture of the mirror system, i.e. no rays miss the opening at the front side.

Camera: We employ a Canon 5D mark II DSLR equipped with a Canon EF 14mm f/2.8 L II USM lens. This lens has a large field of view and comparatively minor radial distortion and chromatic aberrations. As a compromise between light efficiency and depth-of-field, we use an aperture setting of $f/14$. The camera is mounted such that it is close to the entrance of the mirror system while simultaneously avoiding to image parts of the laser projector.

4. System Overview

Here we present the processing pipeline, it will be detailed in Sect. 5, see also Fig. 4.

In practice, we are performing laser stripe scanning. Putting aside the complexities of the mirror system, we can see the system as one consisting of multiple cameras (typically on the order of 200), and multiple laser projectors (about 100), for details see Table 3. The cameras operate in parallel while the laser projectors project their stripes in sequence. A plot of the view point distribution is shown in Fig. 3. In order to interpret our system as a multi-camera/multi-projector system, we first have to compute the camera chambers $C_c^n, n = 1..N$ and projector chambers $C_p^m, m = 1..M$, Sect. 5.2, which, in turn is based on the geometric calibration of our system, Sect. 5.1. For

proper reflectance sampling, it is also necessary to perform a radiometric calibration of the camera and the mirrors.

Once the virtual projectors $V_p^m, m = 1..M$ are known, we determine a sequence of projector *segments* for capturing. Because the segments are limited by the valid projector coordinates, they have the property that their corresponding 3D planes intersect the scanned object from a single direction only. The capturing process then consists of taking HDR images for the segment list of each virtual projector. Each acquired image is thus a full spherical light field view of the intersection of a world space illuminating plane with the object.

We use the resulting data for geometry and reflectance estimation. The geometry is computed by multi-view triangulation of the laser stripe, Sect. 5.4. The reflectance is estimated in a reflectance sharing approach, Sect. 5.6. We generate the reflectance samples by first assembling a reflectance field from the stripe data, Sect. 5.5, and then sampling it using the estimated object geometry to resolve occlusion and self-occlusion both for the viewing and the lighting directions. The result of this approach is an object geometry with per-surface point BRDF described as a linear combination of basis BRDFs that have been estimated from the data as e.g. in [20, 18].

5. Implementation

5.1. Calibration

Geometric Calibration: We perform geometric calibration by first estimating the camera intrinsics using Bouguet's MATLAB calibration toolbox [1]. We then estimate the camera position and orientation as well as the mirror planes of the three main mirrors by imaging a checkerboard placed at different depths inside the mirror system [14]. Next, the laser projector is calibrated for its internal parameters using the model and procedure of Manakov et al. [11]. The remaining task is to improve the position and orientation estimate of the laser projector with respect to the mirror system and the camera. For this, we place a checkerboard inside the mirror system and illuminate a set of projector coordinates, taking an image for every illuminated point. Since the camera/mirror system is calibrated already, the checkerboard allows for the computation of the 3D coordinates of the illuminated point. These are usually in disagreement with the prediction produced by the initial calibration of the laser system. We perform an optimization on the laser intrinsic and extrinsic parameters to improve this prediction. In a final step, we perform a bundle adjustment on the parameters of all system components. For successful calibration, the whole "atmosphere" of the "planet" produced by the mirror system should be sampled by checkerboard positions and projected laser points. We typically use a checkerboard at 5 different depths and illuminate on the order of 30 projector chambers with 5 points each.

The back mirror is calibrated for every scanned object independently. This is necessary since it must be removed in order to insert a different object into the system. We estimate its position after an object is inserted. For this purpose, our mirror has a number of small marks in the corners
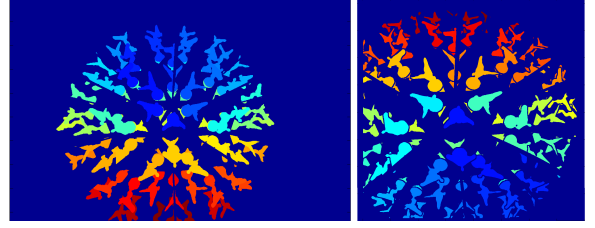

Figure 5. Visualization of the camera chambers (left) and projector chambers (right). Different colors correspond to different chambers. The plot simultaneously shows the viewpoint and lighting direction coverage of our system.

which can be triangulated by multi-view triangulation.

