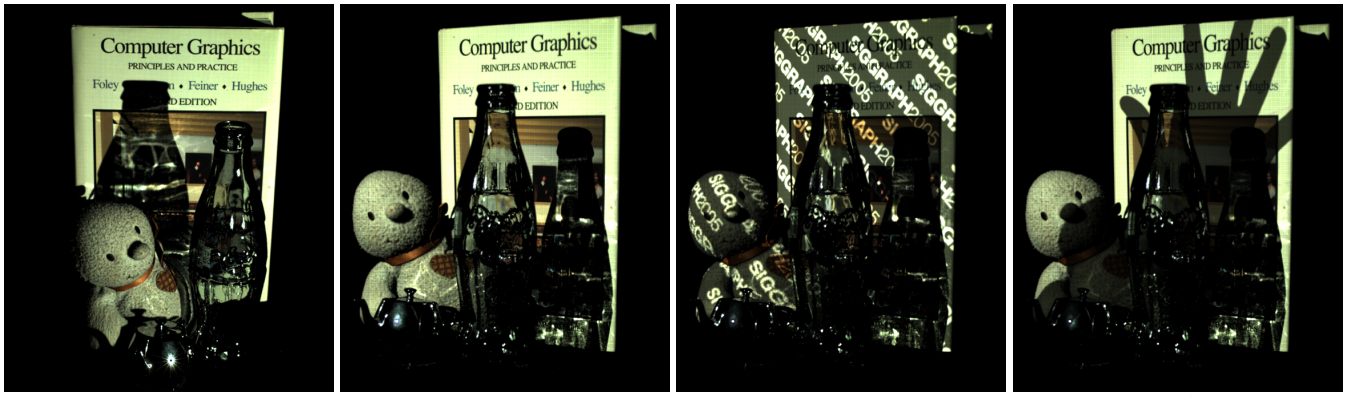
Dual Photography

Pradeep Sen* Billy Chen* Gaurav Garg* Stephen R. Marschner†
Mark Horowitz* Marc Levoy* Hendrik P. A. Lensch*

*Stanford University †Cornell University



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

Figure 1: (a) Conventional photograph of a scene, illuminated by a projector with all its pixels turned on. (b) After measuring the light transport between the projector and the camera using structured illumination, our technique is able to synthesize a photorealistic image from the point of view of the projector. This image has the resolution of the projector and is illuminated by a light source at the position of the camera. The technique can capture subtle illumination effects such as caustics and self-shadowing. Note, for example, how the glass bottle in the primal image (a) appears as the caustic in the dual image (b) and vice-versa. Because we have determined the complete light transport between the projector and camera, it is easy to relight the dual image using a synthetic light source (c) or a light modified by a matte captured later by the same camera (d).

Abstract

We present a novel photographic technique called dual photography, which exploits Helmholtz reciprocity to interchange the lights and cameras in a scene. With a video projector providing structured illumination, reciprocity permits us to generate pictures from the viewpoint of the projector, even though no camera was present at that location. The technique is completely image-based, requiring no knowledge of scene geometry or surface properties, and by its nature automatically includes all transport paths, including shadows, inter-reflections and caustics. In its simplest form, the technique can be used to take photographs without a camera; we demonstrate this by capturing a photograph using a projector and a photo-resistor. If the photo-resistor is replaced by a camera, we can produce a 4D dataset that allows for relighting with 2D incident illumination. Using an array of cameras we can produce a 6D slice of the 8D reflectance field that allows for relighting with arbitrary light fields. Since an array of cameras can operate in parallel without interference, whereas an array of light sources cannot, dual photography is fundamentally a more efficient way to capture such a 6D dataset than a system based on multiple projectors and one camera. As an example, we show how dual photography can be used to capture and relight scenes.

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

Keywords: dual photography, image synthesis, image-based relighting

*e-mail:{psen|billyc|ggaurav|horowitz|levoy|lensch}@stanford.edu
†e-mail:srm@cs.cornell.edu

1 Introduction

Helmholtz reciprocity, the idea that the flow of light can be effectively reversed without altering its transport properties, is exploited in many graphics applications to reduce computational complexity (e.g. in ray-tracing systems [Whitted 1980]). In the graphics literature, this reciprocity is typically summarized by an equation describing the symmetry of the radiance transfer between incoming and outgoing directions ω_i and ω_o: $f_r(\omega_i \rightarrow \omega_o) = f_r(\omega_o \rightarrow \omega_i)$, where f_r represents the BRDF of the surface. Although this general form is often attributed to Helmholtz, his original treatise on optics makes this claim only for specular interactions [1856]. Rayleigh later extended reciprocity to include non-specular reflection [1900]. For a more complete discussion on reciprocity, interested readers are referred to Veach [1997].

In this paper we introduce the concept of *dual photography* which is based on Helmholtz reciprocity. In particular, we present techniques for efficiently capturing the light transport between a camera and a projector, including effects such as mirrored reflections, caustics, diffuse inter-reflections and subsurface scattering. More precisely, we acquire the transport coefficients between every pair of camera and projector pixels at the resolution of both devices. Once this information has been captured, it can be used for scene relighting [Masselus et al. 2003], where a virtual pattern is used to illuminate the scene as a post-process. Furthermore, since we have resolved the light transport down to the pixel level of the projector,

we can interchange the role of camera and projector and generate a full resolution *dual image* from the point of view of the projector. This dual image can then be relit by an adjustable light source located at the original camera position. Because this light can be of high spatial resolution, we can achieve complex lighting effects, such as casting sharp shadows.

This paper has three main contributions. First, in Section 2 we demonstrate how Helmholtz reciprocity can be used to generate a dual image from a viewpoint where no camera was present. In Section 3, we develop efficient algorithms to capture the light transport in parallel by projecting patterns that adapt to the scene content. These improvements allow for practical capture of the transport matrix needed for dual photography. Finally, in Section 4 we show that dual photography can be applied to the problem of scene relighting. In particular, we show that we can efficiently capture a 6D slice of the full 8D reflectance function that describes the relationship between the 4D incoming lightfield and the 4D outgoing light field [Debevec et al. 2000]. This 6D slice allows us to relight a scene using a 4D incident light field while keeping the viewpoint fixed. By combining the speed-up from the algorithm we have developed with the fundamental efficiency of dual photography, we can acquire our reflectance data at one order of magnitude higher effective resolution and capture it two orders of magnitude faster than the best previous results, e.g. [Masselus et al. 2003] or [Goesele et al. 2004].

2 Dual Photography

We explain the principle of dual photography with reference to the imaging configuration shown in Figure 2. We have a projector of resolution $p \times q$ shining light onto a scene and a camera of resolution $m \times n$ capturing the reflected light. Since light transport is linear, the transport from the projector through the scene and into the camera can be expressed as the following simple equation:

$$\mathbf{c}' = \mathbf{T}\mathbf{p}' \tag{1}$$

The column vector $\mathbf{p}'$ is the projected pattern (size $pq \times 1$), and $\mathbf{c}'$ (size $mn \times 1$) represents the image captured by the camera. Matrix $\mathbf{T}$ (size $mn \times pq$) is the transport matrix that describes how light from each pixel of $\mathbf{p}'$ arrives at each pixel of $\mathbf{c}'$, taking into account reflections, refractions, and all other light transport paths. For intuition on the composition of $\mathbf{T}$, readers are referred to Figure 3.

We use the prime superscript ($'$) to indicate that we are working in the primal space to distinguish it from its dual counterpart, which we will introduce in a moment. Helmholtz reciprocity states that the light sources and cameras in a scene can be interchanged without changing either the path taken by the light or the transfer of energy from one to the other. As we show in Appendix A, this means that we can represent the dual of Equation 1 as follows:

$$\mathbf{p}'' = \mathbf{T}^T\mathbf{c}'' \tag{2}$$

In this equation the transport matrix $\mathbf{T}$ of the scene is the same as before except that we have now transposed it to represent light going from the camera to the projector. We shall refer to Equation 1 as the "primal" equation and Equation 2 as the "dual" equation. In the dual space, $\mathbf{p}''$ represents the virtual image that would be visible at the projector if the camera were "projecting" pattern $\mathbf{c}''$. We call the process of transposing the transport matrix and multiplying by the desired lighting *dual photography*. Since the two representations are equivalent, the $\mathbf{T}$ matrix can be acquired in either space and then transposed to represent transport in the other space. This is a relatively large matrix, so we develop algorithms to accelerate its acquisition in Section 3. Also note that the two equations are not mathematical inverses of each other (i.e. $\mathbf{T}\mathbf{T}^T \neq \mathbf{I}$). This is because energy is lost in any real system through absorption or scattering.

