A Theory of Inverse Light Transport

Steven M. Seitz
University of Washington

Yasuyuki Matsushita
Microsoft Research Asia

Kiriakos N. Kutulakos*
University of Toronto

Abstract

In this paper we consider the problem of computing and removing interreflections in photographs of real scenes. Towards this end, we introduce the problem of *inverse light transport*—given a photograph of an unknown scene, decompose it into a sum of n-bounce images, where each image records the contribution of light that bounces exactly n times before reaching the camera. We prove the existence of a set of *interreflection cancelation operators* that enable computing each n-bounce image by multiplying the photograph by a matrix. This matrix is derived from a set of "impulse images" obtained by probing the scene with a narrow beam of light. The operators work under unknown and arbitrary illumination, and exist for scenes that have arbitrary spatially-varying BRDFs. We derive a closed-form expression for these operators in the Lambertian case and present experiments with textured and untextured Lambertian scenes that confirm our theory's predictions.

1 Introduction

Modeling *light transport*, the propagation of light through a known 3D environment, is a well-studied problem in computer graphics. However, the inverse light transport problem—using photos of an unknown environment to infer how light propagates through it—is wide open.

Modeling inverse light transport enables reasoning about shadows and interreflections–two major unsolved problems in computer vision. Aside from the interesting theoretical questions of what about these effects can be inferred from photographs, understanding shadows and interreflections has major practical importance, as they can account for a significant percentage of light in a photograph. Modeling inverse light transport can greatly expand the applicability of a host of computer vision techniques, i.e., photometric stereo, shape from shading, BRDF measurement, etc., that are otherwise limited to convex objects that do not interreflect light or cast shadows onto themselves.

The intensities recorded in an image are the result of a complex sequence of reflections and interreflections, as light emitted from a source will bounce off of the scene's surfaces one or more times before it reaches the camera. In
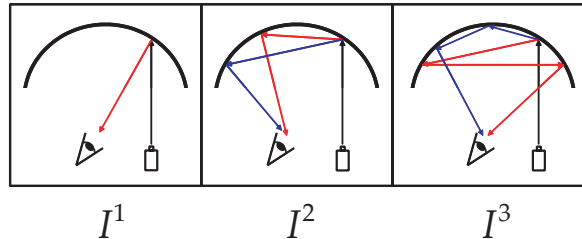

Figure 1: An n-bounce image I^n records the light that bounces exactly n times before reaching the camera.

theory, therefore, every image can be thought of as an infinite sum, $I = I^1 + I^2 + I^3 + \ldots$, where I^n records the contribution of light that bounces exactly n times before reaching the camera (Figure 1). For instance, I^1 is the image we would capture if it were possible to block all indirect illumination from reaching the camera, while the infinite sum $I^2 + I^3 + \ldots$ represents the total contribution of *indirect illumination*. While we can capture the composite image I using a camera, the individual "n-bounce" images are not directly measurable.

In this paper we prove the existence of a set of linear operators that compute the entire sequence of n-bounce images in a photograph of an unknown scene captured under unknown illumination. These operators, which we call *interreflection cancellation operators*, exist under very general conditions—the scene can have arbitrary shape and arbitrary BRDF and be illuminated by an arbitrary illumination field. Moreover, we show that in the special case of scenes with Lambertian reflectance, we can compute these operators by first collecting a sequence of images in which the scene is "probed" with a narrow beam of light that slides over the surface. We emphasize that our procedure requires acquiring *multiple* images of a scene as a preprocessing step in order to cancel interreflected light from any new photograph. The cancellation process is performed simply by multiplying the new photograph by a matrix derived from the pre-acquired images.

By removing the effects of interreflections, these operators can be used to convert photographs into a form more amenable to processing using existing vision algorithms, since many techniques do not account for interreflections.

Techniques for simulating interreflections, and other light transport effects are well-known in the graphics com-

munity (e.g., work on ray tracing [1, 2] and radiosity [1, 3]), and have been studied in the vision community as well [4, 5]. However, relatively little is known about the inverse problem of modeling the effects of interreflections, in images of real scenes. A number of authors [6, 7, 8] have recently proposed methods to capture the *forward* light transport function from numerous photographs taken under different illuminations, implicitly taking into account the global effects of interreflections, shadows, sub-surface scatting, etc. Yu *et al.* [9] and Machida *et al.* [10] describe *inverse global illumination* techniques that model diffuse interreflections to recover diffuse and specular material properties from photographs. In both methods, the geometry and lighting distribution is assumed to be known a priori. None of these methods provides a means, however, for measuring and analyzing the effects of interreflections, in an image where the scene's shape and the incoming light's distribution are both unknown.

Although the focus of our work is cancelling interreflections, it is closely related to the shape-from-interreflections, problem. This problem has received very limited attention, focusing on the case of Lambertian scenes [11, 12] and of color bleeding effects between two differently-colored facets [13, 14, 15, 16]. Closest to our analysis is the inspiring work of Nayar *et al.* [11], who demonstrated an iterative photometric stereo algorithm that accounts for interreflections. Our analysis differs in a number of interesting ways from that work. Whereas [11] assumed uniform directional lighting, we place no restriction on the illumination field. Moreover, in contrast to the iterative approach in [11] for estimating the light transport equations and analyzing interreflections, we derive closed-form expressions for the interreflection cancellation operators, that need be applied only once to an image. On the other hand, a disadvantage of our approach compared to Nayar's is that we require many more input images to model the transport process.

Our approach is based on two key insights. The first one, already known in the graphics literature, is that the forward propagation of light can be described using linear equations. We use this fact to show that the mapping from an arbitrary input image to each n-bounce component can be expressed as a matrix multiplication. The second is that in the case of a Lambertian scene, we can compute this matrix from a set of images in which an individual scene point is illuminated by a narrow beam of light. Intuitively, each such image can be thought of as a "impulse response" that tells us how light that hits a single scene point contributes to the indirect illumination of the rest of the scene.

