Primal-Dual Coding to Probe Light Transport

Matthew O'Toole[*]
University of Toronto

Ramesh Raskar[†]
MIT Media Lab

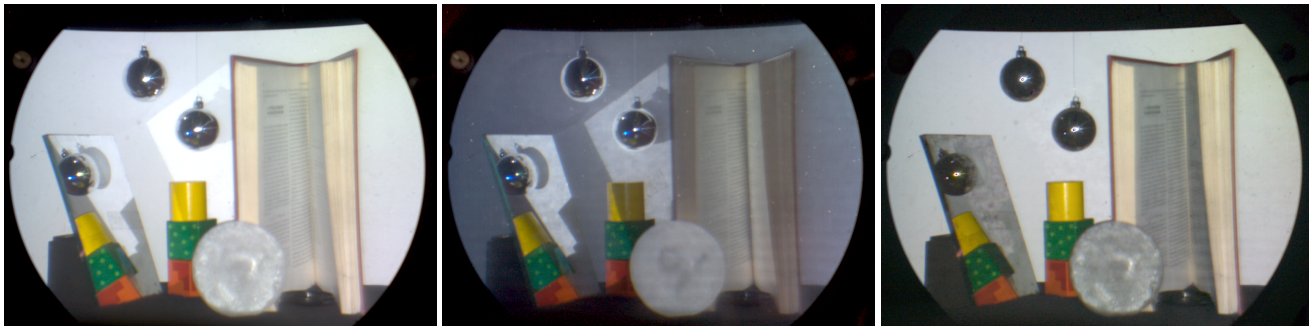Kiriakos N. Kutulakos[*]
University of Toronto

Figure 1: *Left to right: Scene under white illumination, global illumination component captured in just a single RAW photograph (no processing), and direct component from subtracting the middle image from the left image. The global image was captured using 512 primal and dual patterns during one 30-second exposure. Notice that the objects embedded within the wax disk are only visible in the global image.*

Abstract

We present *primal-dual coding,* a photography technique that enables direct fine-grain control over which light paths contribute to a photo. We achieve this by projecting a sequence of patterns onto the scene while the sensor is exposed to light. At the same time, a second sequence of patterns, derived from the first and applied in lockstep, modulates the light received at individual sensor pixels. We show that photography in this regime is equivalent to a *matrix probing* operation in which the elements of the scene's transport matrix are individually re-scaled and then mapped to the photo. This makes it possible to directly acquire photos in which specific light transport paths have been blocked, attenuated or enhanced. We show captured photos for several scenes with challenging light transport effects, including specular inter-reflections, caustics, diffuse inter-reflections and volumetric scattering. A key feature of primal-dual coding is that it operates almost exclusively in the optical domain: our results consist of directly-acquired, unprocessed RAW photos or differences between them.

CR Categories: I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture—Radiometry, imaging geometry; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Color, shading, shadowing and texture.

Keywords: primal-dual coding, light transport, inter-reflections, scattering, coded exposure, coded illumination, optical processing

Links: ◆DL 🅟PDF 🌐WEB

[*]e-mail: {motoole,kyros}@cs.toronto.edu
[†]e-mail: raskar@media.mit.edu

1 Introduction

A common assumption in computer graphics and computer vision is that photos from a standard photographic camera always satisfy the light transport equation [Ng et al. 2003]:

$$\mathbf{p} \,=\, \mathbf{T}\,\mathbf{i} \qquad (1)$$

where $\mathbf{p}$ is a photo represented as a column vector of pixels, $\mathbf{T}$ is the scene's $N \times M$ transport matrix, and $\mathbf{i}$ is a vector representing the scene's incident illumination. This relation underlies a large body of work, from precomputed radiance transfer [Sloan et al. 2002] to image-based relighting [Debevec et al. 2000; Wang et al. 2009], computational photography [Sen et al. 2005] and computational light transport [Chandraker et al. 2010; O'Toole and Kutulakos 2010].

In this paper we observe that a much more general imaging regime exists for photography with off-the-shelf cameras. This regime, which we call *primal-dual coding,* enables direct acquisition of photos that violate the scene's light transport equation. These photos have a bilinear, rather than a linear, relationship to conventional photographs and appear as though the transport matrix of the scene itself has been manipulated (Figure 1). Unlike techniques that rely on exotic technologies [Kirmani et al. 2011], primal-dual coding is realized with standard components—a camera, a projector and an LCD panel.

More specifically, primal-dual coding acquires photos governed by the following *transport probing equation* (Figure 2):

$$\mathbf{p} \,=\, (\mathbf{\Pi} \odot \mathbf{T})\,\mathbf{1} \qquad (2)$$

where the symbol $\odot$ denotes the element-wise multiplication of two equal-sized matrices. According to this equation, a photo is formed by multiplying element-wise the scene's transport matrix with a *probing matrix* $\mathbf{\Pi}$ and then multiplying the result with a constant vector of all ones. Intuitively, the probing matrix can be thought of as a generalized illumination condition: it is under the complete control of the image acquisition process, it is known, and can be arbitrary. In contrast to conventional photography where there are M degrees of freedom for controlling the output photo (*i.e.,* the size of the illumination vector $\mathbf{i}$), the probing matrix has $N \times M$ degrees of freedom (the number of elements it contains). This gives
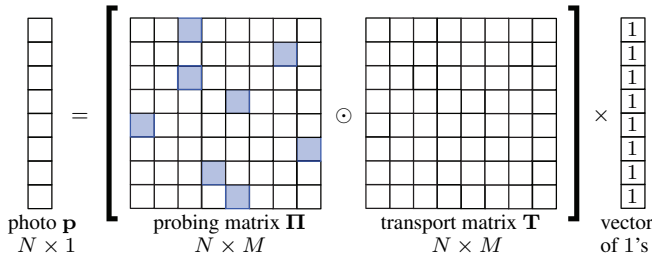
Figure 2: *The transport probing equation.*

tremendous flexibility over image acquisition, well beyond what is implied by the light transport equation.

To achieve this, primal-dual coding exerts simultaneous control over two aspects of the image formation process: it controls the scene's illumination (the *primal domain*) and it modulates the light arriving at individual pixels on the camera's sensor (the *dual domain*). To acquire a photo, we project a sequence of patterns onto the scene while the sensor is exposed to light. At the same time, a second sequence of patterns, derived from the first and applied in lockstep, modulates the light received at individual sensor pixels (Figure 3). Capturing a photo by primal-dual coding therefore involves two basic choices—what primal-domain patterns to use for scene illumination and how to convert them into dual-domain patterns for modulating the sensor. As we will see later, the former determines the quality of the captured photo whereas the latter implicitly determines the probing matrix. Most of our paper is about a detailed analysis of these two choices.

Primal-dual coding is most closely related to confocal imaging [Corle and Kino 1996], aperture correlation [Wilson et al. 1996] and differential spinning disk microscopy [Mertz 2011]. These methods are a special case of primal-dual coding where the primal- and dual-domain patterns are binary, and either coincide or are complements of each other. To our knowledge, they have not been used for one-shot imaging in standard photography settings, although "synthetic" implementations of aperture correlation have been demonstrated by Levoy and colleagues [Levoy et al. 2004; Fuchs et al. 2008]. These implementations keep optical processing to a minimum and rely on substantial image acquisition and computational processing to synthesize individual photos. In contrast, we rely almost exclusively on optical processing; analyze the general case, where primal and dual codes may differ; and introduce the transport probing equation as a general image formation model for primal-dual coding. This equation characterizes the space of possible photos and their relationship to a scene's transport matrix.

Primal-dual coding can also be thought of as a combination of illumination coding, which operates exclusively in the primal domain [Veeraraghavan et al. 2011; Schechner et al. 2007; O'Toole and Kutulakos 2010] and coded-exposure photography, which operates in the dual [Hitomi et al. 2011; Nayar et al. 2004; Wetzstein et al. 2010]. These coding techniques have been applied very successfully in recent years but have not been combined for one-shot photo acquisition. Like many coded-exposure methods, we use relay optics and a display panel to modulate the light arriving at sensor pixels without having to access the camera's interior.

