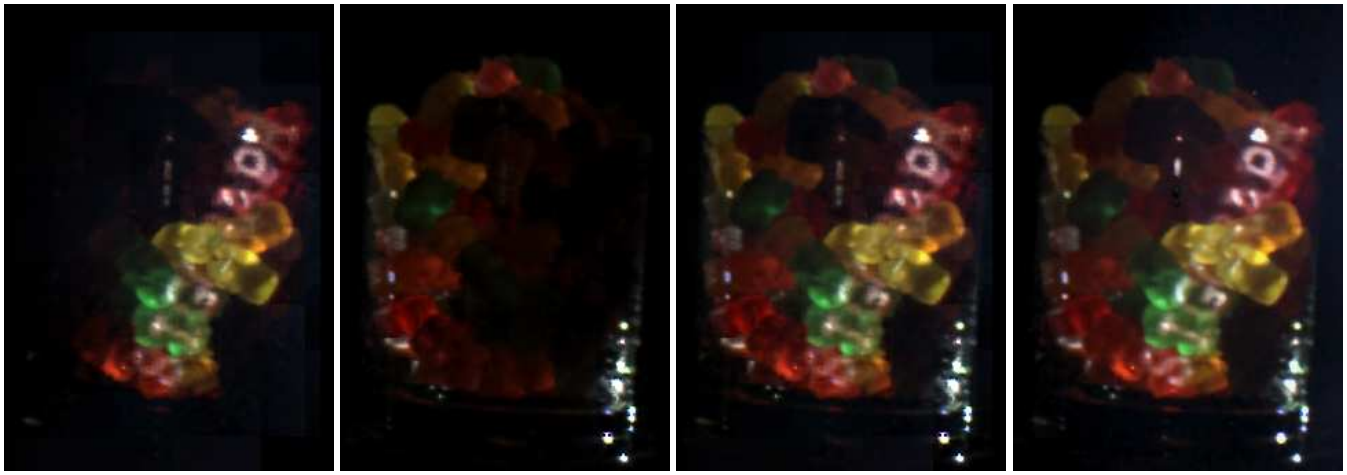
Symmetric Photography : Exploiting Data-sparseness in Reflectance Fields

Gaurav Garg* Eino-Ville Talvala* Marc Levoy* Hendrik P. A. Lensch*

Stanford University



(a) (b) (c) (d)

Figure 1: The reflectance field of a glass full of gummy bears is captured using two coaxial projector/camera pairs placed 120° apart. (a) is the result of synthetically relighting the scene from the front projector, which is coaxial with the presented view, with a high resolution "SIGGRAPH" matte. Note that due to their sub-surface scattering property, even a single beam of light that falls on a gummy bear illuminates it completely. In (b) we simulate homogeneous backlighting from the second projector. (c) combines (a) and (b). For validation, a ground-truth image (d) was captured by loading the same projector patterns into the real projectors. Our approach is able to faithfully capture and reconstruct the complex light transport in this scene.

Abstract

We present a novel technique called symmetric photography to capture real world reflectance fields. The technique models the 8D reflectance field as a transport matrix between the 4D incident light field and the 4D exitant light field. It is a challenging task to acquire a full transport matrix due to its sheer size. We observe that the transport matrix is data-sparse and symmetric. This symmetry enables us to measure the light transport from two sides simultaneously, from the illumination directions and the view directions. This provides the basis for our acquisition algorithm, which exploits the data-sparseness in the matrix, resulting in tremendous speed-up during acquisition. We introduce the use of hierarchical tensors as the underlying data structure to represent data-sparseness by local rank-1 factorizations of the transport matrix. Besides providing an efficient representation for acquisition and storage, it enables fast rendering of images from the captured transport matrix. Our prototype acquisition system consists of an array of coaxial projectors and mirrors. We demonstrate the effectiveness of our system with scenes rendered from reflectance fields that were captured by our system. In these renderings we can change the viewpoint freely as well as relight using arbitrary incident light fields.

CR Categories: I.3.3 [Computer Graphics]: Picture Image Generation—Digitizing and scanning; I.3.6 [Computer Graphics]: Methodology and Techniques—Graphics data structures and data types; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Color, shading, shadowing, and texture

Keywords: light fields, reflectance fields, symmetric photography, transport matrix, hierarchical tensors, image-based rendering, image-based relighting

*e-mail: {ggaurav|talvala|levoy|lensch@graphics.stanford.edu}

1 Introduction

The most complete image-based description of a scene for computer graphics applications is its 8D reflectance field [Debevec et al. 2000]. When available, a reflectance field can be used to render images of the scene from any viewpoint under arbitrary lighting. The resulting images capture all global illumination effects such as diffuse inter-reflections, shadows, caustics and sub-surface scattering without the need for an explicit physical simulation.

The 8D reflectance field can also be described as a matrix that describes the transfer of energy between a bundle of incoming rays (the illumination) and a bundle of outgoing rays (the view) in a scene, each of which are 4D. Viewing and illuminating with light fields of 10 view positions and a resolution of 100×100 in each image require a transport matrix containing about 10^{10} entries. If constructed by measuring the transport coefficients between every pair of incoming and outgoing light rays, it could take days to compute, making it intractable.

This paper introduces *symmetric photography* - a technique for acquiring 8D reflectance fields efficiently. It relies on two key observations. First, the reflectance field is data-sparse in spite of its high dimensionality. Second, the transport matrix is symmetric, due to Helmholtz reciprocity. This symmetry enables simultaneous measurements for both dimensions of the transport matrix. We

use these measurements to develop a hierarchical acquisition algorithm that can exploit the data-sparseness. To facilitate this, we have built a symmetrical capture setup, which consists of a coaxial array of projectors and cameras. In addition, we introduce the use of hierarchical tensors as an underlying data structure to represent reflectance fields. The hierarchical tensor representation turns out to be a natural data structure for the acquisition algorithm and provides compact factorized representation for storing a data-sparse transport matrix. Further, hierarchical tensors provide fast computation during rendering.

The rest of the paper is organized as follows. Section 2 presents a brief survey of prior research in the area. We explain the concept of data-sparseness in a matrix in Section 3 and contrast it with sparseness and smoothness. This is followed by an explanation of the data-sparseness and symmetry of the transport matrix in Section 4. In Section 5 we describe our complete acquisition scheme. Experimental results follow in Section 6. We conclude and describe some ideas for future work in Section 7.

2 Related Work

The measurement of reflectance fields is an active area of research in computer graphics. However, most of this research has focused on capturing various lower dimensional slices of the reflectance field. For instance, if the illumination is fixed and the viewer allowed to move, the appearance of the scene as a function of outgoing ray position and direction is a 4D slice of the reflectance field. This is also called the (exitant) light field [Levoy and Hanrahan 1996] or the lumigraph [Gortler et al. 1996]. By extracting appropriate 2D slices of the light field, one can virtually fly around a scene but the illumination cannot be changed. If the viewpoint is fixed and the illumination is provided by a set of point light sources, one obtains another 4D slice of 8D reflectance field. Various researchers [Debevec et al. 2000; Malzbender et al. 2001; Hawkins et al. 2005] have acquired such data sets where the captured images can be combined using a weighted sum to obtain re-lit images from a fixed viewpoint only. Since point light sources radiate light in all directions, it is impossible to cast sharp shadows onto the scene using a spatially varying illumination pattern with this technique.

