Kernel Nyström Method for Light Transport

Jiaping Wang[*]   Yue Dong[†][*]   Xin Tong[*]   Zhouchen Lin[*]   Baining Guo[*][†]

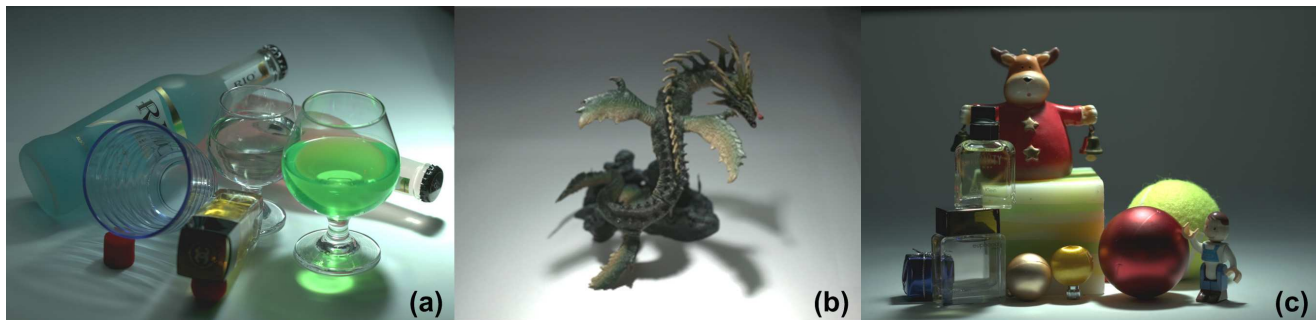[*]Microsoft Research Asia   [†]Tsinghua Univeristy

Figure 1: Relighting results using the light transport matrix reconstructed by our method. Complex light transport effects, including caustics (a), complex occlusions (b), and a mixture of caustics, complex occlusions, inter-reflections, and subsurface scattering (c) are all faithfully reproduced.

Abstract

We propose a kernel Nyström method for reconstructing the light transport matrix from a relatively small number of acquired images. Our work is based on the generalized Nyström method for low rank matrices. We introduce the light transport kernel and incorporate it into the Nyström method to exploit the nonlinear coherence of the light transport matrix. We also develop an adaptive scheme for efficiently capturing the sparsely sampled images from the scene. Our experiments indicate that the kernel Nyström method can achieve good reconstruction of the light transport matrix with a few hundred images and produce high quality relighting results. The kernel Nyström method is effective for modeling scenes with complex lighting effects and occlusions which have been challenging for existing techniques.

1 Introduction

The goal of image-based relighting is to directly capture the light transport of a real-world scene so that it can be rendered with new illumination [Debevec et al. 2000; Wenger et al. 2005; Peers et al. 2009]. Mathematically, image-based relighting can be formulated as the following equation [Ng et al. 2003; Peers et al. 2009]:

$$\mathbf{b} = \mathbf{T} \cdot \mathbf{l}, \qquad (1)$$

where $\mathbf{T}$ is the $m \times n$ light transport matrix that describes the light transport from n light sources to m camera pixels, $\mathbf{l}$ is the illumination condition represented by a vector of incident radiance from n

[*]email:{jiapw,xtong,zhoulin,bainguo}@microsoft.com

light sources, and $\mathbf{b}$ is the outgoing radiance observed in a camera image with m pixels. Our task is to find the matrix $\mathbf{T}$.

Image-based relighting offers an unparalleled advantage of realistic rendering without scene modeling, which is often an arduous task. However, to generate high quality results, the existing techniques often require tens of thousands of images for accurate reconstruction of the light transport matrix (e.g., see [Debevec et al. 2000; Wenger et al. 2005]). Several approaches have shown promise in reducing the number of required images [Zongker et al. 1999; Matusik et al. 2004; Peers and Dutré 2005; Peers et al. 2009; Sen and Darabi 2009]. However, these techniques are either dedicated to specific light transport effects [Zongker et al. 1999] or mainly effective with scenes of simple geometry configurations [Matusik et al. 2004; Peers and Dutré 2005]. Applying these techniques to scenes with complex lighting effects and occlusions still requires a large number of input images and expensive reconstruction methods [Peers et al. 2009].

In this paper we propose a *kernel Nyström* method for reconstructing the light transport matrix from a small number of images. We first acquire a small number of rows and columns of the light transport matrix of a real world scene and then reconstruct the entire matrix from these sparse samples. Our work is inspired by the Nyström method proposed by Williams and Seeger [2000] for reconstructing a low rank symmetric matrix using a linear combination of sparsely sampled rows and columns. In this paper, we apply a generalized form of the Nyström method [Goreinov et al. 1997] previously developed in the linear algebra community for reconstructing asymmetrical matrices.

The main contribution of this paper is the introduction of the light transport kernel so that the nonlinear coherence of the light transport matrix can be exploited to enhance the effectiveness of the Nyström method. The kernel Nyström method refers to the Nyström method with the light transport kernel incorporated. The power of the Nyström method essentially comes from its ability to exploit the *linear* coherence of the light transport matrix. By further exploiting the *nonlinear* coherence, the kernel Nyström method becomes much more powerful and the number of sampled rows and columns (and hence acquired images) needed for the matrix reconstruction is greatly reduced. The light transport kernel works

as follows. First, we design a data specific kernel function that can be estimated from the sparse row and column samples. Then by mapping these samples with the kernel function, we reduce the rank of the light transport matrix so that a high quality matrix can be reconstructed from the kernel-mapped sparse samples via the Nyström method. Finally, the light transport matrix is obtained from elements of the reconstructed matrix by inverse kernel mapping.

A novel aspect of the kernel Nyström method is that it exploits the data coherence during the matrix reconstruction process, before the whole matrix is known. It has been recognized in the past that the data coherence in light transport can be used for data compression after the whole matrix is known [Masselus et al. 2004; Mahajan et al. 2007]. The kernel Nyström method takes a different approach and only uses the known row and column samples as an oracle for analyzing the data coherence and reconstructing the matrix. To this end, we develop an adaptive scheme for measuring the sparse row and column images of the light transport matrix and estimating the light transport kernel.

We have evaluated the kernel Nyström method using a variety of test scenes, including those with inter-reflections, caustics, and complex occlusions. We also used the method to measure the light transport between object surface points due to diffuse subsurface scattering inside the object volume. Our experiments indicate that the kernel Nyström method can achieve good reconstruction of the light transport matrix with a few hundred images and produce high quality relighting results. Moreover, the kernel Nyström method is effective for modeling scenes with complex lighting effects and occlusions which have remained challenging for existing techniques.

2   Related Work

The light transport matrix represents discrete samples of the reflectance field [Debevec et al. 2000]. A complete 8D reflectance field, which describes the light transport from the incident light field to the outgoing light field [Levoy and Hanrahan 1996], is difficult to capture and process [Garg et al. 2006]. Therefore, most existing methods only consider simplified 4D [Debevec et al. 2000; Lin et al. ; Matusik et al. 2004; Peers and Dutré 2005] and 6D reflectance fields [Matusik et al. 2002; Masselus et al. 2003; Sen et al. 2005; Wenger et al. 2005]. In this paper, we focus on 4D reflectance fields with a fixed viewpoint and point light sources that lie in a 2D plane.

Light Transport Acquisition. We categorize the existing methods into three classes: brute force, sparsity based, and coherence based methods.