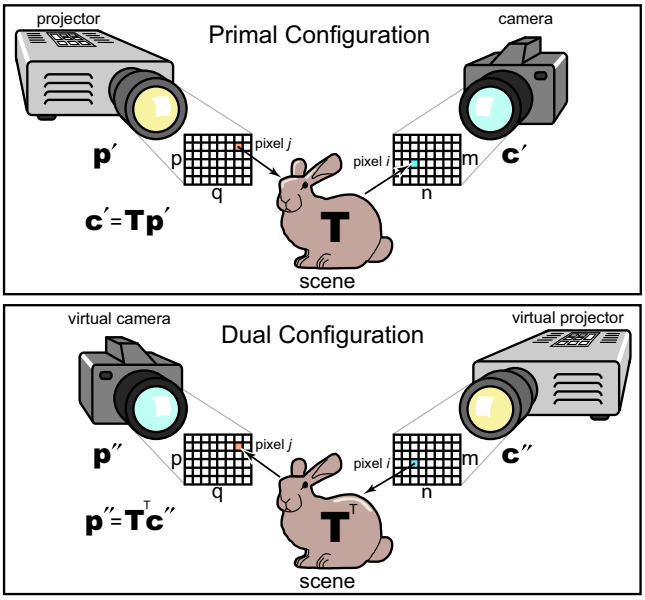


Figure 2: The principle of dual photography. The top diagram shows our primal configuration, with light being emitted by a real projector and captured by a real camera. Matrix $\mathbf{T}$ describes the light transport between the projector and the camera (element $\mathbf{T}_{ij}$ is the transport coefficient from projector pixel j to camera pixel i). The bottom diagram shows the dual configuration, with the positions of the projector and camera reversed. Suppose $\mathbf{T}''$ is the transport matrix in this dual configuration, so that $\mathbf{T}''_{ji}$ is the transport between pixel i of the virtual projector and pixel j of the virtual camera. As shown in Appendix A, Helmholtz reciprocity specifies that the pixel-to-pixel transport is equal in both directions, i. e. $\mathbf{T}''_{ji} = \mathbf{T}_{ij}$, which means $\mathbf{T}'' = \mathbf{T}^T$. As explained in the text, given $\mathbf{T}$, we can use $\mathbf{T}^T$ to synthesize the images that would be acquired in the dual configuration.

Therefore, if we measure $\mathbf{c}'$ after applying $\mathbf{p}'$, we cannot put this back in $\mathbf{c}''$ and expect the resulting $\mathbf{p}''$ to equal the original $\mathbf{p}'$.

We can extend this idea to a configuration of multiple cameras and projectors. Suppose first we had a single projector and an array of $k \times l$ cameras all aimed at the scene. Since each camera sees a unique 2D slice of the 8D reflectance field, it will observe a unique transport function with respect to the projector. These individual transport matrices can be concatenated together to form a larger $klmn \times pq$ matrix that represents the 6D transfer between a single spatially varying light source (the projector) and multiple cameras. Applying the principle of duality we can turn this data into a matrix representing the transport between multiple projectors and one camera. We demonstrate this idea in Section 4. Measuring this transport for multiple projector positions (in primal space) would capture the full 8D reflectance field, but this has not yet been implemented.

Photography without an imaging element. We now demonstrate a simple example of dual photography by synthesizing a dual image of a scene without using an imaging device. Suppose that instead of a camera we have a light sensor that integrates $\mathbf{c}'$ over some solid angle to produce a scalar value c'. This turns the $\mathbf{T}$ matrix in Equation 1 into a row vector that for any input lighting $\mathbf{p}'$ results in only a scalar output. Thus Equations 1 and 2 become $c' = \begin{bmatrix} t_1 \cdots t_{pq} \end{bmatrix} \mathbf{p}'$ and $\mathbf{p}'' = \begin{bmatrix} t_1 \cdots t_{pq} \end{bmatrix}^T c''$ respectively. Thus the dual image $\mathbf{p}''$ we would get at the projector is simply this column vector scaled by

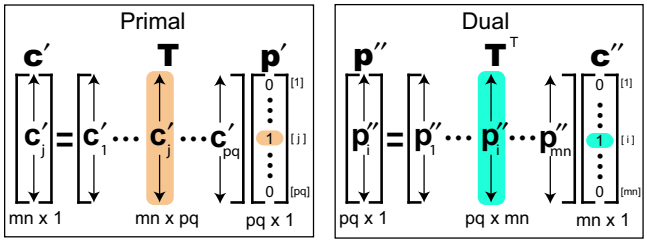


Figure 3: Makeup of the T matrix. The columns of the **T** matrix represent the pictures that would be taken at $\mathbf{c}'$ when the appropriate pixel at $\mathbf{p}'$ is lit. Thus, we can think of **T** as a concatenation of camera images $\mathbf{c}'_1$ through $\mathbf{c}'_{pq}$ in column vector form. In a similar way, the columns of $\mathbf{T}^T$ in the dual space represent images at the projector $\mathbf{p}''$ when a single pixel at $\mathbf{c}''$ is illuminated.

the new virtual light source c''. This new light source cannot be spatially modulated (since c'' is a scalar) and it shines in every direction (up to the viewing angle of the photo-resistor) in proportion to the sensitivity of the photo-resistor in that direction.

To generate the dual image, we must first find the values of $\mathbf{t} = [t_1, \ldots, t_{pq}]$. A simple way to do this is to perform a pixel scan with the projector by displaying $p \times q$ different patterns each with only one element lit up at a time. We shall refer to this technique as the "brute-force" pixel scan. When projected into the scene, each of these basis patterns extracts a single component of the **t** vector which will be measured by the photo-resistor as value c'. By putting these measurements back together in the correct order, the **t** vector can be constructed and used to synthesize an image from the point of view of the projector.

An image constructed using this approach is presented in Figure 4. This image was acquired by scanning a 3×3-pixel square on the screen and recording the measurement by the photo-resistors at each position. The area subtended by each projector square is integrated to form one pixel in the final image, so larger squares will yield coarser images with respect to the scene. However, this integration has the beneficial side effect of antialiasing the image using a box filter as large as the pixel square. A 3×3-pixel square was chosen for two reasons. First, the pixel-by-pixel scan is slow so the square had to be appropriately sized to the scene to finish within reasonable time (90 minutes for this example). Second, it is difficult to get a good signal with less than a 3×3-pixel square because of the limited sensitivity of the photo-resistor. Typical DLP projectors also suffer from limited contrast which creates a "dark level" and acts as a source of noise in our measurements.

Comparison to Previous Work. The experiment of the last subsection might remind readers of the "flying-spot" camera, a precursor to the modern television camera. First demonstrated successfully by John Baird in 1926, this camera used a light source modulated by a perforated disk (known as a Nipkow disk) to scan light spots across a scene. As in our experiment, the outgoing light was measured by photosensors used to generate the image. There are also modern versions of flying-spot cameras (see e.g. Microvision's confocal microprobe which uses a scanning laser to image). Arguably other scanned beam systems such as sonar, ultrasound, and scanning electron microscopes can be considered dual photographic systems as well. Thus, the experiment shown in the previous section is not new — what is novel is the dual photography framework we present that extends the ideas from scanned beam systems with simple non-imaging sensors to systems with both cameras and projectors. Also new is our demonstration that cameras and projectors can be interchanged when the transport from one to the other is known.