If the illumination is provided by an array of video projectors and the scene is captured as illuminated by each pixel of each projector, but still as seen from a single viewpoint, then one obtains a 6D slice of 8D reflectance field. Masselus et al. [2003] capture such data sets using a single moving projector. More recently, Sen et al. [2005] have exploited Helmholtz reciprocity to improve on both the resolution and capture times of these data sets in their work on dual photography. With such a data set it is possible to relight the scene with arbitrary 4D incident light fields, but the viewpoint cannot be changed. Goesele et al. [2004] use a scanning laser, a turntable and a moving camera to capture a reflectance field for the case of translucent objects under a diffuse sub-surface scattering assumption. Although one can view the object from any position and relight it with arbitrary light fields, the captured data set is still essentially 4D because of their assumption. All these works capture some lower dimensional subset of the 8D reflectance field. A full 8D reflectance field has never been acquired before.

Hierarchical data structures have been previously used for representing reflectance fields. These representations provide greater efficiency both in terms of storage and capture time. A typical setup for capturing reflectance fields consists of a scene under controlled illumination, as imaged by one or more cameras. Peers and Dutré [2003] illuminate the scene with wavelet patterns in order to capture environment mattes (a 4D slice of the reflectance field). A feedback loop determines the next pattern to use based on knowledge of previously recorded photographs. The stopping criteria is based on the error of the current approximation. Although their scheme adapts to the scene content, it does not try to parallelize the capture process. Matusik et al. [2004] use a kd-tree based subdivision structure to represent environment mattes. They express environment matte extraction as an optimization problem. Their algorithm progressively refines the approximation of the environment matte with an increasing number of training images taken under various illumination conditions. However, the choice of their patterns is independent of the scene content. Sen et al. [2005] also use a hierarchical scheme to capture 6D slices of the reflectance field. Their illumination patterns adapt to the scene content, and the acquisition system tries to parallelize the capture process depending on the sparseness of the transport matrix. Their technique reduces to scanning if the transport matrix is dense, e.g. in case of scenes with various diffuse bounces. We observe that the light transport in these cases is data-sparse. Using the symmetry of the transport matrix to exploit data-sparseness, we are able to capture full 8D reflectance fields in reasonable time.

3 Sparseness, Smoothness and Data-sparseness

To efficiently store large matrices, sparseness and smoothness are two ideas that are typically exploited. But the notion of data-sparseness is more powerful than these. A sparse matrix has a small number of non-zero elements in it and hence can be represented compactly. A data-sparse matrix on the other hand can have a lot of non-zero elements but the actual information content in the matrix is small enough that it can still be expressed compactly. A simple example will help understand this concept better. Consider taking the cross product of two vectors. Although the resulting matrix (which is rank-1 by construction) could be non-sparse, we only need two vectors to represent the contents of the entire matrix. Such matrices are data-sparse. More generally, any matrix in which a significant number of sub-blocks can have a low-rank representation is data-sparse. Note that a low-rank sub-block of a matrix need not be smooth and may contain high frequencies. A frequency or wavelet-based technique would be ineffective in compressing this block. Therefore, the concept of data-sparseness is more general and powerful than sparseness or smoothness of the matrix.

Sparseness in light transport has been exploited to accelerate acquisition times in the work of Sen at al. [Sen et al. 2005]. Ramamoorthi and Hanrahan [2001] analyze the smoothness in BRDFs and use it for efficient rendering and compression. A complete frequency space analysis of light transport has been presented by Durand et al. [Durand et al. 2005]. The idea of exploiting data-sparseness for factorizing high dimensional datasets into low-rank approximations has been investigated in the context of BRDFs [Kautz and McCool 1999; McCool et al. 2001; Latta and Kolb 2002] and also for light fields and reflectance fields [Vasilescu and Terzopoulos 2004; Wang et al. 2005]. We tie in the factorization with a hierarchical subdivision scheme (see section 5). This hierarchical factorization approach allows us to exploit the data-sparseness locally.

4 Properties of Light Transport

Our acquisition scheme is based on two key observations about light transport in a scene. We observe that the transport matrix is typically data-sparse and is always symmetric. These observations