From a mathematical perspective, both the transport probing equation and our overall analysis involve straightforward extensions of elementary results in *matrix probing*. This is a topic of numerical mathematics concerned with the efficient estimation of the trace or diagonal of very large unobservable matrices [Bekas et al. 2007; Tang and Saad 2012]. The relation of these methods to imaging has apparently not been noted, although they are highly relevant. For
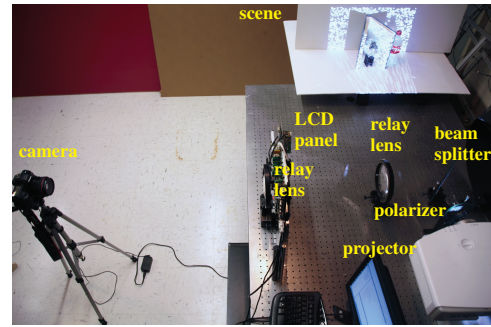

Figure 3: *Photo and diagram of our primal-dual coding prototype. We use a pair of relay lenses and an LCD panel with its backlight removed to modulate the light reaching the camera's sensor. To acquire a primal-dual coded photo, we project a sequence of illumination patterns onto the scene and simultaneously display a sequence of modulation patterns on the LCD panel. The camera's shutter remains open throughout this process.*

instance, aperture correlation implements an efficient stochastic diagonal estimator of the transport matrix.

As an initial feasibility test of the primal-dual coding framework, we present results from a prototype system where a camera and a projector share the same viewpoint. We show how to capture photos where the direct component of light transport is missing, and where components of direct or indirect illumination are selectively attenuated or enhanced. Compared to recent primal-domain methods with similar goals [Nayar et al. 2006; Seitz et al. 2005; Zhang and Nayar 2006], a major advantage of our approach is that it is independent of the frequency content of light transport, the reflectance of the scene's surfaces, and the rank of its transport matrix. It thus applies to fully-general scenes where a mixture of indirect transport phenomena may occur—including specular and diffuse inter-reflections, caustics, sub-surface scattering, *etc.*

Our investigation is limited in two respects, however, one conceptual and one practical. On the conceptual side, while primal-dual coding and the transport probing equation apply to any camera-projector arrangement, all our experimental results except one rely on the camera and projector being coaxial. As such, use of the primal-dual coding framework in a non-coaxial arrangement (*e.g.*, for 3D shape recovery [Gupta et al. 2011]) is not yet well understood. From a practical standpoint, our reliance on off-the-shelf hardware significantly restricted the rate by which primal- and dual-domain patterns change in our prototype. We compensated by capturing photos with an exposure of several seconds, with all the disadvantages this entails (slightly-increased noise, static scenes, *etc.*)

2 Probing Light Transport

We begin by taking a closer look at the transport probing equation and its relation to light transport analysis. The question of how to implement it is considered in the following sections.
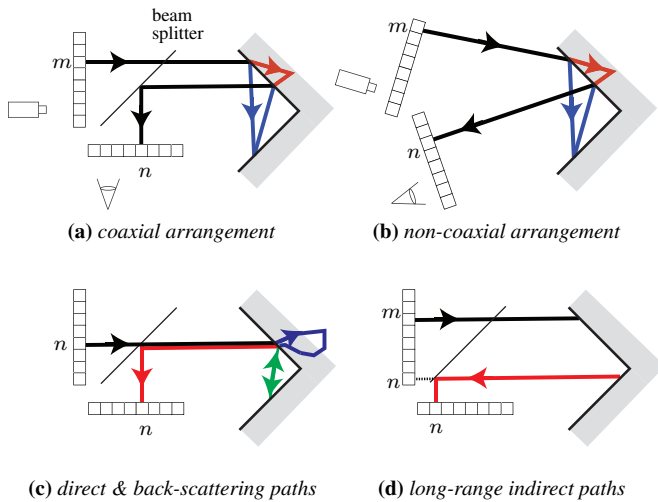
(a) *coaxial arrangement* **(b)** *non-coaxial arrangement*

(c) *direct & back-scattering paths* **(d)** *long-range indirect paths*

Figure 4: *Basic light transport paths. (a), (b) Element $\mathbf{T}_{nm}$ of the transport matrix takes takes into account all light paths that begin at projector pixel m and terminate at sensor pixel n. Two such paths are shown, one due to inter-reflections (blue) and one due to sub-surface scattering (red). (c) For coaxial arrangements, the direct transport path (red) always contributes to the diagonal element, $\mathbf{T}_{nn}$. This element may also include contributions from back-scattering (blue) and retro-reflection paths (green) which also begin and end at the same pixel. (d) In coaxial arrangements it is also possible to distinguish between long- and short-range transport paths by the distance $|n - m|$ between their endpoints.*

In a typical photography setting we may know very little—or nothing at all—about how the scene transports incident light. This black-box view of the scene and of its transport matrix leaves just one way to get information about them, *i.e.*, by capturing one or more photos. Unfortunately, there is a fundamental dimensionality gap between the transport matrix $\mathbf{T}$ and the photos it can produce: the transport matrix has $N \times M$ elements whereas each photo can only provide N measurements.[1] We therefore have an extremely limited bandwidth for acquiring information about $\mathbf{T}$. The key advantage of the transport probing equation is the control it offers over how the elements of $\mathbf{T}$ are mapped onto a photo's limited pixels.

To gain further intuition about the space of allowable mappings, it is helpful to consider the probing equation in the special case when the probing matrix is binary and has only one non-zero element per row (dark squares in Figure 2). Multiplying element-wise such a matrix with $\mathbf{T}$ leaves only one non-zero element on each row, equal to the corresponding element of the transport matrix. Multiplication with a constant vector then transfers it unaltered to the photo. Thus, by choosing different binary probing matrices we can directly acquire different N-dimensional slices of the $N \times M$ transport matrix. Table 1 shows several such examples, along with other probing operations made possible with non-binary probing matrices.

Probing matrices for light transport calculations Observe that we can re-write the transport probing equation as

$$\mathbf{p} = \hat{\mathbf{T}}\,\mathbf{1} \qquad (3)$$

[1]This gap may be reduced by measuring physical properties beyond irradiance such as polarization state [Ghosh et al. 2010], phase [Popoff et al. 2010] and temporal irradiance variation [Kirmani et al. 2011]. We do not consider such measurements here.

by setting $\hat{\mathbf{T}} = \mathbf{\Pi} \odot \mathbf{T}$. This brings it into the familiar form of Equation 1 and suggests a physical interpretation for the acquired photo: $\mathbf{p}$ is the photo we would have captured by conventional photography if the scene had transport matrix $\mathbf{\Pi} \odot \mathbf{T}$ and was illuminated uniformly. Viewed this way, $\mathbf{\Pi}$ re-scales the elements of the transport matrix individually for the purpose of image acquisition.

Each element of the transport matrix describes the transport of radiance along a distinct collection of light paths from the illumination source to the sensor (Figures 4 and 5). By re-scaling these elements, the probing matrix gives the ability to directly acquire photos in which light transport along specific paths has been enhanced, attenuated, or blocked entirely. See Table 2 for several examples of imaging tasks, along with their associated probing matrices.

3 Photography by Primal-Dual Coding

Let us now look at how to capture photos governed by the transport probing equation. The key ingredient is an ability to simultaneously modulate light at its source (the projector) and its destination (the sensor). We first consider a naive algorithm that serves as an existence proof and illustrates the basic properties of the approach.

3.1 Naive Approach: Path Isolation

Since we have complete control over which projector pixels emit light and which sensor pixels receive it, we can control in a very precise way the light transport paths that contribute to a photo. In particular, by turning on a single projector pixel and unmasking a single sensor pixel we guarantee that only the transport paths between those two pixels will contribute to the photo. The naive algorithm, shown in Figure 7, applies this basic idea in a time-multiplexed fashion: it allocates a time slice τ within the exposure period to every possible combination of projector pixel m and sensor pixel n, and uses $\mathbf{\Pi}_{nm}$ for the projector pixel's intensity during that time slice. This ensures that transport paths that are "active" during a particular time slice will contribute irradiance to the final photo precisely as described in Equation (2).

