Multiplexed Acquisition of Bidirectional Texture Functions for Materials

Dennis den Brok, Heinz C. Steinhausen, Matthias B. Hullin and Reinhard Klein
Institut für Informatik II, Universität Bonn, Germany

ABSTRACT
The bidirectional texture function (BTF) has proven a valuable model for the representation of complex spatially-varying material reflectance. Its image-based nature, however, makes material BTFs extremely cumbersome to acquire: in order to adequately sample high-frequency details, many thousands of images of a given material as seen and lit from different directions have to be obtained. Additionally, long exposure times are required to account for the wide dynamic range exhibited by the reflectance of many real-world materials.

We propose to significantly reduce the required exposure times by using illumination patterns instead of single light sources (“multiplexed illumination”). A BTF can then be produced by solving an appropriate linear system, exploiting the linearity of the superposition of light. Where necessary, we deal with signal-dependent noise by using a simple linear model derived from an existing database of material BTFs as a prior. We demonstrate the feasibility of our method for a number of real-world materials in a camera dome scenario.

Keywords: digital material appearance, bidirectional texture functions, illumination multiplexing

1. INTRODUCTION
Analytical or physically-based reflectance models such as the bidirectional reflectance distribution function (BRDF) and its spatially varying sibling (SVBRDF) generally fail to accurately represent the noticeable non-local effects such as interreflections and self-shadowing that can be observed on many common real-world materials. The image-based bidirectional texture function (BTF) accounts for these effects, even in real-time applications. This image-based nature, however, makes material BTFs extremely cumbersome to acquire: in order to adequately sample high-frequency details, many thousands of images of a given material with different lighting and viewing positions have to be obtained. Parallel setups such as camera domes and kaleidoscopes significantly accelerate this process, but they still suffer from long exposure times for high dynamic range (HDR) capture and the sheer number of samples required.

We propose to significantly reduce the required exposure times by using multiplexed illumination: instead of taking images for each light source separately, we use patterns consisting of many light sources at once, greatly increasing the amount of light on the material sample. Exploiting the linearity of the superposition of light, it is then possible to reconstruct the corresponding single-light images by solving an appropriate linear system. In the presence of signal-dependent noise such as photon noise, this process may yield very noisy images, in particular if the brightness of the scene varies greatly.

Figure 1. Our acquisition setup displaying an S-matrix pattern.
We demonstrate that for rather simple materials, this effect is not noticeable in practice, and for moderately complex materials where it is, it is possible to denoise the resulting BTFs by using a simple linear model derived from an existing database of material BTFs as a prior.

We evaluate the proposed method on a number of real-world materials in a camera dome scenario and show that it is possible to reduce acquisition times by about 75–95%.

2. RELATED WORK

2.1 Plenoptic multiplexing
An in-depth theoretical background on plenoptic multiplexing in general has been given by Hartwit and Sloane. They prove or conjecture the optimality of Hadamard patterns and their binary derivates with respect to several measures. They also cover the various noise sources in optical systems and how they influence demultiplexing, albeit quite abstractly so.

As far as the authors are aware, Wenger et al. were among the first to investigate multiplexing for the purpose of capturing some kind of appearance; in this case, time-varying light fields of human faces. However, they observed an intolerable amount of noise when using Hadamard patterns that they were unable to reduce to a tolerable level through simple filtering.

In contrast, Schechner et al. deal with the special case of illumination multiplexing using digital photo cameras and various types of light sources. They explain in great detail the implications of this scenario regarding noise and derive a formula useful for judging whether multiplexing is beneficial in a specific situation, taking into account the setup’s relevant intrinsic parameters.

Ratner et al. provide an optimization method which produces illumination patterns that take the noise characteristics of a given setup into account. They assume a one-dimensional affine noise model and a nearly diffuse scene, assumptions generally violated in BTF acquisition. Furthermore, they show that in the presence of overexposure it is preferable to reduce the number of light sources instead of shutter times.

Mitra et al. take this even further and compute illumination patterns using an optimization based on image priors. In contrast to previous methods, their method is able to handle large amounts of light, but it still relies on the assumption of a one-dimensional affine noise model. They show how to extend their method to low-resolution light fields, which is, however, computationally very expensive and therefore likely prohibitive in the case of BTFs.

