Data-driven Local Coordinate Systems for Image-Based Rendering

Gero Müller †, Ralf Sarlette ‡ and Reinhard Klein §

Institute for Computer Science II, University of Bonn

Abstract

Image-based representations of an object profit from known geometry. The more accurate this geometry is known, the better corresponding pixels in the different images can be aligned, which leads to less artifacts and better compression performance. For opaque objects the per-pixel data can then be interpreted as a sampling of the BRDF at the respective surface point. In order to parameterize this sampled data a coordinate frame has to be defined. In previous work this coordinate frame was either the global frame or a local frame derived from the base geometry. Both approaches lead to misalignments between sample vectors: Features of basically very similar BRDFs will be shifted to different regions in the sample vector leading to poor compression performance. In order to improve alignment between the sampled BRDFs in image-based rendering, we propose an optimization algorithm which determines consistent coordinate frames for every sample point on the object surface. This way we efficiently align the features even of anisotropic reflection functions and reconstruct approximate local coordinate frames without performing an explicit 3D-reconstruction. The optimization is calculated efficiently by exploiting the Fourier-shift theorem for spherical harmonics. In order to deal with different materials in a scene, the technique is combined with a clustering algorithm. We demonstrate the utility of our method by applying it to BTFs and 6D surface reflectance fields.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Picture/Image Generation]: Digitizing and scanning I.3.7 [Three-Dimensional Graphics and Realism]: Color, shading, shadowing, and texture

1. Introduction

Due to the advent of sophisticated acquisition devices and compression algorithms data-driven techniques became increasingly popular in the computer graphics community during the last decade. For example the digitization of the 3D object shape using laser scanning or other 3D-acquisition techniques is currently a major part of the 3D production pipeline. Of particular interest are image-based rendering (IBR) methods which offer photorealistic renderings of a complex object under arbitrary illumination based on differently illuminated photographs of that object. The idea is to represent the complex appearance of an object by sampling the plenoptic function [AB91] under varying lighting condi-

tions [DHT*00] instead of modeling the appearance which can be a complex and sometimes even impossible task.

A key issue in IBR is the parametrization of the plenoptic function, i.e., the parametrization of rays of light defined by position and direction in space, which is needed to interpolate novel views from the samples. From the early stages of IBR it has been known that the effectiveness of image-based representations concerning acquisition effort, compression performance and rendering quality is strongly related to the chosen parametrization.

Typically a parametrization is induced by a base geometry which encloses the captured object. The rays are then parameterized by a point on the base surface and the camera position. Using this parametrization the accuracy of the available base geometry effects the number of images needed to obtain an anti-aliased rendering [CCST00] and thus the possible achievable compression rates for the measured data. In the special case where the base geometry coincides with the

© The Eurographics Association and Blackwell Publishing 2006. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

[†] e-mail: gero@cs.uni-bonn.de † e-mail: sarlette@cs.uni-bonn.de

[§] e-mail: rk@cs.uni-bonn.de

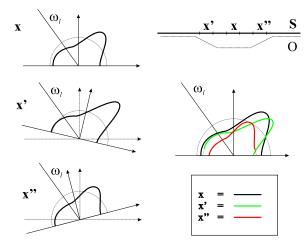


Figure 1: Small deviations of the base geometry **S** from the real surface O (upper right) lead to misalignments in the angular dimension: The actual BRDF slice for fixed lighting direction ω_l is rotated and might change its shape, e.g., because the Fresnel effect causes an increased specular peak for grazing angles (left). If the data is put into the common coordinate system derived from the base surface **S** the data at the points \mathbf{x} , \mathbf{x}' and \mathbf{x}'' differ even if the underlying BRDF is the same (center right).

exact geometry and the reflection is Lambertian, one texture image suffices to reconstruct the appearance of the object during rendering. Keeping the base geometry exact but allowing for uniform non-Lambertian reflection, novel views under varying light directions can only be generated from the geometry and a single BRDF if a local coordinate system parameterizing this BRDF is defined in each surface point. But to achieve this kind of compression in image-based rendering the local coordinate system must be derived from the measurements. For example for a material like brushed aluminium the tangent direction has to be defined in such a way that the anisotropic reflection behavior caused by the brushing is consistent from pixel to pixel, i.e., that the tangent coincides with the orientation of the brushing. Of course, in reality the normal is also not known, since that would require an exact geometry reconstruction. In this case data misalignments occur as illustrated in Figure 1.

To improve upon this situation we propose a general and efficient algorithm which estimates local coordinate systems for every surface point in order to improve the alignment of image-based data before compression. This strategy is illustrated in Figure 2. In particular this paper introduces the following contributions:

 data-driven estimation of local coordinate systems – no assumptions about the underlying reflectance data are made, i.e., the method does not rely on a specific reflectance model for efficient computation the Fourier-shift theorem for spherical harmonics is exploited – it enables extremely fast computation of spherical correlations and might also be of relevance for further applications like 3D texture reconstruction

Inspiration for this work came from the field of BRDF reparameterization [Rus98]. These techniques have been successfully applied to improve the compression of sampled BRDFs for real-time BRDF rendering [KM99]. Our work can be interpreted as a generalization of this approach to spatially varying reflectance data.

The remainder of this paper is organized as follows. After discussing previous work in Section 2, we discuss the details of our technique in Sections 3 and 4. We demonstrate the usefulness of our method by applying it to Bidirectional Texture Functions (BTF) and general 6D surface reflectance fields and show the results in Section 5.