We start by proving the existence of linear cancellation operators that compute the n-bounce images under very general conditions (Section 2). For the case of Lambertian scenes and a fixed camera viewpoint, we then show how to construct these operators from a set of input images (Section 3). Section 4 presents experimental results.

2 Inverting Light Transport

The appearance of a scene can be described as a light field $L_{out}(\omega_\mathbf{x}^\mathbf{y})$, representing radiance as a function of outgoing ray $\omega_\mathbf{x}^\mathbf{y}$ [17, 18, 19]. While various ray representations are possible, we specify a ray $\omega_\mathbf{x}^\mathbf{y}$ by its 3D point of origin $\mathbf{x}$ and direction from $\mathbf{x}$ to another point $\mathbf{y}$. We restrict $\mathbf{x}$ to lie on scene surfaces. Since the set of rays that touch the scene is 4D (assuming the scene is composed of a finite set of 2D surfaces), we may speak of a light field as being a 4D function. To simplify our analysis, we consider light of a single wavelength.[1] Scene illumination may also be described by a light field $L_{in}(\omega_{\mathbf{x}'}^\mathbf{x})$ describing the light emitted from source points $\mathbf{x}'$ to surface points $\mathbf{x}$. We use the terms *outgoing light field* and *incident light field* to refer to L_{out} and L_{in}, respectively.

An outgoing light field L_{out} is formed by light from one or more emitters that hits objects in the scene and gets reflected. Some of this reflected light hits other objects, which in turn hits other objects, and the process continues until an equilibrium is reached. We can therefore think of L_{out} as being composed of two components: light that has bounced a single time off of a scene point (*direct* illumination), and light that has bounced two or more times (*indirect* illumination), i.e.,

$$L_{out}(\omega_\mathbf{x}^\mathbf{y}) = L_{out}^1(\omega_\mathbf{x}^\mathbf{y}) + L_{out}^{2,3,\cdots}(\omega_\mathbf{x}^\mathbf{y}) . \qquad (1)$$

The first component, $L_{out}^1(\omega_\mathbf{x}^\mathbf{y})$, is determined by the BRDF of $\mathbf{x}$ and the emitters that illuminate $\mathbf{x}$. The second component depends on interreflected light that hits $\mathbf{x}$ from other points in the scene. Eq. (1) can be expressed as an integral equation, known as the *light transport equation* or the *rendering equation* [20, 3] in the computer graphics literature:

$$L_{out}(\omega_\mathbf{x}^\mathbf{y}) = L_{out}^1(\omega_\mathbf{x}^\mathbf{y}) + \int_{\mathbf{x}'} A(\omega_\mathbf{x}^\mathbf{y}, \omega_{\mathbf{x}'}^\mathbf{x}) L_{out}(\omega_{\mathbf{x}'}^\mathbf{x}) d\mathbf{x}' .$$
$$(2)$$

The function $A(\omega_\mathbf{x}^\mathbf{y}, \omega_{\mathbf{x}'}^\mathbf{x})$ defines the proportion of irradiance from point $\mathbf{x}'$ to $\mathbf{x}$ that gets transported as radiance towards $\mathbf{y}$. As such, it is a function of the scene's BRDF, the relative visibility of $\mathbf{x}$ and $\mathbf{x}'$ and of light attenuation effects [20, 3].[2] When $\mathbf{x} = \mathbf{x}'$, $A(\omega_\mathbf{x}^\mathbf{y}, \omega_{\mathbf{x}'}^\mathbf{x})$ is 0.

If we assume that the scene is composed of a collection of small planar facets and if we discretize the space of rays, Eq. (2) can be expressed in a discrete form [3, 11] as

$$\mathbf{L}_{out}[i] = \mathbf{L}_{out}^1[i] + \sum_j \mathbf{A}[i, j] \mathbf{L}_{out}[j] , \qquad (3)$$

where $\mathbf{L}_{out}$ is a discrete light field represented by a set of sampled rays, $\mathbf{L}_{out}^1$ is the corresponding 1-bounce light field, and $\mathbf{A}[i, i] = 0$. Rewriting Eq. (3) as a matrix equation yields

$$\mathbf{L}_{out} = \mathbf{L}_{out}^1 + \mathbf{A}\mathbf{L}_{out} . \qquad (4)$$

---

[1]Color may be treated by considering each wavelength separately.
[2]Formally, Eq. (2) interprets the light reflected by $\mathbf{x}$ from external light sources as if it came directly from $\mathbf{x}$, i.e., $\mathbf{x}$ is treated as an emitter.

Eq. (4) defines, in a discrete form, how light is transported through a scene. A direct solution is obtained by solving Eq. (4) for $\mathbf{L}_{out}$, obtaining [20, 3]

$$\mathbf{L}_{out} = (\mathbf{I} - \mathbf{A})^{-1} \mathbf{L}_{out}^1 . \tag{5}$$

This equation, well known in the graphics community, shows that the global effects of light transport can be modeled by a linear operator $(\mathbf{I} - \mathbf{A})^{-1}$ that maps a light field containing only direct illumination to a light field that takes into account both direct *and* indirect illumination.

2.1 Cancelling Interreflections

Consider the following operator:

$$\mathbf{C}^1 \stackrel{\text{def}}{=} \mathbf{I} - \mathbf{A} . \tag{6}$$

From Eq. (5), it follows that

$$\mathbf{L}_{out}^1 = \mathbf{C}^1 \mathbf{L}_{out} . \tag{7}$$

It is therefore possible to cancel the effects of interreflections in a light field by multiplying with a matrix $\mathbf{C}^1$. We call $\mathbf{C}^1$ an *interreflection cancellation operator*. Hence $\mathbf{C}^1$ exists for general BRDFs, and is linear. Note that this result is a trivial consequence of the rendering equation—while we do not believe the cancellation operator has been exploited previously in computer vision, its existence is implicit in the derivations of forward light transport [20, 3, 11].