The *brute force methods* directly measure the light transport matrix from the scene, where each column is an image of the scene lit by a single light source in the incident light domain. Debevec et. al. [2000] developed a light stage device for capturing 4D reflectance fields for a fixed viewpoint and distant lighting by moving a point light source around the object. They later improved the device for fast capturing [Wenger et al. 2005]. Hawkins et. al. [2005] exploited the Helmholtz reciprocity to capture the reflectance field of highly reflective objects. To obtain dense samples in the incident light domain, rows of the light transport matrix are captured by shooting rays from the viewpoint and capturing the high resolution image of the scene projected over the incident light domain. Reciprocity is also exploited in [Sen et al. 2005] for acquiring 6D reflectance fields. All these methods require tens of thousands of images for modeling a high quality light transport matrix. For similar quality, our approach only requires a few hundred images.

The *sparsity based methods* model the light transport matrix with a set of basis functions defined over the incident light domain and assume that each row of the light transport matrix can be approximated by a linear combination of a sparse set of basis functions. Thus the light transport matrix can be reconstructed by deriving the sparse basis and their weights for each row from a set of images captured under special lighting conditions. Environment matting [Zongker et al. 1999] models the reflectance of specular or refractive objects by representing the light transport of each pixel (i.e., a row of transport matrix) with a single 2D box function. It was later extended for modeling glossy objects by replacing the box function with an oriented Guassian kernel [Chuang et al. 2000]. Matusik et. al. [2004] modeled the light transport matrix with hierarchical rectangular basis functions. An adaptive algorithm is developed for deriving the sparse basis and their weights for each pixel from images of the scene captured under various natural illumination conditions. Peers et al. [2005] modeled the light transport matrix with wavelets and inferred the light transport matrix from images of the scene illuminated by carefully designed wavelet noise patterns. Both approaches apply a greedy strategy to find a suboptimal sparse basis for each pixel, which only works well for scenes with simple occlusions. Recently a compressive sensing approach [Peers et al. 2009] was proposed which computes the solution for each pixel from images captured from a scene illuminated by patterned lighting. However, the number of images needed for reconstruction depends on both the row length and the number of basis used for each row, which becomes quite large for scenes with complex occlusions. The reconstruction process is also time consuming.

Unlike the above sparsity based approaches, our method exploits the coherence in the matrix for reconstruction. It can handle well scenes with complex occlusions and caustics. The number of images needed for reconstruction is only proportional to the rank of the light transport matrix and independent on the size of the matrix. Moreover, our reconstruction algorithm consists of a set of matrix operations, which is simple and fast.

Coherence based methods acquire the light transport matrix by exploiting the coherence of reflectance field data. Masselus et. al. [2004] explored the interpolation and compression of reflectance field. Fuchs et. al. [2007] proposed an adaptive sampling scheme for sampling the 4D reflectance field. The spatial coherence of the reflectance field in the incident domain is exploited for accelerating the acquisition process. Starting from a set of images taken with a sparse set of regularly distributed lighting directions, their algorithm analyzes the observed data and then captures more images in each iteration with the new lighting directions where the reflectance field is not smooth. Note that the smoothness of reflectance data among neighboring pixels is also exploited in [Matusik et al. 2004; Peers and Dutré 2005; Peers et al. 2009] for improving the result quality.

While these methods only exploit the coherence in either rows or columns, our method exploits the data coherence in the entire light transport matrix for reconstruction. Since our approach makes no assumptions about the smoothness of the sampled reflectance field, it can handle well sharp variations of the light transport, such as shadows, caustics, and surface textures.

Low Rank Matrix Approximation. In addition to the reconstruction of low rank symmetric matrices, the Nyström method is also widely used in the machine learning community for approximately computing the eigenvalues and eigenvectors of a symmetric matrix from sparse matrix samples [Platt 2005]. For asymmetric matrices, Goreinov et. al. [1997] presented a pseudoskeleton approximation for reconstructing a complete matrix from a sparse collection of its rows and columns. In graphics research, An et. al. [2008]

applied the Nyström method to accelerate appearance edit propagation by approximating the dense symmetric distance matrix with sparsely sampled rows and columns. For given illumination, Hasan et. al. [2007] applied similar techniques to efficiently render synthetic scenes. In their approach, columns of the matrix are clustered in a small number of groups according to their values in the sparsely sampled rows. A representative column for each group is then sampled and weighted for approximating other columns in the same group. This approach only uses the coherence between the columns for approximating the matrix. Coherence between the rows is not exploited. Moreover, applying this approach to the light transport matrix may generate temporal artifacts under animated light, as noted in [Hašan et al. 2008].

3 The Kernel Nyström Method

In this section, we first review the generalized Nyström method for asymmetric matrices [Goreinov et al. 1997] and then introduce our kernel Nyström method. Here we assume that a sparse set of rows and columns of the light transport matrix is known. We will discuss how to capture individual rows and columns from a real world scene in the next section.

Throughout this paper, we use the following notational convention: the matrix is denoted by bold uppercase letters, e.g. $\mathbf{T}$, a vector is denoted by bold lowercase letters, e.g. $\mathbf{l}$, and a scalar or a scalar function is denoted by lowercase italic letters, e.g. f. Given a matrix $\mathbf{T}$, its element at row i and column j is denoted as $\mathbf{T}_{ij}$, while $f(\mathbf{T})$ denotes a matrix obtained by applying f to each element of the matrix $\mathbf{T}$.

3.1 Asymmetric Generalization

The Nyström method in [Williams and Seeger 2000] reconstructs a low rank symmetric matrix from sparsely sampled columns. As shown in Figure 2(a), an unknown $(n+k) \times (n+k)$ symmetric matrix $\mathbf{T}$ with k sampled rows $[\ \mathbf{A}\ \mathbf{C}^T\ ]$ can be approximated as

$$\mathbf{T} = \begin{bmatrix} \mathbf{A} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{B} \end{bmatrix} \approx \begin{bmatrix} \mathbf{A} & \mathbf{C}^T \\ \mathbf{C} & \mathbf{C}\mathbf{A}^{-1}\mathbf{C}^T \end{bmatrix}. \qquad (2)$$

The reconstruction is accurate when the symmetric matrix $\mathbf{T}$ has a rank $d \le k$, except that the sampled rows $[\ \mathbf{A}\ \mathbf{C}^T\ ]$ are of a rank smaller than d.

For asymmetric light transport matrices in which the image pixels (columns) and light sources (rows) are sampled in different spaces, the generalized Nyström method [Goreinov et al. 1997] can be applied for reconstruction. As illustrated in Figure 2(b), we assume r rows $[\ \mathbf{A}\ \mathbf{R}\ ]$ and c columns $[\ \mathbf{A}^T\ \mathbf{C}^T\ ]^T$ are known out of an $(r+m) \times (c+n)$ asymmetric matrix $\mathbf{T}$. If the rank of $\mathbf{T}$ equals to that of $\mathbf{A}$: $d = rank(\mathbf{T}) = rank(\mathbf{A})$, we can expect the following relationship:

$$[\ \mathbf{C}\ \ \mathbf{B}\ ] = \mathbf{P}[\ \mathbf{A}\ \ \mathbf{R}\ ] \quad \text{and} \quad \begin{bmatrix} \mathbf{R} \\ \mathbf{B} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{C} \end{bmatrix} \mathbf{Q}, \quad (3)$$

where $\mathbf{P}$ and $\mathbf{Q}$ are matrices of appropriate sizes, which implies that $\mathbf{C} = \mathbf{PA}$, $\mathbf{R} = \mathbf{AQ}$ and $\mathbf{B} = \mathbf{PR} = \mathbf{CQ}$. Thus the missing portion $\mathbf{B}$ in the matrix $\mathbf{T}$ can be reconstructed as:

$$\mathbf{B} = \mathbf{PAQ} = \mathbf{PAA}^+\mathbf{AQ} = \mathbf{CA}^+\mathbf{R}, \qquad (4)$$

where $\mathbf{A}^+$ denotes the Moore-Penrose pseudoinverse of $\mathbf{A}$, which has the property $\mathbf{AA}^+\mathbf{A} = \mathbf{A}$. So the reconstruction is

$$\mathbf{T} = \begin{bmatrix} \mathbf{A} & \mathbf{R} \\ \mathbf{C} & \mathbf{CA}^+\mathbf{R} \end{bmatrix}. \qquad (5)$$
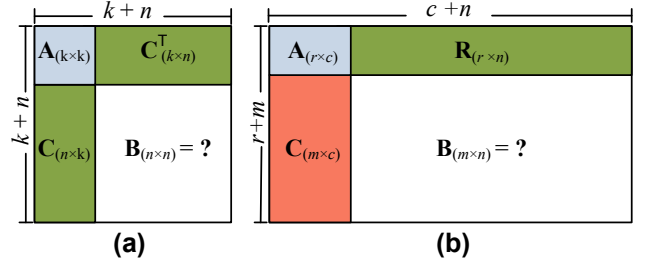

Figure 2: Matrix reconstruction from sparsely sampled columns and rows. (a) Symmetric case. (b) Asymmetric case.