2. Previous work

2.1. Local coordinate systems in IBR

In most previous work local coordinate systems were simply derived from the base geometry, which does neither capture the normal variation due to fine scale surface height variations nor does it give any information about the orientation of small-scale features causing anisotropic reflection.

The typical solution is to arrange the reflectance data in spatially close blocks like in the work on surface light fields of Chen et al. [CBCG02], where the reflectance data is considered per triangle. It is assumed that the local coordinate systems are only gradually varying across the triangles, such that one coordinate system is sufficient for the whole triangle. Then the reflectance data of each triangle is compressed independently. Consequently, the coherence between the reflectance data on different triangles is not exploited.

An approach more closely related to our method was introduced by Wood et al. [WAA*00]. They proposed to reflect the surface light field data around the normal in order to increase the alignment of the reflective peak. Unfortunately, this technique does not suffice in the case of general reflectance data and the effectiveness of the method also depends on the accuracy of the available normals.

Lensch et al. [LKG*03] presented the first technique that estimates per-point local coordinate systems in order to improve fitting performance. In particular, they included the estimation of per-point normals into their data-fitting procedure. Their fitting procedure assumes isotropic reflection behavior and relies on an analytic reflectance model. In contrast, our general alignment method can be combined with any kind of compression technique.

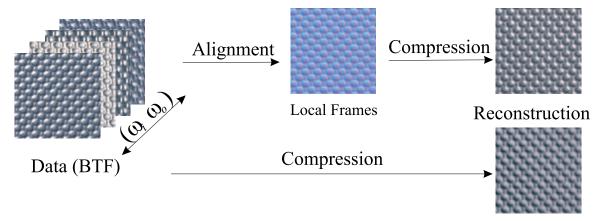


Figure 2: The principle behind our algorithm is the computation of consistent local frames (illustrated by the rgb-encoded normal map in the top row). This is done by aligning the data along the angular dimension. In the shown example (plastic BTF with knobs) the data is then compressed. As can be seen on the right the reconstruction using our technique is much sharper (for equal compressed data size).

2.2. Compression for IBR

After alignment, the data is more amenable to compression techniques. The most popular techniques for compression in image-based rendering are Vertex Quantization (VQ) and basis decomposition techniques like wavelets or Matrix Factorization (MF) based on SVD or PCA. More recently Local PCA, which can be thought of as a combination of VQ and MF, and Tensor Factorization (TF) have become popular. Analytical reflectance models originally developed for BRDF modeling (e.g., [War92] or [LFTG97]) have also been used, but their known deficiencies restrict them to special cases like dealing with sparse data [LKG*03] or real-time rendering [MLH02].

VQ is a very general tool and was already used in the seminal paper of Levoy and Hanrahan about light field rendering [LH96]. Another prominent application is surface light field rendering [WAA*00]. The principle behind VQ is to approximate similar datavectors by only one constant representative using clustering. It is easy to implement but usually leads to visible artifacts since it generates sharp transitions between clusters and also destroys smooth transitions in the data.

Basis decompositions like the Discrete Cosine Transform (DCT) or Wavelet Transforms (WT) are well known from image compression. A generalization of the WT to 4D light field data was presented by Magnor et al. [MEG00]. Ho et al. [HWL05] combined linear basis decomposition using PCA with DCT to compress the resulting coefficient maps. MF in general is probably the most widely used compression technique in IBR. It has also been applied to surface light fields [WAA*00, NSI99, CBCG02], reflectance fields [MPN*02] and BTFs [SSK03,LHZ*04]. In these techniques the data is projected onto a given basis which in the

case of MF is computed from the data itself. Such a basis should represent the main features of the data well to achieve good results. The artifacts introduced by basis decomposition techniques depend on the type of basis functions used and range from smoothing (MF) to the well known block artifacts (DCT, WT).

Combinations of VQ and MF like Local PCA have found their way into computer graphics recently [SHHS03, MMK03]. At the expense of higher fitting complexity these techniques offer a much better reconstruction quality than using VQ or MF alone, since artifacts of one technique (smoothing, cluster borders) are compensated by the other.

While MF requires the image data to be arranged in a 2D matrix, a more general arrangement would be a higher order tensor. As demonstrated by several authors [FKIS02, VT04, WWS*05] this allows to exploit correlations, e.g., not only between images but also between image rows and columns. As a result, these techniques allow for higher compression rates at the expense of higher reconstruction costs. Nevertheless, also these techniques profit from an improved alignment of the input data as the one we present in this paper.

3. Data representation

We formulate our algorithm for the general case of bidirectional reflectance data where both view and lighting direction vary. Such data is typically captured using a setup as sketched in Figure 3. Surface light fields and image-based relighting are popular special cases of this arrangement.