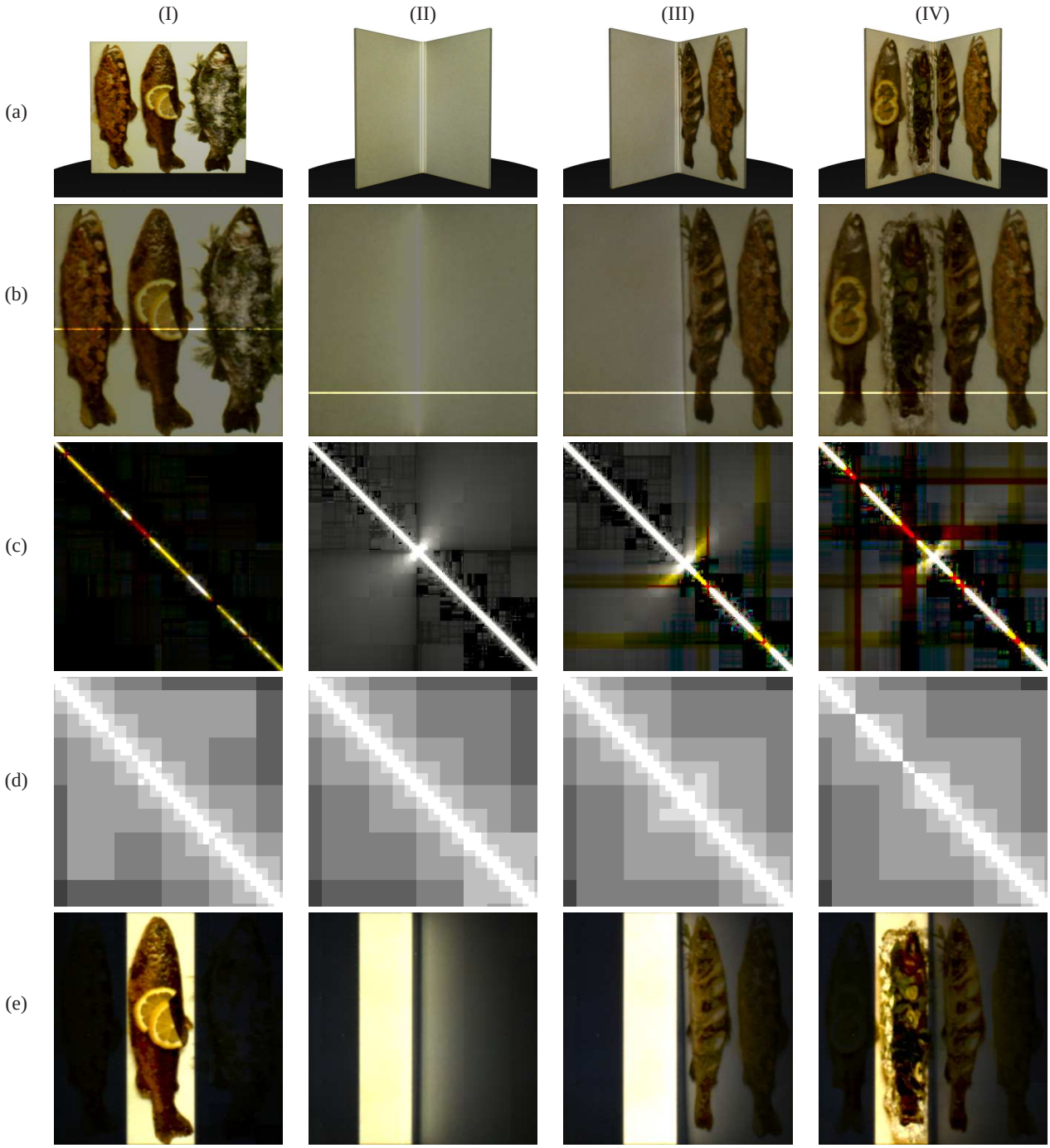


Figure 2: Understanding the transport matrix. To explain the intrinsic structure of reflectance fields we capture the transport matrix for 4 real scenes shown in row (a) with a coaxial projector/camera pair. The scenes in different columns are: (I) a diffuse textured plane, (II) two diffuse white planes facing each other at an angle, (III) a diffuse white plane facing a diffuse textured plane at an angle, and (IV) two diffuse textured planes facing each other at an angle. Row (b) shows the images rendered from the captured transport matrices under floodlit illumination. A 2D slice of the transport matrix for each configuration is shown in row (c). This slice describes the light transport between every pair of rays that hit the brightened line in row (b). Note that the transport matrix is symmetric in all 4 cases. Since (I) is a flat diffuse plane, there are no secondary bounces and the matrix is diagonal. In (II), (III) and (IV) the diagonal corresponds to the first bounce light and is therefore much brighter than the rest of the matrix. The top-right and bottom-left sub-blocks describe the diffuse-diffuse light transport from pixels on one plane to the other. Note that this is smoothly varying for (II). In case of (III) and (IV), the textured surface introduces high frequencies but the sub-blocks describing the secondary bounce light transport are still data-sparse and can be represented using rank-1 factors. The top-left and bottom-right sub-blocks correspond to the energy from 3rd-order bounces in our scenes. Because this energy is around the noise threshold in our measurements we get noisy readings for these sub-blocks. Row (d) is a visualization of the level in the hierarchy when a block is classified as rank-1. White blocks are leaf nodes, while darker shades of gray progressively represent lower levels in the hierarchy. Finally, row (e) shows the result of relighting the transport matrix with a vertical bar. Note the diffuse glow on the right plane in (II), (III) and (IV). Since the left plane is textured in (IV) the diffuse glow is dimmer than in (III).

form the basis of our acquisition algorithm and guide the design of our acquisition setup.

4.1 Data-Sparseness

The flow of light in a scene can be described by a light field. Light fields, which were introduced in the seminal work of Gershun [1936], are used to describe the radiance at each point $\mathbf{x}$ and in each direction ω in a scene. This is a 5D function which we denote by $\widetilde{L}(\mathbf{x},\omega)$. Under this paradigm, the appearance of a scene can be completely described by an outgoing radiance distribution function $\widetilde{L}_{out}(\mathbf{x},\omega)$. Similarly, the illumination incident on the scene can be described by an incoming radiance distribution function $\widetilde{L}_{in}(\mathbf{x},\omega)$. The relationship between $\widetilde{L}_{in}(\mathbf{x},\omega)$ and $\widetilde{L}_{out}(\mathbf{x},\omega)$ can be expressed by an integral equation, the well known rendering equation [Kajiya 1986]:

$$\widetilde{L}_{out}(\mathbf{x},\omega) = \widetilde{L}_{in}(\mathbf{x},\omega) + \int_V \int_\Omega K(\mathbf{x},\omega;\mathbf{x}',\omega')\widetilde{L}_{out}(\mathbf{x}',\omega')d\mathbf{x}'d\omega' \tag{1}$$

The function $K(\mathbf{x},\omega;\mathbf{x}',\omega')$ defines the proportion of irradiance from $(\mathbf{x}',\omega')$ that gets transported as radiance to $(\mathbf{x},\omega)$. It is a function of the BSSRDF, the relative visibility of $(\mathbf{x}',\omega')$ and $(\mathbf{x},\omega)$ and foreshortening and light attenuation effects. Eq. (1) can be expressed in discrete form as:

$$\widetilde{\mathbf{L}}_{out}[i] = \widetilde{\mathbf{L}}_{in}[i] + \sum_j \mathbf{K}[i,j]\widetilde{\mathbf{L}}_{out}[j] \tag{2}$$

where $\widetilde{\mathbf{L}}_{out}$ and $\widetilde{\mathbf{L}}_{in}$ are discrete representations of outgoing and incoming light fields respectively. We can rewrite eq. (2) as a matrix equation:

$$\widetilde{\mathbf{L}}_{out} = \widetilde{\mathbf{L}}_{in} + \mathbf{K}\widetilde{\mathbf{L}}_{out} \tag{3}$$

Eq. (3) can be directly solved [Kajiya 1986] to yield:

$$\widetilde{\mathbf{L}}_{out} = (\mathbf{I}-\mathbf{K})^{-1}\widetilde{\mathbf{L}}_{in} \tag{4}$$

The matrix $\widetilde{\mathbf{T}} = (\mathbf{I}-\mathbf{K})^{-1}$ describes the complete light transport between the 5D incoming and outgoing light fields as a linear operator[1]. Heckbert [1991] uses a similar matrix in the context of radiosity problems and shows that such matrices are not sparse. This is also observed by Börm et al. [2003] in the context of linear operators arising from an integral equation such as eq. (1). They show that even though the kernel $\mathbf{K}$ might be sparse, the resulting matrix $(\mathbf{I}-\mathbf{K})^{-1}$ is not. But it is data-sparse. In global illumination computations the kernel is sparse because of occlusions. But due to multiple scattering events one can observe light transport between any pair of points in the scene, resulting in a dense $\widetilde{\mathbf{T}}$. We observe that the indirect bounces affect large portions of the scene similarly. Therefore, large portions of the transport matrix, e.g. those resulting from inter-reflections of diffuse and glossy surfaces, are *data-sparse*. One can exploit the data-sparseness by using local low-rank approximations for sub-blocks of $\widetilde{\mathbf{T}}$. We choose a rank-1 approximation.

Figure 2 illustrates the data-sparseness for a few example transport matrices that we have measured and also demonstrates the local rank-1 approximation. To gain some intuition, let us look at the light transport between two homogeneous planar patches. The light transport between the two is smooth and can be easily factorized. It can be seen in the top-right and bottom-left sub-blocks of the transport matrix for scene (II). Even if the surfaces are textured, it only results in appropriate scaling of either the columns or rows of the transport matrix as shown in (III) and (IV). This will not change the factorization. If a blocker was present between the two patches, it will introduce additional diagonal in the matrix sub-blocks. We can still handle it by subdividing the blocks and factorizing at a finer level as explained in section 5.