Radiometric Calibration: Since our laser source is only roughly white, we compute an affine RGB correction matrix using a ColorCheckerSG chart positioned inside the mirror system and illuminated by laser "white light". This color transformation is applied to all captured images. We estimate the attenuation factor of the mirrors by imaging a Spectralon target, inside the system. To produce hemispherical illumination for this task, we mount diffusive paper at the front entrance of the mirror system and illuminate it with a strong diffuse light source. The inter-reflections inside the mirror system distribute the light over the full hemisphere. Since the mirroring sequence for every virtual camera is known from the system geometry, the individual mirrors' attenuation coefficients can be estimated from observed products of these factors on the Spectralon patch as it is visible in different camera chambers. The three main mirrors were estimated to have a reflectance of $0.84, 0.86$, and 0.85, respectively. The rear mirror has an attenuation coefficient of 0.9. These values are constant for the different color channels.

5.2. Pre-Processing

Once the system is calibrated, we proceed to compute the chambers for the virtual cameras and projectors. The computation is based on the segmentation of a flood-lit image of the object positioned inside the mirror system. Using this image, the visual hull of the object can be computed [14]. Its geometry is then used to determine the camera and projector chambers. We perform the actual computation of the chambers by ray-tracing until the ray intersects the visual hull, recording the mirror intersections that occur along the way. Pixels with equal mirror sequences belong to a common chamber and are assigned an ID. This implies that a pixel can only belong to one camera or projector chamber, respectively. The procedure is the same for the virtual cameras and the virtual projectors. The result can be visualized in an image as shown in Fig. 5.

5.3. Capture

In preparation of a capture session, we scan-convert the projector chambers into horizontal scan-lines which we refer to as segments. Due to the geometry of our galvanometric scanning system, this ensures that the scanned laser sheets correspond to three-dimensional plane segments. Restriction of the scan-lines to projector chambers guaran-

tees that object points are illuminated only from a single direction and that the plane segment does not split up regardless of the mirroring sequence traversed until the object is hit. Each virtual projector is assigned a list of segments. We capture one HDR image $I_{m,s}(x,y)$ for each segment s of all virtual projectors V_p^m. The segments simply consist of the start and end horizontal projector coordinates $[x_{start}, x_{end}]$ and one vertical coordinate which allows for the computation of the corresponding 3D projector plane. The total number of segments is typically on the order of 20.000, see Table 3. Each HDR exposure requires 4 images. We use a sensor resolution of 2817×1876 pixels, utilizing the RGGB Bayer pattern of our camera as one larger pixel.

5.4. Geometry Computation

Our triangulation scheme solves two problems: First, it allows for triangulation of the object surface from M views simultaneously, and second, it integrates the multi-view measurements into a coherent watertight surface reconstruction.

Posed more abstractly, our task is to infer an unknown piece of geometry $S \subset \mathbb{R}^3$ from images $I_{m,s}$ that show the intersection of the geometry by laser sheets. This is a complex problem, as our unknown is a whole piece of geometry S, which is characterized by a high-dimensional parameter space (in our case: a voxelized indicator function within a cuboid). In addition, we have to account for noise: Uncertainty in calibration, varying angles between virtual cameras and lasers, as well as varying distances to the object (between $1.5m - 6m$) yield information on the surface position of strongly varying accuracy, which needs to be fused appropriately in the overall estimate of S.

We employ a statistically motivated reconstruction technique that operates in two steps: First, we compute an occupancy volume from the images that encodes the likelihood and uncertainty of the presence of surfaces in space. In a second step, we employ the segmentation algorithm by Unger et al. [17] to compute a globally optimal minimal surface that separates the interior and exterior of the object.

Occupancy volume: The occupancy volume assigns for every point $\mathbf{x} \in \mathbb{R}^3$ a likelihood that the measured images $I_{m,s}$ are compatible with the assumption that $\mathbf{x} \in S$. We denote this function by $p(I|\mathbf{x} \in S)$, which is the marginal of the likelihood $p(I|S)$ after projection of the space of all S onto $\mathbb{R}^3$. The variable I stands for the collection of all segment images $I_{m,s}$. We model the term as product of an image intensity and a noise term:

$$p(I|\mathbf{x} \in S) = p_{int}(\mathbf{x}) \cdot p_{noise}(\mathbf{x}) \qquad (1)$$

$p_{int}(\mathbf{x})$ accounts for the recorded image *radiance values* and $p_{noise}(\mathbf{x})$ describes the spatial uncertainty, which is obtained via classical ray-plane intersections.