The general framework is the 8D reflectance field $RF_{\mathbf{S}}(\mathbf{x}_{\nu}, \boldsymbol{\omega}_{\nu}, \mathbf{x}_{l}, \boldsymbol{\omega}_{l})$ [DHT*00] which describes the appearance of an object as seen under arbitrary viewing directions and illuminated by arbitrary lighting for fixed time

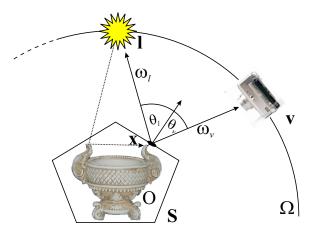


Figure 3: Our work is put in the context of this typical image-based data acquisition setting: The object O is captured from a set of view positions $\mathbf{v} \in \mathbf{V}$ and lit from a set of light positions $\mathbf{l} \in \mathbf{L}$ on the sphere Ω around the object. The data is parameterized across the base surface \mathbf{S} which should enclose the object but is not required to coincide with its surface. The data at a point \mathbf{x} on \mathbf{S} is parameterized by the directions ω_l and ω_v given in the local coordinate system at \mathbf{x} . Please note that in this setting the camera also records interreflections and subsurface scattering, i.e., light that travels between different points on the surface (dotted arrow).

and wavelength (and neglecting fluorescence and phosphorescence effects). Here, \mathbf{x}_{v} and \mathbf{x}_{l} are points on the parameterized surface \mathbf{S} which is not required to coincide with but should enclose the real object surface and ω_{v} , ω_{l} are the directions to the viewer and the light source given in the local coordinate system at \mathbf{x}_{v} and \mathbf{x}_{l} respectively.

Since RF_S is an 8D entity and still beyond reach of current acquisition systems and interactive rendering algorithms, we describe now a common simplification to 6D. It is based on omitting the spatial variation of the incoming light field. Formally this simplification can be interpreted as absence of interreflections and subsurface scattering (no light travels between the points \mathbf{x}_V and \mathbf{x}_I) or infinitely distant (directional) lighting. Since both assumptions are clearly not valid in reality we propose to formulate the projection down to 6D as the following integral:

$$RF_{\mathbf{S}}^{6D}(\mathbf{x}, \omega_{\nu}, \omega_{l}) = \int_{\mathbf{S}} RF_{\mathbf{S}}(\mathbf{x}, \omega_{\nu}, \mathbf{x}_{l}, \omega_{l}(\mathbf{x}_{l})) d\mathbf{x}_{l}$$
 (1)

The term $\omega_l(\mathbf{x}_l)$ expresses the direction from \mathbf{x}_l to point light source \mathbf{l} in the local coordinate system at \mathbf{x}_l . $RF_{\mathbf{S}}^{6D}(\mathbf{x}, \omega_v, \omega_l)$ now gives the light reflected from point \mathbf{x} in direction ω_v while the object is illuminated by the point light source "I". A measurement of $RF_{\mathbf{S}}^{6D}$ using a setting as

illustrated in Figure 3 is now performed by taking images from a set V of camera positions while lighting the scene from a set L of light positions. Then such a measurement can be tabulated in a large matrix D:

$$\mathbf{D} = \left\{ \left\{ RF_{\mathbf{S}}^{6D}(\mathbf{x}, \omega_{\nu}(\mathbf{x}), \omega_{l}(\mathbf{x})) \right\}_{\nu \in \mathbf{V}, \ l \in \mathbf{L}} \right\}_{\mathbf{x} \in \tilde{\mathbf{S}}}$$
(2)

The columns of \mathbf{D} are indexed by surface positions \mathbf{x} (assuming the surface \mathbf{S} has been discretized into some sort of grid $\tilde{\mathbf{S}}$) while the rows are indexed by view and light direction which allows the interpretation of the columns of \mathbf{D} as sampled BRDFs. Figure 4 shows a visualization of \mathbf{D} for a synthetic data set where the rows and columns in the image correspond the rows and columns of \mathbf{D} .

In practice, the set of view and light directions varies in every column and the data might be inconsistent due to occlusions and shadowing from the large scale geometry. For the presentation of our algorithm in the following section we assume that the data has been resampled into a common set $\tilde{\Omega}^+$ of directions of the upper hemisphere using the local coordinate systems derived from base geometry S and a standard resampling technique [CBCG02], and that inconsistent data has been removed or completed.

4. Algorithm

In this section we present our algorithm for determination of consistent local coordinate frames from image-based reflectance data sets. The main idea is to find a local orientation that maximizes the correlation between the reflectance data of different surface points.

4.1. Aligning local frames for objects with uniform material

In the IBR community per-point reflectance data (a column of **D**) is typically termed apparent BRDF (ABRDF) [WHON97] because it has the formal structure of a BRDF but also captures influence from the neighborhood like occlusions and interreflections. We measure the quality of alignment between two ABRDFs **a** and **b** by the following correlation term $C_{\mathbf{a},\mathbf{b}}$:

$$\textit{C}_{\mathbf{a},\mathbf{b}} = \sum_{\omega_{\nu} \in \tilde{\Omega}^{+}} \sum_{\omega_{l} \in \tilde{\Omega}^{+}} \mathbf{a}(\omega_{\nu}, \omega_{l}) \mathbf{b}(\omega_{\nu}, \omega_{l})$$

If we assume that **a** and **b** are noisy measurements of the same BRDF but with a misaligned local frame the likelihood for the rotation $\mathbf{R}(\Phi) = R_z(\phi) \cdot R_y(\theta) \cdot R_z(\psi)$ parameterized by the zyz-Euler angles $\Phi = (\phi, \theta, \psi)$ to align both frames is given by

$$C_{\mathbf{a},\mathbf{b}}(\Phi) = \sum_{\omega_{\nu} \in \tilde{\Omega}^{+}} \sum_{\omega_{l} \in \tilde{\Omega}^{+}} \mathbf{a}(\omega_{\nu}, \omega_{l}) \mathbf{b}(\mathbf{R}^{T}(\Phi)\omega_{\nu}, \mathbf{R}^{T}(\Phi)\omega_{l}).(3)$$

Please note that this formulation requires the ABRDFs to

be defined on the lower hemisphere as well. In our current implementation we simply set values from the lower hemisphere to zero. Maximizing $C_{\mathbf{a},\mathbf{b}}(\Phi)$ now corresponds to determining the Euler angles $\bar{\Phi}$ which maximize the correlation between \mathbf{a} and the rotated \mathbf{b} :

$$\bar{\Phi} = \arg\max_{\Phi} C_{\mathbf{a},\mathbf{b}}(\Phi) \tag{4}$$

The alignment transformation $\mathbf{R}(\tilde{\Phi})$ then defines an approximate local coordinate frame at the corresponding surface position.