Comparing Equation 2 and 5, the traditional Nyström method is a special case of the generalized Nyström method when $\mathbf{R} \equiv \mathbf{C}^T$.

To compute the Moore-Penrose pseudoinverse of $\mathbf{A}$, we apply the singular value decomposition (SVD) to $\mathbf{A}$ and obtain

$$\mathbf{A} = \mathbf{U_A}\mathbf{\Sigma_A}\mathbf{V_A}^T, \qquad (6)$$

where $\mathbf{U_A}$, $\mathbf{\Sigma_A}$ and $\mathbf{V_A}^T$ are of sizes $r \times d$, $d \times d$ and $d \times c$, respectively, and the columns of $\mathbf{U_A}$ and $\mathbf{V_A}$ are orthonormal: $\mathbf{U_A}^T\mathbf{U_A} = \mathbf{I}$ and $\mathbf{V_A}^T\mathbf{V_A} = \mathbf{I}$. The matrix $\mathbf{A}^+$ is then computed as

$$\mathbf{A}^+ = \mathbf{V_A}\mathbf{\Sigma_A}^{-1}\mathbf{U_A}^T. \qquad (7)$$

With $\mathbf{A}^+$, we can reconstruct the light transport matrix $\mathbf{T}$ using Equation 5.

3.2 Kernel Extension

Both the traditional and the generalized Nyström methods rely on the assumption that the ranks of $\mathbf{T}$ and $\mathbf{A}$ are identical, in order to achieve an exact reconstruction. In reality, this assumption may be violated, resulting in some reconstruction error. One possible way to make the Nyström method more effective is to apply a transformation to the entries in the matrix so that this low rank assumption is better satisfied, hence the reconstruction error can be expected to be minimized. As linear transforms do not change the rank of a matrix, nonlinear transforms are necessary.

In the machine learning literature, the "kernel trick" [Cristianini and Shawe-Taylor 2000] is a standard approach for enhancing the performance of algorithms based on nonlinear transformations of the input data. The kernel trick is to map vectors in the data space to a (usually) higher dimensional feature space. Then the same procedures of the original algorithm done in the data space are transferred to the feature space. The key to the success of the kernel trick is that the mapping function need not be explicitly specified. Rather, a kernel function is sufficient for computing the inner products in the feature space.

Inspired by the success of the kernel trick, we consider using a nonlinear function f to change the values of the entries in light transport matrix $\mathbf{T}$, such that the rank assumption can be better fulfilled, i.e., the rank of $f(\mathbf{T})$ is as close to that of $f(\mathbf{A})$ as possible. We name this nonlinear function f the light transport kernel. After reconstructing $f(\mathbf{T})$ using the generalized Nyström method, i.e.,

$$f(\mathbf{T}) \approx \mathbf{K} = \begin{bmatrix} f(\mathbf{A}) & f(\mathbf{R}) \\ f(\mathbf{C}) & f(\mathbf{C})(f(\mathbf{A}))^+ f(\mathbf{R}) \end{bmatrix}, \qquad (8)$$

the original $\mathbf{T}$ can be recovered by an inverse mapping with f^{-1}: $\mathbf{T} \approx f^{-1}(\mathbf{K})$.

To see that the above nonlinear mapping process is a kind of the kernel method, one may regard $\mathbf{T}$ as an inner product matrix:

$$\mathbf{T}_{ij} = \phi_i \cdot \psi_j, \qquad (9)$$

where $\{\phi_i\}$ and $\{\psi_j\}$ are two different point sets in a particular space[1]. This is slightly different from the traditional kernel method which requires that $\{\phi_i\}$ and $\{\psi_j\}$ are identical. We use different point sets because here $\mathbf{T}$ is asymmetric. The standard kernel method then uses an implicit mapping Γ to map the point sets to another space. Then the kernel matrix $\mathbf{K}$ in the mapped space, a.k.a. the inner product matrix of the mapped point sets, is:

$$\mathbf{K}_{ij} = \Gamma(\phi_i) \cdot \Gamma(\psi_j). \qquad (10)$$

To compute $\mathbf{K}$, one only has to prepare an explicit kernel function g such that $g(\phi_i, \psi_j) \equiv \Gamma(\phi_i) \cdot \Gamma(\psi_j)$, rather than explicitly specifying Γ, which is much more difficult. One of the most frequently used kernel functions is the polynomial kernel: $g(\phi_i, \psi_j) = (1 + \phi_i \cdot \psi_j)^p$. So one can choose the kernel g in such a form: $g(\phi_i, \psi_j) = f(\phi_i \cdot \psi_j)$. With this choice of the kernel function, we have that

$$\mathbf{K}_{ij} = f(\phi_i \cdot \psi_j) = f(\mathbf{T}_{ij}), \text{ i.e., } \mathbf{K} = f(\mathbf{T}). \qquad (11)$$

3.3 Estimating the Light Transport Kernel f

To make the kernel Nyström method work, one has to specify the light transport kernel f. As the space of all monotonic smooth functions is of infinite dimension, one has to assume its form in order to narrow down the search space. In this paper, we simply assume that f is a power function: $f(x) = x^\gamma$ as this family of functions has only one parameter and hence the optimal function is easy to find. Moreover, our experiments show that such a choice of the light transport kernel can indeed produce a greatly enhanced reconstruction performance. We leave the problem of whether there are even better light transport kernel functions to future work.

As argued before, to reduce the reconstruction error, one has to make the rank of $f(\mathbf{T})$ as close to that of $f(\mathbf{A})$ as possible. However, we do not have the full information on $\mathbf{T}$. So the real rank of $f(\mathbf{T})$ is actually unknown. To overcome this difficulty, we choose to minimize the rank of $f(\mathbf{A})$ instead. The underlying philosophy is that if the rank of $f(\mathbf{A})$ is *much less* than $\min(r,c)$, then it is very likely that the rank of $f(\mathbf{T})$ does not exceed $\min(r,c)$. So the generalized kernel Nyström method can be effective. This leads to the rank minimization problem [Fazel 2002], which is usually formalized as minimizing the nuclear norm of a matrix. We seek a light transport kernel f that minimizes the rank factor

$$e_r = \frac{\|f(\mathbf{A})\|_*}{\|f(\mathbf{A})\|_2}, \qquad (12)$$

where the nuclear norm is defined as $\|\mathbf{X}\|_* = \sum_i \sigma_i$, the spectral norm is defined as $\|\mathbf{X}\|_2 = \max_i\{\sigma_i\}$ and σ_i's are the singular values of the matrix $\mathbf{X}$. Note that here we normalize the nuclear norm with the largest singular value because we don't want to reduce the rank by mapping the entries to be close to zeros (which corresponds to a small $\|f(\mathbf{A})\|_2$). Rather, we want to reduce the rank of $f(\mathbf{A})$ by enhancing the linear coherence of the rows/columns of $f(\mathbf{A})$.