Even though Eq. (7) provides an existence proof of an interreflection cancellation operator, $\mathbf{C}^1$ is defined in terms of shape and BRDF quantities (contained in the entries of $\mathbf{A}$) instead of image quantities. We now provide an interpretation in terms of image quantities, as follows.

Consider emitting unit radiance along ray i towards the scene (e.g., using a laser beam or a projector). The resulting light field, which we denote $\mathbf{t}_i$, captures the full transport of light in response to an *impulse* illumination. We call $\mathbf{t}_i$ an Impulse Scatter Function (ISF) .[3] Now concatenate all the ISFs into an ISF matrix $\mathbf{T}$:

$$\mathbf{T} \stackrel{\text{def}}{=} [\mathbf{t}_1 \ \mathbf{t}_2 \ \dots \ \mathbf{t}_m] . \tag{8}$$

Because $\mathbf{T}$ is made up of ISFs , it is possible in principle to measure $\mathbf{T}$ in the laboratory using controlled illumination. Although capturing a full dense set of ISFs would be extremely time- and storage-intensive, previous authors have explored the problem of capturing two-dimensional forms of $\mathbf{T}$ [6, 7, 8].

Because light is linear, any light field $\mathbf{L}_{out}$ can be described as a linear combination of ISFs , enabling applications such as synthetic relighting of scenes [6]. In particular, we can express any outgoing light field as a function of the illumination $\mathbf{L}_{in}$ by

$$\mathbf{L}_{out} = \mathbf{T} \mathbf{L}_{in} . \tag{9}$$

---

[3]It has also been referred to as the *impulse response* in [7].

Applying the cancellation operator to an ISF yields

$$\mathbf{t}_i^1 \stackrel{\text{def}}{=} \mathbf{C}^1 \mathbf{t}_i , \tag{10}$$

where $\mathbf{t}_i^1$ is the component of $\mathbf{t}_i$ due to 1-bounce reflection. By defining $\mathbf{T}^1 = [\mathbf{t}_1^1 \ \mathbf{t}_2^1 \ \dots \ \mathbf{t}_m^1]$ we get the matrix equation

$$\mathbf{T}^1 = \mathbf{C}^1 \mathbf{T} \tag{11}$$

and therefore

$$\mathbf{C}^1 = \mathbf{T}^1 \mathbf{T}^{-1} . \tag{12}$$

Eq. (12) provides an alternative definition of the cancellation operator $\mathbf{C}^1$ in terms of image quantities. Intuitively, applying $\mathbf{C}^1$ to a light field $\mathbf{L}_{out}$ has the effect of first computing the scene illumination field ($\mathbf{L}_{in} = \mathbf{T}^{-1} \mathbf{L}_{out}$) and then re-rendering the scene with a 1-bounce model ($\mathbf{L}_{out}^1 = \mathbf{T}^1 \mathbf{L}_{in}$).

Note that while $\mathbf{T}$ is measurable, $\mathbf{T}^1$ is generally not. Hence, the derivations in this section provide only an existence proof for $\mathbf{C}^1$. Also note that Eq. (12) is valid only when $\mathbf{T}$ is invertible. $\mathbf{C}^1$ is well-defined even when $\mathbf{T}$ is not invertible, however, since Eqs. (6) and (7) always hold.

2.2 *N*-bounce Light Fields

Suppose we wanted to compute the contribution of light due to the *second* bounce of reflection. More precisely, suppose light from the light source first hits a point $\mathbf{p}$, then bounces to point $\mathbf{q}$, and then is reflected toward the camera (see image I^2 of Figure 1). How can we measure the contribution of this light to the image intensity at $\mathbf{q}$'s projection? While this problem has a straightforward solution when the scene's shape, BRDF, and illumination are known [1], we are not aware of any solutions to this problem for unknown shapes and illuminations. Beyond removing interreflections, $\mathbf{C}^1$ offers a simple way of solving this problem.

In particular, given a light field $\mathbf{L}_{out}$ of the scene, the portion of light due purely to interreflections is given by $\mathbf{L}_{out} - \mathbf{C}^1 \mathbf{L}_{out}$. Since the direct illumination has been removed, we can treat the indirect illumination coming from each visible point $\mathbf{p}$ as if it were coming directly from a light source located at $\mathbf{p}$. Hence, the light field $\mathbf{C}^1(\mathbf{L}_{out} - \mathbf{C}^1 \mathbf{L}_{out})$ is the component of $\mathbf{L}_{out}$ that is due to the second bounce of light. More generally, the nth-order interreflection cancellation operator and the n-bounce light field are given by

$$\mathbf{C}^n \stackrel{\text{def}}{=} \mathbf{C}^1 (\mathbf{I} - \mathbf{C}^1)^{n-1} \text{ and}$$
$$\mathbf{L}_{out}^n \stackrel{\text{def}}{=} \mathbf{C}^n \mathbf{L}_{out} ,$$

where $\mathbf{L}_{out}^n$ defines a light field due to the nth bounce of light. This light has hit exactly n scene points between the light source and the camera. We can therefore "unroll" the individual steps of light transport as it propagates through

the scene by expressing the outgoing light field $\mathbf{L}_{out}$ as a sum of individual n-bounce light fields:

$$\mathbf{L}_{out} = \sum_{n=1}^{\infty} \mathbf{L}_{out}^n.$$

3 The Lambertian Case

While the results in Section 2 place no restriction on the BRDF or range of camera viewpoints, they provide only an existence proof of inverse light transport operators. We now turn to the problem of computing $\mathbf{C}^i$. To this end, we show that if the scene is Lambertian, we can compute these operators from images taken at a single viewpoint.