The question now arises how to align a set of ABRDFs $\mathcal{A} = \{\mathbf{a}_i\}_{0 < i \leq N}$. If we again assume that \mathcal{A} is a set of noisy misaligned instances of the same BRDF all we need is a target ABRDF \mathbf{c} to which all other ABRDFs are registered. The result will be a set $\mathcal{R} = \{\mathbf{R}_i\}$ of alignment transformations. Since \mathbf{c} is not known in the beginning we propose a simple expectation-maximization procedure to find \mathbf{c} and \mathcal{R} simultaneously (Algorithm 1).

return Set of alignment rotations
$$\mathcal{R} = \{\mathbf{R}_i\}$$

 $j = 0$
 $\mathbf{c}_0 = \mathbf{a}_i \ i \in \{1..N\}$ {choose \mathbf{c}_0 from \mathcal{A} randomly}
repeat

Algorithm 1 Alignment of ABRDFs $A = \{a_i\}_{0 < i < N}$

$$\begin{aligned} & \textbf{for } i = 1 \text{ to } N \textbf{ do} \\ & \Phi_i = \operatorname{argmax}_{\Phi} C_{\mathbf{a}_i, \mathbf{c}_j}(\Phi) \\ & \mathbf{R}_i = \mathbf{R}(\Phi_i) \\ & \textbf{end for} \\ & \mathbf{c}_{j+1}(\omega_{\nu}, \omega_{l}) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{a}_i(\mathbf{R}_i^T \omega_{\nu}, \mathbf{R}_i^T \omega_{l}) \end{aligned}$$

until $\varepsilon < \tau$ { ε is the increase in accumulated correlation $\sum_{i=1}^{N} C_{\mathbf{a}_i, \mathbf{c}_i}(\Phi_i)$ between two iterations}

The method is inspired by curve alignment techniques where the alignment curve is computed as the mean of the curves to be aligned.

4.2. Making it fast

The most interesting and costly statement of Algorithm 1 is the maximization step (Equation (4)). Simply calculating the correlation for a dense set of grid points in the space of the rotation group SO(3) and returning the point with maximum correlation will be horribly slow. E.g., if we assume a desired accuracy of one degree in all three dimensions we actually have to calculate the correlation for $360 \times 90 \times 360$ rotations which is infeasible. Please note also that the rotated directions usually will not be in the set $\tilde{\Omega}^+$ of sampled directions necessitating a costly resampling. Therefore, the general strategy will be to evaluate the correlation only at a sparse grid of rotations and to use the location of the found maximum as initialization for a non-linear optimization step. Of course we then could end up in local minima, but even a non optimal alignment will be better than no alignment.

4.2.1. Spherical correlation

In order to speed up the computation of Equation (3) itself we apply techniques from Fourier theory over the sphere. It is convenient for the derivation of this optimization to formulate the correlation term for 2D reflection functions and for the continuous domain of the unit sphere S^2 . A generalization to 4D ABRDFs will be given afterwards. The modified correlation term reads as follows:

$$C_{\mathbf{a},\mathbf{b}}^{S^2}(\Phi) = \int_{\omega \in S^2} \mathbf{a}(\omega) \overline{\mathbf{b}(\mathbf{R}^T(\Phi)\omega)} \, d\omega \tag{5}$$

This integral has the form of a spherical convolution such that the generalized Fourier shift theorem is applicable [KR03]. Indeed, we can decompose bandlimited $\bf a$ and $\bf b$ with maximum bandwidth $\bf L$ into the Spherical Harmonics (SH) basis:

$$\mathbf{a}(\mathbf{\omega}) = \sum_{l=0}^{L} \sum_{|m| \leq l} \hat{a}_{m}^{l} Y_{m}^{l}(\mathbf{\omega})$$

$$\mathbf{b}(\omega) = \sum_{l=0}^{L} \sum_{|m| \leq l} \hat{b}_{m}^{l} Y_{m}^{l}(\omega).$$

The generalized Fourier shift theorem now stems from the fact that a rotated SH-basis function can be written as a linear combination of SH-basis functions of the same band:

$$Y_m^l(\mathbf{R}^T(\Phi)\omega) = \sum_{|k| \le l} D_{km}^l(\Phi) Y_k^l(\omega)$$
 (6)

Here $D_{km}^{l}(\Phi)$ denotes a Wigner-D function. From (6) follows

$$\mathbf{b}(\mathbf{R}^{T}(\mathbf{\Phi})\mathbf{\omega}) = \sum_{l=0}^{L} \sum_{|k| \le l} \left(\sum_{|m| \le l} \hat{b}_{m}^{l} D_{km}^{l}(\mathbf{\Phi}) \right) Y_{k}^{l}(\mathbf{\omega}) \quad (7)$$

which by exploiting the orthogonality of the SH basis immediately leads to

$$C_{\mathbf{a},\mathbf{b}}^{S^2}(\Phi) = \sum_{l=0}^{L} \sum_{|k| \le l} \sum_{|m| \le l} \hat{a}_k^l \overline{\hat{b}_m^l D_{km}^l(\Phi)}$$
(8)