Path isolation is an extremely inefficient approach for two reasons. Firstly, every element of the probing matrix gets equal time even though its value may be zero. Since many probing matrices of interest are sparse (*e.g.*, see Tables 1 and 2), the sensor would not be integrating any light at all for much of the exposure time. An obvious improvement is to allocate time proportionally to $\mathbf{\Pi}_{nm}$. This solution, for the special case where the $\mathbf{\Pi}$ is the identity matrix, describes the basic confocal microscopy technique[2] [Corle and Kino 1996]. A second—much more serious—issue is that only one pixel integrates light at any given point in time. Given that a typical image may contain millions of pixels, the short time slices mean that light received at a pixel would fall well below the additive noise floor under typical photography conditions.

3.2 Optical Matrix Probing

Acquisition schemes that make efficient use of the sensor are highly desirable. For this, we turn to work in numerical linear algebra [Bekas et al. 2007]. Let $\{\mathbf{i}^k\}$ and $\{\mathbf{m}^k\}$ be two sequences of column vectors corresponding to a decomposition of the probing matrix into a sum of rank-1 matrices:

$$\mathbf{\Pi} = \sum_{k=1}^{K} \mathbf{m}^k (\mathbf{i}^k)^{\mathrm{T}} \qquad (4)$$

[2]The coaxial arrangements used in confocal imaging guarantee that the diagonal of the transport matrix contains all direct (*a.k.a.* in-focus) paths.
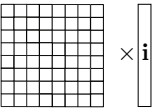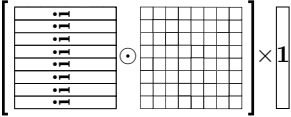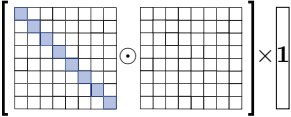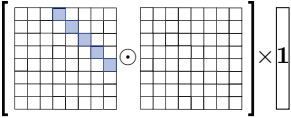
Table 1
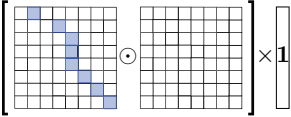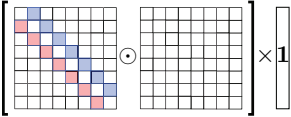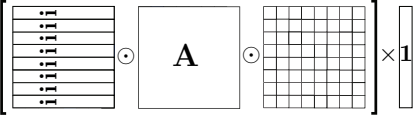
Probing Objective	Transport Probing Equation	Expression for Element Π_{nm}	Expression for Photo Pixel $\mathbf{p}_n$
Capture a conventional photo under illumination $\mathbf{i}$ (Equation 1)	$\times\,\mathbf{i}$		$\sum_m \mathbf{T}_{nm}\,\mathbf{i}_m$
Capture a conventional photo under illumination $\mathbf{i}$ (Equation 2)	$\left[\;\odot\;\right]\times\mathbf{1}$	$\mathbf{i}_m$	$\sum_m \mathbf{\Pi}_{nm}\,\mathbf{T}_{nm}$ $=\;\sum_m \mathbf{i}_m\,\mathbf{T}_{nm}$
Capture the diagonal of $\mathbf{T}$	$\left[\;\odot\;\right]\times\mathbf{1}$	$\delta(n-m)$	$\sum_m \mathbf{\Pi}_{nm}\,\mathbf{T}_{nm}$ $=\;\mathbf{T}_{nn}$
Capture the wth off-diagonal of $\mathbf{T}$	$\left[\;\odot\;\right]\times\mathbf{1}$	$\delta(n-m+w)$	$\mathbf{T}_{n(n+w)}$
Sample one element from each row of $\mathbf{T}$ according to a given sampling vector $\mathbf{a}$	$\left[\;\odot\;\right]\times\mathbf{1}$	$\delta(\mathbf{a}_n - m)$	$\mathbf{T}_{n\mathbf{a}_n}$
Capture a linear combination of two elements in each row, according to sampling vectors $\mathbf{a}$ and $\mathbf{b}$	$\left[\;\odot\;\right]\times\mathbf{1}$	$a\,\delta(\mathbf{a}_n - m)$ $+\;b\,\delta(\mathbf{b}_n - m)$	$a\,\mathbf{T}_{n\mathbf{a}_n}\;+\;b\,\mathbf{T}_{n\mathbf{b}_n}$
Capture a photo that simulates conventional photography under illumination $\mathbf{i}$ of a scene whose transport matrix is $\mathbf{A}\odot\mathbf{T}$	$\left[\;\odot\;\mathbf{A}\;\odot\;\right]\times\mathbf{1}$	$\mathbf{A}_{nm}\,\mathbf{i}_m$	$\sum_m (\mathbf{A}_{nm}\,\mathbf{T}_{nm})\,\mathbf{i}_m$

Table 1: *Probing the transport matrix. **Rows 1 and 2:** Relation between the light transport equation and the transport probing equation. The last column shows that the two equations coincide when $\mathbf{\Pi}_{nm} = \mathbf{i}_m$. **Rows 3-7:** Probing makes it possible to compute general linear mappings from transport matrix elements to camera pixels without having to capture (or even approximate) the full transport matrix first.*

Acquisition Goal	Expression for Element Π_{nm}
Capture direct plus retro-reflection and back-scattering (Figure 4c)	$(n=m)$
Capture all other indirect (Figure 4a)	$(n \neq m)$
Capture long-range indirect (Figure 4d)	$(|n-m| > w)$
Capture short-range indirect (Figure 4d)	$(|n-m| \leq w) \wedge (n \neq m)$
Transport masking: block all light paths from object $\mathcal{B}$ to object $\mathcal{A}$ but leave other paths unaffected	$(n \notin \mathcal{A}) \vee (m \notin \mathcal{B})$
Local de-scattering: capture direct plus high-frequency retro-reflection and back-scattering	$(n=m) - (n=m+1)$
Transport-domain filtering: use a filter kernel $\mathbf{f}$ to modulate transport paths according to their distance from the direct component	$(n=m+w)\mathbf{f}_w$