A defining property of Lambertian scenes is that each point radiates light equally in a hemisphere of directions. We may therefore reduce the 4D outgoing light field to a 2D set of outgoing beams, one for each point on the surface. A second property of Lambertian scenes is that if we illuminate a point from two different directions, the pattern of outgoing radiance (and hence interreflections) is the same up to a scale factor. We may therefore reduce the 4D incident light field $\mathbf{L}_{in}$ to a 2D subset.

Because the incident and outgoing light fields are both 2D for the Lambertian case, we can capture an ISF matrix by scanning a narrow beam of light over the surface and capturing an image from a fixed camera for each position of the beam. Each ISF $\mathbf{t}_i$ is constructed by concatenating the pixels from the i-th image into a vector and normalizing to obtain a unit-length vector (thereby eliminating the scale dependence on incoming beam direction).

We now assume without loss of generality that there is a one-to-one correspondence between m scene points, m image pixels, and m incident light beams, i.e., incident beam i hits scene point i which is imaged at pixel i.[4]

We assume that only points which reflect light (i.e., have positive albedo) are included among the m scene points. Finally, to simplify presentation, we assume that all points that contribute reflected light to the image, direct or indirect, are included among the m points. We relax this last assumption in Section 3.1.2.

Our formulation of the Lambertian ISF matrix, together with the interreflection equations, leads to a closed-form and computable expression for the interreflection operator:

Lambertian Interreflection Cancellation Theorem: Each view of a Lambertian scene defines a unique $m \times m$ interreflection cancellation operator, $\mathbf{C}^1$, given by the expression

$$\mathbf{C}^1 = \mathbf{T}^1 \mathbf{T}^{-1}, \qquad (13)$$

where $\mathbf{T}^1$ is a diagonal $m \times m$ matrix containing the reciprocals of the diagonal elements of $\mathbf{T}^{-1}$.

---

[4]This is achieved by including only beams that fall on visible scene points and removing pixels that are not illuminated by any beam.

Proof: From Eqs. (6) and (12) we have

$$\mathbf{C}^1 = \mathbf{T}^1 \mathbf{T}^{-1},$$

where $\mathbf{C}^1[i, i] = 1$, and $\mathbf{C}^1[i, j] = -\mathbf{A}[i, j]$. Since only one point $\mathbf{p}_i$ is illuminated in each ISF and $\mathbf{p}_i$ appears only at pixel i, it follows that $\mathbf{T}^1$ must be a diagonal matrix. Since $\mathbf{T}^1$ is diagonal and $\mathbf{C}^1$ has ones along the diagonal, it follows that

$$\mathbf{T}^1[i, i] = \frac{1}{\mathbf{T}^{-1}[i, i]} .$$

QED

This closed-form expression for $\mathbf{C}^1$ provides explicit information about surface reflectance and relative visibility. Specifically, from [11] we obtain

$$\mathbf{C}^1 = \mathbf{I} - \mathbf{PK} , \qquad (14)$$

where $\mathbf{P}$ is a diagonal matrix with $\mathbf{P}[i, i]$ specifying the albedo for point i divided by π, and $\mathbf{K}$ is the matrix of diffuse *form factors* [21]. Moreover, Eq. (14) implies that our method can handle variation in surface texture.

Since $\mathbf{T}$ is formed from impulse images, its diagonal elements correspond to points that were directly illuminated by the laser. Hence, the diagonal of $\mathbf{T}$ dominates the other elements of the matrix. In practice, we have found that $\mathbf{T}$ is well conditioned and easy to invert.

3.1 Practical Consequences

3.1.1 General 4D incident lighting

A key property of the Lambertian cancellation operator $\mathbf{C}^1$ is that it cancels interreflections for the given viewpoint under **any unknown 4D incident light field**. Because the space of ISFs for the Lambertian case is 2D and not 4D, it follows that any outgoing light field $\mathbf{L}_{out}$ (under any 4D incident illumination) can be expressed as a linear combination of 2D ISFs . The coefficients of this linear combination determine an equivalent 2D illumination field (along the m rays defined by the captured ISFs) that produces a light field identical to $\mathbf{L}_{out}$.

It follows that $\mathbf{C}^1$ will correctly cancel interreflections even in very challenging cases, including illumination from a flashlight or other non-uniform area source; from a video projector sending an arbitrary image; or from an unknown surrounding environment. See Section 4 for demonstrations of some of these effects.

3.1.2 Interactions with occluded points

To facilitate our proof of the interreflection theorem, we assumed that an ISF was captured for every point that contributes reflected light to the image (direct or indirect). Stated more informally, every point that plays a role in the light transport process must be visible in the image. This assumption is not as restrictive as it sounds because our formulation allows for multi-perspective images—for instance, we can create an "image" by choosing, for every
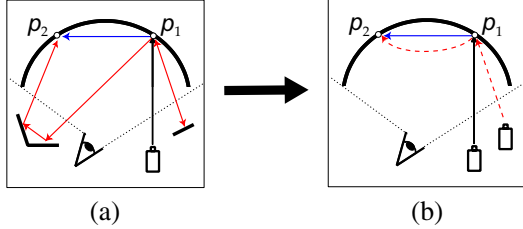
(a)   (b)

Figure 2: (a) Light hitting $\mathbf{p}_1$ interreflects off of occluded points that are not within the field of view of the camera (dotted lines), causing additional light to hit $\mathbf{p}_1$ and $\mathbf{p}_2$. (b) An alternative explanation is that there are no occluded points, but additional light flowed from $\mathbf{p}_1$ directly to $\mathbf{p}_2$, and from external light sources to $\mathbf{p}_1$ (dashed lines).

point on the surface of the object, a ray that ensures the point is visible.