[1] Note that our derivation is similar to that of Seitz et al. [2005], but they only derive the formula for light transport between the first bounce 4D light field and outgoing 4D light field whereas our derivation is for complete 5D radiance transfer.

4.2 Symmetry of the Transport Matrix

Capturing full transport matrix is a daunting task. However, $\widetilde{\mathbf{T}}$ is highly redundant, since the radiance along a line is constant unless the line is blocked. So, if one is willing to stay outside the convex hull of the scene to view it or to illuminate it, the 5D representation of the light field can be reduced to 4D [Levoy and Hanrahan 1996; Gortler et al. 1996; Masselus et al. 2003]. We will be working with this representation for the rest of the paper. Let us represent the 4D incoming light field by $\mathbf{L}_{in}(\theta)$ and the 4D outgoing light field by $\mathbf{L}_{out}(\theta)$ where θ parameterizes the space of all possible incoming or outgoing directions on a sphere [Masselus et al. 2003]. The light transport can then be described as:

$$\mathbf{L}_{out} = \mathbf{T}\mathbf{L}_{in} \tag{5}$$

$\mathbf{T}[i,j]$ represents the amount of light received along outgoing direction θ_i when unit radiance is emitted along incoming direction θ_j. Helmholtz reciprocity [von Helmholtz 1856; Rayleigh 1900] states that the light transport between θ_i and θ_j is equal in both directions, i.e. $\mathbf{T}[i,j] = \mathbf{T}[j,i]$. Therefore, $\mathbf{T}$ is a *symmetric* matrix. Also, note that since we are looking at a subset of rays (4D from 5D), $\mathbf{T}$ is just a sub-block of $\widetilde{\mathbf{T}}$. Therefore, $\mathbf{T}$ is also data-sparse.

We will use the two properties, *data-sparseness* and *symmetry*, to develop our method for acquiring $\mathbf{T}$ in the next section.

5 Data Acquisition

In order to measure $\mathbf{T}$, we use projectors to provide the incoming light field and cameras to measure the outgoing light field. Thus, the full $\mathbf{T}$ matrix can be extremely large, depending on the number of pixels in our acquisition hardware. A naive acquisition scheme would involve scanning through individual projector pixels and concatenating the captured camera images to construct $\mathbf{T}$. This could take days or even months to complete. Therefore, to achieve faster acquisition, we would like to illuminate multiple projector pixels at the same time. If we do this, the captured camera image would be the sum of columns in the transport matrix that correspond to illuminated projector pixels. Because of the symmetry of the transport matrix, this would also be the sum of corresponding rows in the matrix. We can use this to exploit data-sparseness in the matrix. Consider a sub-block of the matrix that can be approximated by a rank-1 factorization. By just shining two projector patterns we can capture images such that one provides the column sum and the other provides the row sum of this sub-block. After appropriate normalization of these two measurements we directly get a rank-1 factorization for this sub-block. Thus the whole sub-block can be constructed using just two illumination patterns. This is the key idea behind our algorithm. The algorithm tries to find sub-blocks in $\mathbf{T}$ that can be represented as a rank-1 approximation by a hierarchical subdivision strategy. For an effective hierarchical acquisition we need an efficient data structure to represent $\mathbf{T}$. We will describe our data structure now.

5.1 Hierarchical Tensors for Representing T

We introduce a new data structure called *hierarchical tensors* to represent light transport. Hierarchical tensors are a generalization of hierarchical matrices (or $\mathscr{H}$-matrices), which have been introduced by Hackbush [1999] in the applied mathematics community to represent data-sparse matrices. The key idea behind $\mathscr{H}$-matrices is that a data-sparse matrix can be represented by an adaptive subdivision structure and a low-rank approximation for each node. At each level of the hierarchy, sub-blocks in the matrix are subdivided into 4 children (as in a quadtree). If a sub-block at any level in the tree can be represented by a low-rank approximation, then it is not subdivided any further. Thus, the leaf nodes in the tree contain low-rank approximation for the corresponding sub-block, which reduces to just a scalar value at the finest level in the hierarchy.

Consider the 4D reflectance field that describes the light transport for a single projector/camera pair. We have a 2D image representing the illumination pattern and a resulting 2D image captured by the camera. The connecting light transport can therefore be represented by a 4th-order tensor. One can work by flattening out the 2D image into a vector but that would destroy the spatial coherency present in a 2D image [Wang et al. 2005]. To preserve coherency we represent the light transport by a 4th-order hierarchical tensor. Instead of 4, a node in the hierarchical tensor is divided into 16 children at each level of the hierarchy. Thus, we call the hierarchical representation for a 4th-order tensor, a *sedectree*[2]. Additionally, we use a rank-1 approximation for representing data-sparseness in the leaf nodes of the hierarchical tensor. This means that a leaf node is represented by a tensor product of two 2D images, one from the camera side and the other from the projector side.

5.2 Hierarchical Acquisition Scheme

Our acquisition algorithm follows the structure of the hierarchical tensor described in the previous section. At each level of the hierarchy we illuminate the scene with a few projector patterns. We use the captured images to decide which nodes of the tensor in the previous level of hierarchy are rank-1. Once a node has been determined to be rank-1, we do not subdivide it any further as its entries are known. The nodes which fail the rank-1 test are subdivided and scheduled for investigation during the next iteration. The whole process is repeated until we reach the pixel level. We initiate the acquisition by illuminating with a floodlit projector image. The captured image provides a possible rank-1 factorization of the root node of the hierarchical tensor. We schedule this node for investigation in the first iteration of the algorithm.

The first step in the algorithm is to decide what illumination patterns to use. At each level in the hierarchy, we start off with a list of tensor nodes that need to be investigated. We divide each tensor node into 16 children and the 8 blocks in the projector corresponding to this subdivision are accumulated in a list $\mathbf{B} = \{B_1, B_2, ..., B_n\}$. In order to speed-up our acquisition, we need to minimize the number of patterns we use. To achieve this, our algorithm determines the set of blocks which can be illuminated in the same pattern. We can only illuminate two blocks B_i and B_j in parallel if the light transport between them has been measured at a previous level in the hierarchy. This is because we can use the measured light transport to subtract the contribution of B_i to B_j and B_j to B_i from the camera images obtained upon illuminating with the pattern.

Two blocks B_i and B_j cannot be illuminated in parallel if the light transport between them has not been measured at a previous level of the hierarchy. This is because if both B_i and B_j are illuminated, we would not only measure the contribution from B_j to B_i but also the contribution of B_i to itself. The same holds for block B_j. Therefore, for all possible block pairs for which the light transport has not been resolved yet there is a direct conflict and we generate a conflict set $\mathbf{C} = \{(B_i, B_j) : B_i, B_j \in \mathbf{B}\}$. Given these two sets, we define an undirected graph $\mathbf{G} = (\mathbf{B}, \mathbf{C})$, where $\mathbf{B}$ is the set of vertices in the graph and $\mathbf{C}$ is the set of edges. Now let us look at two blocks B_p and B_q which do not have a direct conflict with each other and hence do not have an edge between them in the graph. If both these blocks have a direct conflict with a common block B_r, we still cannot illuminate them in parallel. This is because in the measurement for block B_r we will receive contributions from both B_p and B_q, thereby causing an indirect conflict between B_p and B_q. Such blocks correspond to vertices at a distance two from each other in our graph $\mathbf{G}$. In order to capture these conflicts as direct edges in a graph, we construct another graph $\mathbf{G}^2$ which is the square of graph $\mathbf{G}$ [Harary 2001]. The square of a graph contains an edge between any two vertices which are at most distance two away from each other in the original graph. Thus, in the graph $\mathbf{G}^2$, any two vertices which are not connected can be scheduled together. We use a graph coloring algorithm on $\mathbf{G}^2$ to obtain conflict free subsets of $\mathbf{B}$ which can be illuminated in parallel. The obtained camera images are then corrected for intra-block light transport due to blocks which are illuminated in the same pattern as described in the previous paragraph.