Table 2: *Probing matrices useful for light transport calculations in **coaxial arrangements**. We slightly abuse notation in the expressions for Π_{nm}: logical operations (equality, inequality, set membership) are treated as functions returning 1 if true and 0 if false. For transport masking, we assume that the objects $\mathcal{A}$ and $\mathcal{B}$ are specified as sets of pixels on the image plane. The first five operations can be thought of as applying a binary mask to the transport matrix (e.g., the matrix shown in Figure 5b). The last two operations involve cross-correlations along rows of the matrix. Note that unlike methods that rely on assumptions about the frequency content of light transport [Nayar et al. 2006], none of the above operations do so, with the exception of local de-scattering. The de-scattering expression was proposed by Fuchs et al. [2008] to remove back-scattering contributions that are constant in the neighborhood of the direct component. To our knowledge, transport masking and filtering have not been considered previously. For **non-coaxial arrangements**, the same relations apply with one modification: m must be equal to $\mathbf{a}_n$, where vector $\mathbf{a}$ encodes stereo correspondences. These relations are less useful when stereo correspondences are unknown.*
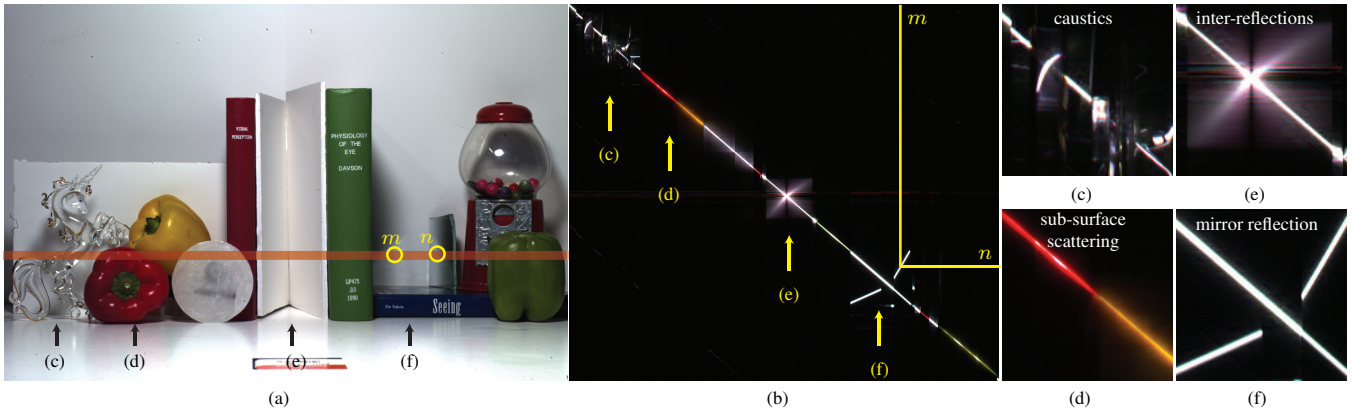
Figure 5: *Transport matrix of a complex scene. (a) A photo of the scene. From left-to-right, the scene contains a glass unicorn, a pair of bell peppers, a translucent wax disk, a V-wedge between a pair of books, a mirror reflecting light onto the wall, a gumball dispenser and a green bell pepper. (b) Since the complete transport matrix is too large to capture and visualize, we show a 2D slice of the matrix for the row highlighted in (a). The slice was acquired with a coaxial arrangement using the method of Schechner et al. [2007]. Element $\mathbf{T}_{nm}$ in this slice represents light paths that originate from projector pixel m and reach camera pixel n in the highlighted row, shown as yellow circles in (a). Note that non-zero elements in this slice mostly concentrate around the diagonal. This indicates that most light is transported between nearby camera/projector pixels. Nevertheless, the matrix does include several interesting structures, magnified in (c)-(f). (c) Caustics from the glass unicorn appear as "streaks" in the transport matrix. (d) Sub-surface scattering from the yellow and red bell peppers produces a thick band along the diagonal, with isolated white specularities on the diagonal itself. (e) Inter-reflections from the V-wedge appear as fans of light. Dark regions in these fans are due to occlusion, which blocks inter-reflection paths in certain directions. (f) Mirror reflections produce high-frequency, off-diagonal structures.*
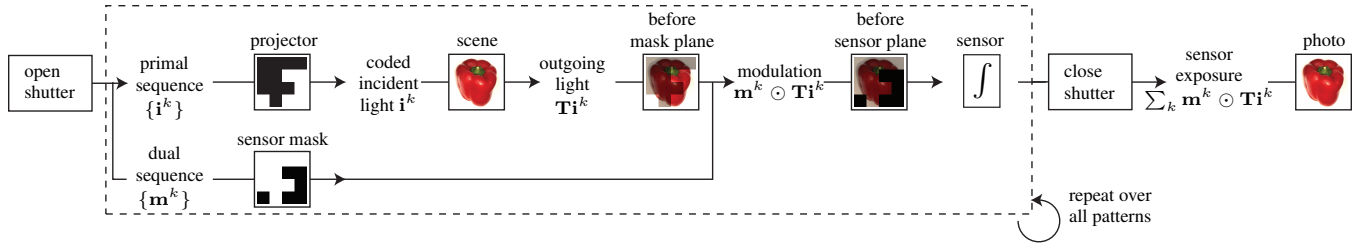

Figure 6: *Image formation pipeline for a single photo captured by primal-dual coding. The primal and dual sequences, $\{\mathbf{i}^k\}$ and $\{\mathbf{m}^k\}$, are computed in a preprocessing step. Patterns in the primal and dual domains are changed synchronously.*

This decomposition allows us to re-write the transport probing equation as a sum identical to the pipeline in Figure 6:

$$(\mathbf{\Pi} \odot \mathbf{T})\,\mathbf{1} \;=\; \sum_{k=1}^{K}\sum_{m=1}^{M}\mathbf{m}^k \mathbf{i}_m^k \odot \mathbf{T}_m \;=\; \sum_{k=1}^{K}\mathbf{m}^k \odot \mathbf{T}\,\mathbf{i}^k \quad (5)$$

Here we used $\mathbf{T}_m$ to denote the m-th column of $\mathbf{T}$. It follows that every decomposition of the probing matrix into a sum of rank-1 matrices provides a candidate sequence of illumination patterns $\{\mathbf{i}^k\}$ and masks $\{\mathbf{m}^k\}$ for primal-dual coding.

Stochastic diagonal estimators   Not all rank-1 decompositions are equally efficient. Bekas *et al.* [2007] study this problem for the case where the probing matrix is the identity. This matrix allows us to acquire the diagonal of $\mathbf{T}$ (third row of Table 1 and first row of Table 2). The main idea is to consider rank-1 decompositions of the identity matrix created by an infinite sequence of random vectors $\{\mathbf{i}^k\}$ drawn from a distribution with mean zero and unit variance:

$$(\mathbf{I} \odot \mathbf{T})\,\mathbf{1} \;=\; \lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}\mathbf{i}^k \odot \mathbf{T}\mathbf{i}^k \quad (6)$$

They show that the provably-optimal convergence rate is achieved by vectors drawn from the Rademacher distribution, whose ele-

ments have a 50% chance of being $+1$ or -1.[3] Moreover, as shown in the Supplemental Material, the variance of a K-term estimate of diagonal element $\mathbf{T}_{nn}$ is

$$\frac{1}{K}\sum_{m=1,n\neq m}^{M}\mathbf{T}_{nm}^2 \quad (7)$$

Therefore, to double the accuracy of our estimate of $\mathbf{T}_{nn}$ we must use twice as many vectors. Importantly, Equation (7) tells us that this accuracy depends on the transport matrix itself and thus on the type of global transport within a scene: it is low when the total intensity of non-zero elements along a row of $\mathbf{T}$ is distributed over many elements (*e.g.*, diffuse inter-reflections) and high when it is concentrated at a few isolated ones (*e.g.*, caustics).

Stochastic estimators for general probing   Generalizing the above estimator to the case of a general probing matrix $\mathbf{\Pi}$ is trivial in some respects and non-trivial in others. On the trivial side, extracting the wth off-diagonal of $\mathbf{T}$, as shown in the fourth row of

---

[3]Although not described in these terms, aperture correlation [Wilson et al. 1996] uses this sequence for extracting the diagonal of $\mathbf{T}$. This was proposed well before its optimality was established.

Path isolation:

In: exposure time E, probing matrix $\mathbf{\Pi}$

Out: photo equal to $(\mathbf{\Pi} \odot \mathbf{T})\,\mathbf{1}$

1: $\tau = E/NM$

2: open camera shutter

3:    **for** $n = 1$ to N
4:      unmask pixel n
5:      **for** $m = 1$ to M
6:        turn on projector pixel m for time τ with intensity $\mathbf{\Pi}_{nm}$
7:      mask all pixels

8: close shutter

9: **return** captured photo

Optical matrix probing:

In: exposure time E, probing matrix $\mathbf{\Pi}$, K illumination vectors $\{\mathbf{i}^k\}$

Out: photo equal to $(\mathbf{\Pi} \odot \mathbf{T})\,\mathbf{1}$

1: $\tau = E/K$

2: open camera shutter

3:    **for** $k = 1$ to K
4:      apply mask $\mathbf{m}^k = \mathbf{\Pi}\,\mathbf{i}^k$
5:      project vector $\mathbf{i}^k$ for time τ

6: close shutter

7: **return** captured photo

Figure 7: *Basic algorithms for primal-dual coding photography.*

Table 1, simply requires a mask that is a w-shifted version of the illumination pattern:

$$(\mathbf{\Pi} \odot \mathbf{T})\,\mathbf{1} \;=\; \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \text{shift}(\mathbf{i}^k, w) \odot \mathbf{T}\mathbf{i}^k \qquad (8)$$

More generally, left-multiplying both sides of Equation (6) by $\mathbf{\Pi}$ gives the equivalent expression for general probing matrices:

$$(\mathbf{\Pi} \odot \mathbf{T})\,\mathbf{1} \;=\; \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \mathbf{\Pi}\mathbf{i}^k \odot \mathbf{T}\mathbf{i}^k \qquad (9)$$

This expression is straightforward to implement optically (see Figure 7). Unfortunately, it is not known if the optimality guarantees of the Rademacher distribution extend to general probing matrices.