2.2 Sparse acquisition
Sparse acquisition is related with the proposed method not directly in terms of methodology, but in that its goal is faster acquisition, and to that end usually some kind of model is used as a prior as well. We thus briefly review a number of articles on sparse acquisition.

Matusik et al. perform singular value decomposition (SVD) on a database of 100 measured BRDFs of a wide range of isotropic materials to obtain a linear model.

The same authors introduce two methods for sparse reconstruction of isotropic BRDFs: The first method is based on a wavelet analysis of their BRDF database. A set of basis wavelets termed common wavelet basis is determined and used to reconstruct previously unseen BRDFs with approximately 1.5 million samples from approximately 70000 measurements. The second method uses the entire BRDF database itself as a linear model for reconstruction of fully measured BRDFs from as little as 800 out of the original approximately 1.5 million samples, at the cost of slightly increased reconstruction errors and the required availability of the BRDF database. Samples are chosen using a simple optimization algorithm such that the linear system to be solved for reconstruction is well-conditioned. They do not investigate how well their methods generalize to more complex reflectance such as anisotropic BRDFs.

Filip et al. propose a vector quantization of BTFs for the purpose of compression, guided by a psychophysically validated metric. They conclude that as little as 10 – 35 % of the original textures are sufficient to maintain the same visual appearance in renderings. It would be interesting to investigate whether there is a common quantization for all materials, and if so, whether it could be used for sparse acquisition. A large user study would be needed in order to adapt the metric to a bigger BTF database.
Peers et al.9 introduce compressed sensing10 to the acquisition of reflectance fields, assuming both 2D outgoing (here: fixed viewing direction) and incident light fields. Their algorithm uses a hierarchical, multi-resolution Haar wavelet basis, taking spatial coherence into account. It is not clear how to extend this approach to a multi-view setup. Common BTF acquisition setups only have a very limited number of light sources, where the advantage of compressed sensing might be negligible. Due to these light sources’ brightness, we expect shot noise to become a problem.

Dong et al.11 reconstruct a material’s SVBRDF from a sparse measurement using a manifold constructed from analytical BRDFs fit to fully measured BRDFs of manually selected representative points on the material’s surface. The algorithm is unlikely to scale to BTFs because of the typically much higher intrinsic dimensionality of the manifold of per-texel reflectance distribution functions. A generalization to previously unseen materials is not obvious, albeit conceivable.

Marwah et al.12 use sparsity-based methods related to compressed sensing in order to sparsely acquire 4D light fields with an angular resolution of $5 \times 5$. They compute a dictionary of what they call light field atoms – $11 \times 11$ spatial light field patches which allow for a sparse representation of natural light fields. Such a dictionary does generally not exist in the case of BTFs; their dimensionality is likely too high.

Inspired by the work of Matusik et al. we13 demonstrated that there exists linear models that lend themselves to sparse reconstruction of BTFs. We proposed to fit linear models to small BTF patches in order to account for non-local effects, with separate models per (manually chosen) classes of materials to constrain the dimensionality of that data. Straight-forward models fit to per-texel reflectance distribution functions proved insufficient for our purpose.

3. BACKGROUND

3.1 Bidirectional texture functions

BTFs have been introduced by Dana et al.15 as an image-based approach to spatially varying appearance. Like SVBRDFs, they are 6-dimensional functions of the form

$$B(x, \omega_i, \omega_o),$$

where $\omega_{i,o} \in \mathbb{R}^2$ are the incoming and outgoing light directions, respectively, and $x \in \mathbb{R}^2$ is the position on a parameterized surface $V$. In the case of material BTFs, $V$ is typically flat; it does not need to coincide with the material’s actual surface geometry. It is generally assumed that light sources are directional and have the same spectrum. In particular, effects such as phosphorescence, fluorescence and subsurface scattering cannot be captured accurately.

The fundamental difference from SVBDRFs is that the function $B(x, -)$ need not be BRDF-valued: the corresponding per-texel reflectance function does not need to adhere to Helmholtz reciprocity and conservation of energy and is, therefore, capable of capturing non-local effects such as interreflections and self-shadowing. Moreover, because $V$ does not necessarily coincide with the actual surface, the per-texel reflectance functions may also describe parallax effects. For these reasons, the term apparent BRDF (ABRDF) has been suggested by Wong et al.16 for this kind of functions. Conversely, the values of the function $B(-, \omega_i, \omega_o)$ are just 2D textures corresponding to specific pairs of incoming and outgoing light directions.