Since the right hand side of Equation (8) has the form of the Fourier decomposition of $C_{\mathbf{a},\mathbf{b}}^{S^2}(\Phi)$ in the rotation group SO(3) the generalized Fourier shift theorem follows immediately by comparing coefficients:

$$\hat{c}_{mn}^l = \hat{a}_m^l \overline{\hat{b}_n^l} \tag{9}$$

which now (analogue to convolutions in the planar domain) allows to compute the Fourier coefficients \hat{c}_{mn}^l of $C_{\mathbf{a},\mathbf{b}}^{S^2}(\Phi)$ by multiplying the Fourier coefficients of $\mathbf{a}(\omega)$ and $\mathbf{b}(\omega)$ (in fact, an outer product has to be computed per band).

The direct evaluation of Equation (8) has complexity $O(L^3)$. But fortunately, there exists an (inverse) Fast Fourier Transform (FFT) [KR03] which evaluates Equation (8) for a grid of $(2L+1)^3$ rotations with complexity $O(L^3\log^2 L)$ compared to $O(L^6)$ for the direct evaluation.

L	Eqn. (5) $ \tilde{\Omega}^+ = 151$	Eqn. (8) direct	Eqn. (8) FFT
4	0.325	0.685	0.029
8	1.873	27.673	0.336
16	15.144	>1000	5.326

Table 1: Average computation times (in milliseconds) for spherical correlation on an $O(L^3)$ grid for varying L on an AMD Athlon 3200+.

We implemented both the direct evaluation of Equation (8) and its efficient evaluation by applying inverse FFT using the freely available C-library SOFT [SOF]. Furthermore, we implemented the "brute-force" evaluation of Equation (5) using numerical integration. Table 1 shows timing results for different values of *L*.

4.2.2. Generalization to 4D

To generalize the results from the previous section to the ABRDF correlation in Equation (3) we follow a tensor product approach, i.e., we represent an ABRDF in the tensor product basis of Spherical Harmonics:

$$\mathbf{a}(\omega_{v},\omega_{l}) = \sum_{l_{1},l_{2}}^{L} \sum_{m_{1},m_{2}} \hat{a}_{m_{1}m_{2}}^{l_{1}l_{2}} Y_{m_{2}}^{l_{2}}(\omega_{v}) Y_{m_{1}}^{l_{1}}(\omega_{l})$$

Repeating the steps from the previous section (we omit them here for brevity) leads to the following lengthy sum

$$C_{\mathbf{a},\mathbf{b}}^{S^2 \times S^2}(\Phi) = \sum_{l_1 l_2}^{L} \sum_{m_1 m_1'} \sum_{m_2 m_2'} \hat{a}_{m_1 m_2}^{l_1 l_2} \hat{b}_{m_1' m_2'}^{l_1 l_2} \hat{b}_{m_2 m_2'}^{l_1 l_2}(\Phi) D_{m_1 m_1'}^{l_1}(\Phi)$$
(10)

After a simple reordering of terms the direct evaluation of this sum has complexity $O(L^5)$ which would leave us with an expensive $O(L^8)$ algorithm. Applying the inverse FFT in a tensor product fashion will result in a still expensive $O(L^6log^2L)$ algorithm and returns the correlation for arbitrary combinations of rotations $(\Phi_1, \Phi_2) \in SO(3) \times SO(3)$ while we require only those elements where $\Phi_1 = \Phi_2$. We currently have not found a way to exploit this.

Another problem of the tensor product representation is the memory requirement. For L=8 each ABRDF consists of 4096 SH coefficients which easily sums up to hundreds of megabytes of data for typical datasets. Therefore, we propose to factorize the ABRDFs ${\bf a}, {\bf b}$ into a sum of products of 2D functions which reduces both the memory requirements and computational complexity. Specifically, we find representations of the form

$$\mathbf{a}(\omega_{\nu}, \omega_{l}) \approx \sum_{i}^{K} a_{1,i}(\omega_{\nu}) a_{2,i}(\omega_{l})$$
$$\mathbf{b}(\omega_{\nu}, \omega_{l}) \approx \sum_{i}^{K} b_{1,i}(\omega_{\nu}) b_{2,i}(\omega_{l})$$

using for example SVD. Substituting these representations

L	Eqn. (3)	Eqn. (10)	Eqn. (11) K=4	K=6
4	291.2	26.86	1.145	2.591
8	>1500	>3500	10.683	23.803
16	N.N.	N.N.	195.937	445.118

Table 2: Average computation times (in milliseconds) for correlation of ABRDFs on an $O(L^3)$ grid for varying L on an AMD Athlon 3200+. Direct approaches are not feasible for L > 4.

for a and b leads to a modified correlation term

$$\tilde{C}_{\mathbf{a},\mathbf{b}}^{S^2 \times S^2}(\Phi) = \sum_{i}^{K} \sum_{j}^{K} C_{a_{1,i},b_{1,j}}^{S^2}(\Phi) C_{a_{2,i},b_{2,j}}^{S^2}(\Phi)$$
(11)

which reduces the 4D correlation to a sum of products of 2D correlations. Using again the inverse FFT for spherical correlations we now have an $O(K^2L^3\log^2L)$ algorithm. In our experiments we set $4 \le K \le 8$ but typically 4 components suffice without degrading the results (see Figure 7), since the main features relevant for the alignment are captured in this representation.