It is easy to imagine that mapping all the entries to a constant can trivially reduce the rank of $f(\mathbf{A})$ to 1 or even 0. However, such a trivial mapping causes a problem in inverting from $\mathbf{K}$ to $\mathbf{T}$ by using the inverse function f^{-1}. To reduce the reconstruction error, we have to further make sure that this inversion is numerically robust. The robustness of inversion can be measured by the slope of f^{-1}: the steeper f^{-1} is, the less robust the inversion is. So we have to

---
[1] As we will show, it is not necessary to specify what these two point sets are.


```
AdaptiveCapture( r', c', ε )
    finished = false
    capturing r' rows and c' columns
    r = r'; c = c'
    While(finished == false)
        estimating f from A
        estimating rank factor e_r of f(A)
        If( e_r/min(r,c) < ε )
            finished = true
        Else
            capturing r' rows and c' columns
            adding samples in sample set
            updating A
            r = r + r'; c = c + c'
```

Figure 3: The pseudo code of adaptive capturing scheme.

further minimize

$$e_s = \int_0^1 (f^{-1})'(x)p(x)\mathrm{d}x = \int_0^1 \frac{1}{f'(x)}p(x)\mathrm{d}x, \qquad (13)$$

where the identity $(f^{-1})'(x) = \frac{1}{f'(x)}$ is used and $p(x)$ is the distribution density of the values in $\mathbf{A}$, which is estimated from the histogram of entry values of $\mathbf{A}$. $p(x)$ is assumed to be identical to that in $\mathbf{T}$, due to the random sampling of $\mathbf{A}$ from $\mathbf{T}$. We weight the slope of f^{-1} by $p(x)$ in order to achieve the best overall robustness for all entry values in $\mathbf{T}$.

Combining the above two criteria, our algorithm selects the light transport kernel function f whose parameter γ minimizes the objective function

$$g(\gamma) = e_r \cdot e_s = \frac{\|f(\mathbf{A})\|_*}{\|f(\mathbf{A})\|_2} \int_0^1 \frac{1}{f'(x)}p(x)\mathrm{d}x. \qquad (14)$$

We use the golden section search [Press et al. 1992] to search for the optimal γ within $[0.001, 1000]$ in logarithmic space. As $\mathbf{A}$ is of a relatively small size ($r \times c$), this optimization takes only a few seconds.

In summary, our kernel Nyström method works as follows. Given the matrices of sampled rows and columns ($\mathbf{A}$, $\mathbf{C}$, and $\mathbf{R}$), we estimate the light transport kernel f from $\mathbf{A}$ and map matrices of sparse samples to their kernel version $f(\mathbf{A}), f(\mathbf{C})$, and $f(\mathbf{R})$ respectively. We then apply the Nyström method to reconstruct the kernel light transport matrix $\mathbf{K} = f(\mathbf{T})$ with the Moore-Penrose pseudoinverse of $f(\mathbf{A})$ using Equation 5. Finally, we obtain the light transport matrix $\mathbf{T}$ via the inverse kernel mapping $\mathbf{T} = f^{-1}(\mathbf{K})$.

4 Adaptive Light Transport Measurement

The kernel Nyström method requires a sparse set of rows and columns of the light transport matrix as input for reconstruction. However, without any knowledge about the light transport in the scene, it is difficult to determine the sample number in advance. To address this issue, we design an adaptive scheme for capturing sparse rows and columns from the scene, where the sample number is determined from the captured data. As shown in Figure 3, after a batch of rows and columns is sampled from the scene, we estimate the light transport kernel f from the matrix $\mathbf{A}$ of the current sample set and compute the rank factor e_r of $f(\mathbf{A})$ using Equation 12. For a sample set that has r rows and c columns, if $e_r/min(r,c)$ is smaller than a pre-defined threshold ε, the rank of $f(\mathbf{A})$ is much less than $\min(r,c)$. Based on the the same philosophy used in the kernel estimation, it is very likely the rank of $f(\mathbf{T})$ does not exceed $\min(r,c)$. Thus the sampled data are sufficient for reconstructing
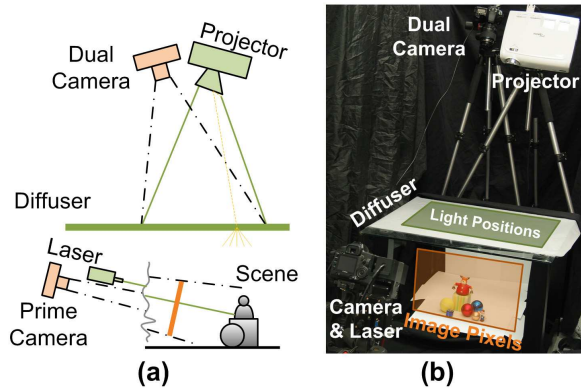
Figure 4: Device setup for capturing sparse columns and rows of the light transport matrix. (a) Illustration. (b) Photograph.
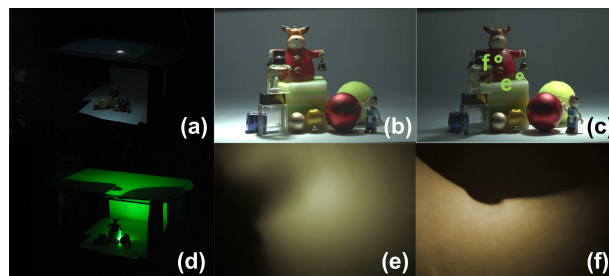

Figure 5: Sampling columns and rows of the light transport matrix. (a) Photograph of the scene in column sampling. (b)(c) Two column sampling images. (d) Photograph of the scene in row sampling. (e)(f) Two row sampling images. The corresponding pixels are marked in (c).

the $f(\mathbf{T})$ and the capturing is finished. Otherwise, we capture a new batch of rows and columns and repeat the above steps with the extended sample set.

A device setup is designed for acquiring a batch of row and column samples from the scene. As shown in Figure 4, we focus on the light transport from point light sources on a 2D plane to image pixels captured from a fixed view point. In our setup, a column of the light transport matrix is directly obtained from the image of the scene under the a virtual point light source, while a row of the matrix is measured via a dual setup by exploiting the reciprocity of light transport [Hawkins et al. 2005; Sen et al. 2005]. To obtain a batch of rows and columns, we first capture the columns and then acquire the rows of the light transport matrix. A stratified sampling scheme is applied for sampling rows of the matrix based on sparsely sampled column values.

Device Setup and Calibration As shown in Figure 4(b), our setup consists of a Optoma HD73 DLP projector, three laser emitters that can generate red, blue and green laser beams, two Canon 20D cameras and a flat vegetable parchment paper (i.e., diffuser) placed between the projector and the scene. The reflection and refraction of the paper are assumed to be diffuse and spatially uniform. When the projector shoots a light beam (5×5 pixels in our implementation) onto a point of the diffuser, partial light is refracted from the other side of the paper and illuminates the scene as a point light source. With this setup, we model the light transport from the point light sources that correspond to all projector pixels (i.e., all possible shifts of 5×5 pixels) to image pixels captured by the prime camera, which is represented by a very large transport matrix with nearly a million of columns and rows.

Before capturing, we calibrate the color and intensity of each point light source. In our implementation, we first calibrate the point light sources sampled on 40×40 regular grids by capturing the images of the diffuser plane lit with each of the 1600 point light sources from the scene side. We take images of the diffuser plane from the dual camera simultaneously to build the correspondence between the point light source positions and pixels of the dual camera image, in which the images captured by the dual camera are downsampled to 1024×768 to match the projector's resolution. We then calibrate the other point light sources within the regular grid by interpolating the calibration results of the neighbor point light sources on the grid.