Negative values Up to now we have ignored the fact that the illuminations and masks may contain negative values which are not physically realizable. Negative values may occur because of the random vector sequence being used or because $\mathbf{\Pi}$ itself contains negative values. Correct treatment of such cases requires capturing at most two photos followed by a (computational) subtraction. In particular, by expressing $\mathbf{i}^k$ and $\mathbf{m}^k$ as a difference of two non-negative vectors, it is possible to break each term in Equation (4) into four terms, two of which are positive and two of which are negative. The first two terms are used during the exposure period of the positive photo and the last two of the negative one:

$$\overbrace{[\;\mathbf{m}_+^k(\mathbf{i}_+^k)^{\mathrm{T}} \;+\; \mathbf{m}_-^k(\mathbf{i}_-^k)^{\mathrm{T}}\;]}^{\text{positive terms}} \;-\; \overbrace{[\;\mathbf{m}_-^k(\mathbf{i}_+^k)^{\mathrm{T}} \;+\; \mathbf{m}_+^k(\mathbf{i}_-^k)^{\mathrm{T}}\;]}^{\text{negative terms}} \qquad (10)$$
$$\underbrace{\phantom{\mathbf{m}_+^k(\mathbf{i}_+^k)^{\mathrm{T}}}}_{\#1} \quad \underbrace{\phantom{\mathbf{m}_-^k(\mathbf{i}_-^k)^{\mathrm{T}}}}_{\#2} \quad \underbrace{\phantom{\mathbf{m}_-^k(\mathbf{i}_+^k)^{\mathrm{T}}}}_{\#3} \quad \underbrace{\phantom{\mathbf{m}_+^k(\mathbf{i}_-^k)^{\mathrm{T}}}}_{\#4}$$

where $\mathbf{i}^k = \mathbf{i}_+^k - \mathbf{i}_-^k$ and $\mathbf{m}^k = \mathbf{m}_+^k - \mathbf{m}_-^k$. It follows that implementing a K-vector primal-dual coding sequence with both positive and negative values requires capturing two photos using a $2K$-long sequence during each photo's exposure.

Fortunately, useful probing matrices exist that produce strictly non-negative illuminations and masks. This makes it possible to probe efficiently with just one photo. Two important examples are capturing an indirect-only photo of the scene (second row in Table 2) and capturing the direct component plus one-half the indirect (*i.e.*, contrast-enhancing the direct component). See the Supplemental Materials for a brief derivation of how Rademacher sequences can be replaced by non-negative Bernoulli sequences in such cases.[4]

---
[4]According to Equation (4), if $\mathbf{\Pi}$'s rank is low applying non-negative factorization also yields efficient non-negative illuminations and masks.

4 Implementation

Hardware Figure 3 illustrates our hardware setup. We use an Epson PowerLite G5000 3-LCD projector with a linear polarizer placed in front, a disassembled Barco E-2320 PA monochrome LCD panel positioned between two linear polarizing sheets, a 50R/50T Edmund Optics beam splitter, two spherical plano-convex relay lenses, a Canon EOS Rebel XSi camera, a 2.67 GHz Intel Core i7 workstation with 3 GB of RAM and two NVIDIA GeForce 9500GT graphics cards. We operate the projector at resolution 1024×768 and use two vsync-enabled borderless OpenGL windows to control the projector and LCD at a 60 Hz refresh rate. Pattern generation and display are done through MATLAB.

The rightmost relay lens in Figure 3 focuses light from the scene onto the plane of the LCD display. The display attenuates this incident light using a 1600×1200-resolution pattern. The leftmost relay lens, along with the camera's own 50mm lens, is then used to capture a focused image of the LCD display. The focal length of the leftmost relay lens was 330mm in all experiments; the other relay lens was 330mm in all experiments except one (Figure 9), where a 220mm lens was used. These relay lenses were slightly larger than the active area of the LCD display (15cm versus 10cm) and positioned asymmetrically, with the leftmost one much closer to the display. We align the projector and camera using a beamsplitter to ensure a coaxial arrangement.

Calibration All experiments except the one in Figure 11 require a coaxial camera and projector and aligned projector and LCD patterns. We do this as follows. We temporarily place a diffuser on the LCD panel; focus the image of the scene onto the diffuser by adjusting the position of the rightmost relay lens; align the projector to produce a coaxial arrangement; and move and resize the OpenGL windows so that the pixels of the projector and LCD panel overlap from the camera's perspective. Since we relied on binary patterns for our experiments, no radiometric calibration was required. For more general probing experiments, where illuminations and/or masks are not binary, radiometric calibration of both the LCD panel and projector becomes necessary.

Codes robust to misalignment Even after calibration, the illumination and mask patterns in our prototype are not aligned with pixel accuracy. Moreover, despite driving the projector and the LCD panel with the same vertical synchronization signal, the two devices do not refresh in perfect synchrony. These misalignments forced us to operate at a reduced spatio-temporal resolution. To account for temporal misalignment, we interleave black images on the LCD panel before and after displaying each mask pattern, reducing the effective primal-dual coding display rate to 20 Hz. To deal with pixel misalignment, we rely on reduced-resolution, 64×48 primal-dual coding sequences and use four times as many codes (included in our code counts for experiments). In particular, we replace every pair $\mathbf{i}^k, \mathbf{m}^k$ in the primal-dual coding sequence with a four-code sequence $\mathbf{i}_l^k, \mathbf{m}_l^k$, that eliminates artifacts due to misaligned pixel boundaries. This sequence exposes the center $\frac{1}{4}$-th of reduced-resolution pixels and then shifts both the illumination pattern and the mask three times to expose the complete coarse-resolution pixels. The sum of the four mask patterns, $\mathbf{m}_l^k$, produces the original dual code $\mathbf{m}^k$. Together with the black masks, this results in the following twelve-code sequence:

Primal:	$\mathbf{i}_1^k$	$\mathbf{i}_1^k$	$\mathbf{i}_1^k$	$\mathbf{i}_2^k$	$\mathbf{i}_2^k$	$\mathbf{i}_2^k$	$\mathbf{i}_3^k$	$\mathbf{i}_3^k$	$\mathbf{i}_3^k$	$\mathbf{i}_4^k$	$\mathbf{i}_4^k$	$\mathbf{i}_4^k$
Dual:	$\mathbf{0}$	$\mathbf{m}_1^k$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{m}_2^k$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{m}_3^k$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{m}_4^k$	$\mathbf{0}$

See the Supplementary Material for an example. For high code resolutions, the center $\frac{1}{4}$-th of each pixel may no longer mask the misaligned pixels. This imposes a maximum working resolution to any primal-dual coding setup (ours is 64×48). This means that all light paths within a coarse pixel are clumped together in the photos we capture.