During measurement, a finite discretization of the measured material’s BTF $B$ is obtained. After rectification of the acquired images, the discrete BTF has a natural representation as a matrix $B \in \mathbb{R}^{n \times m}$ with the columns representing the $m$ discrete ABRDFs, each entry corresponding to some pair $(\omega_i, \omega_o)$ of incoming and outgoing angle, and the rows representing the $n$ rectified textures (cf. figure 2).

3.2 BTF acquisition

Several setups for the acquisition of BTFs have been proposed. We briefly review the most prominent paradigms, as our method benefits all of them to a greater or lesser extent. An up-to-date, in-depth overview is given in the recent survey by Schwartz et al.17
3.2.1 Gonioreflectometer

In what is historically the first BTF acquisition setup, proposed by Dana et al., the material sample is placed on a turntable, and a camera and a light source held by robot arms are moved across the hemisphere above the sample to capture images of the sample under different lighting and viewing conditions. The gonioreflectometer is very flexible in terms of possible samplings of the hemisphere, but measurement times are excessive – on the order of weeks for a moderate angular resolution – due to the little amounts of light sources and sensors and the movable parts’ low speeds.

Due to the absence of multiple light sources, our method does not readily apply to traditional setups. If the gonioreflectometer’s total flexibility is not a necessity for the envisioned scenario, additional light sources can be added.

3.2.2 Kaleidoscope

Han et al. introduced an intriguing parallel setup: The sample is placed underneath a tapered kaleidoscope, lit and captured from a projector and a camera placed at the other end, which allows for a number of lighting and viewing conditions to be measured in a single camera shot. By appropriately arranging the mirrors, the angular and spatial resolution can be adjusted; however, both are typically rather low, and increasing one leads to a decrease of the other, so there is always a tradeoff to be made.

3.2.3 Camera domes

Camera domes as proposed, for instance, by Müller et al. and Schwartz et al. ideally provide a highly parallel means to acquire BTFs: A number of cameras is spread across the hemisphere above the sample holder. Their flashes or separate LEDs are used as light sources. Parallelism may be traded for fewer cameras and lower cost by placing the sample on a turntable in order to achieve a similarly dense sampling of the hemisphere. Due to the number of cameras, data transfer times become a new bottleneck.

In all of the above setups, it is usually necessary to capture the same scene several times with different shutter times in order to obtain HDR data.

3.3 Illumination multiplexing

Multiplexed measurement is based on the observation that by using appropriate patterns, the amount of signal-independent noise in the demultiplexed measurements will be lower than without multiplexing. In the case of imaging systems, the resulting linear system to be solved is

$$\mathbf{M} \cdot \mathbf{I}_{\text{single}} = \mathbf{I}_{\text{multiplexed}}$$

where $\mathbf{M} \in \mathbb{Z}^{n \times n}$, $\mathbf{I}_{\text{single}} \in \mathbb{R}^{n \times (w \cdot h)}$ with each row a $w \times h$ image of the scene lit by an individual light source, and $\mathbf{I}_{\text{multiplexed}} \in \mathbb{R}^{n \times (w \cdot h)}$ where each row is a $w \times h$ image of the scene lit by an individual illumination pattern.

Typically, $\mathbf{M}$ is supposed to be binary, corresponding to light sources of equal power that can only be on or off. It has been shown that binary illumination patterns given by a certain kind of matrices called $S$-matrices...
are at least very close to minimizing the average mean square error, and that the associated signal-to-signal-independent-noise ratio is increased by a factor of $\sqrt{\frac{n}{2}}$ for large $n$. It has, however, also been shown that the presence of signal-dependent noise such as photon noise strongly counteracts this advantage.\textsuperscript{13}

There are several ways to produce $S$-matrices $M \in \mathbb{N}^{n \times n}$ of a given order $n \in \mathbb{N}$;\textsuperscript{1} we chose to implement the most straightforward one, the so-called \textit{quadratic residue construction}: Let $n = 4p + 3$, $n, p \in \mathbb{N}$ prime, be the desired order, and $Q$ the set of quadratic residues of $0, \ldots, \frac{n-1}{2}$ modulo $n$. Then

$$m_{1j} = \begin{cases} 1, & j \in Q \\ 0, & \text{else} \end{cases}$$

The remaining $n - 1$ rows of $M$ are obtained by cyclically shifting the preceding row to the left by one.