For computing $p_{int}(\mathbf{x})$, we project $\mathbf{x}$ to those images that are potentially covered by the laser line. A correct point should project to many bright positions for different combinations of virtual camera and laser position. However, since the object geometry is unknown at this point, there is a good chance that the point is occluded in many of
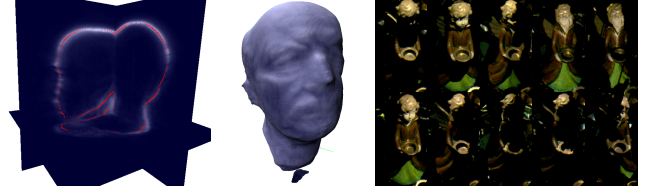

Figure 6. *Left:* Probability distribution for the Max Planck dataset. The red line shows the extracted iso-contour. *Middle:* The extracted iso-surface. *Right:* Example of a partial reflectance field. We show one virtual camera view under ten different lighting conditions, i.e. for ten different virtual projectors.

these views (typically, about 50%). We model this intuition by a robust likelihood term

$$p_{int}(\mathbf{x}) = \prod_{m,s} \prod_{V_c^n} ((1-\epsilon)\, p\left(I_{m,s}, \pi_c^n(\mathbf{x})\right) + \epsilon), \qquad (2)$$

that is, point $\mathbf{x}$ is projected into the virtual camera V_c^n via the projection π_c^n. The segment image $I_{m,s}$ is then sampled at this location. The intensity is interpreted as a three-dimensional probability distribution, denoted here by $p(I_{m,s}, \pi_c^n(\mathbf{x}))$. The uniform term, ϵ, models the potential for occlusion and the chance that the point is an outlier.

The term p_{noise} is modeling the inaccuracies of the system calibration. To obtain candidate points for triangulation, we perform a space-time peak detection over the noise-filtered segment images $I_{m,s}$ for each virtual projector V_p^m individually, i.e. m is fixed in the following. For every pixel (x, y) that has a non-negligible space-time maximum we compute $t = \max_s I_{m,s}$. The maximum segment index t identifies the virtual projector plane that was used to illuminate the object. In addition, we determine the virtual camera V_c^n that observed the pixel by checking if the pixel belongs to its corresponding chamber $n = \{i | (x, y) \in C_c^i\}$. The ray corresponding to pixel (x, y) is then given by $(\pi_c^n)^{-1}(x, y)$. Similarly, the projector plane is given by $(\pi_p^m)^{-1}(x, t), \ x \in [x_{start}, x_{end}]$.

Triangulation results in the mean 3D position of the surface point $\mathbf{x}_k$. We aim at estimating its uncertainty with respect to the system calibration. We do so by determining the region of uncertainty by intersecting the rays $(\pi_c^n)^{-1}(x \pm 2\sigma_c, y \pm 2\sigma_c)$ with the planes $(\pi_p^m)^{-1}(x, t \pm 2\sigma_p)$, where σ_c is the standard deviation of the uncertainty in the camera calibration and σ_p the corresponding uncertainty in projector plane calibration. This triangulation results in eight 3D points surrounding a sheared three-dimensional Gaussian blob

$$p_{noise}^k(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{3}{2}}\sqrt{|\Sigma|}} \exp{-\frac{1}{2}(\vec{p}-\vec{\mu})\Sigma^{-1}(\vec{p}-\vec{\mu})}, \qquad (3)$$

where $\vec{p}$ is a 3-vector containing the x- and y-components of $\pi_c^n(\mathbf{x})$, and the y-component of $\pi_p^m(\mathbf{x})$, $\vec{\mu} = (x, y, t)$, and $\Sigma = \text{diag}(\sigma_c, \sigma_c, \sigma_p)$. Eq. 3 describes the uncertainty of an individual triangulated point $\mathbf{x}_k$ in three-space.

We obtain the final three-dimensional probability distribution for the surface by summing the negative

log-likelihoods for all individual points

$$p(I|\mathbf{x} \in S)(\mathbf{x}) = -\sum_k \log\left(p_{int}(\mathbf{x}) \cdot p_{noise}^k(\mathbf{x})\right). \quad (4)$$

For computational efficiency, we evaluate it only inside the bounding box provided by the eight 3D points surrounding the $\pm\sigma$ region of every triangulated point instead of the full volume.