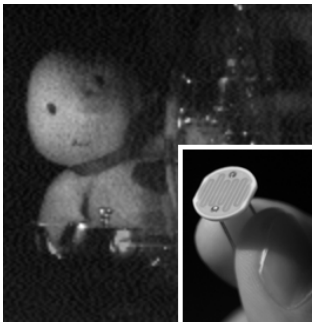


Figure 4: Photography without an imaging sensor. This image was generated using a projector and two photo-resistors like the one shown in the inset. This is a dual image of the scene in Figure 1 and is a view of the scene from the projector's location as illuminated by point light sources at the locations of the two photo-resistors.

The light transport matrix we use in our work appears in similar form in several recent relighting papers. Our work was inspired by Masselus et al. [2003] who also used a camera and a projector for scene relighting. By moving the projector to multiple positions, they acquired a 6D reflectance field with coarse projector resolution (16×16 samples). A reflectance field with higher spatial resolution but no angular resolution has been captured by Goesele et al. [2004] in order to relight objects with heterogeneous diffuse subsurface scattering properties. In their approach, a laser projector scanned the object surface while an HDR video camera took up to one million images, resulting in a moderate effective projector resolution. In this paper, we demonstrate techniques to capture a reflectance field with higher spatial resolution using significantly fewer images.

In the relighting work by Debevec et al. [2000] distant point light sources are assumed, which means that there is no spatial variation within each light source. Therefore, such a technique cannot cast sharp shadows onto the scene, something that is possible with our virtual projector. On the other hand, they can relight using hundreds of light sources, while in our experiments we have used a much smaller number. Environment matting [Zongker et al. 1999] measures the light transport between a monitor and a camera, resulting in high angular but no spatial resolution for the incident light. In Section 3 we will extend Peers' and Dutré's [2003] wavelet environment matting approach for deriving the acquisition scheme, which can be applied to capture environment mattes as well as other slices of the 8D reflectance field with high efficiency. Opacity hulls [Matusik et al. 2002b; Matusik et al. 2002a] combine reflectance fields with environment mattes. They achieved very high angular resolution, which allows for relighting even of refractive objects. However, they cannot relight their scenes with a spatially varying light source as well.

None of the previous relighting approaches have made use of Helmholtz reciprocity. However, there has been some previous work in the computer vision community that takes advantage of Helmholtz reciprocity. Zickler et al. used reciprocity to reconstruct the geometry of surfaces with arbitrary BRDFs in what they call Helmholtz stereopsis [2002]. The authors observed that by interchanging light source and camera during acquisition, they can use Helmholtz reciprocity to guarantee that points on the surface would have exactly the same transfer characteristic in both configurations. This simplifies stereo matching, even for surfaces with complex BRDFs. A similar approach was taken by Tu et al. [2003], who utilized reciprocity for the task of 3D to 2D registration. All these techniques contain an additional cosine term to account for the foreshortening of the incident illumination, which is not needed in dual photography. A formal proof of this is shown in Appendix A.

Finally, reciprocity has been used in a manner analogous to our own in work by Zotkin et al. [2004] in order to model sound transfer through a human head. In this work, the authors noted that an array

of microphones and a single speaker can be used to replace the more common setup of a single microphone and multiple speakers. This allows the data to be acquired in parallel in a similar manner to our technique for acquiring 6D reflectance.

3 Efficient Acquisition of the T Matrix

In the case of a projector and one or more cameras, the full **T** matrix is extremely large (on the order of $10^6 \times 10^6$ elements for conventional projector/camera resolutions). Thus, we must develop algorithms that can acquire it efficiently. In addition, the contrast between brightest and dimmest entries in this matrix can be large. This means that we must (1) devise a set of patterns $\mathbf{p}'$ that relate the measured values to the original projector pixels efficiently and (2) develop infrastructure to deal with the high dynamic range of scenes that exhibit both specular and diffuse inter-reflection.

To address the first issue, we developed an efficient hierarchical adaptive algorithm. However, this algorithm requires a considerable amount of infrastructure to support the feedback necessary to "adapt." For this reason, we begin our discussion by describing a straightforward, fixed-pattern algorithm that can be implemented in a short amount of time and yields reasonable results. This technique has some drawbacks, however, which are addressed in subsection 3.2 where we introduce a simple adaptive algorithm. The shortcomings of this simplified adaptive algorithm are then addressed by our full hierarchical adaptive scheme in subsection 3.3. In the last subsection (3.4), we discuss the technical details of implementing the HDR pipeline required to measure the **T** matrix faithfully.

3.1 Fixed Pattern Scanning

In order to accelerate the acquisition we would like to parallelize the patterns, illuminating multiple pixels at the same time. With a fixed scheme, this is possible only if we make the assumption that each projector pixel affects a small, localized region of the scene from the point of view of the camera (i.e. the j^{th} projector pixel only affects a few elements of $\mathbf{c}'_j$). This means that only scenes with direct illumination can be handled properly and is similar to the assumption made by Masselus et al. [2003] to accelerate their acquisition. To understand how parallelization can accelerate the acquisition, suppose the non-zero values of camera image vector $\mathbf{c}'_i$ are in different positions than the non-zero values of vector $\mathbf{c}'_j$ (i.e. $\mathbf{c}'^T_i \mathbf{c}'_j = 0$). In this scenario, we shall say that $\mathbf{c}'_i$ and $\mathbf{c}'_j$ are not in conflict. If these positions are known for both i and j, we can project a single pattern $\mathbf{p}' = \mathbf{p}'_i + \mathbf{p}'_j$ to capture $\mathbf{c}' = \mathbf{c}'_i + \mathbf{c}'_j$ and then separate image vectors $\mathbf{c}'_i$ and $\mathbf{c}'_j$ as a post-process.

Since projector pixels typically affect more than one pixel in the camera, we must illuminate only those pixels that do not overlap in their projected region (i.e. no conflict). Furthermore, to reduce acquisition time we must turn on as many pixels as possible at a time. Our algorithm achieves these two goals by breaking up the projector image into square blocks within each of which only one pixel is illuminated at a time. This ensures a minimum distance between neighboring pixels while maximizing the number of displayed pixels. For this work, a block size of 8×8 was used.

Using this decomposition, the pixels within a block are scanned sequentially. This covers the entire projector image in 64 exposures. The captured images will be composed of the sum of the contributions of every pixel that was illuminated (i.e. $\mathbf{c}' = \sum_{j \in S} \mathbf{c}'_j$ where S is the set of pixels illuminated). The problem now is to find the correspondence between lit projector pixels and pixels of $\mathbf{c}'$ (i.e. we need to find out which elements of $\mathbf{c}'$ are being affected by which projector pixels). To do this, we label each block with a unique binary number. We display a temporal sequence of patterns from the projector that "spell" out this number for every block one bit at a time. For example, for the seventh block we would display a "0 1 1 1" pattern, so the camera pixels seeing this pattern would know they are observing contribution from the seventh projector block.

There are two issues we must address to improve the quality of acquisitions using this technique. First, our technique incorrectly limits the contribution of pixels near block boundaries because their influence often extends beyond the given block edge and is mislabeled. To alleviate this problem, we shine offset block patterns shifted by half a block horizontally, vertically, and diagonally. This ensures that every pixel in the projector will have a block centered around it to within a few pixels so that all of its contribution is labeled correctly. Second, sometimes the blocks can be mislabeled because of noisy measurement. To reduce this problem, the block labels contain extra bits to add redundancy for error correction. Specifically, we add a sequence of Hamming error-correcting codes that allow us to correct single bit errors that might occur after the block labeling phase. We have found that these two things generally fix mislabeled pixels in the image.

Results from an acquisition that used this technique are shown shown in Figure 15. Note that by implicitly finding correspondences between projector blocks and camera pixels, we are essentially doing something similar to stripe-based rangefinding, for which numerous other techniques exist (e.g. Rusinkiewicz et al. [2002]). However, any direct correspondence technique has the disadvantage that it only works if the correspondence between projector pixels and camera pixels is one-to-one. In our application, this means that our technique fails to capture global illumination effects, such as diffuse illumination, that map many projector pixels to the same camera pixel. We address this problem in the next section.