Clearly, the number of light sources in a given acquisition setup need not necessarily be prime of the required form. In that case, one can choose a suitable higher order and truncate the resulting $S$-matrix such that the number of columns equals the number of light sources, and solve the resulting linear system in the least-squares sense.\textsuperscript{1}

4. MULTIPLEXED BTF ACQUISITION

4.1 Experimental setup

We chose 12 materials from 3 classes – \textit{cloth}, \textit{leather}, and \textit{wood} – each. Per class, we picked 4 materials for the purpose of verifying the proposed method which we measured both conventionally and using the proposed method.

The measurement device we used in our experiments is a camera dome with 11 industrial-grade cameras with a maximum frame-rate of 8 Hz and 198 LED light sources, 10 of which are placed opposing the lower-most 10 cameras to produce direct reflections. The quadratic residue construction requires the $S$-matrix order to be prime of the form $4p + 3$, $p$ prime. The smallest $S$-matrix order of that form greater or equal the number of LEDs in our setup is $n = 199$, which we chose for our experiments.

In order to speed up the acquisition of the ground truth data, the single-light measurements were obtained using a camera gain of 10 dB, while the multiplexed measurements were obtained with a camera gain of 0 dB. Additionally, we measured 4 small patches (approximately $1 \text{ cm} \times 1 \text{ cm}$) of different materials of the same class at once. We argue that, even in practice, this is a reasonable trade-off to be made: If required, larger BTFs for each single sample can be produced quickly by measuring them again using the proposed method, or by extrapolation as described \textit{e.g.} by Steinhausen \textit{et al.}\textsuperscript{21}

The materials are placed on a turntable, which during measurement is rotated 12 times by $30^\circ$, in order to achieve a dense sampling of viewing directions. For each turntable position, the cameras take pictures of the material lit by each of the LEDs separately using several exposure times determined manually before the actual acquisition. The test materials were additionally measured using $S$-matrix patterns of order $n = 199$ before proceeding to the next turntable position.

After measurement, we combine the low dynamic range (LDR) raw images into high dynamic range (HDR) images and subsequently demosaic and rectify them. We use LED calibration data in order to account for variations in the LEDs’ spectra. For further details on the various post-processing steps, \textit{cf.} Schwartz \textit{et al.}\textsuperscript{17}

Finally, the resulting images are arranged and stored as matrices, which in the case of multiplexed measurements are then demultiplexed. For rendering, the BTFs are furthermore resampled in the angular domain, such that the light and view hemispheres are the same for each texel, and compressed using truncated SVD.
4.2 Linear models and denoising

For the purpose of denoising, we fit linear models separately to the per-color channel ABRDFs of each of the three classes, converted to log(Y) U/Y V/Y color space, without the test materials. The fitting is done using a truncated SVD

\[ \mathbf{D}_{\text{class}} \approx \mathbf{U} \Sigma \mathbf{V}^t, \]

which is known by the Eckhart-Young theorem to be optimal in terms of L^2 error,\(^{22}\) where \( \mathbf{D}_{\text{class}} \) is a matrix with the ABRDFs of the BTFs of class \( \text{class} \) as columns. We have shown\(^{13}\) that these model generalize to materials \( \mathbf{B} \) that do not belong to the particular database; i.e.

\[ \min_{\mathbf{C}_\mathbf{B}} \| \mathbf{UC}_\mathbf{B} - \mathbf{B} \| < \varepsilon. \]

For our experiments, we used 768 basis vectors for log(Y) and 128 for both U/Y and V/Y, which, according to our previous experiments,\(^{13}\) is sufficient for the materials classes considered in the present article. Under the assumption that demultiplexing noise cannot be represented in these bases, denoising can now be performed by simply projecting a demultiplexed BTF onto the corresponding basis:

\[ \mathbf{B}_{\text{denoised}} = \mathbf{U} \cdot (\mathbf{U}^t \cdot \mathbf{B}_{\text{demultiplexed}}). \]