In the next step, we use these measurements to test if the tensor nodes in the previous level of the hierarchy can be factorized using rank-1 approximation. We have a current rank-1 approximation for each node from the previous level in the hierarchy. The camera images corresponding to 8 projector blocks of a node are used as test cases to validate the current approximation. This is done by rendering estimate images for the 8 projector blocks using the current rank-1 approximation. The estimated images are compared against the corresponding camera images and an RMS error is calculated for the node. A low RMS error indicates our estimates are as good as our measurements and we declare the node as rank-1 and stop any further subdivision on this node. If on the other hand the RMS error is high, it is subdivided into 16 children. The 8 camera images are used to provide rank-1 estimates for the child nodes. These nodes are scheduled for investigation in the next iteration.

A tensor node containing just a scalar value is trivially rank-1. Therefore, the whole process terminates when the size of the projector block reduces to a single pixel. Upon finishing, the scheme directly returns the hierarchical tensor for the reflectance field of the scene.

5.3 Setup and Pre-processing

In order to experimentally validate our ideas we need an acquisition system that is capable of simultaneously emitting and capturing along each ray in the light field. This suggests having a coaxial array of cameras and projectors. Figure 3 shows the schematic of such a setup. Our actual physical implementation is built using a single projector, a single camera, a beam-splitter and an array of planar mirrors. The projector and the camera are mounted coaxially using the beam splitter on an optical bench as shown in Figure 4 and the mirror array divides the projector/camera pixels into 16 coaxial pairs. Once the optical system has been mounted it needs to be calibrated. First, the center of projection of the camera and projector is aligned. The next task is to find the per pixel mapping between the projector and camera pixels. We use a calibration scheme similar to that used by Han and Perlin [2003] and Levoy et al. [2004] in their setup to find this mapping.

[2] Derived from *sedecim*, Latin equivalent of 16.

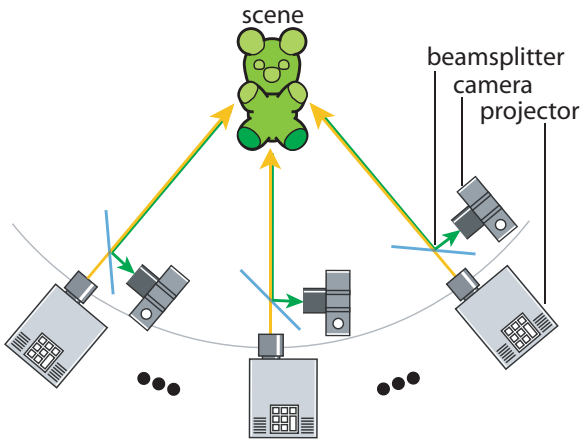


Figure 3: Schematic of symmetric photography setup. A coaxial array of projectors and cameras provides an ideal setup for symmetric photography. The projector array illuminates the scene with an incoming light field. Since the setup is coaxial, the camera array measures the corresponding outgoing light field.

The dynamic range of the scenes that we capture can be very high. This is because the light transport contains not only the high direct bounce effects but also very low secondary bounce effects. In order to capture this range completely, we take multiple images of the scene and combine them into a single high dynamic range image [Debevec and Malik 1997; Robertson et al. 1999]. Additionally, before combining the images for HDR, we subtract the black level of the projector from our images. This accounts for the stray light coming from the projector even when it is shining a completely black image. Also, upon illuminating the scene with individual projector pixels, we notice that the captured images appear darker and have a significantly reduced contrast. This is because an individual projector pixel would be illuminating very few pixels on the Bayer mosaiced sensor of the camera, leading to an error upon interpolation during demosaicing. This problem is also noticed by Sen et al. [2005]. To remove these artifacts, we employ a solution similar to theirs, i.e. the final images are renormalized by forcing the captured images to sum up to the floodlit image.

6 Results

We capture reflectance fields of several scenes. We will explain them now. For reference, Table 1 provides statistics (size, time and number of patterns required for acquisition) for each of these datasets.

In Figure 2, we present the results of our measurement for four simple scenes consisting of planes. This experiment has been designed to elucidate the structure of the **T** matrix. A coaxial projector/camera pair is directly aimed at the scene in this case. The image resolution is 310×350 pixels. Note the storage, time and number of patterns required for the four scenes (listed in Table 1). A brute-force scanning to acquire these **T** matrices would take at least 100 times more images. Also, since the energy in the light after an indirect bounce is low, the camera would have to be exposed for longer time interval to achieve good SNR during brute-force scanning. On the other hand in our scheme, the indirect bounce light transport is resolved earlier in the hierarchy, see rows (c) and (d) in Figure 2. At lower levels of the hierarchy, we are illuminating with bigger projector blocks (and hence throwing more light into the scene than just from a single pixel), therefore we are able to

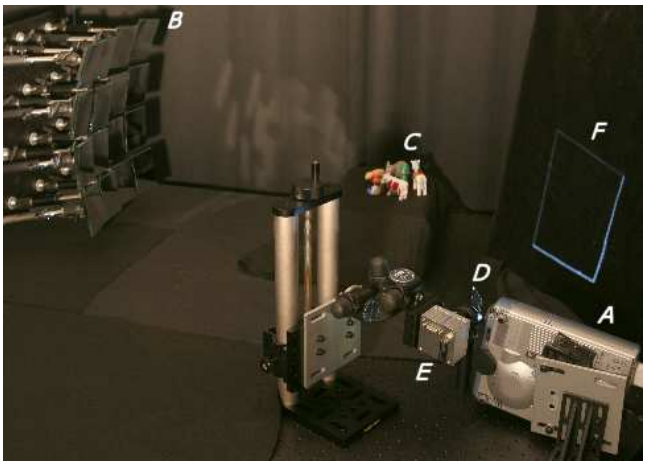


Figure 4: Coaxial setup for capturing 8D reflectance fields. A pattern loaded into projector at A illuminates a 4×4 array of planar mirrors at B. This provides us with 16 virtual projectors which illuminate our scene at C. The light that returns from the scene is diverted by a beam-splitter at D towards a camera at E. Any stray light reflected from the beam-splitter lands in a light trap at F. The camera used is an Imperx IPX-1M48-L (984×1000 pixels) and the projector is a Mitsubishi XD60U (1024×768 pixels). The setup is computer controlled, and we capture HDR images every 2 seconds.

get good SNR even with small exposure times. Also, note that the high frequency of the textures does not affect the data-sparseness of reflectance fields. The hierarchical subdivision follows almost the same strategy in all four cases as visualized in row (d). In row (e), we show the results of relighting the scene with a vertical bar. The smooth glow from one plane to the other in column (II), (III) and (IV) shows that we have measured the indirect bounce correctly.