3.2 Adaptive Multiplexed Illumination

In this section, we address the limitation of the fixed pattern scheme with an algorithm that determines adaptively which pixels of the projector can be illuminated in parallel to relax the "direct illumination" assumption while still accelerating acquisition. Unlike Schechner et al.'s work that uses codes to increase redundancy and reduce SNR [2003], our codes reduce redundancy to improve acquisition efficiency. Specifically, our adaptive algorithm tries to acquire the transport matrix with as few patterns as possible while ensuring that projector pixels affecting the same camera pixel are never illuminated simultaneously. We avoid such conflicts by subdividing the projector space adaptively; starting with the floodlit projector image, we subdivide it into four blocks, which are refined recursively until we reach the pixel level. Whenever we subdivide a block, we illuminate the four children in sequence. A walk-through example of the algorithm is given in Figure 5.

Two blocks can be investigated/subdivided in parallel if no camera pixel received contribution from both blocks. At each level of the subdivision, we determine for each camera pixel k the blocks $\mathbf{B}_k = \{B_0, \ldots, B_n\}$ which illuminate pixel k either indirectly or directly. For all possible pairs of blocks contributing to the same camera pixel we generate a conflict set $\mathbf{C}_k = \{(B_i, B_j) : B_i, B_j \in \mathbf{B}_k\}$. In the next iteration, the only blocks that need to be subdivided are in $\mathbf{B} = \bigcup_k \mathbf{B}_k$, i.e. only the blocks that contributed to any camera pixel. In this manner, blocks that do not contribute to the final image in any way are immediately culled away.

Given the set of all generated conflicts across all camera pixels $\mathbf{C} = \bigcup_k \mathbf{C}_k$, we define a graph $(\mathbf{B}, \mathbf{C})$. A graph coloring scheme is used to determine conflict-free subsets of **B** which can be investigated in parallel. While there might be a conflict for two blocks in one iteration, further subdivision might allow sub-blocks to be parallelized.

At the final subdivision level, each block is the size of a pixel. However, unlike in the fixed block pattern algorithm from the last sec-

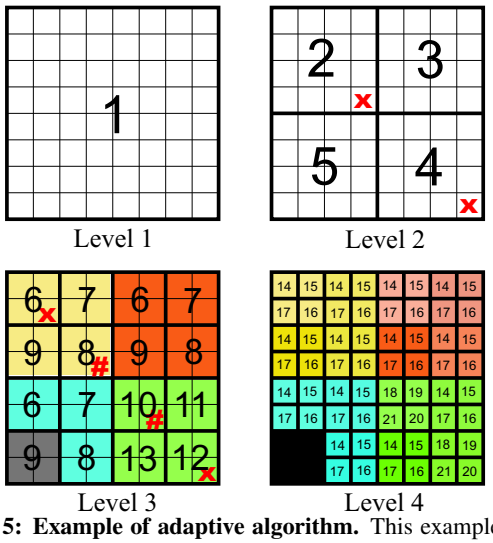


Figure 5: Example of adaptive algorithm. This example shows patterns we would display on an 8×8 pixel projector for each level of subdivision. The numbers in each block indicate the frame when it is lit. In the first frame (level 1), all pixels are on. We subdivide it in level 2 into four children, which are acquired sequentially in clockwise order (frames 2-5). In this example, assume some camera pixels respond to both blocks 2 and 4, e.g. due to inter-reflection within the scene. We denote the conflict between these blocks with red X's. In level 3 (frames 6-13), we ensure that the children of these two blocks are not scheduled for acquisition during the same frame. While acquiring level 3, we discover two additional conflicts: (6 and 12) and (8 and 10). Suppose also that block 9 in the lower left measured nothing, so it is culled. We now schedule level 4, avoiding scheduling the children of conflicting blocks together, which brings us to frame 21. Thus, we can acquire the transport matrix in this example with only 21 frames when 64 would be required with the brute-force scan. Had there been no conflicts, the number of frames would be the number of subdivision levels times four (4 children at each level) plus one for the floodlit image. This gives us $4 \times \log_4(pq) + 1 = 13$ for this example.

tion, we can guarantee that no two projector pixels have a conflict in the camera image because these pixels would not have been scheduled in the same pattern otherwise. Because we know the history of the subdivisions for that pixel, we can exactly determine the correspondence between projector and camera pixels. Using this fact, we can fill in the entries of the **T** matrix with the values measured at the camera. Figure 1 is an example of an image acquired with this technique. Note the capability of the algorithm to capture the complex refraction through the bottle. This image is 578×680 pixels and was acquired in a little over 2 hours. In contrast, a brute-force pixel scan would take almost 11 days at the same resolution (assuming 25 HDR images/min). Figure 6 shows the projection of one of the patterns onto the scene. The large number of white projector pixels highlights the efficient parallelization of the acquisition.

Although the adaptive parallelization algorithm just described works on most scenes, it may perform poorly in scenes where diffuse inter-reflections or subsurface scattering dominates the appearance. These scenes are particularly challenging because the energy emitted by a single projector pixel might be spread over large areas in the scene. In an extreme case, this overlap might cause the algorithm to schedule every pixel of the projector in a separate frame, thereby degenerating into the brute-force scan algorithm.

The adaptive scheme just presented can also fail to capture all of the energy measured by the projector. In certain scenes, a point in

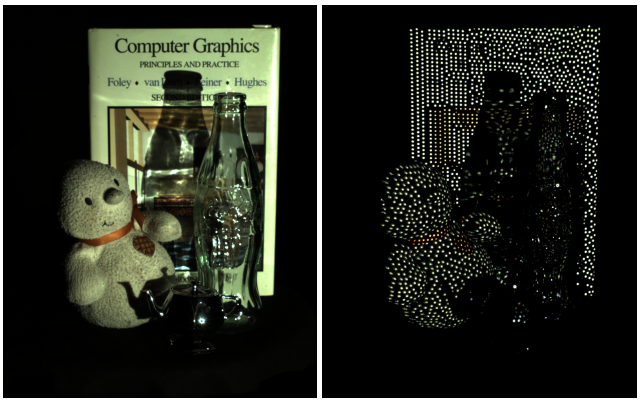


Figure 6: Adaptively parallelized patterns. This figure demonstrates that how our algorithm adapts to the scene content. Because of the complex light transport within the bottle visible in the left image, only a few pixels can be investigated in parallel in this region. Thus, the bottle remains relatively dark when we project an adaptive pattern as shown on the right.

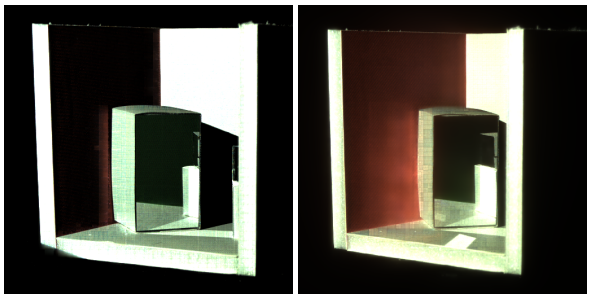


Figure 7: The problem of capturing diffuse inter-reflections. Applying the adaptive parallelization algorithm described in section 3.2 and subdividing down to the pixel level, we produce the dual image on the left. The contribution of the diffuse inflections between the box and the red wall are nearly lost in the camera noise. Using the hierarchical assembly of the transport matrix described in section 3.3, we preserve the energy from higher levels in the subdivision, leading to the improved dual image on the right.

the scene might reflect only a small fraction of its energy towards the camera. If this contribution is below the noise threshold of the camera, some blocks may be erroneously culled and their energy lost. This causes the technique to fail to capture diffuse-diffuse inter-reflections, as shown in Figure 7. We show in the next section that by modifying the adaptive algorithm to store the energy in a hierarchical fashion we can avoid this problem.