In practice, it may be difficult to capture an ISF for every point, and it is convenient to work with single-perspective images that contain occlusions. It is therefore important to consider what happens in the case of interactions with occluded points, i.e., points that are not represented in the columns of $\mathbf{T}$. Fortunately, the cancellation theorem also applies to such cases because of the following observation. Suppose that light along a beam from a visible point $\mathbf{p}_i$ bounces off of k invisible points before hitting the first visible point $\mathbf{p}_j$ (Figure 2a). We can construct a different interpretation of the same image that does not involve invisible points by treating all of this light as if it went directly from $\mathbf{p}_i$ to $\mathbf{p}_j$, i.e., the k bounces are *collapsed* into a single bounce (Figure 2b). It is easy to see that the transport equation, Eq. (3), still applies—there is just more light flowing between $\mathbf{p}_i$ and $\mathbf{p}_j$. The collapsed rays have the effect of (1) increasing $\mathbf{A}[i,j]$ to take into account the additional light that flows "directly" from $\mathbf{p}_i$ to $\mathbf{p}_j$ and (2) increasing the apparent amount of "direct" illumination $\mathbf{L}_{out}^1[i]$ and, hence, $\mathbf{t}_i^1$.

It follows that $\mathbf{C}^1$ applies as before, with the modification that light which $\mathbf{p}_i$ sends to itself via any number of intermediate bounces off of invisible surfaces is treated as *direct* illumination and therefore not cancelled. Similarly, $\mathbf{C}^2$ will not cancel light that $\mathbf{p}_i$ sends to $\mathbf{p}_j$ through any number of intermediate bounces off of invisible surfaces.

4 Experimental Results

In order to confirm our theoretical results, we performed experiments with both real and synthetic scenes. Our emphasis was on computing cancellation operators and n-bounce images, and comparing them to ground-truth from simulations.

Our experimental system consisted of a Canon EOS-20D digital SLR camera, a 50mW collimated green laser source, and a computer-controlled pan-tilt unit for directing the laser's beam. To minimize laser speckle, we used a Canon wide-aperture, fixed focal length (85mm) lens and acquired images using the widest possible aperture, $F1.2$.
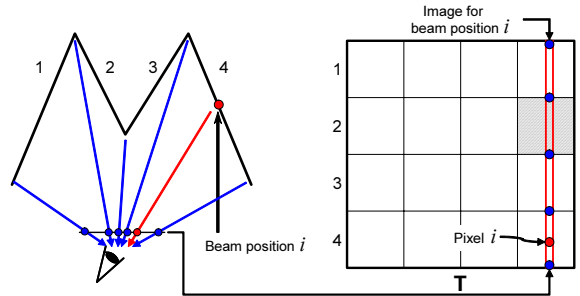

Figure 3: Graphical representation of the ISF matrix for the synthetic "M" scene. The image captured at the i-th beam position becomes the i-th column of the ISF matrix $\mathbf{T}$. The pixel receiving direct illumination (position i) is mapped to element $\mathbf{T}(i,i)$. Because the scene contains four facets, $\mathbf{T}$ has $4^2 = 16$ "blocks:" the block shown in gray, for instance, describes the appearance of points on facet 2 when the beam is illuminating a point on facet 4.
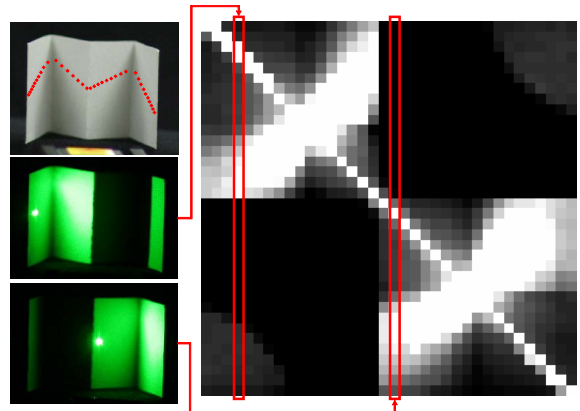

Figure 4: *Left-top:* image of the "M" scene. Spots in top-left image indicate the pixels receiving direct illumination as the beam panned from left to right. *Left-middle:* One of the captured images, corresponding to the 4th position of the laser beam. The 4th column of the ISF was created by collecting the pixels indicated by the spots at top-left and assembling them into a column vector. *Left-bottom:* image of the scene corresponding to another beam position. *Right:* The complete ISF matrix.

To exploit the extreme dynamic range offered by the laser (over 145dB), we acquired HDR images with exposures that spanned almost the entire range of available shutter speeds—from $1/15s$ to $1/8000s$ with an additional image with an aperture of $F16$ and a speed of $1/8000$. All images were linearized with Mitsunaga *et al*'s radiometric calibration procedure [22]. We used only the green channel by demosaicing the raw CCD data.