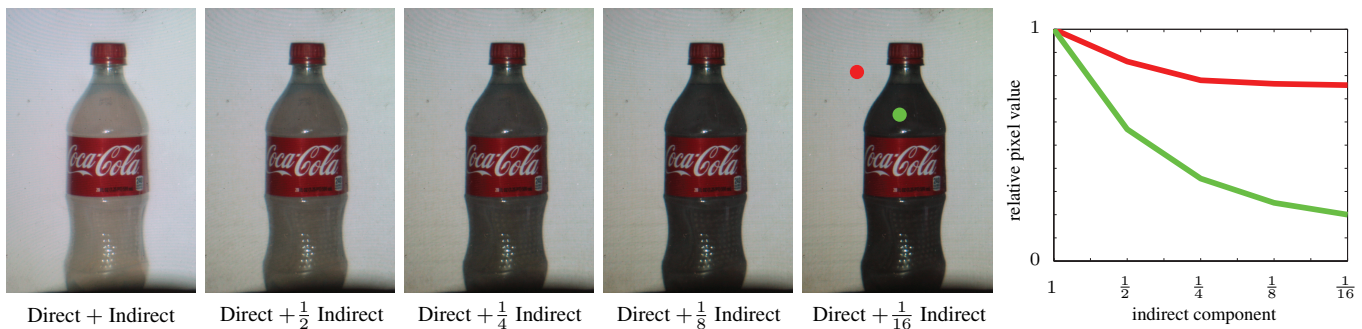| Direct + Indirect | Direct + $\frac{1}{2}$ Indirect | Direct + $\frac{1}{4}$ Indirect | Direct + $\frac{1}{8}$ Indirect | Direct + $\frac{1}{16}$ Indirect |

Figure 8: *Contrast-enhancing the direct component of a coke bottle filled with milky water, which strongly scatters light.* **Left to right:** *A photo of the scene lit by white illumination contains full direct and indirect; the scene with* 50% *indirect illumination uses* 64 *primal-dual codes;* 25% *indirect illumination uses* 128 *codes,* 12.5% *indirect uses* 256 *codes, and* 6.25% *indirect uses* 512 *codes. The plot on the right examines the intensity of two points in the scene, a point lit mostly by indirect (green), and a point lit mostly by direct (red). As we reduce the percentage of indirect light, the intensity of the point on the background decreases slightly, indicating a small indirect component. In contrast, the intensity of the point on the bottle diminishes, an indication of the strong indirect transport effects within the bottle.*

Image capture and adjustment We capture 4272×2848-resolution RAW photos and decode them using DCRAW. We then rotate, flip, tone map and white balance the photos. We use gamma correction with $\gamma = 2.2$ for tone mapping and manually choose a white object in each photo to set the white point through MATLAB.

5 Results

Single-photo enhancement of direct component We first consider the problem of capturing single photos in which the direct component has been contrast-enhanced. Figure 8 shows the captured RAW images for a scene containing a coke bottle filled with milky water. The milky medium scatters light entering within the coke bottle, producing a strong indirect lighting effect.

To do this, we generate a sequence of primal-dual codes using the Bernoulli distribution as described in the Supplementary Materials. These codes satisfy $\mathbf{i}^k = \mathbf{m}^k$, have elements that are 1 with probability p and 0 with probability $1 - p$, and converge to a probing matrix with value p along the diagonal and p^2 everywhere else. Since this matrix attenuates the direct component by a factor of p, the mean brightness of direct-enhanced photos changes with p too. To keep it approximately constant in Figure 8, we compensate for reductions in p by increasing exposure time, *i.e.*, using more codes.

Separating indirect and direct components in the presence of high-frequency indirect transport Figure 12 shows several examples of capturing photos that contain indirect- or direct-only contributions. Each scene contains a variety of effects, including sub-surface scattering, caustics, specular reflections, and diffuse inter-reflections. Note that identifying the high-frequency indirect component in a scene is not possible with existing methods—regardless of the number of input images—because they assume that indirect transport is a low-frequency phenomenon [Nayar et al. 2006]. In contrast, here we correctly capture such effects in one shot. Direct images are computed by subtracting the global image from an image of the scene under white illumination.

The primal-dual codes are similar to those used for direct enhancement: $\mathbf{i}^k$ is drawn from the Bernoulli distribution and $\mathbf{m}^k = \mathbf{1} - \mathbf{i}^k$, resulting in a probing matrix that has 0 along the diagonal and $p(p - 1)$ everywhere else. We use $p = 0.5$ in our experiments because it maximizes the magnitude of the probing matrix.

De-scattering from two shots Fuchs *et al.* [2008] describe an approach for imaging through scattering media that combines con-focal imaging and de-scattering. It requires scanning a line across the scene and capturing, storing and analyzing many photos. We do this in just two shots by implementing their procedure optically.

For this experiment, we replace the 330mm lens with a 220mm lens and focus the camera and projector onto a resolution chart through a tank of milky water (Figure 9). We then take two photos in order to implement the probing matrix for local de-scattering in Table 2. The first photo captures the direct-enhanced component, corresponding to the direct plus $\frac{1}{16}$th indirect, as in Figure 8. The second captures the first off-diagonal of the transport matrix plus $\frac{1}{16}$th of contributions not on that off-diagonal (Equation 8). The difference of the two photos corresponds to a matrix with 1 on the diagonal and -1 on the first off-diagonal. Note that this difference greatly enhances the contrast relative to the floodlit image.

Just like in our direct enhancement experiment, we control the exposure level of our photos by the number of primal-dual codes used to capture them. Since we expect the diagonal entries of the transport matrix to be smaller than those of Figure 8 because of scattering, we double the total exposure time using 1024 codes.

Separating low- and high-frequency indirect transport We can use our probing technique to analyze the transport matrix while blocking the direct component. The indirect-only probing matrix (Table 2) removes any influence of the diagonal on images from the scene. To further analyze the indirect component, we apply the approach of Nayar *et al.* [2006] *only to the indirect component.* To do this, we use probing to optically "simulate" conventional photography on the indirect-only transport matrix, as described in the last row of Table 1. Capturing the indirect component under illumination $\mathbf{i}$ requires slightly modifying the primal-dual codes of the indirect-only probing matrix; we modulate the primal patterns by the illumination pattern $\mathbf{i}$.

This allows us to separate high-frequency indirect light (specularities and caustics) from low-frequency indirect light (inter-reflections and sub-surface scattering) from a total of four photos of a shifting high-frequency illumination pattern, plus four "black" photos used to counteract the transmissivity of the LCD panel.

Figure 10 shows that our algorithm successfully separates the high-frequency specular reflections of the disco ball from the low-frequency inter-reflections and sub-surface scattering of the book and the scattering wax disk, respectively. The number of the primal-dual codes for this experiment was 128 per photo.
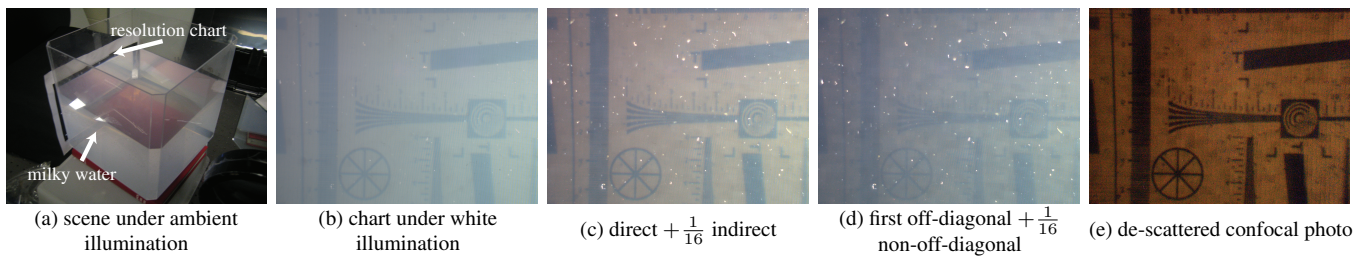
| (a) scene under ambient illumination | (b) chart under white illumination | (c) direct $+ \frac{1}{16}$ indirect | (d) first off-diagonal $+ \frac{1}{16}$ non-off-diagonal | (e) de-scattered confocal photo |