Shape priors and segmentation: The evaluation of Eq. 4 results in a 3D neg-log volume as shown in Fig. 6 (left). It encodes the evidence on S as extracted from the images only. In order to obtain a "reasonable shape", we have to make assumptions $p(S)$ on the space of shapes S has been drawn from. In particular, we have to extract a closed surface that fills potential holes occurring, e.g., due to self-occlusion of the object. We follow the TV segmentation approach as proposed by Unger et al. [17]. The segmentation algorithm favors boundaries running along strong evidence while simultaneously minimizing surface area. The algorithm is especially convenient because of its convex energy functional resulting in a globally optimal solution. An example of an extracted boundary is shown in Fig. 6 (left, red line), whereas in the middle a complete iso-surface is shown. The proposed algorithm is robust against triangulation noise and results in watertight surfaces of arbitrary topology. The algorithm requires a specification of regions definitely inside and outside the object. We use the visual hull to define the outside and manually paint a few coarse strokes into the visualization shown in Fig. 6 (left) to fix the inside region. The result of the algorithm is a binary indicator function of inside and outside voxels, which we triangulate using marching cubes and an additional surface fairing step to remove discretization artifacts.

5.5. Reflectance Fields

In addition to estimating geometry, we also use the segment images $I_{m,s}$ for the computation of the reflectance field of the object. Consider the images for a fixed virtual projector V_p^m: the image $\sum_s I_{m,s}$ forms a spherical light field view of the single illumination direction associated with V_p^m. We can assemble one such image for every virtual projector. Each of these images can be decomposed into its constituting camera chambers. We thus have assembled a spherical reflectance field $R_{m,n}$. The reflectance field images are radiometrically compensated by removing mirror attenuation and a quadratic fall-off due to the laser acting as a point light source. An example of a fixed virtual camera and a varying virtual projector is shown in Fig. 6 (right).

Simply summing the images to obtain the reflectance field corresponds to taking a long exposure shot while the laser is scanning a complete projector chamber. We can avoid summing most of the noise by performing a peak detection along the s-coordinate of the stack $I_{m,s}$: For direct illumination without global illumination effects, there should occur a single peak for all camera pixels that see the laser line at some point during the chamber scan of projector V_p^m. Since the laser has a non-Dirac profile, we use the integral of the measurement values over the

dataset	capture	pre	geometry	ref field	ref share
Angel	93h24m	≈3s/seg	≈10s/seg	≈3s/seg	98min
Emily	104h24m	≈3s/seg	≈10s/seg	≈3s/seg	26min
Max Planck	71h40m	≈3s/seg	≈10s/seg	≈3s/seg	48min

Table 1. Timing for different parts of our pipeline.

neighboring segment positions s as the reflectance value after subtracting the noise floor of the image as determined by background regions, i.e. image regions that do not belong to any camera chamber. Using the brightest peak for this detection often corresponds to choosing the directly illuminated part of the surface. Under some circumstances, indirect reflections can be brighter than the direct illumination, e.g. for grazing angles. To avoid estimation bias in these cases, we add a confidence map to the reflectance field. If there is a single, well defined peak, a pixel has a high confidence, if there are several peaks, or the peak is not well defined, we lower the confidence value. This simple measure implements a form of direct/global separation [13].

5.6. Reflectance Sharing

After geometry and reflectance field computation, we have the means to perform reflectance sampling. The geometry is used to determine visibility of a surface point to a virtual camera V_c^n and a virtual projector V_p^m. If the point is both visible and lit, the reflectance sample is given by $R_{m,n}$. Similar to previous work [9, 6], we exclude samples near depth discontinuities and those at grazing angles.

For reflectance sharing, we implement the approach of Weistroffer et al. [18]. As in their paper, we compute a low-dimensional basis of the SVBRDF directly from the data by approximating it with a set of radial basis functions. This is alternated with the estimation of per-surface point linear weights. The alternating computation of the blending weights and the basis BRDFs is iterated for a fixed number of iterations and is applied to each color channel. We implemented the algorithm in graphics hardware which enables the reconstruction of large data sets. Whereas Weistroffer et al. are restricted to ≈ 30.000 reflectance samples, we can deal with a three order of magnitude larger number, see Table 3. We still need to sub-sample our geometry for estimation of the BRDF basis functions. Our models consist of $\approx 150K$ surface points with an average of 120 view/illumination direction pairs, see also Table 3, i.e. we have about 20 million reflectance samples per object. This is an order of magnitude larger than previous work [9]. We estimate the basis BRDFs from ≈ 15000 surface points determined by Poisson-disc sampling. We use individual estimates for the three color channels.