To capture a scene's ISF, we moved the laser beam to a predetermined set of m directions and captured a high-dynamic range image for each direction. Since our analysis assumes a known one-to-one correspondence between the m illuminated scene points and the pixels they project to, we first determined the pixel that received direct illumination. To do this, we found the centroid of the laser spot in the shortest-exposure image for each beam direction.These
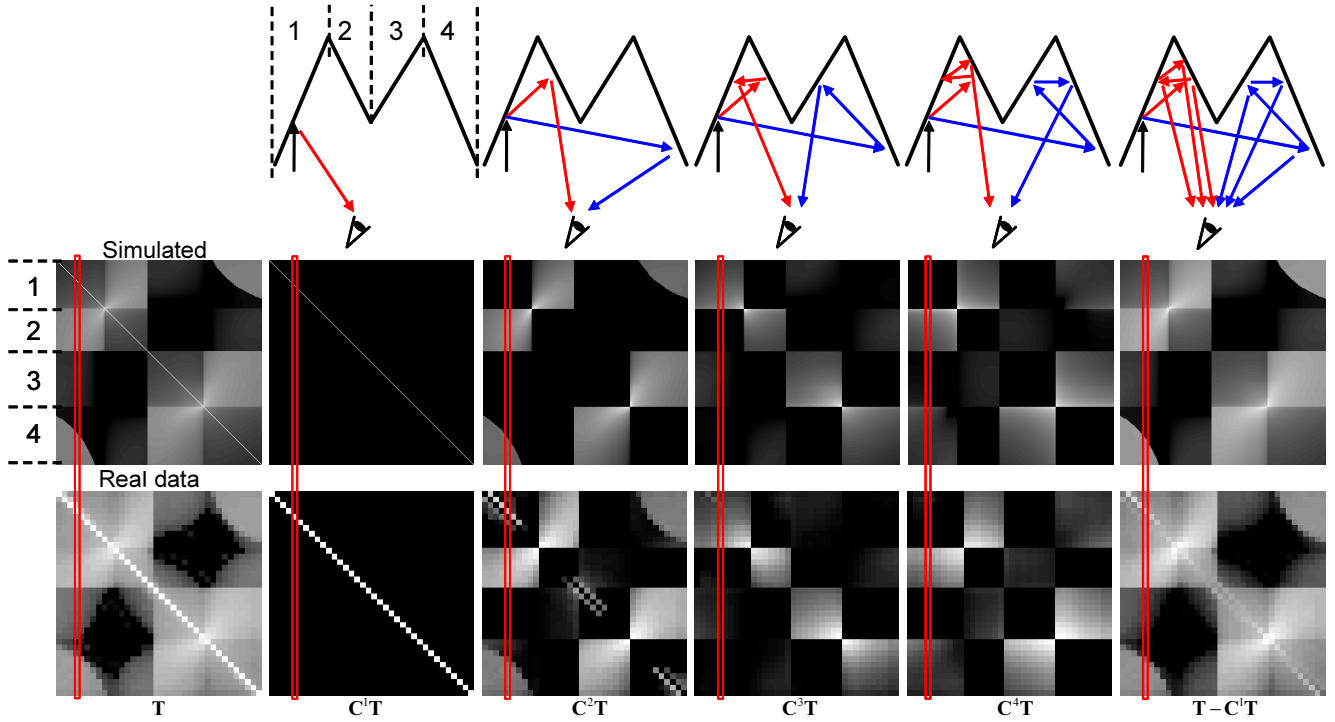
Figure 5: Inverse light transport computed from the ISF matrix **T** for the "M" scene. Top row shows typical light paths, middle row shows simulation results, and bottom row shows results from real-world data. *From left to right:* the ISF matrix, direct illumination (1-bounce), the 2-, 3- and 4-bounce interreflection components of the ISF , and total indirect illumination. Images are log-scaled for display purposes.
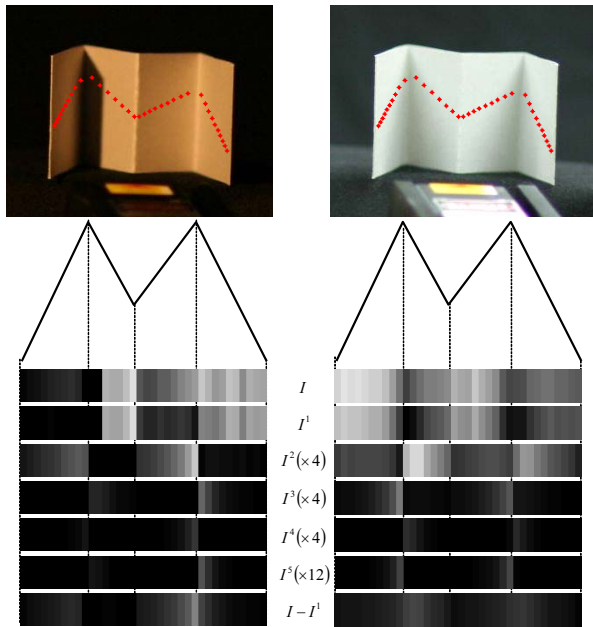

Figure 6: Inverse light transport applied to images I captured under unknown illumination conditions. I is decomposed into direct illumination I^1 and subsequent n-bounce images I^n, as shown. Observe that the interreflections have the effect of increasing brightness in concave (but not convex) junctions of the "M". Image intensities are scaled linearly, as indicated.

centroids were used to sample all m input images. Hence, each image provided m intensities, corresponding to a column of the $m \times m$ ISF matrix.

Computing k-bounce operators in 1D As the first experiment, we simulated a scene containing a Lambertian surface shaped like the letter "M." To build its ISF matrix, we simulated an incident beam that moves from left to right (Figure 3). For each position of the beam, we rendered an image by forward-tracing the light as it bounces from facet to facet, for up to 7 bounces. These images formed the individual columns of the ISF matrix. At the same time, our renderer computed the contribution of individual light bounces in order to obtain the "ground truth" n-bounce images for the scene.

The top row of Figure 5 shows the result of decomposing the simulated ISF matrix into direct, 2- through 4-bounce, and indirect illumination components. To assess this decomposition qualitatively, consider how light will propagate after it hits a specific facet. Suppose, for instance, that we illuminate a point on the scene's first facet (red column of the ISF matrix in Figure 5). Since only one point receives direct illumination, cancelling the interreflections in that column should produce an image with a single non-zero value located at the point's projection. The actual cancellation result, indicated by the red column in $\mathbf{C}^1\mathbf{T}$, matches this prediction. Now, light reflected from facet 1 can only illuminate facets 2 and 4. This implies that the 2-bounce image should contain non-zero responses only for points on

those two facets. Again, application of the 2-bounce operator, $\mathbf{C}^2\mathbf{T}$, yields the predicted result. More generally, light that reaches a facet after n bounces cannot illuminate that facet in the $(n+1)-$th bounce. We therefore expect to see alternating intensity patterns in the 4×4 blocks of $\mathbf{C}^n\mathbf{T}$ as n ranges from 2 to 4. The computed n-bounce matrices confirm this behavior. These cancellation results are almost identical to the ground truth rendered n-bounce images, with squared distances between corresponding (normalized) columns that range from 3.45e-34 for the 1-bounce image to 8.59e-13 for the 3rd bounce.