3.3 Hierarchical Assembly of the Transport Matrix

To address the problem of signal loss for scenes with significant non-localized light transport, we employ a hierarchical representation of the transport matrix. This method is related to the wavelet environment matting technique by Peers and Dutré [2003]. It is also similar to the hierarchical technique of Matusik et al. [2004] but ours is adaptive. Specifically, we follow the subdivision scheme of the previous section, but build a finer and finer representation of **T** at every subdivision level. At level 1, our approximation of **T** is simply a column vector of length mn representing the image $\mathbf{c}'$ captured while illuminating a floodlit image $\mathbf{p}'$. We call this approximation $\mathbf{T}_1$. Intuitively, $\mathbf{T}_1$ represents the light transport between the camera and a one-pixel projector. At the next level, our approximation $\mathbf{T}_2$ contains four columns, one for each of the four subdivided regions. This continues down to the pixel level where the matrix $\mathbf{T}_k$ with 4^{k-1} columns matches the resolution of the original **T**. The

energy for each element of the matrix is stored at only one level (at the highest possible resolution that still returns a measured value) since we do not want to double-count the energy. For a complete overview of the algorithm, readers are referred to the pseudocode in Appendix B.

The benefit of this hierarchical representation is that the energy is stored at the last level where it can still be accurately measured. A threshold is used to decide whether to subdivide a specific block or to store its contribution at the current level of the hierarchical structure. If the contribution is stored, we terminate subdivision of that block. The threshold is set empirically and depends on the characteristics of the measurement system.

To synthesize the dual image, the contribution of the different levels of the hierarchy need to be added together for each pixel in the final image. This reconstruction can be expressed in the following mathematical form:

$$\mathbf{p}'' = \sum_k f(\mathbf{T}_k{}^T \mathbf{c}'')$$

We will explain the elements of this equation as we describe our rendering algorithm. First, the desired light pattern $\mathbf{c}''$ is applied to the $\mathbf{T}$ approximation at each level to illuminate the scene for that level of the hierarchy. This is expressed by the dot product $\mathbf{T}_k{}^T\mathbf{c}''$ which results in a $4^{k-1} \times 1$ column vector. Referring to equation 2, we see that this vector represents the 4^{k-1}-pixel image that would be viewed at the projector under lighting $\mathbf{c}''$. Figure 14 shows a visualization of $f(\mathbf{T}_k{}^T\mathbf{c}'')$ of levels 1 through 10 for one of our scenes.

We must now add up the energy at each level to generate the final image. Because each of these vectors is of different dimensions, they must be resized to the final resolution of the image in order to be added up. We represent this resizing operation by the function f. The function f is needed to generate a continuous waveform from the discrete samples. In our present implementation, we use bi-cubic interpolation for f. Once the vectors have been resized, we simply add them to get the final image.

The technique presented allows us to efficiently capture the transport matrix $\mathbf{T}$ of a scene and measure many global illumination effects using only a moderate number of patterns and images. Figure 8 shows two more scenes that were acquired using this hierarchical technique. To show that algorithm accelerates our acquisition and results in a manageable size of the $\mathbf{T}$ matrix, we list the relevant data for various scenes in the table below. We compare it against calculated values for a brute-force pixel scan acquisition, assuming a capture rate of approximately 25 patterns/minute. The data is stored as three 32-bit floats for each matrix element. We can see that our technique is several orders of magnitude more efficient in both time and storage space, although further compression is still possible.

	Brute-force scan		**Our adaptive algorithm**		
SCENE	SIZE (TB)	TIME (days)	SIZE (MB)	TIME (min)	# PATTERNS
Fig. 1	5.4	10.9	272	136	3397
Fig. 14	3.7	7.3	179	14	352
Fig. 8 (top)	1.6	8.3	56	19	501
Fig. 8 (bot.)	1.4	8.3	139	15	369
Fig. 11 (80 positions)	114	362	6,675	1,761	19,140

To characterize the effect of projector resolution on our hierarchical adaptive algorithm, we plot the number of acquired frames against projector resolution in Figure 9 for the box scene (Fig. 14) and cover scene (Fig. 1). As we increase the resolution exponentially the curves approximate a straight line. This shows that the adaptive multiplexed illumination approach operates in $O(\log pq)$ time where pq is the projector resolution.

Figure 8: Sample scenes. The acquired primal image is on the left, the synthesized dual on the right. Note for example the detail on the pillar in the dual image of the bottom row which is barely visible in the primal due to foreshortening.

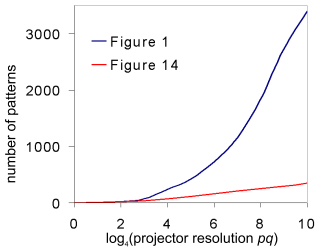


Figure 9: Logarithmic behavior of adaptive algorithm. Plot of the number of required patterns for the scenes in Figure 1 and Figure 14 against the projector resolution reveals that the algorithm operates in $O(\log pq)$ time.

3.4 Image Capture and Pre-processing

Accurate measurement of $\mathbf{T}$ required attention to several physical aspects of our experimental procedure. Since the transport matrix entries we care most about have low energies, high dynamic range image capture [Debevec and Malik 1997; Robertson et al. 1999] was used for all the algorithms described in this section. We used a Canon EOS 20D for the fixed pattern algorithm of Section 3.1 and a Basler A504kc video camera for the adaptive acquisition algorithm of Sections 3.2 and 3.3. The reason for this switch was that while the Canon camera offered excellent resolution, it was too slow to be used in the adaptive scheme. The Basler video camera, on the other hand, had lower resolution but allowed us to capture a full HDR image every 1.5 seconds with an extra second for processing. The scenes were illuminated by a Mitsubishi X480U DLP projector with a native resolution of 1024×768. The shortest exposure time in the HDR sequence had to be slow enough to not alias against the color wheel of the DLP projector, which was determined empirically to be around 30 frames per second.

Before combining the images for HDR, we subtracted the projector's black level from our captured images, which was acquired by photographing the scene while projecting a black image. Subtracting this black-level image from every input image reduced the contribution by stray projector light to our matrix entries, as well as partially compensated for fixed-pattern noise in the camera. Experimentation found the black level to vary slightly with the number of pixels illuminated at a time, but this was not a problem for our acquisitions.

Figure 10: Scene Relighting. A beam of light is virtually projected into the dual image of the scene from Figure 1. Note that specular reflections and caustics from the glass bottle move consistently with the position of the beam.

A second aspect of our measurement procedure that required care was the impact of the Bayer color mosaic in both cameras. These depend on having enough samples at the CCD to be able to properly interpolate the color components from the pixel values. We found that this introduced artifacts when illuminating the scene with individual projector pixels. A focused projector beam can illuminate very few pixels on the camera CCD, yielding errors when the samples are interpolated during demosaicing. When this happened, color contrast was significantly reduced and the images appeared darker than they should.

To remove these darkening and desaturation artifacts, we renormalized the final images by forcing the individual images to add up to the floodlit image. First, all color values of the individual images were summed up so that we could see per-pixel how much of the total energy each image contained. Then the color energy of the floodlit image was distributed to the individual images in proportion to their contribution to the total energy.

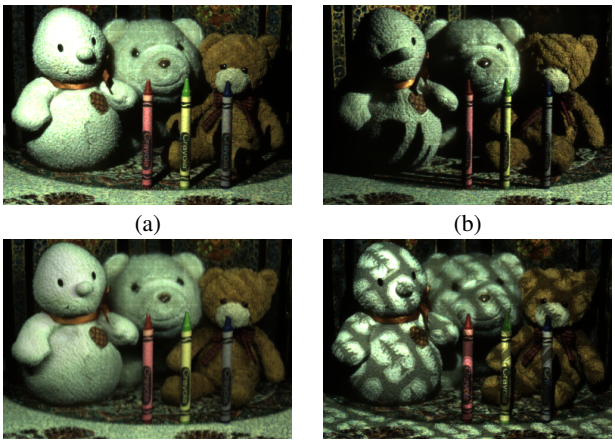
4 Application to Scene Relighting

The transport matrix $\mathbf{T}$ between the projector and camera must be acquired to the resolution of the two devices in order to perform dual photography. This means that we also have the information needed to relight the primal and dual images by multiplying $\mathbf{T}$ and $\mathbf{T}^T$ by the desired illumination vectors $\mathbf{p}'$ and $\mathbf{c}''$ respectively.