6. Results

We have tested our system on the three objects *Angel*, *Emily*, and *Max Planck* shown in Fig. 8. Two of the objects, Emily and Max Planck, are almost single material, metallic objects, whereas the Angel data set shows a larger number of different materials.

In Table 1 we show the capture and computation times for the different stages of our pipeline. The capture times are rather long which is mainly due to the slow scan speed of our galvanometer and the low output power of our
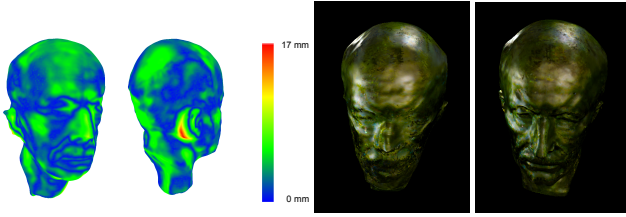
Figure 7. *Left:* Geometric quality of our recovered geometry as compared to laser scanned geometry (object coated with white powder). Most of the errors are below $6mm$, The gap behind the ear could not be resolved. *Right:* A comparison of reflectance sharing computed and displayed with our approximate geometry (left) and with laser scanned ground truth geometry (right). Details appear less blurred in the accurate geometry version.

dataset		min	max	$\sqrt{max}$	avg	$\sqrt{avg}$
Angel	cam	1	28274	168	7718	88
	proj	32	35947	190	9299	96
Emily	cam	10	20963	145	5744	76
	proj	3	29946	173	6536	81
Max Planck	cam	1	23732	154	7838	89
	proj	19	32283	180	9344	97

Table 2. Number of camera and projector pixels for the virtual cameras and projectors. This statistic counts object pixels only, i.e. background is ignored. We give the minimum, the maximum, and the average numbers. The square roots give the side length of an equivalent square of pixels.

laser system. Preprocessing, geometry estimation, and reflectance field computation are mainly I/O bound. The raw size of our HDR data sets is between 350 and 400 GB. We therefore report processing times per image disregarding the I/O part for single threaded computation. The timing for the geometry computation step includes the computation of the surface probability volume only. The segmentation part of the pipeline, implemented with CUDA running on a Nvidia GTX 580, requires about 3 min for one data set.

The reflectance sharing step has a significantly lower memory footprint since it is dealing with the sampled reflectance field data, requiring between 200 and 400 MB of memory per data set. The two steps of the reflectance sharing procedure, see Sect. 5.6, have different computational requirements. First, the computation of the spatial weights requires the solution of many small linear systems and can be solved in a parallel fashion. We use a multi-threaded CPU implementation running on a Xeon 5650, 2.6 GHz dual-hexacore CPU. Second, the computationally most demanding part for basis BRDF estimation is the setup of a large matrix that cannot be determined independently, requiring several passes over all reflectance samples. We implement this part in CUDA, again running on a single Nvidia GTX 580 GPU.

For the Max Planck data set, a high resolution range scan is available from the Aim@Shape repository. It was acquired by coating the object with white powder and range scanning it with a Minolta vi910 which is at least an order of magnitude more accurate than our system and can be considered as ground truth. In order to evaluate our reconstruction pipeline we registered the geometry recovered by our technique with the ground truth scan using rigid ICP. The error distribution is shown in Fig. 7. The accuracy of

dataset	#cam	#proj	# segs	avg	max	total samples
Angel	246	144	20177	223	901	31.9 million
Emily	235	139	18863	133	815	20.8 million
Max Planck	239	134	14483	123	606	17.7 million

Table 3. System properties for the different data sets. #cam and #proj are the number of virtual cameras and projectors in the system, avg samples is the average number of view and lighting directions per surface point. Max samples reports the surface point with the maximum number of reflectance samples. Total samples shows the overall number of samples used for reflectance estimation.

dataset	#pts basis	#pts refl	#basis BRDFs	#RBFs
Angel	12612	143800	3×3	1146
Emily	15532	156232	3×1	382
Max Planck	9803	144150	3×1	382