We repeated the same experiment in our laboratory with a real scene whose shape closely matched the scene in our simulations (Figure 4). The result of decomposing the captured ISF matrix is shown in the bottom row of Figure 5. The computed n-bounce matrices are in very good agreement with our simulation results. The main exceptions are near-diagonal elements in the 2-bounce matrix, $\mathbf{C}^2\mathbf{T}$. These artifacts are due to lens flare in the neighborhood of the directly-illuminated pixel. Lens flare increases the intensity at neighboring pixels in a way that cannot be explained by interreflections and, as a result, the intensities due to flare cannot be cancelled by $\mathbf{C}^2$.

Inverting light transport in 1D A key feature of our theory is that it can predict the contribution of the n-th light bounce in images taken under unknown and completely arbitrary illumination. Figure 6 shows the results of an experiment that tests this predictive power. We took two photos of the "M" scene while illuminating it with a flashlight and with room lighting. The resulting images are quite complex, exhibiting a variety of soft and sharp shadows. Despite these complexities, the computed n-bounce images successfully isolate the contribution of direct illumination, whose direction of incidence can be easily deduced from the location of the sharp shadows. Moreover, the higher-order operators allow us to track the propagation of light through the scene even up to the 5th bounce.

Inverting light transport in 2D To test our theory further, we computed n-bounce operators for the 2D scene configuration shown in Figure 7. The scene consisted of three interior faces of a box. We chose laser beam directions that allowed us to sample points on all three faces of the scene. Figure 8 shows the scene's ISF matrix. Each column in this matrix represents the intensity of a 2D set of image pixels, in scanline order. The figure also shows the computed decomposition of the matrix into direct, indirect, 2-bounce, 3-bounce and 4-bounce of reflections. Figure 9 shows results from inverting the light transport process in images of the scene taken under three different illumination conditions. While ground truth information for this scene was not available, the alternating intensity patterns visible in the computed n-bounce images are consistent with the expected behavior, in which light that illuminates a specific face after n bounces cannot illuminate it in the next bounce.
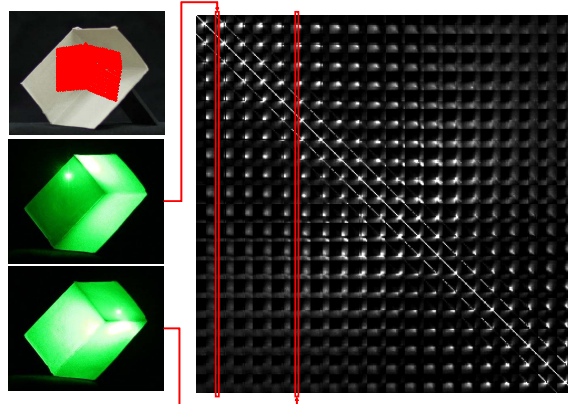

Figure 7: *The 2D scene and its ISF matrix* $\mathbf{T}$. *One column of the ISF matrix represents the resampled image captured at a corresponding laser point.*

5 Conclusions

This paper addressed the problem of computing and removing interreflections in photographs of real scenes. We proved the existence of operators that cancel interreflections in photographs when the geometry, surface reflectance, and illumination are all unknown and unconstrained. For the case of Lambertian scenes we demonstrated that such operators can be computed, and verified the correctness and viability of the theory on both synthetic and real-world scenes. Problems for future work include devising more efficient methods for capturing ISF matrices and estimating cancellation operators for non-Lambertian scenes.

Acknowledgements

References

[1] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes, *Computer Graphics: Principles and Practice*. Reading, MA: Addison-Wesley, 1990.

[2] H. W. Jensen, *Realistic Image Synthesis Using Photon Mapping*. Natick, MA: AK Peters, 2001.

[3] D. S. Immel, M. F. Cohen, and D. P. Greenberg, "A radiosity method for non-diffuse environments," *Computer Graphics (Proc. SIGGRAPH)*, vol. 20, no. 4, pp. 133–142, 1986.

[4] B. K. P. Horn, "Understanding image intensities," *Artificial Intelligence*, vol. 8, no. 2, pp. 201–231, 1977.

[5] J. J. Koenderink and A. J. van Doorn, "Geometric modes as a general method to treat diffuse interreflections in radiometry," *J. Opt. Soc. Am. A*, vol. 73, no. 6, pp. 843–850, 1983.

[6] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the reflectance field of a human face," in *Proc. SIGGRAPH*, pp. 145–156, 2000.
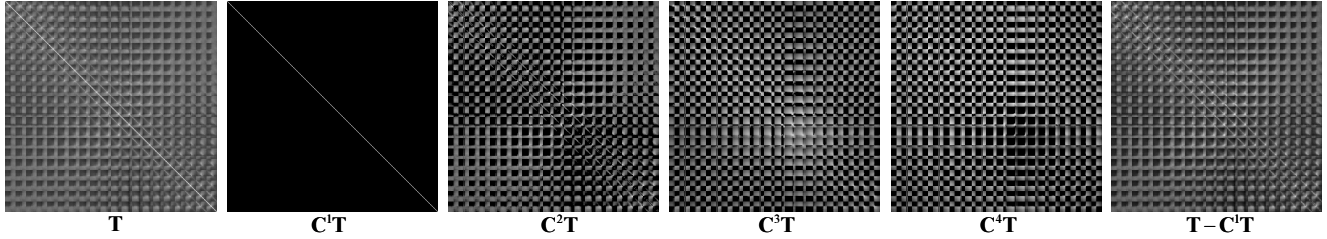
| **T** | **C¹T** | **C²T** | **C³T** | **C⁴T** | **T − C¹T** |