Figure 9: *Imaging a resolution chart through a tank of milky water using a pair of primal-dual coded photos.* **(a)** *A tank filled with water and a small amount of milk scatters light. The objective is to image a resolution chart taped to the back side of the tank using no more than two photos.* **(b)** *Illuminating the scene with a white illumination pattern produces an image with poor contrast because of back-scattering.* **(c)** *A contrast-enhanced direct photo, obtained using the same procedure as in Figure 8, uses* 1024 *primal-dual codes. This produces a photo where the global component is reduced.* **(d)** *Enhancing the first off-diagonal of the transport matrix involves using a shifted version of the primal-dual codes used in (c), according to Equation* (8). **(e)** *By subtracting the off-diagonal-enhanced photo in (d) from the direct-enhanced photo in (c), we obtain a de-scattered confocal photo. This significantly enhances contrast relative to the original floodlit image in (b).*
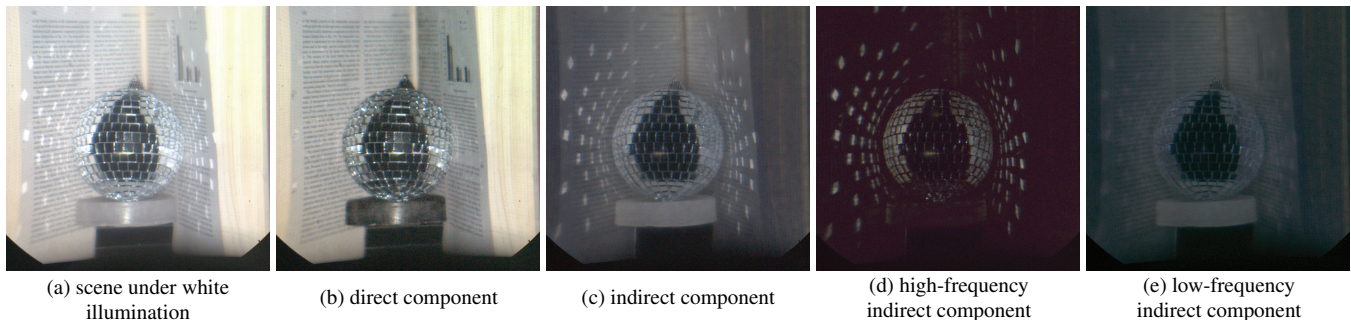

| (a) scene under white illumination | (b) direct component | (c) indirect component | (d) high-frequency indirect component | (e) low-frequency indirect component |

Figure 10: **(a)** *A disco ball ornament placed on top of a wax disk and in front of an open book. The scene is lit by projecting a white pattern. Observe that the open book causes diffuse inter-reflections, the wax disk exhibits strong sub-surface scattering, whereas the disco ball produces specular reflections across the scene.* **(b)** *Primal-dual coding produces a direct component from two photos. This removes light due to inter-reflections, sub-surface scattering and specular inter-reflections even though significant high-frequency light transport effects are present.* **(c)** *We also obtain the indirect component in a single RAW photo; correcting for the black level in this figure requires an additional photo. By capturing the indirect illumination of the scene under shifted high-frequency patterns, it is possible to separate* **(d)** *high-frequency indirect transport (due to specularities) from* **(e)** *low-frequency indirect transport (due to inter-reflections and sub-surface scattering).*

6 Discussion

Limitations of our prototype Our system is fixed to an optical bench, limited to displaying primal-dual codes at 20 Hz, has spherical aberrations, and uses a relatively low-contrast projector and LCD. There is also a small amount of light that transmits through the LCD panel which becomes visible over long exposures; compensating for this light requires a second photo with the mask set to black. Our single-photo experiments assume no ambient light, though ambient light can be eliminated using two photos.

Existing technologies such as digital micro-mirror devices (DMD), liquid crystal on silicon (LCoS) and micro-shutters can speed up primal-dual coding by over two orders of magnitude. DMDs, for instance, can operate at over 10000Hz for binary patterns, opening the possibility of primal-dual coded video. Performing on-chip operations on the camera sensor before read-out would eliminate the need for a mask. Designs for capturing both the light that is blocked by a mask and the light that is unblocked [Heintzmann et al. 2001] can speed up optical processing even more.

Comparison to mask-less multi-image acquisition An alternative way of implementing primal-dual coding is to capture a separate photo for each illumination pattern without using any mask, and then perform masking and integration computationally rather than optically. Such an approach might seem advantageous because it does not require hardware for optical masking and because individually-captured photos might be useful for other purposes. In the Appendix we use an example to show that the utility of such photos is far less clear when acquisition time is taken into account. This is because high-speed multi-image acquisition is limited by *additive sensor noise* whereas primal-dual coding is photon limited, and mainly constrained by the *display rate* of projectors and masks. Looking forward, we expect display technologies to improve much faster than additive sensor noise, conferring further advantage to our all-optical approach. Indeed, using a DMD for masking could enable far higher display rates without moving outside our photon-limited regime.

Primal-dual coding with non-coaxial systems Although our focus has been on analyzing light transport in coaxial systems, primal-dual coding confers new imaging capabilities to non-coaxial systems as well. For instance, in Figure 11 we demonstrate the ability to capture photos where all direct light is blocked from a selected 3D region of physical space. We are currently exploring several such extensions of primal-dual coding.

Future work Finding the optimal patterns for the general probing case is still an unsolved problem. We are exploring optimal decompositions of general probing matrices with optional binary and/or non-negativity constraints. Moreover, although our patterns are currently scene independent, efficiency of numerical matrix probing increases substantially when the structure of the matrix being probed (*e.g.*, its sparsity) is known [Bekas et al. 2007].

On the conceptual side, the basic approach can be extended in several directions. These include extending the basic 2D approach to the case of 4D light fields [Levoy et al. 2006]; exploring its use for 3D reconstruction; using primal-dual coded photos as input
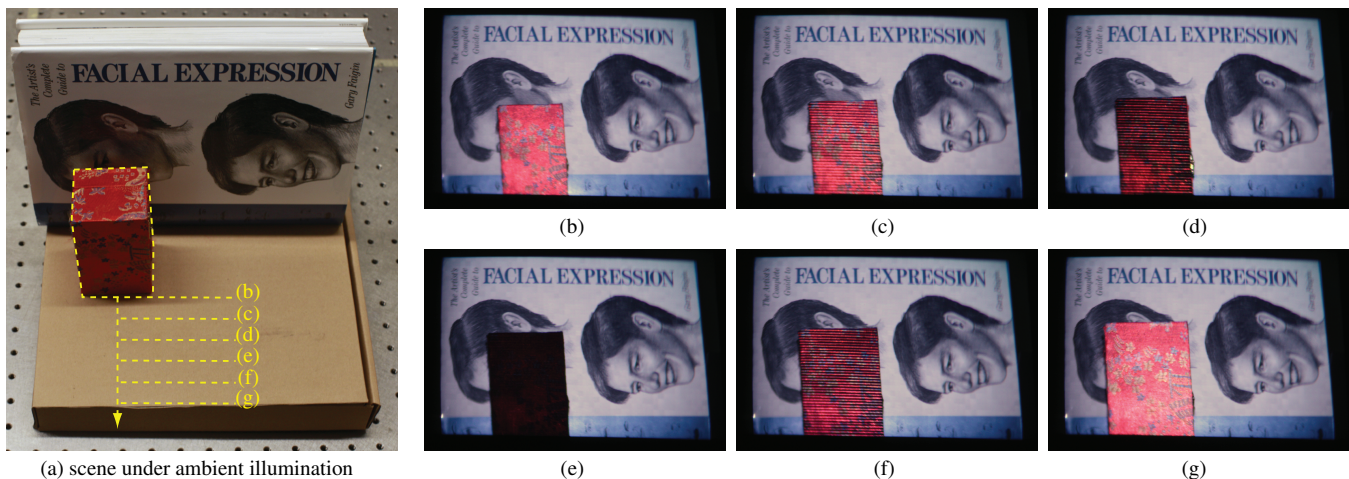
(a) scene under ambient illumination   (b)   (c)   (d)   (e)   (f)   (g)

Figure 11: *Optical z-keying by primal-dual coding. In non-coaxial configurations, the transport matrix encodes both geometry and light transport. By blocking light from a selected range of stereo disparities (Figure 4b) we can capture photos that receive no contribution from objects within the corresponding range of distances. **(a)** An overview of the scene. Dashed lines (b)-(g) correspond to the position of the front plane of the red box in photos (b)-(g). Before taking these primal-dual coded photos, we computed the homography that maps projector pixels to camera pixels for a vertical plane positioned at (e). We then proceeded to probe the transport matrix that maps homography-warped projector pixels to camera pixels (this was done by pre-warping the primal codes before displaying them). The diagonal of this matrix, corresponding to a stereo disparity of 0, represents direct transport paths due to points on the vertical plane at (e). To block light from these paths, we probe with a matrix that excludes the diagonal,* i.e., *the same probing matrix we used to capture the indirect-only photos in Figures 1 and 12. **(b)-(g)** Primal-dual coded RAW photos taken with exactly the camera settings and the same probing matrix but with the red box at different distances from the camera: **(b)** The box is far from the camera and its front plane is fully visible. **(c), (d)** As the box approaches the zero-disparity plane it appears increasingly "vignetted" and disappears completely at (e). **(f), (g)** It reappears again after moving away from the zero-disparity plane. Note that this was achieved without any information about the geometry or appearance of the red box.*

to vision algorithms; and using the basic transport probing technique to efficiently acquire transport matrices for relighting applications [Wang et al. 2009; Sen and Darabi 2009; Peers et al. 2009; Garg et al. 2006].