Table 4. Reflectance sharing statistics, "#pts basis" are the number of surface points used to estimate the basis BRDFs. We estimate "#basis BRDFs" basis materials for each color channel. "#RBFs" refers to the generic basis functions used to estimate the basis BRDFs [18]. "#pts refl" is the number of points in the object geometry. We estimate individual weights of the basis BRDFs for each of them.

our system, with an average error of about $5 - 6$mm, is low by current standards. However, taking into account the actual resolution achieved by our virtual multi-view/multi-projection system, the results can be interpreted as quite satisfactory, see Table 2: Multiplexing onto a 5 MPixel sensor results in a very low number of pixels that are available for every virtual view and projector. The equivalent square numbers are provided to give an impression of the image size of the object in each virtual view. The fill rate of our sensor with useful information is between 20 and 25%.

In another experiment, we investigated the impact of the low resolution geometry as output by our system. We registered the available ground truth geometry of the Max Planck data set with our reflectance field. As can be seen in the images, more detail is preserved with the more accurate geometry. This effect can be attributed to parallax effects if the geometry is expected in the wrong position. Thus, improving the geometry, e.g. by separately scanning it as in [15] might be a viable option for reflectance scanning with a kaleidoscope.

7. Discussion and Conclusions

We have presented a multiplexing system for surround projection and imaging without moving the acquisition system or the object. We have demonstrated the principal feasibility of the approach. In terms of practical applicability it is, however, still not competitive with alternative approaches.

In the process of developing the system we have encountered a number of problems that yet limit the practical use of the system. The most fundamental of these is the limited resolution of todays imaging and projection systems; a significant increase in resolution is required for a system with spatial resolution comparable to the best existing alternatives. Another limitation of our practical setup is the use of a low scan speed laser-galvanometric system. The low output power of the laser leads to very long acquisition times and significant amounts of noise in the data. The use
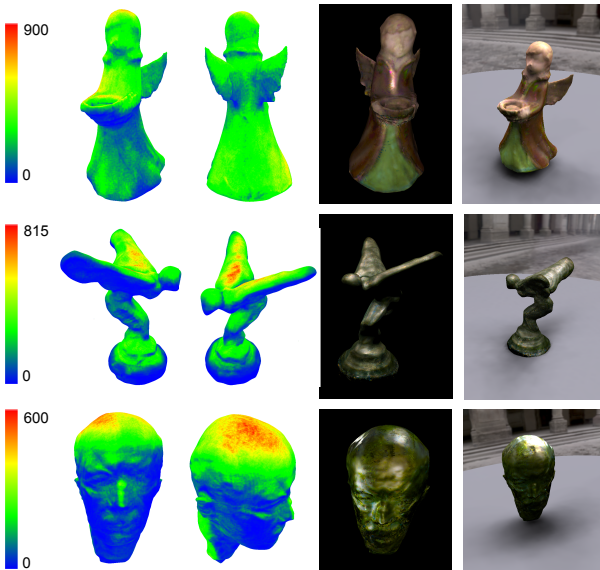
Figure 8. Results for our data sets *Angel* (top row), *Emily* (middle row), and *Max Planck* (bottom row). The left two columns show the distribution of the number of unoccluded reflectance samples (view/light pairs) on the object surface. The center right column shows the objects illuminated from a single direction (Angel and Emily) and from four directions for the Max Planck data set. In the right column, we show the objects illuminated by environment lighting.

of high-end laser cinema projection systems might mitigate this practical limitation. A fast scan speed could also allow for using coded structured light instead of laser stripe scanning, reducing the required scan time once more. An interesting future research direction is addressing the light super-position problem computationally: In this paper, we have opted for sequentializing the acquisition of different illumination directions. However, using appropriate over-lapping patterns and a corresponding decoding step could yield significant speed-ups.

In the short term, the system might be used for a comparatively inexpensive acquisition of reflectance fields using a standard projector, scanning the geometry by some other means. In this case, the illumination only has to fill complete chambers with floodlit illumination, thereby reducing noise issues and dramatically shortening acquisition times. In the long term, the system can be seen as a way of generating three-dimensional illumination distributions which could e.g. be used for spatial coding or display. Obviously, a trade-off between the directional and spatial resolution is inherent to our approach, but designs with a few additional cameras and augmented mirror configurations might provide different compromises in hardware costs versus attainable resolution.