Column Sampling After calibration, we place the scene objects under the diffuser plane and start to capture the columns of the light

transport matrix with the projector and the prime camera. Without any knowledge on the light transport in the scene, we acquire a set of columns from the prime camera with r point light sources uniformly distributed on the diffuser plane. For each sampled point light source l_j, eight images of different exposures are taken from the prime camera and then down-sampled and fused into HDR images as in [Debevec and Malik 1997]. The result is a set of column vectors $\mathbf{T}_{\cdot,j}$ of the matrix $\mathbf{T}$.

Row Sampling A dual setup that includes the three laser emitters and the dual camera is used for row sampling. The laser emitters are placed close to the prime camera, their orientations being controlled by a computer; the scene is lit by the laser beams. The reflectance of the scene is then projected onto the diffuser, refracted uniformly on the other side of the diffuser, and recorded by the dual camera neighboring the projector.

We utilize the sampled column values to guide the row sampling. Similar to [Hašan et al. 2007], we pack the r sparse elements in each row to a vector $\hat{\mathbf{r}}_i = (T_{i,j_1}, T_{i,j_2}, \ldots, T_{i,j_r})$, where $j_1, \ldots, j_r$ are indices of the sampled columns. We then employ k-means to cluster the $\{\hat{\mathbf{r}}_i\}$ into r clusters. For each cluster, the vector closest to the cluster center vector is selected as the sampling pixel.

For each row (i.e., pixel location) to be sampled, we adjust the direction of three laser beams so that they focus on the same surface point whose projection on the prime camera falls in the sampled pixel position. Since the prime camera and the laser emitters are relatively far away from the scene, the angle difference between the camera's pixel ray and the laser beams is small and ignored in capturing. A HDR image of the diffuse plane is then acquired from the dual camera. With the correspondence between the dual camera pixels and the point light source positions, we can easily obtain the sampled row values from the downsampled HDR images. We repeat this process until all sampled rows are measured. Finally, we scale the sampled columns so that the matrix element values captured by both rows and columns become the same. Figure 5 shows the sampled column and row images.

5   Results and Discussions

We implemented our kernel Nyström method on a PC with Intel Core™2 Duo 3.2GHz CPU and 4GB of memory. In our implementation, we capture a batch of 10 rows and 10 columns in each adaptive capture step and experimentally set the threshold as $\varepsilon = 0.05$. A typical acquisition session (including image acquisition, HDR reconstruction, and kernel estimation) takes about 135 minutes for capturing 150 rows and 150 columns from the scene. The image

resolution is 1752×1168, while the light sampling resolution is 1024×768. To reconstruct the light transport matrix for relighting, we store the sampled matrices ($\mathbf{C}, \mathbf{R}, \mathbf{A}$ and $\mathbf{A}^+$) in memory and reconstruct all rows of the matrix during rendering. With 150 rows and 150 columns, our kernel Nyström method takes less than 5 minutes to estimate the kernel and reconstruct all rows of the light transport matrix of the scene. For samples that do not fit in memory, we store one matrix ($\mathbf{C}$ in our implementation) on disk and the other matrices in memory. An out-of-core implementation, which is dominated by the disk I/O, then takes about 30 minutes for reconstruction.

We tested our method on a variety of scenes exhibiting different light transport effects. As illustrated in Figure 6, the caustics scene includes a set of transparent objects, where the light transport is dominated by caustics created by light refraction. In the shadow scene, the fine geometry of the dragon model results in complex occlusion and produces detailed shadows. The interreflection scene is designed to illustrate the strong inter-reflections between different objects and the color bleeding effects. In the general scene, objects with different material properties are placed together and present a mixture of different light transport effects, including reflection, refraction, shadows, and inter-reflections, subsurface scattering, and caustics. The light transport kernel coefficient and the sample number used in reconstruction for each scene are noted in the figure.

5.1 Method Validation

We validate the kernel Nyström method with the light transport matrices acquired from the general scene and the shadow scene, in which the light sources are sampled on regular 35×35 grids on the diffuser. We choose these two scenes for validation because the general scene includes a mixture of different kinds of light transport, while the shadow scene requires a large number of samples for reconstruction. We simulate the adaptive sampling scheme for each scene. In each iteration, we randomly pick 10 rows and 10 columns that are not in the sample set from the acquired full matrix and add them in the sparse sample set. After that, we estimate the light transport kernel f and compute the rank of the kernel mapped matrix $f(\mathbf{A})$ of the sparse samples. The rank of a matrix is determined by the number of eigenvalues of the matrix that can preserve 95.0% energy of the original matrix. We also reconstruct the light transport matrix from the sparse sample set via the kernel
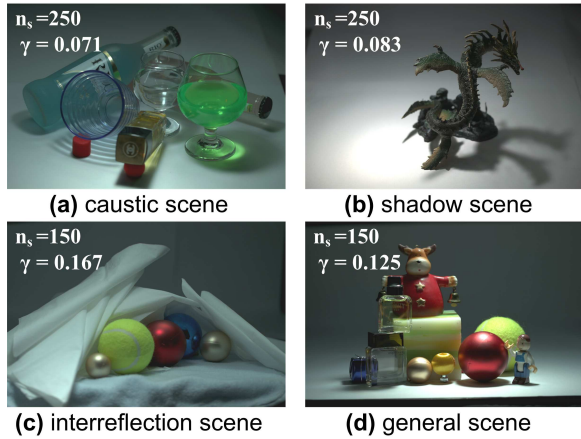

(a) caustic scene **(b)** shadow scene

(c) interreflection scene **(d)** general scene

Figure 6: Four test scenes from our experiments: (a) caustic scene (b) shadow scene (c) interreflection scene (d) general scene. The number of rows n_s and the light transport kernel coefficient γ for each scene are noted in the images.
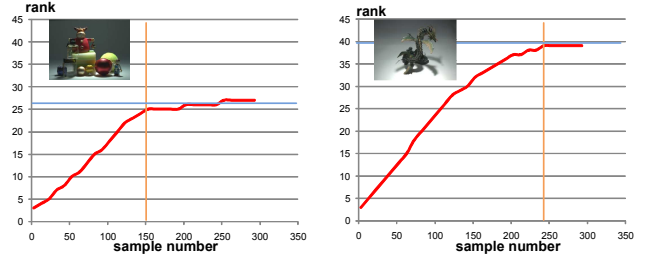

Figure 7: Ranks of the kernel mapped matrices $f(\mathbf{A})$ of the sparse sample sets with different numbers of samples. The orange line indicates the sample number determined by the adaptive capturing scheme, while the blue line indicates the rank of the kernel mapped matrix $f(\mathbf{T})$.

Nyström method. We repeat this process until the number of sparse samples achieves 300, which is more than the number of sparse samples n_s determined by the adaptive sampling scheme with the pre-defined ε.

Figure 7 shows plots of the ranks of kernel mapped matrices $f(\mathbf{A})$ of sparse sample sets with a different number of samples, where the number of sparse samples n_s determined by the adaptive sampling scheme is marked by an orange line. The blue line indicates the rank of the kernel mapped matrix $f(\mathbf{T})$, in which the kernel function is estimated from the sparse sample set with n_s samples. For both scenes, the ranks of $f(\mathbf{A})$ of sparse samples grow as the number of samples increases and are close to the $f(\mathbf{T})$ as the number of samples is close to n_s, which leads to a good reconstruction.