In Figure 10, we demonstrate this by relighting the dual image of the scene from Figure 1 with a sweeping plane of light and can see that the caustics generated by the glass bottle vary spatially with the changing illumination. This example of relighting only uses a single projector and a single camera, so the incident lighting in this case is only 2D.

In previous work in relighting (e.g. Masselus et al. [2003]), scenes were relit with incident 4D light fields by acquiring the 6D reflectance function of the scene. They did this by keeping the camera static with respect to the scene and repositioning the projector while doing measurements. This is equivalent to using a single camera and an array of projectors. Dual photography allows us to acquire this 6D reflectance field in the dual domain with a single projector and an array of cameras, which has two advantages. First, because cameras are passive devices, we can take measurements from each of them in parallel without interference. This can significantly accelerate the acquisition of the reflectance field. Second, there are physical and economic advantages of using a camera array versus a projector array. Projectors are generally heavier, larger, and more costly than cameras. They can also be more difficult to pack densely, align, and calibrate.

In Figure 15 we demonstrate the capture of the 6D reflectance field using a mirror array to simulate an array of cameras. A single



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

Figure 11: Scene relit from multiple directions. Using a camera mounted on a computerized gantry, we acquired the transport matrix for this scene at 80 different camera positions. We can then relight the dual image with a light source located at these positions, e.g. (a) one on the left or (b) on the right. By combining the contributions of these lights, we can illuminate the scene with an area light source (c) that casts soft shadows. Finally, because each transport is captured at high resolution, we can relight the scene using a high resolution matte as shown in (d).

Canon EOS 20D camera was aimed into a 4×4 array of planar mirrors, yielding 16 virtual cameras with 800×600 pixels of resolution each (see Figure 15a). The captured transport matrix can be used to relight the dual image as if it were illuminated by up to 16 point light sources with fine angular control (i.e. by 16 virtual projectors). This is sufficient to simulate soft shadows cast from an area light source (Figure 15b). Since the positions of the virtual projectors can be determined through calibration, one can synthesize shadows cast by the virtual actor (in Figure 15c) that interact correctly with the scene. While the full size of the reflectance field for this example would be $4 \times 4 \times 800 \times 600 \times 1024 \times 768$ (66 TB total size), it was acquired in only two hours using the fixed block pattern scan of Section 3.1 that took 144 HDR images. Because the fixed scheme assumes localized light transport which in turn makes the acquired matrix extremely sparse, a simple compression of only storing the $\mathbf{T}$ matrix elements above a certain threshold allowed us to store this data in only 876 MB.

Since our adaptive algorithm is fast, we can also use a single camera to acquire the 6D reflectance field of a scene by moving it like Masselus et al. moved their projector. Figure 11 shows the relighting of a scene which was acquired in this manner by mounting the Basler camera on a computerized gantry.

5 Discussion and Conclusions

In this paper, we introduced the concept of dual photography, whereby the cameras and light sources in a scene can be interchanged due to Helmholtz reciprocity. This allows us to measure the 6D light transport through a scene by using a projector an array of cameras, which is faster than an array of projectors because of the passive nature of cameras. To further accelerate acquisition, we have developed an adaptive algorithm that looks for regions of the scene whose transport paths do not interact. This permits us to project many beams into the scene at once, letting us measure multiple entries of the matrix in parallel.

Once we have acquired the transport matrix, we showed that we can synthesize a view from the projector's location by simply transposing the matrix. We can also relight the scene by multiplying the

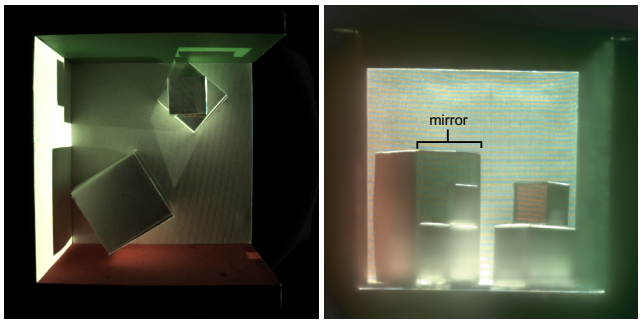


Figure 12: Limits of the hierarchical assembly of the T matrix. In this example, the camera is above the box looking down while the projector is on the right illuminating from the front, as can be seen in the primal image (left). Because of the large difference in angle between the camera and projector viewpoints, there are many regions in the scene where there is no direct light transport. In these regions we are unable to resolve the transport matrix to the full resolution, and thus have to fall back to higher levels of the hierarchical tree. This results in a blurred dual image (right). Nevertheless, it is remarkable that the mirror reflection is captured, even though the mirror is barely visible in the primal image.

transposed matrix by the appropriate vector. This allows us to illuminate a scene with a point light with directional control (for 4D transport) or an arbitrary light field (for 6D transport).

There are some limitations to our technique, however. Scenes with significant global illumination would reduce the parallelism that the adaptive algorithm exploits. Since multiple projector pixels can affect overlapping regions in the camera after several bounces, they would be scheduled in separate passes. In the limit, the technique degenerates to a brute force scan.

Measuring these inter-reflected transport paths accurately is also a problem as discussed in section 3.4, especially when the angle of view of the projector and camera is great with respect to the scene. Figure 12 shows an example of a difficult scene to relight because the camera and projector were at right angles to one another. The blurring problem visible in the dual image might be reduced by using a higher contrast projector and a capturing device with better signal to noise ratio.

To demonstrate that it is still possible to accurately measure diffuse-diffuse reflection, we conducted the experiment shown in Figure 16 where the dual photograph would be recognizable only if this global illumination could be captured accurately. We show that we are still able to extract the signal despite having the light undergo two diffuse bounces before reaching the camera.

At the same time, it might also not always be desirable to substitute a camera by a projector. Cameras have a greater depth of field, ease of focus, and other controls that are often important in imaging. The limited depth of field of a projector could result in dual images that are out of focus. In the scenes that we tested in this paper, however, the depth of field was not a problem due to a reasonably long standoff with respect to scene depth.

The adaptive algorithm accelerates the capture of the transport matrix, but it still takes several minutes to capture the information needed for a high-resolution dual image. We might be able to reduce this time while obtaining a dual image of higher resolution than the projector by placing coaxial projectors and cameras at both the primal and dual positions using beamsplitters. This could allow us to produce a high-resolution view of the scene from the intended observer position under some fixed illumination even though the remainder of the transport matrix could be at lower resolution for relighting. The availability of projectors and cameras on both ends would also mean that both $\mathbf{T}$ and $\mathbf{T}^T$ might be acquired simultaneously, leading to improved acquisition algorithms.

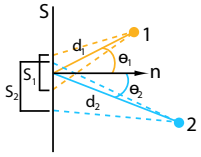
Another extension of this work would be to increase the number of cameras. Since these cameras will be turned into virtual light sources, they need not be of high resolution. This suggests doing dual photography using a large array of inexpensive cameras [Wilburn et al. 2005]. By aiming this array at two different scenes, one can capture the light field emitted by one scene and, flipping the images appropriately, use it to relight a second scene. Alternatively, by applying image flow-based interpolation to the transport matrix, one should be able to move a virtual light smoothly along a path that spans the array in much the same way view-interpolation can be done with camera arrays. This has been an open problem in the relighting community.

Finally, although we focus in this paper on scene relighting, the concept of dual photography is very general, and has other uses. For example, by replacing the non-imaging photosensor in Figure 4 with an imaging spectrophotometer, one can record a spectral power distribution curve for every pixel in the dual image. We also believe that dual photography brings us closer to the goal of measuring the entire 8D reflectance function for a scene, which has never been measured in its entirety due to its overwhelming size.

Acknowledgments We wish to thank Pat Hanrahan for his support during this research. Augusto Román lent us the Basler camera and Vaibhav Vaish offered advice through out the project. Elizabeth Chen made many of the box models shown in the results. Yasamin Mostofi and John Owens proofread drafts of the paper. The animated model for Figure 15 was provided by Mark Kilgard. This project was funded by an NVIDIA fellowship, NSF contract IIS-0219856-001, DARPA contract NBCH-1030009, a Reed-Hodgson Stanford Graduate Fellowship, and a grant from the Max Planck Center for Visual Computing and Communication (BMBF-FKZ 01IMC01). Finally, we are grateful to the anonymous reviewers whose feedback significantly improved the paper.