Figure 8: *The inverse light transport computed over the 2D ISF matrix* **T**. *From left to right, the ISF matrix, direct illumination (1-bounce), indirect illumination, the 2-bounce, 3-bounce and 4-bounce of interreflections (log-scaled).*
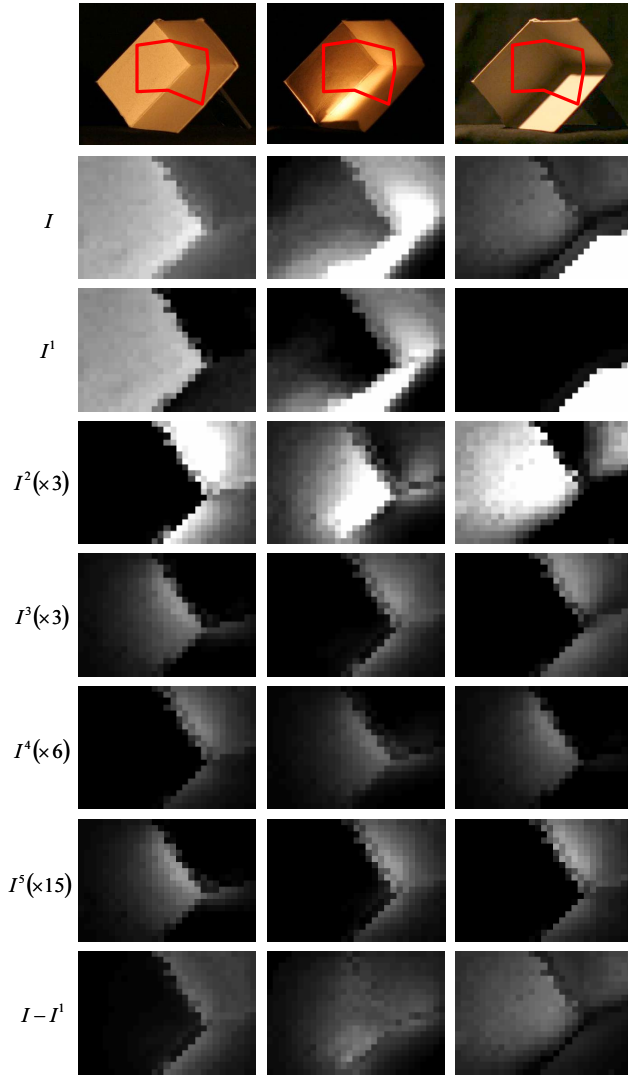

Figure 9: *Inverse light transport applied to images captured under unknown illumination conditions: input images I are decomposed into direct illumination I^1, 2- to 5-bounce images I^2–I^5, and indirect illuminations $I - I^1$.*

[7] M. Goesele, H. P. A. Lensch, J. Lang, C. Fuchs, and H.-P. Seidel, "DISCO: acquisition of translucent objects," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, vol. 23, no. 3, pp. 835–844, 2004.

[8] P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. P. Lensch, "Dual photography," *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 2005. to appear.

[9] Y. Yu, P. E. Debevec, J. Malik, and T. Hawkins, "Inverse global illumination: Recovering reflectance models of real scenes from photographs," in *Proc. SIGGRAPH*, pp. 215–224, 1999.

[10] T. Machida, N. Yokoya, , and H. Takemura, "Surface reflectance modeling of real objects with interreflections," in *Proc. Int. Conf. on Computer Vision*, pp. 170–177, 2003.

[11] S. K. Nayar, K. Ikeuchi, and T. Kanade, "Shape from interreflections," *Int. J. of Computer Vision*, vol. 6, no. 3, pp. 173–195, 1991.

[12] D. Forsyth and A. Zisserman, "Reflections on shading," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 13, no. 7, pp. 671–679, 1991.

[13] S. A. Shafer, T. Kanade, G. J. Klinker, and C. L. Novak, "Physics-based models for early vision by machine," in *Proc. SPIE 1250: Perceiving, Measuring, and Using Color*, pp. 222–235, 1990.

[14] R. Bajcsy, S. W. Lee, and A. Leonardis, "Color image segmentation with detection of highlights and local illumination induced by inter-reflections," in *Proc. Int. Conf. on Pattern Recognition*, pp. 785–790, 1990.

[15] B. V. Funt, M. S. Drew, and J. Ho, "Color constancy from mutual reflection," *Int. J. of Computer Vision*, vol. 6, no. 1, pp. 5–24, 1991.

[16] B. V. Funt and M. S. Drew, "Color space analysis of mutual illumination," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, no. 12, pp. 1319–1326, 1993.

[17] L. McMillan and G. Bishop, "Plenoptic modeling: An image-based rendering system," in *Proc. SIGGRAPH 95*, pp. 39–46, 1995.

[18] M. Levoy and P. Hanrahan, "Light field rendering," in *Proc. SIGGRAPH 96*, pp. 31–42, 1996.

[19] S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, "The lumigraph," in *Proc. SIGGRAPH 96*, pp. 43–54, 1996.

[20] J. T. Kajiya, "The rendering equation," *Computer Graphics (Proc. SIGGRAPH 86)*, vol. 20, no. 4, pp. 143–150, 1986.

[21] A. S. Glassner, *Principles of digital image synthesis*. San Francisco, CA: Morgan Kaufmann Publishers, 1995.

[22] T. Mitsunaga and S. Nayar, "Radiometric self calibration," in *Proc. Computer Vision and Pattern Recognition Conf.*, pp. 374–380, 1999.