Figure 8 shows plots of the relative errors of the light transport matrices reconstructed from a different number of samples, where the relative reconstruction error is computed as

$$\varepsilon = \sqrt{\frac{\sum_{i,j} |B_{i,j} - \tilde{B}_{i,j}|^2}{\sum_{i,j} |B_{i,j}|^2}}. \tag{15}$$

Here $B_{i,j}$ is the ground truth submatrix element that is not in the sparse sample set and $\tilde{B}_{i,j}$ is the same submatrix element reconstructed using the kernel Nyström method. To further explore the impact of different sample sets to the reconstruction quality, we execute the above experiment 16 times, each time with different rows and columns randomly selected from the matrix. As shown in Figure 8, the relative error decreases quickly as the number of sparse samples increases. With the same number of sparse samples as we used in our capturing, the relative error of the reconstructed light transport matrix is below 5%. Also note that the variance of the error under the same number of samples reduces quickly with the
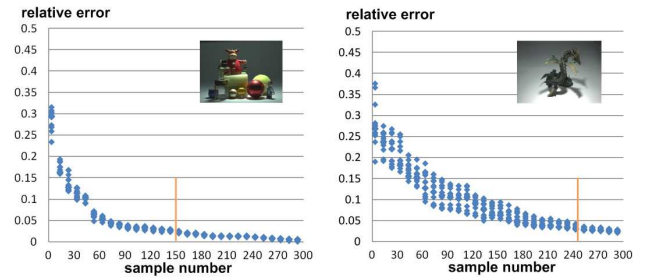

Figure 8: Relative reconstruction errors of the light transport matrix reconstructed from sparse sample sets with different number of samples. The orange line indicates the sample number determined by the adaptive capturing scheme.
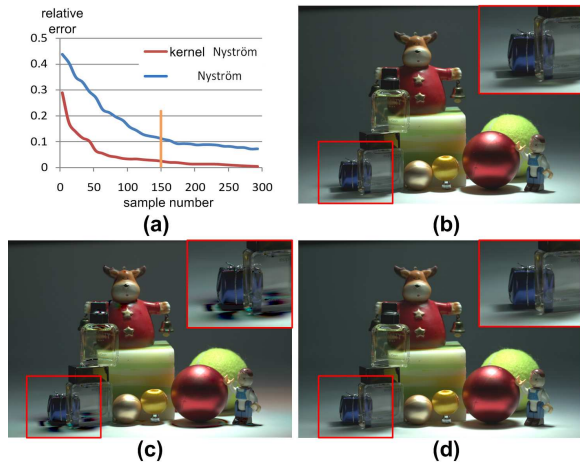
Figure 9: Reconstruction results with the kernel Nyström method and with the Nyström method. The relative reconstruction errors with respect to the different numbers of samples are compared in (a). (b) The ground truth image. (c) Image rendered with matrix reconstructed from 150 samples with the Nyström method. (d) Image rendered with matrix reconstructed from 150 samples with the kernel Nyström method.

increasing number of samples, which guarantees the quality of the light transport matrix reconstructed from the random sampled rows and columns.

Figure 9 compares the performance of the Nyström method and the kernel Nyström method for the general scene. For each sparse sample set, we reconstruct the light transport matrices using both the Nyström method and the kernel Nyström method. Their relative errors are shown in Figure 9 (a). Without the kernel extension, the Nyström method requires roughly five to six times the number of samples to achieve the same reconstruction quality as the kernel Nyström method does. Figure 9 (b) shows a visual comparison, where the same number of samples are used in both methods for reconstruction. The artifacts in the results generated with the Nyström method are clearly visible.

Figure 10 compares the performance of the kernel Nyström method with different light transport kernels (i.e., different γ values) for the shadow scene. For each γ value, we reconstruct the light transport matrix from the same set of sparse samples (250 samples) with the kernel Nyström method. Figure 10(a) shows plots of the objective function $g(\gamma)$ and the relative error of the reconstructed light transport matrices as a function of γ. Note that the relative reconstruction error is minimal when the optimal γ value derived by our approach is used for reconstruction. Also, the light transport matrix is well reconstructed by the kernel Nyström method using the optimal γ value. However, the light transport matrices reconstructed with the kernel Nyström method using other γ values exhibit larger relative errors and visible artifacts. As shown in Figure 10(e)(g), the large γ value leads to visible artifacts concentrated in the region with high frequency features, while the small γ value results in the noise like artifacts scattered across the entire result image.

5.2 Relighting Results

Figure 12 compares images rendered using the reconstructed light transport matrix to ground truth images showing illumination under a single point light that was not considered during the reconstruction process. The rendering results generated from the reconstructed light transport matrix are visually indistinguishable from the ground truth images.

We used the reconstructed light transport matrix for image-based relighting. Figure 13 exhibits several results rendered with different lighting conditions. Note that complex light transport effects are faithfully reproduced, providing realistic results. See the companion video for relighting results under dynamic lighting.

5.3 Subsurface Scattering

We also applied the kernel Nyström method for modeling the light transport due to subsurface scattering. The light transport matrix in this case is an $n \times n$ symmetric matrix, each element of which describes the diffuse BSSRDF R_d between the two surface points x_i and x_j as $\mathbf{T}_{ij} = R_d(x_i, x_j)$ [Goesele et al. 2004]. Since this matrix is symmetrical, we only need to sample the columns for reconstruction.

In our implementation, we use a device setup similar to the one used in [Goesele et al. 2004] for capturing. The three laser emitters are close to each other and shoot three color beams to points over the object surface. We control the laser beams to make sure that they focus on the same surface point. HDR images of the object surface are captured from a fixed viewpoint. We repeat this process by shooting the laser beams to a random set of surface points. With the calibrated camera position and known geometry of the object, we map the image pixels onto the object surface. Figure 11(d) shows an image of the marble used in our experiments.

With 50 sampled images, we reconstruct the light transport matrix of subsurface scattering using the kernel Nyström method. The coefficient of the light transport kernel used in reconstruction is $\gamma = 0.0625$. The resolution of surface points in the light transport matrix is the same as the image resolution. Figure 11 compares the rendering result of the reconstructed light transport matrix, the ground truth image captured from the object under the same lighting condition, and the result rendered by directly interpolating the nearby sample images. Note that the detailed spatial patterns and anisotropic subsurface scattering in the real material are well preserved with the reconstructed light transport matrix, while the results generated by interpolation clearly exhibit artifacts. Also note that to capture the light transport effects with a similar resolution, brute force methods [Goesele et al. 2004; Peers et al. 2006] need dense light sampling, which is prohibitively expensive and time consuming. We show relighting results of the reconstructed light transport matrix in the companion video.

5.4 Discussions

The light transport matrix of a scene exhibits both data coherence and sparsity. While the sparsity based approaches exploit the data sparsity for capturing the light transport matrix, the kernel Nyström method exploits the coherence in both rows and columns of the light transport matrix for the same task. For scenes with low frequency light transport effects, such as inter-reflections, subsurface scattering and glossy shading, their light transport matrices always exhibit strong coherence in both rows and columns, which can be efficiently reconstructed using the kernel Nyström method with dozens of images. In a scene that includes more sharp light transport variations, such as occlusions, specular reflections and refractions, the coherence in the column vectors may be reduced. By exploiting the coherence in the row vectors, the kernel Nyström method can still achieve good results with hundreds of image samples. In both cases, the kernel Nyström methods require much less images than the sparsity based methods. For comparison, we project the light transport matrices of the general scene and

the shadow scene onto the Haar wavelet basis. In each case, for reconstructing the entire matrix with the similar 5% relative error, we need to retain thousands of wavelet coefficients, which means that more than thousands of images are required for reconstructing the matrix with the sparsity based methods.

In the extreme cases where the light transport matrix exhibits little coherence in both rows and columns, the kernel Nyström method would need all rows and columns of the matrix for good reconstruction. In this case, our approach has no advantage over the brute force methods and the sparsity based methods. An example is a scene of a mirror ball, where the light transport matrix exhibits good sparsity but little coherence. In this scenario, the sparsity based approaches may provide a good solution.