Figure 1 demonstrates that our technique works well for acquiring the reflectance fields of highly sub-surface scattering objects. The image (240×340 pixels) reconstructed from relighting with a spatially varying illumination pattern (see Figure 1(c)) is validated against the ground-truth image (see Figure 1(d)). We also demonstrate the result of reconstructing at different levels of the hierarchical tensor for this scene in Figure 8. This figure also explains the difference between our hierarchical tensor representation and a wavelet based representation.

Figure 5 describes the result of an 8D reflectance field acquired using our setup. The captured reflectance field can be used to view from any position (see Figure 5(b)) and also to relight from any position (see Figure 5(a)). The resolution of the reflectance field for this example is about $3 \times 3 \times 130 \times 200 \times 3 \times 3 \times 130 \times 200$. The total size of this dataset would be 610 GB if three 32-bit floats were used for each entry in the transport matrix. Our hierarchical tensor representation compresses it to 1.47 GB. A brute force approach would require 233,657 images to capture it. Our algorithm only needs 3,368 HDR images and takes around 8 hours to complete. In our current implementation though, the processing time is a significant amount when compared against the actual image capture time. We believe that the acquisition times can be reduced even further by implementing a parallelized version of our algorithm.

We can also relight the scenes with arbitrary illumination patterns. In Figure 6 we illuminate the scene with two spotlights. The spotlights have a spatially varying illumination distribution that falls off gradually as one moves radially outwards from their center. The specular highlights, caustics and shadows indicate that light transport paths were correctly measured during our acquisition.

Scene	Brute-force	Symmetric photography			Number of leaf nodes (rank-1 blocks) at different levels									
	# Patterns	Size (MB)	Time (min)	# Patterns	1	2	3	4	5	6	7	8	9	10
Fig. 1	111,354	236	103	2,209	16	64	256	1168	5360	16736	54144	119152	116096	610208
Fig. 2(I)	108,500	255	44	809	12	45	237	826	3327	7512	27683	104364	124551	4176784
Fig. 2(II)	108,500	371	70	1,085	12	42	276	1004	2619	9436	36560	127736	228335	6048016
Fig. 2(III)	108,500	334	65	1,081	12	44	244	966	3279	8882	33187	118112	176948	4930752
Fig. 2(IV)	108,500	274	46	841	12	40	312	910	3277	7392	26603	97660	135383	3327632
Fig. 5	233,657	1,470	484	3,368	0	175	639	7979	32080	109197	330760	1109974	2096115	29551312
Fig. 6	133,004	825	227	2,853	8	86	419	3117	11664	40097	146703	529257	855441	11128304

Table 1: Table of relevant data (size, time and number of patterns) for different example scenes captured using our technique. Note that our algorithm requires about 2 orders of magnitude fewer patterns than the brute-force scan. The number of nodes that are represented as rank-1 factors at different levels in the hierarchical tensor are also presented. The greater the number of rank-1 nodes at lower levels of the hierarchy, the faster the acquisition and the smaller the size of our dataset.

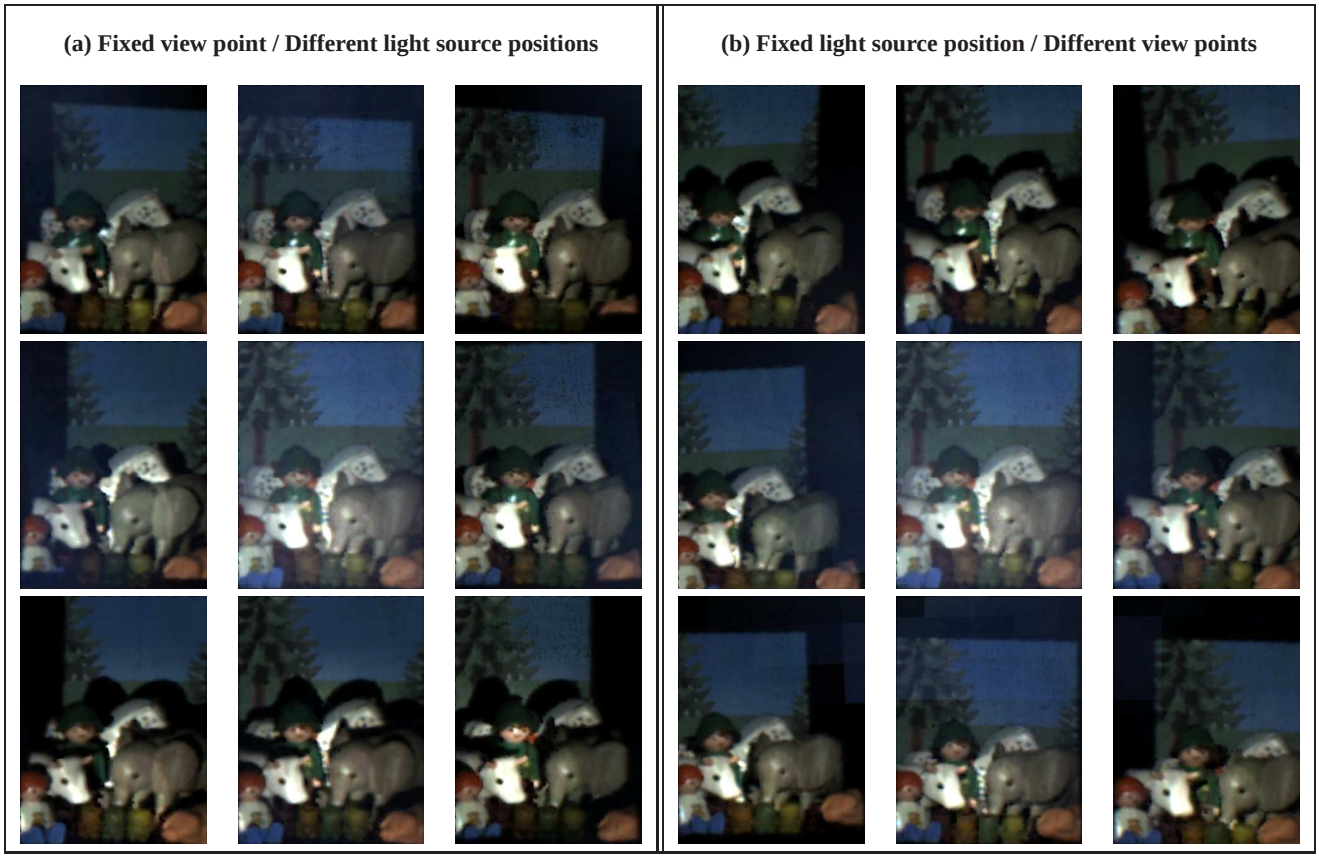


Figure 5: 8D reflectance field of an example scene. This reflectance field was captured using the setup described in Figure 4. A 3×3 grid of mirrors was used. In (a) we see images rendered from the viewpoint at the center of the grid with illumination coming from 9 different locations on the grid. Note that the shadows move appropriately depending upon the direction of incident light. (b) shows the images rendered from 9 different viewpoints on the grid with the illumination coming from the center. In this case one can notice the change in parallax with the viewpoint.