As in the 2D case we implemented the direct evaluation of Equation (10) and simple numerical integration for comparison. The huge performance benefits gained from Equation (11) are illustrated in Table 2. Using the settings L=8 and K=4 we align about one hundred ABRDFs per second for a sufficiently dense grid of rotations. This adds up to computation times of about one hour for the alignment of a complete dataset of 256^2 ABRDFs using four iterations of Algorithm 1 and including the preprocessing time required for factorization and SH projection.

4.3. Aligning ABRDFs for objects with different materials

Most objects are made of different materials which may vary across the surface. Applying the approach from Section 4.1 will then align ABRDFs which have been measured from probably very different BRDFs which may lead to at least sub-optimal results. Our solution is to apply a k-means clustering algorithm with an orientation invariant distance metric. The metric is based on a normalized version of the correlation term (3):

$$d(\mathbf{a}, \mathbf{b}) = 1 - \max_{\mathbf{\Phi}} \bar{C}_{\mathbf{a}, \mathbf{b}}(\mathbf{\Phi})$$
 (12)

 $d(\mathbf{a}, \mathbf{b})$ defines a so-called semi-metric (triangle inequality does not hold) and we use it as an orientation invariant distance metric for the k-means clustering. By using this metric an ABRDF will be assigned to the cluster where the correlation between the cluster center and the suitable rotated ABRDF is maximized.

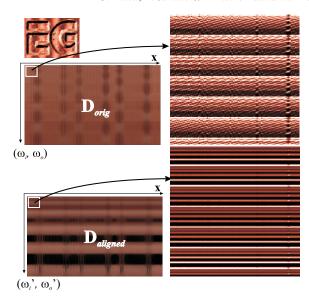


Figure 4: The original data matrix \mathbf{D}_{orig} is turned into an aligned data matrix $\mathbf{D}_{aligned}$ by our algorithm. The images on the right are zoom-ins and show how well the features are aligned along the angular dimension: the diagonal features have been turned into almost perfect horizontal features.

5. Results

To test our algorithm we applied it to three kinds of data sets: synthetically generated data, measured BTF data and measured 6D surface reflectance fields. The measured data sets consist of 22801 images captured from 151 different view directions and 151 lighting directions. They were measured with a multi-camera dome developed at the university of Bonn [MBK05]. The storage requirement of such uncompressed bidirectional datasets in high-dynamic range easily exceed tens of gigabytes.

To evaluate the improvement in compression performance achieved by the method we applied SVD to the datasets before and after alignment. We used online SVD [Bra03] to decompose the data matrices since their size easily exceeds the maximum size of a single memory block.

5.1. Synthetic Data

To begin the tests of our algorithm in a controlled environment we generated a synthetic BTF by generating a map of local coordinate frames using a normal map (124×90 pixels) calculated from the Eurographics logo and a cosine modulated tangent map. We then generated an anisotropic BRDF using the Ashikhmin-Shirley BRDF model [AS00] and rotated it for every tangent frame given in the map. We sampled this synthetic BTF at $81 \times 81 = 6561$ directions and applied the alignment algorithm described in the previous section. Using three iterations and the settings L=8

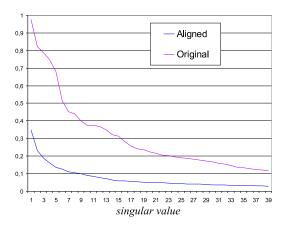


Figure 5: The decay of singular values for the decomposition of \mathbf{D}_{orig} and $\mathbf{D}_{aligned}$ respectively.



Figure 6: Left: original normal map (RGB encoded) used to generate the synthetic BTF. Right: normal map resulting from our alignment algorithm.

and K=5 the alignment took about five minutes. Figures 4-6 show the results. Note how well the angular features of the ABRDFs are aligned and the dramatic variance reduction that follows from that alignment. Without alignment 20 components are needed to achieve the same variance reduction as achieved with 2 components for the aligned data. Note also the high quality of the recovered normal map which suggests that our algorithm might be useful for small-scale 3D-reconstruction from reflectance measurements.

5.2. Bidirectional texture functions

BTFs can be understood as 6D surface reflectance fields with planar base geometry. Therefore, no special care has to be taken to handle large scale occlusions or shadows and the local coordinate system is usually defined by the plane of the sample holder and the up direction of the texture. Nevertheless, in many cases it is worthwhile to derive per-pixel coordinate systems as can be seen for the two car interior materials shown in Figure 8. The variance reduction achieved for the *plastic knobs* material is shown in Figure 7. From our experience the alignment technique is most useful for BTFs when extreme compression rates are required like, for example, in real-time rendering. In this case the fitted local coor-