6 Conclusion

We have presented the kernel Nyström method for reconstructing the light transport matrix from a relatively small number of acquired
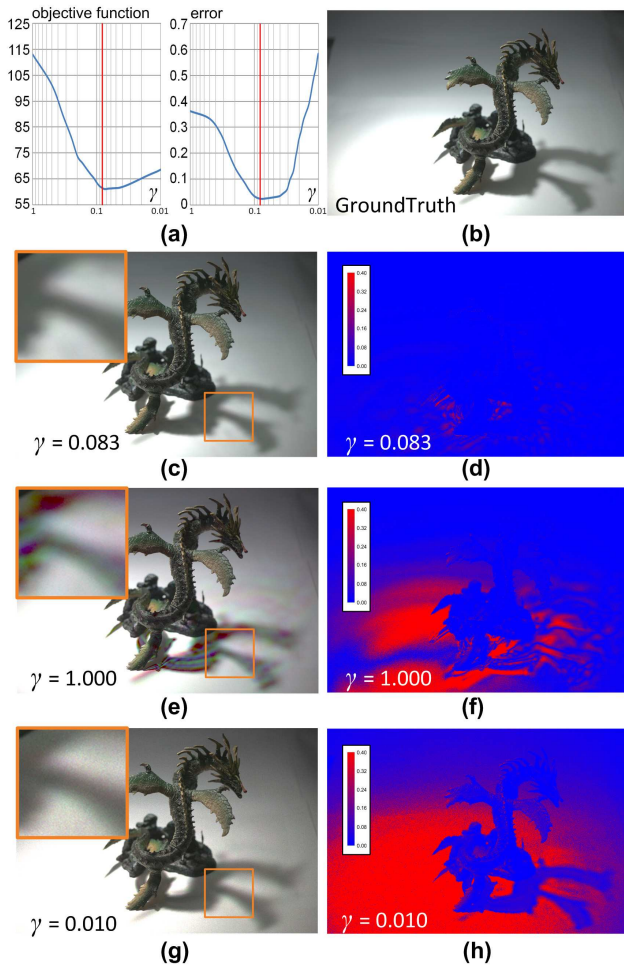

Figure 10: Reconstruction results of the kernel Nyström method with different light transport kernels. (a) The objective function values and the relative reconstruction errors with respect to different γ values. (b) The ground truth image. (c)(e)(g) Images rendered with the matrices reconstructed with the kernel Nyström method using different γ values. The γ value used in (c) is the optimal one. (d)(f)(h) The difference images between the rendering results of the reconstructed matrices (c)(e)(g) and the ground truth image (b).
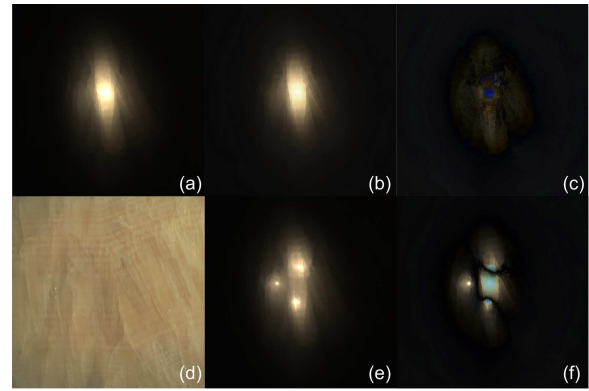

Figure 11: Modeling subsurface light transport with the kernel Nyström method. (a) The image of marble lit by laser beam at one point. (b) Image rendered with reconstructed light transport matrix with light on the same point. (c)The difference between (a) and (b), intensity of which is scaled by 10. (d) Photograph of the flat marble sample. (e) Image rendered by interpolating the nearby sample images. (f) The difference between (a) and (e), intensity of which is scaled by 10.

images. While existing techniques typically require tens of thousands of images for accurate reconstruction of a light transport matrix, the kernel Nyström method can achieve a good reconstruction with a few hundred images and produces high quality relighting results. The kernel Nyström method is able to capture complex lighting effects and occlusions which are particularly challenging for many existing techniques, especially various sparsity-based methods. The effectiveness of the kernel Nyström method comes from its ability to exploit both the linear and nonlinear coherence in the light transport matrix from a relatively small number of rows and columns of the matrix.

In future work, we are interested in investigating better kernel estimation methods and designing new kernel functions for light transport matrix reconstruction. We also plan to apply the kernel Nyström method to model surface reflectance. Finally, we want to explore ways to extend the kernel Nyström method to handle high-dimensional tensor data.

7 Acknowledgements

References

AN, X., AND PELLACINI, F. 2008. Appprop: All-pairs appearance-space edit propagation. *ACM Transactions on Graphics 27*, 3 (Aug.), 40:1–40:9.

CHUANG, Y.-Y., ZONGKER, D. E., HINDORFF, J., CURLESS, B., SALESIN, D. H., AND SZELISKI, R. 2000. Environment matting extensions: Towards higher accuracy and real-time capture. In *Proceedings of ACM SIGGRAPH 2000*, Computer Graphics Proceedings, Annual Conference Series, 121–130.

CRISTIANINI, N., AND SHAWE-TAYLOR, J. 2000. An introduction to support vector machines and other kernel-based learning methods. *Cambridge University Press.*

DEBEVEC, P. E., AND MALIK, J. 1997. Recovering high dynamic range radiance maps from photographs. In *Proceedings of SIGGRAPH 97*, Computer Graphics Proceedings, Annual Conference Series, 369–378.

DEBEVEC, P., HAWKINS, T., TCHOU, C., DUIKER, H.-P., SAROKIN, W., AND SAGAR, M. 2000. Acquiring the reflectance field of a human face. In *Proceedings of ACM SIGGRAPH 2000*, Computer Graphics Proceedings, Annual Conference Series, 145–156.

FAZEL, M. 2002. Matrix rank minimization with applications. *PhD thesis, Stanford University*.

FUCHS, M., BLANZ, V., LENSCH, H. P. A., AND SEIDEL, H.-P. 2007. Adaptive sampling of reflectance fields. *ACM Transactions on Graphics 26*, 2 (June), 10:1–10:18.

GARG, G., TALVALA, E.-V., LEVOY, M., AND LENSCH, H. P. A. 2006. Symmetric photography: Exploiting data-sparseness in reflectance fields. In *Rendering Techniques 2006: 17th Eurographics Workshop on Rendering*, 251–262.

GOESELE, M., LENSCH, H. P. A., LANG, J., FUCHS, C., AND SEIDEL, H.-P. 2004. Disco: acquisition of translucent objects. *ACM Transactions on Graphics 23*, 3 (Aug.), 835–844.

GOREINOV, S., TYRTYSHNIKOV, E. E., AND ZAMARASHKIN, N. L. 1997. A theory of pseudo-skeleton approximations. *Linear Algebra and Applications 261*, 1–21.

HAŠAN, M., PELLACINI, F., AND BALA, K. 2007. Matrix row-column sampling for the many-light problem. *ACM Transactions on Graphics 26*, 3 (July), 26:1–26:10.

HAŠAN, M., VELAZQUEZ-ARMENDARIZ, E., PELLACINI, F., AND BALA, K. 2008. Tensor clustering for rendering many-light animations. *Computer Graphics Forum (Proc. Eurographics Rendering 2008) 27*, 4, 1105–1114.

HAWKINS, T., EINARSSON, P., AND DEBEVEC, P. 2005. A dual light stage. In *Rendering Techniques 2005: 16th Eurographics Workshop on Rendering*, 91–98.