7 Discussion and Conclusions

In this paper we have presented a framework for acquiring 8D reflectance fields. The method is based on the observation that reflectance fields are data-sparse. We exploit the data-sparseness to represent the transport matrix by local rank-1 approximations. The symmetry of the light transport allows us to measure these local rank-1 factorizations efficiently as we can obtain measurements corresponding to both rows and columns of the transport matrix simultaneously. We have also introduced a new data structure called hierarchical tensor that can represent these local low-rank approximations. Based on these observations we have developed a hierarchical acquisition algorithm which looks for regions of data-sparseness in the matrix. Once a data-sparse region has been measured we can use it to parallelize our acquisition resulting in tremendous speedup.

There are limitations in our setup (Figure 4) that can corrupt our measurements. To get a coaxial setup we use a beam-splitter. Although we use a 1mm thin plate beam-splitter, it produces slight double image inherent to beam-splitters. This along with the light reflected back off the light trap reduces the SNR in our measurements. The symmetry of our approach requires projector and cam-

Figure 6: Relighting with 2 spot lights. Image rendered from a reflectance field captured by our system. The egg and the wooden toy are lit from 2 spotlights (with spatially varying intensity distribution), one shining from the right and the other from left. Another faint floodlight adds to the overall brightness of the scene. Note that the caustics from the glass, specularities on the egg, and the toy, and shadows are all faithfully reconstructed.

era to be pixel aligned. Any slight misalignment adds to the measurement noise. Cameras and projectors can also have different optical properties. This can introduce non-symmetries such as lens flare, resulting in artifacts in our reconstructed images (see Figure 7). We hope that future advances in hardware technology would result in symmetric projector/camera *combo* devices that can avoid the need for beam-splitters and calibration.

In order to keep our implementation simple, we use a 4th order hierarchical tensor. This means that we are flattening out 2 of the 4 dimensions of the light field, thereby not exploiting the full coherency in the data. An implementation based on 8th order tensor should be able to exploit it and make the acquisition more efficient.

Since we use a 4×4 array of planar mirrors, the resolution of our incoming and outgoing light fields is low. Therefore, the reflectance fields that we can capture are sparse. Techniques have been proposed for interpolating slices of the reflectance fields before, both from the view direction [Chen and Williams 1993] and from the illumination direction [Chen and Lensch 2005] but the problem of interpolating reflectance fields is still open. By applying similar flow based techniques to the transport matrix, one should be able to create more densely sampled reflectance fields. One can also imagine directly sampling incoming and outgoing light fields more densely by pointing the projector/camera pair at a bumpy mirrored sheet. This will increase the number of viewpoints in the light field but at the cost of image resolution.

Finally, although we introduce the hierarchical tensor as a data structure for storing reflectance fields, the concept has implications for other high dimensional datasets as well. It can be used for representing any dataset which is data-sparse. The hierarchical representation also has some other benefits. It provides constant time access to the data during evaluation or rendering. At the same time it maintains the spatial coherency in the data, making it attractive for parallel computation.

References

BÖRM, S., GRAYSEDYCK, L., AND HACKBUSH, W. 2003. Introduction to Hierarchical Matrices with Applications. *Engineering Analysis with Boundary Elements 27*, 5, 405–422.

CHEN, B., AND LENSCH, H. P. A. 2005. Light Source Interpolation for Sparsely Sampled Reflectance Fields. In *Vision Modeling ang Visualization (VMV'05)*, 461–468.

CHEN, S. E., AND WILLIAMS, L. 1993. View interpolation for Image Synthesis. In *SIGGRAPH '93*, 279–288.

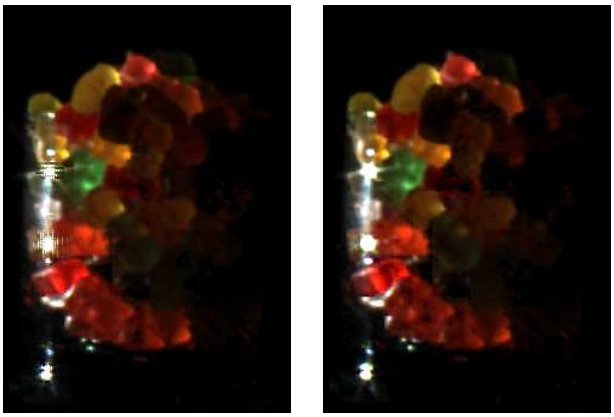


Figure 7: Artifacts due to non-symmetry in measurement. The lens flare around the highlights (right image) is caused by the aperture in the camera. Since this effect does not occur in the incident illumination from the projector, the measurements are non-symmetric. Applying a strong threshold for the rank-1 test subdivides the region very finely and produces a corrupted result in the area of the highlights (left image). If the inconsistencies in measurement are stored at a higher subdivision level by choosing a looser threshold for rank-1 test, these artifacts are less noticeable (right image) .

DEBEVEC, P., AND MALIK, J. 1997. Recovering High Dynamic Range Radiance Maps from Photographs. In *SIGGRAPH '97*, 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 *SIGGRAPH '00*, 145–156.

DURAND, F., HOLZSCHUCH, N., SOLER, C., CHAN, E., AND SILLION, F. X. 2005. A Frequency Analysis of Light Transport. In *SIGGRAPH '05*, vol. 24, 1115–1126.

GERSHUN, A. 1936. The Light Field, Moscow. *Journal of Mathematics and Physics (1939)*. Vol. XVIII, MIT. Translated by P. Moon and G. Timoshenko.

GOESELE, M., LENSCH, H. P. A., LANG, J., FUCHS, C., AND SEIDEL, H.-P. 2004. DISCO: Acquisition of Translucent Objects. In *SIGGRAPH '04*, 835–844.

GORTLER, S. J., GRZESZCZUK, R., SZELISKI, R., AND COHEN, M. F. 1996. The Lumigraph. In *SIGGRAPH '96*, 43–54.

HACKBUSCH, W. 1999. A Sparse Matrix Arithmetic based on $\mathscr{H}$-Matrices. Part I: Introduction to $\mathscr{H}$-matrices. *Computing 62*, 2, 89–108.

HAN, J. Y., AND PERLIN, K. 2003. Measuring Bidirectional Texture Reflectance with a Kaleidoscope. In *SIGGRAPH '03*, 741–748.

HARARY, F. 2001. *Graph Theory*. Narosa Publishing House.

HAWKINS, T., EINARSSON, P., AND DEBEVEC, P. E. 2005. A Dual Light Stage. In *Eurographics Symposium on Rendering*, 91–98.

HECKBERT, P. S. 1991. *Simulating Global Illumination Using Adaptive Meshing*. PhD thesis, University of California, Berkeley.

KAJIYA, J. T. 1986. The Rendering Equation. In *SIGGRAPH '86*, 143–150.