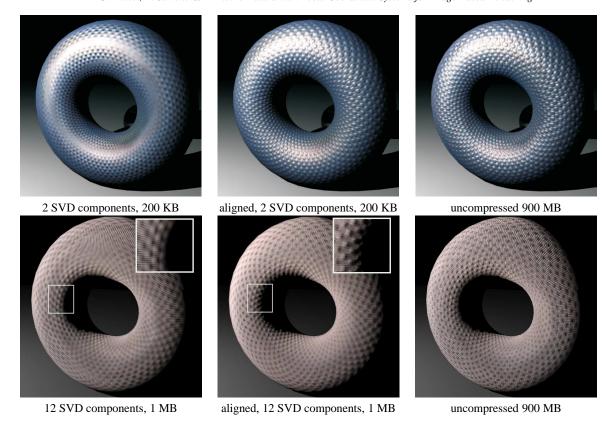


Figure 8: Extreme compression of the plastic knobs and grey textile BTF datasets. The structure is drastically smoothed or even disappears (see bottom left image, upper part of torus) for SVD-only compression. After alignment much more structure is preserved.

dinate systems already capture much of the material's structure, so that the remaining available memory can be used to reconstruct the correct shape of the BRDF. It has to be mentioned that the additional rendering costs introduced by using per-pixel local coordinate systems are negligible.

5.3. 6D surface reflectance fields

In the case of general 6D reflectance fields the problem of inconsistent coordinate systems is even more pressing than for BTFs. While BTFs are usually textures with small scale variations which are repeated across the sample, a general 3D object typically contains much more variance in surface appearance. Furthermore, additional variance is introduced by large scale occlusions and shadows. As in [CHLG05] we applied online SVD with missing values [Bra03] to reconstruct the missing data.

The geometry of the captured object was reconstructed using a visual hull approach. Then the resulting geometry was parameterized using the Maya texturing tool and the images where projected onto this geometry and resampled into the texture domain. The resulting Lempel-Ziv compressed dataset has a size of 2 GB. After applying our alignment algorithm with 16 clusters we compressed the data using Local PCA with 16 clusters and 10 components per cluster, which resulted in a datasize of 22 MB. Figure 9 shows the normal map computed by our algorithm and compares the rendering from the compressed to a rendering generated from the uncompressed data. Apart from a small color degradation and some artifacts at the handle the difference is hardly noticeable.

Figure 10 shows a few examples of the photorealism that can be achieved by combining 6D reflectance fields with image-based lighting. The overall data size used to generate the images in Figure 10 was 34 MB.

6. Conclusions

In this paper we proposed an algorithm which estimates data-driven local coordinate systems for image-based rendering. These local coordinate systems were chosen as the orientation which maximizes the correlation between the ap-

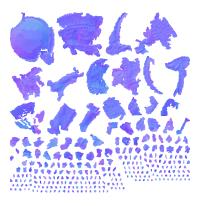






Figure 9: The precious vase data set. Left: the normal map computed by our algorithm. Middle: rendering from compressed representation (Local PCA - 22 MB). Right: rendering from losslessly compressed data (Lempel-Ziv - 2 GB)







Figure 10: The precious vase dataset and a textile BTF combined with image-based lighting: Arbitrary viewpoints under arbitrary lighting with photo-realistic appearance.

parent BRDFs of an image-based data set. To compute this maximization efficiently we evaluated the correlation term in Fourier space by expanding the ABRDFs in the Spherical Harmonics basis. We then applied a fast inverse Fourier transformation for functions defined on the rotation group, SO(3). As shown by several examples, the data alignment achieved by our algorithm leads to significant variance reduction in the data which improves the compression performance for BTFs and 6D surface reflectance field.

For future work we plan to use our algorithm to enable high-quality real-time rendering of complex 6D surface reflectance fields. We will investigate if the method can be used for the faithful 3D reconstruction of textures. In that context it will be also interesting to compare and even combine our approach with 3D reconstruction techniques for small scale geometry like that of Wang et al. [WTS*05].

References

[AB91] ADELSON E. H., BERGEN J. R.: The plenoptic function and the elements of early vision. *Computational Models of Visual*

Processing (1991).

[AS00] ASHIKHMIN M., SHIRLEY P.: An anisotropic Phong BRDF model. *Journal of Graphics Tools: JGT 5*, 2 (2000), 25–32.

[Bra03] BRAND M.: Fast online svd revisions for lightweight recommender systems. In SIAM International Conference on Data Mining (SDM) (May 2003).

[CBCG02] CHEN W.-C., BOUGUET J.-Y., CHU M. H., GRZESZCZUK R.: Light field mapping: Efficient representation and hardware rendering of surface light fields. *Proceedings of ACM SIGGRAPH 2002* (2002).

[CCST00] CHAI J.-X., CHAN S.-C., SHUM H.-Y., TONG X.: Plenoptic sampling. In SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques (New York, NY, USA, 2000), ACM Press/Addison-Wesley Publishing Co., pp. 307–318.

[CHLG05] COOMBE G., HANTAK C., LASTRA A., GRZESZCZUK R.: Online construction of surface light fields. In Eurographics Symposium on Rendering (2005).

[DHT*00] DEBEVEC P., HAWKINS T., TCHOU C., DUIKER H.-

© The Eurographics Association and Blackwell Publishing 2006.

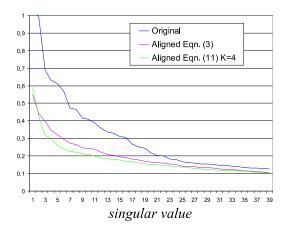


Figure 7: The decay of singular values for the plastic knobs BTF data set before and after alignment. The improvement is not as dramatic as for the synthetic dataset but still significant. Using the optimization techniques described in section 4 does not degrade the result compared to "brute-force" numerical integration.