Appendix A: Proof and Experimental Validation of Dual Photography

Here we prove that the pixel-to-pixel transport from the projector to the camera is the same in both directions. Assume we have a flat surface patch S with arbitrary BRDF f_r viewed/illuminated by a camera/projector pair, as shown in the figure on the right. In the primary configuration, the projector is at point 1 with distance d_1, far enough from the surface so that the rays within a patch can be assumed to be parallel at angle θ_1. Similarly, the camera is at point 2 with distance d_2 and angle θ_2. We call the area illuminated by the single projector pixel S_1, and the region viewed by the camera pixel S_2.



If the projector produces radiant intensity I, the reflected radiance in direction θ_2 from a point in S_1 due to the projector pixel is:

$$L_s = I\frac{f_r(\theta_1 \rightarrow \theta_2)\cos\theta_1}{d_1^2}$$

The irradiance received by the camera pixel at position 2 is the integral of this radiance over the solid angle subtended by the intersection of S_1 and S_2:

$$E_{12} = \frac{L_s|S_1 \cap S_2|\cos\theta_2}{d_2^2} = \frac{I f_r(\theta_1 \rightarrow \theta_2)\cos\theta_1|S_1 \cap S_2|\cos\theta_2}{d_1^2 d_2^2} \quad (3)$$

We specify the area of the intersection of S_1 and S_2 (denoted by $|S_1 \cap S_2|$) because the transfer of energy between the projector and camera pixel only happens in the region of intersection. This defines the transfer of energy between one pixel of the projector in position 1 and one pixel of the camera in position 2. Note that there exists a relationship between the surface area covered by a pixel and its solid angle (Ω), as given by the following equations for the projector and camera respectively:

$$\Omega_p = \frac{|S_1|\cos\theta_1}{d_1^2} \quad \text{and} \quad \Omega_c = \frac{|S_2|\cos\theta_2}{d_2^2}$$

Note that Ω_p and Ω_c are constant for our given projector and camera — they represent the solid angle for the pixel of each device. We now define a projection operator Π:

$$S_1 = \Pi_S^1 \Omega_p \quad \text{and} \quad S_2 = \Pi_S^2 \Omega_c$$

so that S_1 is the projection onto S from position 1 of the solid angle Ω_p, for example. We can now rewrite S_1 and S_2 as:

$$|S_1| = |\Pi_S^1 \Omega_p| = \frac{\Omega_p d_1^2}{\cos\theta_1} \quad \text{and} \quad |S_2| = |\Pi_S^2 \Omega_c| = \frac{\Omega_c d_2^2}{\cos\theta_2} \tag{4}$$

Thus Equation 3 can be rewritten as:

$$E_{12} = \frac{I f_r(\theta_1 \rightarrow \theta_2)\cos\theta_1 |\Pi_S^1 \Omega_p \cap \Pi_S^2 \Omega_c| \cos\theta_2}{d_1^2 d_2^2} \tag{5}$$

We can now see that this equation will be the same in the dual case. The key is to remember that the virtual projector will take on the the camera parameters (in this case the Ω_c) and vice-versa. Thus the transfer of energy in the dual space is given by:

$$E_{21} = \frac{I f_r(\theta_2 \rightarrow \theta_1)\cos\theta_2 |\Pi_S^2 \Omega_c \cap \Pi_S^1 \Omega_p| \cos\theta_1}{d_2^2 d_1^2} \tag{6}$$

Because $f_r(\theta_1 \rightarrow \theta_2) = f_r(\theta_2 \rightarrow \theta_1)$ by Helmholtz reciprocity, we have $E_{12} = E_{21}$. This means that the pixel-to-pixel transport is equal in both directions and we can then perform our technique.

Readers familiar with Zickler's stereopsis [2002] might notice that they have an additional cosine term in their representation. To see where it comes from, we must realize that they physically interchange the projector and the camera when they go into the dual space. Hence, the dual projector and dual camera have the same old parameters Ω_p and Ω_c respectively. This results in the following energy transfer equation in their dual space:

$$E_{21} = \frac{I f_r(\theta_2 \rightarrow \theta_1)\cos\theta_2 |\Pi_S^2 \Omega_p \cap \Pi_S^1 \Omega_c| \cos\theta_1}{d_2^2 d_1^2} \tag{7}$$

First of all, there is no longer a guarantee that $E_{12} = E_{21}$. Furthermore, since the projector is a point light in their configuration, Ω_p is much bigger than Ω_c. Thus for both their primal and dual equations [Zickler et al. 2002] the intersection reduces to just the term for Ω_c:

$$E_{12} = \frac{I f_r(\theta_1 \rightarrow \theta_2)\cos\theta_1 |\Pi_S^2 \Omega_c| \cos\theta_2}{d_1^2 d_2^2}$$

$$E_{21} = \frac{I f_r(\theta_2 \rightarrow \theta_1)\cos\theta_2 |\Pi_S^1 \Omega_c| \cos\theta_1}{d_2^2 d_1^2}$$

If we plug in Equation 4 and cancel out the cosine terms we get:

$$E_{12} = \frac{I f_r(\theta_1 \rightarrow \theta_2)\Omega_c \cos\theta_1}{d_1^2}$$

$$E_{21} = \frac{I f_r(\theta_2 \rightarrow \theta_1)\Omega_c \cos\theta_2}{d_2^2}$$

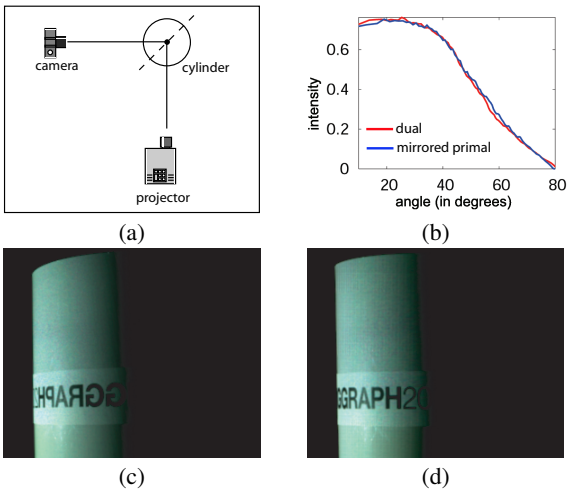


Figure 13: Experimental Validation of Dual Photography. (a) Experimental setup. (b) Plot of intensity vs. surface normal angle for scanlines in the mirrored primal (c) and dual (d) images.

which are equivalent to the equations in the Stereopsis paper, complete with the cosine term to account for the foreshortening of the incident illumination.

We now demonstrate experimentally the radiometric validity of dual photography. The experimental setup is shown in the primal configuration in Figure 13a. The cylinder is mostly diffuse with a slight specular term. The hierarchical adaptive algorithm was used to acquire the **T** matrix for the scene. Because the object is a homogeneous cylinder and the camera and projector are on a plane perpendicular to the axis of the cylinder, the setup is symmetric, meaning that the shading of the cylinder in the primal and dual images should be mirror images of each other. Instead of physically interchanging the camera and projector to compare our synthesized dual with a real photograph from the projector's point of view, this symmetry allows us to simply compare the horizontally-flipped primal image (13c) with our synthesized dual image (13d). For an accurate comparison, we plot (Figure 13b) the intensity averaged over 12 scanlines in the images versus the angle between the normal and the vector between the center of the cylinder and the light source. Because of the proximity of the camera and projector to the cylinder, the range of angles visible by both devices is limited to 10 – 80 degrees. We see that the two curves match.