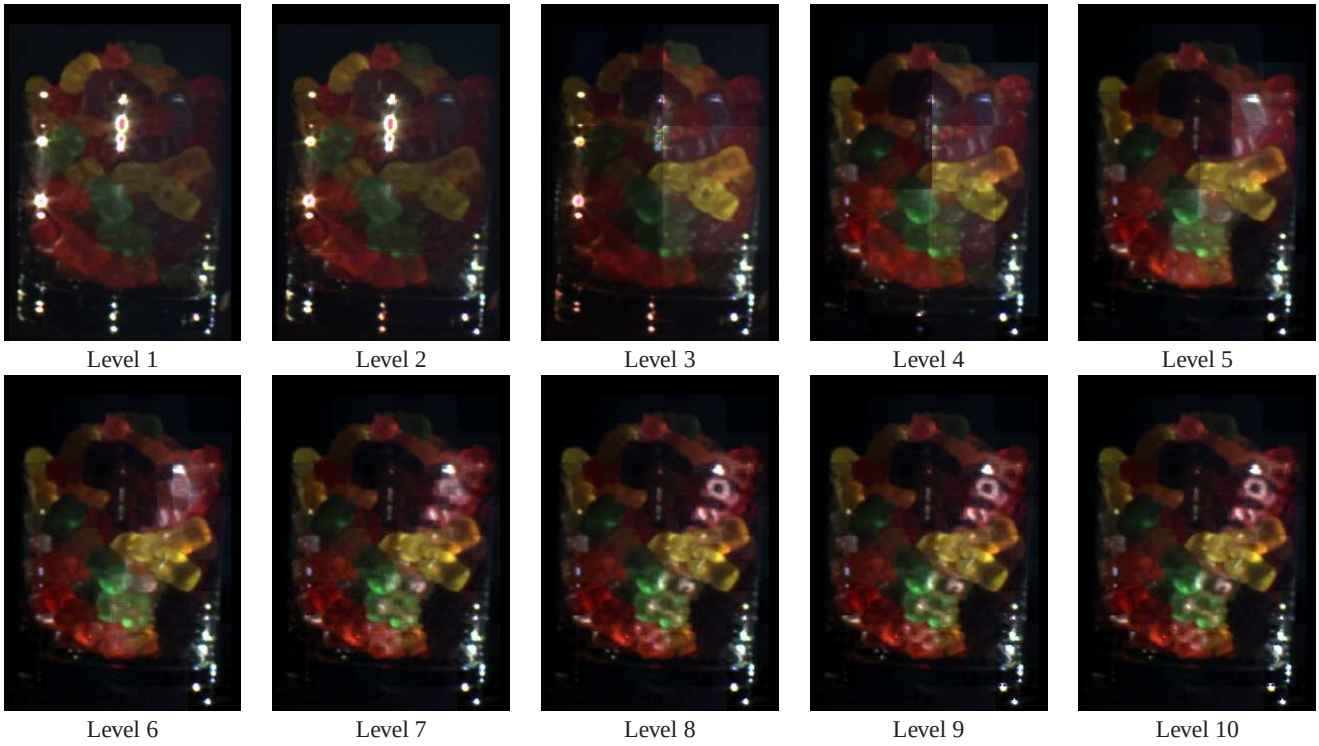


Figure 8: Reconstruction results for different levels of hierarchy. This example illustrates the relighting of the reflectance field of gummy bears scene (Figure 1) with illumination pattern used in Figure 1(c), if the acquisition was stopped at different levels of the hierarchy. Note that at every level, we still get a full resolution image. This is because we are approximating a node in the hierarchy as a tensor product of two 2-D images. Therefore, we sill have a measurement for each pixel in the image, though scaled incorrectly. This is different from wavelet based approaches where a single scalar value is assigned for a node in the hierarchy implying lower resolution in the image at lower levels. Note that at any level, the energy of the projector matte is distributed over the whole block that it is illuminating. This is clear from the intensity variation among blocks, especially in the images at levels 3, 4, and 5.

KAUTZ, J., AND MCCOOL, M. D. 1999. Interactive Rendering with Arbitrary BRDFs using Separable Approximations. In *Eurographics Workshop on Rendering*, 247–260.

LATTA, L., AND KOLB, A. 2002. Homomorphic Factorization of BRDF-based Lighting Computation. In *SIGGRAPH '02*, 509–516.

LEVOY, M., AND HANRAHAN, P. 1996. Light Field Rendering. In *SIGGRAPH '96*, 31–42.

LEVOY, M., CHEN, B., VAISH, V., HOROWITZ, M., MCDOWALL, I., AND BOLAS, M. 2004. Synthetic Aperture Confocal Imaging. In *SIGGRAPH '04*, 825–834.

MALZBENDER, T., GELB, D., AND WOLTERS, H. 2001. Polynomial Texture Maps. In *SIGGRAPH '01*, 519–528.

MASSELUS, V., PEERS, P., DUTRÉ, P., AND WILLEMS, Y. D. 2003. Relighting with 4D Incident Light Fields. In *SIGGRAPH '03*, 613–620.

MATUSIK, W., LOPER, M., AND PFISTER, H. 2004. Progressively - Refined Reflectance Functions for Natural Illumination. In *Eurographics Symposium on Rendering*, 299–308.

MCCOOL, M. D., ANG, J., AND AHMAD, A. 2001. Homomorphic Factorization of BRDFs for High-Performance Rendering. In *SIGGRAPH '01*, 171–178.

PEERS, P., AND DUTRÉ, P. 2003. Wavelet Environment Matting. In *Eurographics Symposium on Rendering*, 157–166.

RAMAMOORTHI, R., AND HANRAHAN, P. 2001. A Signal-Processing Framework for Inverse Rendering. In *SIGGRAPH '01*, 117–128.

RAYLEIGH, J. W. S. B. 1900. On the Law of Reciprocity in Diffuse Reflexion. *Philosophical Magazine 49*, 324–325.

ROBERTSON, M. A., S., AND STEVENSON, R. L. 1999. Dynamic Range Improvement through Multiple Exposures. In *Proceedings of the IEEE Intl. Conference on Image Processing (ICIP'99)*, 159–163.

SEITZ, S. M., MATSUSHITA, Y., AND KUTULAKOS, K. N. 2005. A Theory of Inverse Light Transport. In *Proceedings of the IEEE Intl. Conference on Computer Vision (ICCV '05)*, 1440–1447.

SEN, P., CHEN, B., GARG, G., MARSCHNER, S. R., HOROWITZ, M., LEVOY, M., AND LENSCH, H. P. A. 2005. Dual Photography. In *SIGGRAPH '05*, 745–755.

VASILESCU, M. A. O., AND TERZOPOULOS, D. 2004. TensorTextures: Multilinear Image-Based Rendering. In *SIGGRAPH '04*, vol. 23, 336–342.

VON HELMHOLTZ, H. 1856. *Treatise on Physiological Optics (1925)*. The Optical Society of America. Electronic edition (2001): University of Pennsylvania `http://psych.upenn.edu/backuslab/helmholtz`.

WANG, H., WU, Q., SHI, L., YU, Y., AND AHUJA, N. 2005. Out-of-Core Tensor Approximation of Multi-Dimensional Matrices of Visual Data. In *SIGGRAPH '05*, vol. 24, 527–535.