5. RESULTS

Table 1 shows relative errors

\[ \varepsilon_{\{\text{demult, denoised}\}} = \frac{\| \mathbf{B}_{\{\text{demult, denoised}\}} - \mathbf{B}_{\text{reference}} \|_F}{\| \mathbf{B}_{\text{reference}} \|_F} \]

of the demultiplexed and denoised demultiplexed BTFs, respectively. Errors are computed on the log(Y) color channel to account for human perception.

In the case of wood, the relative errors are relatively small, and indeed there are no obvious differences between the renderings of ground truth, demultiplexed and denoised BTF, even though the relative error is approximately halved by denoising (cf. figure 3).

In contrast, Cloth 1 and Leather 2 exhibit annoying artifacts for grazing viewing angles, as seen at the cylinder’s borders. While the denoised BTF does not look completely identical to the ground truth, it looks much more plausible at the borders (cf. figures 4 and 6). For better comparison, figure 5 shows textures corresponding to a low camera position (\( \theta = 75^\circ \)), extracted from Cloth 1’s BTF. Note that these are rectified images; in an actual rendering, each pixel will be a weighted average of several texels.

Demultiplexed Leather 4, however, exhibits so much noise (yielding a relative error of almost 20%) that our denoising strategy breaks down as well (cf. figure 7), even though the relative error is more than halved.
Figure 4. Renderings of Cloth 1. Left: demultiplexed. Center: ground truth. Right: denoised.

Figure 5. Low camera ($\theta = 75^\circ$) textures extracted from Cloth 1 BTF. Left: demultiplexed. Center: ground truth. Right: denoised.

Figure 6. Renderings of Leather 2. Left: demultiplexed. Center: ground truth. Right: denoised.

Figure 7. Renderings of Leather 4. Left: demultiplexed. Center: ground truth. Right: denoised.

<table>
<thead>
<tr>
<th></th>
<th>single</th>
<th>multiplexed</th>
<th>rel. $\Delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloth 1 – 4</td>
<td>10.8</td>
<td>1.8</td>
<td>−83%</td>
</tr>
<tr>
<td>Leather 1 – 4</td>
<td>23.6</td>
<td>1.2</td>
<td>−95%</td>
</tr>
<tr>
<td>Wood 1 – 4</td>
<td>4.4</td>
<td>1.2</td>
<td>−75%</td>
</tr>
</tbody>
</table>

Table 2. Comparison of acquisition times [h], single light vs. multiplexed.
As can be seen in table 2, the larger amount of light reaching the material samples on each particular image makes it possible to reduce total acquisition times by about 75–95%, even though the ground truth data was obtained with a much higher camera gain. In the case of Leather 1–4 and Wood 1–4, the minimum acquisition time achievable with our setup, caused by several bottlenecks, is reached.

6. CONCLUSION

We demonstrated that it is possibly to produce plausibly looking material BTFs from images taken under multiplexed illumination, possibly by using linear models derived from an existing database of material BTFs as a prior for denoising. The increased amount of light on the material sample can be used to considerably reduce the necessary shutter times, leading to total acquisition times reduced by 75–95%. As expected, we observed that this works best for materials with rather limited dynamic range.

It would be interesting to incorporate the method proposed by Mitra et al.\(^5\), which would, however, require both a suitable noise model for our camera dome setup and a modified optimization algorithm accounting for that model. Moreover, it could be possible to automatically detect particularly noisy textures within a demultiplexed BTF using adequate image statistics. These could then be regarded as missing values, and filled using e.g. our sparse reconstruction method.\(^{13}\)

Finally, it is conceivable that a database to derive linear models for denoising from can be “bootstrapped”, for instance by using smaller S-matrix orders, or by leaving out light sources that over-proportionally increase the BTF’s dynamic range and measuring the corresponding images seperately.

7. ACKNOWLEDGEMENTS

This work was funded by the X-Rite graduate school on Digital Material Appearance. We thank Roland Ruiters and Martin Rump for helpful suggestions.

REFERENCES