The main insight of our paper is that kaleidoscopic multiplexing of omnidirectional acquisition is possible. The resulting trade-off of spatial versus directional resolution offers an additional, novel degree of freedom for the design of acquisition setups.

Acknowledgements

References

[1] J.-Y. Bouguet. Camera Calibration Toolbox for Matlab. (http://www.vision.caltech.edu/bouguetj/calib_doc), 2010. 4

[2] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar. Acquiring the Reflectance Field of a Human Face. In *Proc. SIGGRAPH*, pages 145–156, 2000. 1

[3] M. Fuchs, V. Blanz, H. Lensch, and H.-P. Seidel. Adaptive Sampling of Reflectance Fields. *ACM TOG*, 26(2), 2007. 1

[4] G. Garg, E.-V. Talvala, M. Levoy, and H. P. A. Lensch. Symmetric Photography: Exploiting Data-sparseness in Reflectance Fields. In *Proc. EGSR*, pages 251–262, 2006. 1

[5] J. Y. Han and K. Perlin. Measuring Bidirectional Texture Reflectance with a Kaleidoscope. In *Proc. SIGGRAPH*, pages 741–748, 2003. 1, 3

[6] M. Holroyd, J. Lawrence, and T. Zickler. A Coaxial Optical Scanner for Synchronous Acquisition of 3D Geometry and Surface Reflectance. *ACM TOG*, 29(4):99, 2010. 1, 6

[7] E. P. F. Lafortune, S.-C. Foo, K. E. Torrance, and D. P. Greenberg. Non-linear approximation of reflectance functions. In *Proc. SIGGRAPH*, pages 117–126, 1997. 1

[8] D. Lanman, D. Crispell, and G. Taubin. Surround Structured Lighting: 3-D Scanning with Orthographic Illumination. *CVIU*, 113(11):1107–1117, 2009. 1

[9] H. P. A. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H.-P. Seidel. Image-Based Reconstruction of Spatial Appearance and Geometric Detail. *ACM TOG*, 22(2):234–257, 2003. 1, 6

[10] M. Levoy and P. Hanrahan. Light Field Rendering. In *Proc. SIGGRAPH*, pages 31–42, 1996. 1

[11] A. Manakov, H.-P. Seidel, and I. Ihrke. A Mathematical Model and Calibration Procedure for Galvanometric Laser Scanning Systems. In *Proc. VMV*, pages 207–214, 2011. 3, 4

[12] V. Masselus, P. Peers, P. Dutré, and Y. D. Willems. Relighting with 4D incident light fields. *ACM TOG*, 22(3):613–620, July 2003. 1

[13] S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar. Fast Separation of Direct and Global Components of a Scene Using High Frequency Illumination. In *Proc. SIGGRAPH*, pages 935–944, 2006. 6

[14] I. Reshetouski, A. Manakov, H.-P. Seidel, and I. Ihrke. Three-Dimensional Kaleidoscopic Imaging. In *Proc. CVPR*, pages 353–360, 2011. 2, 4

[15] C. Schwartz, M. Weinmann, R. Ruiters, and R. Klein. Integrated High-Quality Acquisition of Geometry and Appearance for Cultural Heritage. In *Proc. VAST*, pages 25–32, 2011. 1, 7

[16] P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. Lensch. Dual Photography. *ACM TOG*, 24(3):745–755, 2005. 2

[17] M. Unger, T. Pock, W. Trobin, D. Cremers, and H. Bischof. TVSeg - Interactive Total Variation Based Image Segmentation. In *Proc. BMVC*, pages 335–354, 2008. 5, 6

[18] R. P. Weistroffer, K. R. Walcott, G. Humphreys, and J. Lawrence. Efficient Basis Decomposition for Scattered Reflectance Data. In *Proc. EGSR*, pages 207–218, 2007. 1, 4, 6, 7

[19] T. Zickler, P. N. Belhumeur, and D. J. Kriegman. Helmholtz Stereopsis: Exploiting Reciprocity for Surface Reconstruction. *IJCV*, 49(2-3):215–227, 2002. 1

[20] T. Zickler, R. Ramamoorthi, S. Enrique, and P. N. Belhumeur. Reflectance Sharing: Predicting Appearance from a Sparse Set of Images of a Known Shape. *IEEE Trans. PAMI*, 28(8):1287–1302, Aug. 2006. 1, 4