LEVOY, M., AND HANRAHAN, P. M. 1996. Light field rendering. In *Proceedings of SIGGRAPH 96*, Computer Graphics Proceedings, Annual Conference Series, 31–42.

LIN, Z., WONG, T.-T., AND SHUM, H.-Y. Relighting with the reflected irradiance field: Representation, sampling and reconstruction. *International Journal of Computer Vision 49*, 2.

MAHAJAN, D., SHLIZERMAN, I. K., RAMAMOORTHI, R., AND BELHUMEUR, P. 2007. A theory of locally low dimensional light transport. *ACM Transactions on Graphics 26*, 3 (July), 62:1–62:10.

MASSELUS, V., PEERS, P., DUTRÉ, P., AND WILLEMS, Y. D. 2003. Relighting with 4d incident light fields. *ACM Transactions on Graphics 22*, 3 (July), 613–620.

MASSELUS, V., PEERS, P., DUTR0108, P., AND WILLEMS, Y. D. 2004. Smooth reconstruction and compact representation of reflectance functions for image-based relighting. In *Rendering Techniques 2004: 15th Eurographics Workshop on Rendering*, 287–298.

MATUSIK, W., PFISTER, H., NGAN, A., BEARDSLEY, P., ZIEGLER, R., AND MCMILLAN, L. 2002. Image-based 3D photography using opacity hulls. *ACM Transactions on Graphics 21*, 3 (July), 427–437.

MATUSIK, W., LOPER, M., AND PFISTER, H. 2004. Progressively-refined reflectance functions from natural illumination. In *Rendering Techniques 2004: 15th Eurographics Workshop on Rendering*, 299–308.

NG, R., RAMAMOORTHI, R., AND HANRAHAN, P. 2003. All-frequency shadows using non-linear wavelet lighting approximation. *ACM Transactions on Graphics 22*, 3 (July), 376–381.

PEERS, P., AND DUTRÉ, P. 2005. Inferring reflectance functions from wavelet noise. In *Rendering Techniques 2005: 16th Eurographics Workshop on Rendering*, 173–182.

PEERS, P., VOM BERGE, K., MATUSIK, W., RAMAMOORTHI, R., LAWRENCE, J., RUSINKIEWICZ, S., AND DUTRÉ, P. 2006. A compact factored representation of heterogeneous subsurface scattering. *ACM Transactions on Graphics 25*, 3 (July), 746–753.

PEERS, P., MAHAJAN, D. K., LAMOND, B., GHOSH, A., MATUSIK, W., RAMAMOORTHI, R., AND DEBEVEC, P. 2009. Compressive light transport sensing. *ACM Transactions on Graphics 28*, 1 (Jan.), 3:1–3:18.

PLATT, J. C. 2005. Fastmap, metricmap, and landmark mds are all nyström algorithms. In *10th International Workshop on Artificial Intelligence and Statistics*, 261–268.

PRESS, W. H., ET AL. 1992. Numerical recipes in C (second edition). *Cambridge University Press*.

SEN, P., AND DARABI, S. 2009. Compressive Dual Photography. *Computer Graphics Forum 28*, 2, 609 – 618.

SEN, P., CHEN, B., GARG, G., MARSCHNER, S. R., HOROWITZ, M., LEVOY, M., AND LENSCH, H. P. A. 2005. Dual photography. *ACM Transactions on Graphics 24*, 3, 745–755.

WENGER, A., GARDNER, A., TCHOU, C., UNGER, J., HAWKINS, T., AND DEBEVEC, P. 2005. Performance relighting and reflectance transformation with time-multiplexed illumination. *ACM Transactions on Graphics 24*, 3 (Aug.), 756–764.

WILLIAMS, C., AND SEEGER, M. 2000. Using the nyström method to speed up kernel machines. *Advances in Neural Information Processing Systems 13*, 682–688.

ZONGKER, D. E., WERNER, D. M., CURLESS, B., AND SALESIN, D. H. 1999. Environment matting and compositing. In *Proceedings of SIGGRAPH 99*, Computer Graphics Proceedings, Annual Conference Series, 205–214.

Figure 12: Visual comparison of relighting results. Images in the first row are captured from each scene as illuminated by a point light source that is not used when acquiring images for the light transport matrix reconstruction. Images in the second row are rendered from the reconstructed light transport matrices of the scenes under the same lighting condition. The third row shows the difference images, with their intensities scaled up by a factor of 10.
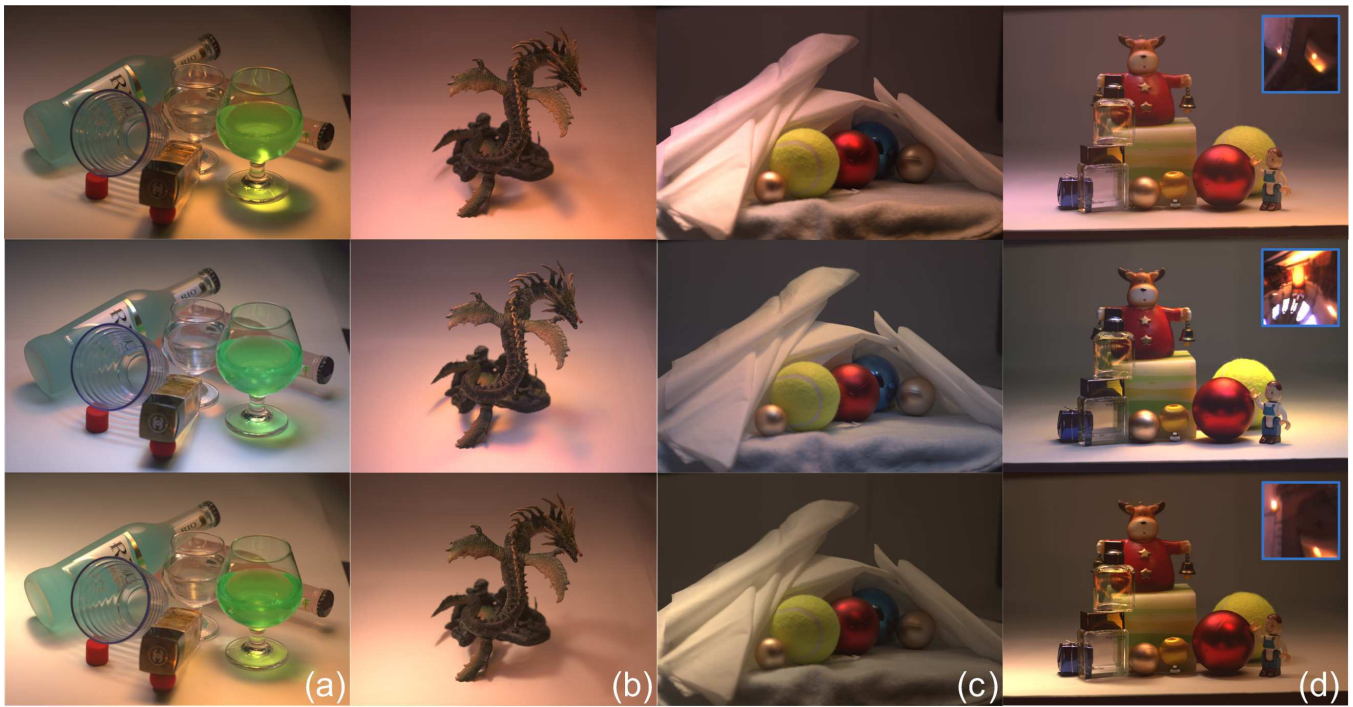

Figure 13: Relighting results under complex lighting. (a) The caustic scene. (b) The shadow scene. (c) The interreflection scene. (d) The general scene. The inset images show the three illumination patterns used for rendering.