Appendix B: Pseudocode of Algorithm

In this section we provide pseudocode to the adaptive hierarchical algorithm described in Section 3.3.

```
Initialization() ;
repeat{
    // construct a conflict-free lists of blocks that can be processed in parallel
    ConstructConflictFreeLists() ;
    // illuminate scene with patterns constructed from each list and acquire with camera
    AcquireImages() ;
    // process images, store results, generate new lists of blocks for next iteration
    ProcessResults() ;
} until lowest level in hierarchy is reached

Initialization() {
    for each camera pixel k {
        // Initially assume every camera pixel is affected by block 0, the floodlit image
        B_k = {0};
    }
    C = empty; // initialize set of conflicts to empty
}
```

```
ConstructConflictFreeLists() {
    // form graph structure
    B = Union(B_k);                                  // nodes B, edges C
    (L[0],...,L[N-1]) = GraphColor(graph(B,C));      // N lists of nodes returned
}
AcquireImages() {
    // we now have N conflict-free lists L[]'s
    for i = 0 to N-1 {
        generate pattern P[i] from L[i];             // light pixels for all blocks in L[i]
        illuminate pattern P[i];
        capture HDR image I[i];
    }
}
ProcessResults() {
    C = empty;
    for each camera pixel k {
        new_B_k = empty;
        for i = 0 to N-1 {
            // find block (if any) that affects current pixel
            current_block = intersect(B_k, L[i]);    // because L[i] was conflict-
                                                     // free, this can be at most
                                                     // one block
            if (current_block = empty) {
                continue;                            // pixel k not affected by L[i]
            }
            else {
                if (pixel k in I[i] = 0) {
                    continue;                        // no value measured, do nothing
                }
                else if (pixel k in I[i] < threshold) or last iteration {
                    // below the threshold so store the energy here.
                    // T() is the hierarchical representation of the matrix
                    // indexed by block in the subdivision tree and camera pixel k
                    T(current_block, k) = pixel k in I[i];
                    continue;                        // no further subdivision
                }
                else {
                    // request subdivision for this block
                    insert 4 children of current_block into new_B_k;
                }
            }
        }
        // set B_k for the next iteration
        B_k = new_B_k;
        // collect conflicts and add to C for next iteration
        for each pair (s,t) where s and t are in B_k and s ≠ t {
            insert (s,t) into C;                     // s and t conflict and can't
                                                     // be measured in parallel
        }
    }
}
```

References

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.

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.

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.

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

MATUSIK, W., PFISTER, H., NGAN, A., BEARDSLEY, P., AND MCMILLAN, L. 2002. Image-Based 3D Photography Using Opacity Hulls. In *SIGGRAPH '02*, 427–437.

MATUSIK, W., PFISTER, H., ZIEGLER, R., NGAN, A., AND MCMILLAN, L. 2002. Acquisition and Rendering of Transparent and Refractive Objects. In *Eurographics Symposium on Rendering*, 277–288.

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

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

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.

RUSINKIEWICZ, S., HALL-HOLT, O., AND LEVOY, M. 2002. Real-time 3D Model Acquisition. In *SIGGRAPH '02*, 438–446.

SCHECHNER, Y. Y., NAYAR, S. K., AND BELHUMEUR, P. N. 2003. A Theory of Multiplexed Illumination. In *Proceedings of the IEEE Intl. Conference on Computer Vision (ICCV '03)*, 808–815.

TU, P., MENDONCA, P., ROSS, J., AND MILLER, J. 2003. Surface Registration with a Helmholtz Reciprocity Image Pair. In *IEEE Workshop on Color and Photometric Methods in Computer Vision*.

VEACH, E. 1997. *Robust Monte Carlo Methods for Light Transport Simulation*. PhD thesis, Stanford University.

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`.

WHITTED, T. 1980. An Improved Illumination Model for Shaded Display. *Commun. ACM 23*, 6, 343–349.

WILBURN, B., JOSHI, N., VAISH, V., TALVALA, E.-V., ANTUNEZ, E., BARTH, A., ADAMS, A., LEVOY, M., AND HOROWITZ, M. 2005. High Performance Imaging using Large Camera Arrays. In *SIGGRAPH '05*.

ZICKLER, T., BELHUMEUR, P., AND KRIEGMAN, D. 2002. Helmholtz Stereopsis: Exploiting Reciprocity for Surface Reconstruction. *Intl. Journal of Computer Vision 49*, 2–3, 215–227.

ZONGKER, D. E., WERNER, D. M., CURLESS, B., AND SALESIN, D. H. 1999. Environment Matting and Compositing. In *SIGGRAPH '99*, 205–214.

ZOTKIN, D., DURAISWAMI, R., GRASSI, E., AND GUMEROV, N. 2004. Fast Head Related Transfer Function Measurement via Reciprocity. Technical Report No. CS-4620, University of Maryland.

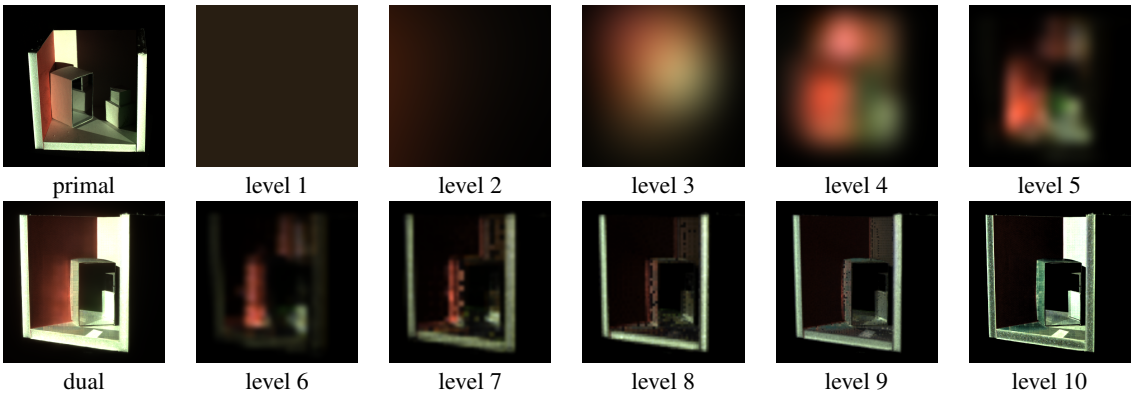


Figure 14: Construction of the dual image with a hierarchical representation. The primal and dual image show diffuse to diffuse inter-reflections which could only be captured by use of the hierarchical acquisition. Energy that might have been lost when further subdividing a block is deposited at a coarse level of the **T** matrix. To synthesize the dual image, the levels are individually reconstructed by applying the appropriate basis functions, then added together to obtain the image on the left. In this figure the intensity of the images for level 1 to 9 has been increased to visualize their contribution.

(a) (b) (c)

Figure 15: Using a mirror array to emulate multiple virtual projectors. (a) A camera was imaged onto the mirror array in order to emulate multiple virtual light positions. A block pattern scan consisting of 144 high dynamic range images was performed to acquire the scene's transport matrix. The region of interest within the projector's field of view was 864×604 pixels, the final resolution of the dual image. Each camera in the mirror array had an approximate resolution of 800×600 pixels, which is the resulting spatial resolution of our virtual lights. (b) The scene is illuminated by 12 point light sources to create soft shadows. (c) An animated character is embedded in the scene and casts shadows onto the scene.

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

Figure 16: Dual photography with indirect light transport. (a) A projector illuminates the front of a playing card while the camera sees only the back of the card and the diffuse page of the book. An aperture in front of the projector limits the illumination only onto the card. The card was adjusted so that its specular lobe from the projector did not land on the book. Thus, the only light that reached the camera underwent a diffuse bounce at the card and another at the book. (b) Complete camera view under room lighting. The back of the card and the page of the book are visible. It seems impossible to determine the identity of the card from this point of view simply by varying the incident illumination. To acquire the transport matrix, a 3×3 white pixel was scanned by the projector and 5742 images were acquired to produce a dual image of resolution 66×87. (c) Sample images acquired when the projector scanned the indicated points on the card. The dark level has been subtracted and the images gamma-corrected to amplify the contrast. We see that the diffuse reflection changes depending on the color of the card at the point of illumination. After acquiring the T matrix in this manner, we can reconstruct the floodlit dual image (d). It shows the playing card from the perspective of the projector being indirectly lit by the camera. No contrast enhancement has been applied. Note that the resulting image has been automatically antialiased over the area of each projector pixel.