- P., SAROKIN W., SAGAR M.: Acquiring the reflectance field of a human face. *Proceedings of ACM SIGGRAPH 2000* (2000).
- [FKIS02] FURUKAWA R., KAWASAKI H., IKEUCHI K., SAKAUCHI M.: Appearance based object modeling using texture database: acquisition, compression and rendering. In Proceedings of EGRW '02 (2002), Eurographics Association, pp. 257–266.
- [HWL05] Ho P.-M., Wong T.-T., Leung C.-S.: Compressing the illumination-adjustable images with principal component analysis. *IEEE Transactions on Circuits and Systems for Video Technology* 15, 3 (2005), 355–364.
- [KM99] KAUTZ J., MCCOOL M.: Interactive Rendering with Arbitrary BRDFs using Separable Approximations. In *Tenth Eu*rographics Workshop on Rendering (1999), pp. 281–292.
- [KR03] KOSTELEC P., ROCKMORE D.: FFTs on the rotation group. Tech. Rep. 03-11-060, Santa Fe Institute's Working Paper Series, 2003.
- [LFTG97] LAFORTUNE E. P. F., FOO S.-C., TORRANCE K. E., GREENBERG D. P.: Non-linear approximation of reflectance functions. In *Proceedings of SIGGRAPH* (1997), ACM Press/Addison-Wesley Publishing Co., pp. 117–126.
- [LH96] LEVOY M., HANRAHAN P.: Light field rendering. Computer Graphics 30, Annual Conference Series (1996), 31–42.
- [LHZ*04] LIU X., HU Y., ZHANG J., TONG X., GUO B., SHUM H.-Y.: Synthesis and Rendering of Bidirectional Texture Functions on Arbitrary Surfaces. *IEEE Transactions on Visualization* and Computer Graphics 10, 3 (2004), 278–289.
- [LKG*03] LENSCH H. P., KAUTZ J., GOESELE M., HEIDRICH W., SEIDEL H.-P.: Image-based reconstruction of spatial appearance and geometric detail. ACM Transactions on Graphics 2, 22 (2003), 234–257.

- [MBK05] MÜLLER G., BENDELS G. H., KLEIN R.: Rapid synchronous acquisition of geometry and btf for cultural heritage artefacts. In *The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST)* (November 2005), Eurographics Association, Eurographics Association, pp. 13–20.
- [MEG00] MAGNOR M., ENDMANN A., GIROD B.: Progressive compression and rendering of light fields. Proc. Vision, Modeling, and Visualization (VMV-2000), Saarbrücken, Germany (Nov. 2000), 199–203.
- [MLH02] MCALLISTER D., LASTRA A., HEIDRICH W.: Efficient rendering of spatial bi-directional reflectance distribution functions. *Graphics Hardware* 2002 (2002).
- [MMK03] MÜLLER G., MESETH J., KLEIN R.: Compression and real-time Rendering of Measured BTFs using local PCA. In Vision, Modeling and Visualisation 2003 (November 2003), pp. 271–280.
- [MPN*02] MATUSIK W., PFISTER H., NGAN A., BEARDSLEY P., ZIEGLER R., MCMILLAN L.: Image-based 3d photography using opacity hulls. In *Proceedings of SIGGRAPH* (2002), ACM Press, pp. 427–437.
- [NSI99] NISHINO K., SATO Y., IKEUCHI K.: Eigen-texture method: appearance compression based on 3d model. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'99) (June 1999), vol. 1, pp. 618 – 624.
- [Rus98] RUSINKIEWICZ S.: A new change of variables for efficient BRDF representation. In *Rendering Techniques* '98 (New York, NY, 1998), Drettakis G., Max N., (Eds.), Springer Wien, pp. 11–22.
- [SHHS03] SLOAN P.-P., HALL J., HART J., SNYDER J.: Clustered Principal Components for Precomputed Radiance Transfer. ACM Transactions on Graphics 22, 3 (2003), 382–391.
- [SOF] http://www.cs.dartmouth.edu/~geelong/soft/.
- [SSK03] SATTLER M., SARLETTE R., KLEIN R.: Efficient and realistic visualization of cloth. Proceedings of the Eurographics Symposium on Rendering 2003 (2003).
- [VT04] VASILESCU M. A. O., TERZOPOULOS D.: Tensortextures: Multilinear image-based rendering. In *Proceedings of SIG-GRAPH* (August 2004).
- [WAA*00] WOOD D. N., AZUMA D. I., ALDINGER K., CUR-LESS B., DUCHAMP T., SALESIN D. H., STUETZLE W.: Surface light fields for 3D photography. In *Proceedings of SIG-*GRAPH (2000), pp. 287–296.
- [War92] WARD G. J.: Measuring and modeling anisotropic reflection. In *Proceedings of SIGGRAPH* (1992), ACM Press, pp. 265–272.
- [WHON97] WONG T.-T., HENG P.-A., OR S.-H., NG W.-Y.: Image-based rendering with controllable illumination. In *Proceedings of EGRW '97* (1997), Springer-Verlag, pp. 13–22.
- [WTS*05] WANG J., TONG X., SNYDER J., CHEN Y., GUO B., SHUM H.-Y.: Capturing and rendering geometry details for btfmapped surfaces. *The Visual Computer 21*, 8-10 (September 2005), 559–568.
- [WWS*05] WANG H., WU Q., SHI L., YU Y., AHUJA N.: Out-of-core tensor approximation of multi-dimensional matrices of visual data. ACM Trans. Graph. 24, 3 (2005), 527–535.