Finally, note that since many probing matrices require a single shot with no processing, primal-dual coding does not strictly need a camera to operate—it can be a see-through device.

7 Concluding Remarks

In this paper, we introduced primal-dual coding as a general imaging technique that offers fine-grain control over how light paths contribute to a photo. We showed that this greatly increases the available degrees of freedom in photography and that it allows direct acquisition of photos that cannot be captured efficiently by conventional methods. Importantly, the technique operates in the optical domain and requires little or no post-processing after acquisition.

Appendix: One-shot primal-dual coding versus maskless multi-image acquisition

We consider the following hypothetical procedures:

- *PDC:* Capture one optimally-exposed photo with 512 primal-dual codes and a 30-second exposure time at ISO 50 (*i.e.*, slightly more time than needed for a 20 Hz display rate).

- *MI:* Capture 512 photos without a mask, each exposed for $30/512$ seconds at ISO 50 or ISO 1600 (*i.e.*, same total exposure time as PDC and therefore the same total incoming photons).

- *PDC*×10 *and MI*×10: Reduce exposure time by a factor of 10 and use ISO 1600 for both PDC and MI.

These procedures represent the idealized cases of a conventional high-speed system that captures and transfers photos with no overhead and of a spatiotemporally-aligned primal-dual coding system with 100% mask transmissivity (about twice that of our LCD, as afforded by DLP masks). We compare their signal-to-noise ratio (SNR) for a specific camera—the Canon EOS 1D MarkII—whose noise characteristics have been studied in detail [Clark 2007].

For a given pixel, the SNR can be expressed as

$$10 \log_{10} \frac{\text{intensity} \times \text{gain}}{\sqrt{\text{intensity} \times \text{gain} + (\text{additive})^2 + (\text{thermal} \times \text{time})^2}}$$

where pixel intensity is measured in digital numbers; gain is measured in electrons per digital number and depends on the camera's ISO setting; and the terms in the denominator represent the variance of photon noise, additive noise and thermal noise, respectively. According to Clark [2007], the gain at ISO 50 is 26.03 electrons per 12-bit digital number, the standard deviation of the pre-gain readout component of additive noise is 30.62 electrons and thermal noise has an approximate upper bound of 0.25 electrons per second.

To calculate the SNR for the PDC procedure, we assume that the mean intensity of the optimally-exposed primal-dual coded photo is 407.66 12-bit digital numbers. This corresponds to 13% of the maximum possible intensity of 3135.9 and is in accordance with general auto-exposure criteria [Iso 2721:1982]. The mean number of electrons is 10611.57 in this case (against an additive floor of 30.62), yielding an SNR equal to 19.9 dB.

Since a sensor pixel is masked by approximately half the primal-dual codes on average, the electrons in a PDC photo are due to 256 of those codes. This means that these electrons are distributed across 256 photos in the MI procedure, or 41.45 electrons per photo (against the additive floor of 30.62 electrons). Thus, the MI procedure is dominated by additive noise and has an extremely low SNR, just 1.22 dB. This makes the captured photos practically unusable on their own. Even at ISO 1600, which has the lowest pre-gain readout noise at 3.9, the SNR is just 7.4 dB.

We now examine the SNR of their masked sum according to Equation 5. Since only 256 photos contribute to this sum on average, we must count the contribution of additive noise only from those photos. The standard deviation of additive noise will therefore be 16 times larger, or 489.92 electrons at ISO 50 and 62.4 at ISO 1600. This yields an SNR of 13.2 dB at ISO 50 and 19.4 dB at ISO 1600. Thus, despite the significant computational burden of acquiring, transferring and storing 512 individual photos, these photos are not very useful on their own and do not offer an SNR gain over a well-engineered primal-dual coding system.

Finally, applying PDC$\times$10 and MI$\times$10 means that ten times fewer electrons contribute to a pixel. While the resulting under-exposed photo has an SNR of 15.1 dB under PDC$\times$10, MI$\times$10's reduction to 11.8 dB for the masked sum is much more severe.

References

BEKAS, C., KOKIOPOULOU, E., AND SAAD, Y. 2007. An estimator for the diagonal of a matrix. *Appl. Numer. Math. 57*, 11-12, 1214–1229.

CHANDRAKER, M., NG, T., AND RAMAMOORTHI, R. 2010. A dual theory of inverse and forward light transport. In *Proc. ECCV*.

CLARK, R. N. 2007. Canon 1D Mark II analysis. In *http://wwwclarkvisioncom/articles/evaluation-1d2/*.

CORLE, T. R., AND KINO, G. S. 1996. *Confocal scanning optical microscopy and related imaging systems*. Academic Press.

DEBEVEC, P., HAWKINS, T., TCHOU, C., DUIKER, H., SAROKIN, W., AND SAGAR, M. 2000. Acquiring the reflectance field of a human face. *ACM SIGGRAPH*, 145–156.

FUCHS, C., HEINZ, M., LEVOY, M., SCIDEL, H.-P., AND LENSCH, H. P. A. 2008. Combining confocal imaging and descattering. *Computer Graphics Forum 27*, 4, 1245–1253.

GARG, G., TALVALA, E.-V., LEVOY, M., AND LENSCH, H. P. A. 2006. Symmetric photography: Exploiting data-sparseness in reflectance fields. In *Proc. EGSR*, 251–262.

GHOSH, A., CHEN, T., PEERS, P., WILSON, C. A., AND DEBEVEC, P. 2010. Circularly polarized spherical illumination reflectometry. *ACM SIGGRAPH Asia*.

GUPTA, M., AGRAWAL, A., VEERARAGHAVAN, A., AND NARASIMHAN, S. 2011. Structured light 3D scanning in the presence of global illumination. In *Proc. CVPR*, 713–720.

HEINTZMANN, R., HANLEY, Q., ARNDT-JOVIN, D., AND JOVIN, T. 2001. A dual path programmable array microscope (PAM): simultaneous acquisition of conjugate and non-conjugate images. *J. Microscopy 204*, 119–135.

HITOMI, Y., GU, J., GUPTA, M., MITSUNAGA, T., AND NAYAR, S. K. 2011. Video from a single coded exposure photograph using a learned over-complete dictionary. In *Proc. ICCV*.

ISO 2721:1982. Photography—Cameras—Automatic controls of exposure.

KIRMANI, A., HUTCHISON, T., DAVIS, J., AND RASKAR, R. 2011. Looking around the corner using ultrafast transient imaging. *Int. J. Computer Vision 95*, 1, 13–28.

LEVOY, M., CHEN, B., VAISH, V., HOROWITZ, M., MCDOWALL, I., AND BOLAS, M. 2004. Synthetic aperture confocal imaging. *ACM SIGGRAPH*, 825–834.

LEVOY, M., NG, R., ADAMS, A., FOOTER, M., AND HOROWITZ, M. 2006. Light field microscopy. *ACM SIGGRAPH*, 924–934.

MERTZ, J. 2011. Optical sectioning microscopy with planar or structured illumination. *Nat. Meth. 8*, 10, 811–819.

NAYAR, S. K., BRANZOI, V., AND BOULT, T. 2004. Programmable imaging using a digital micromirror array. In *Proc. CVPR*, 436–443.

NAYAR, S. K., KRISHNAN, G., GROSSBERG, M. D., AND RASKAR, R. 2006. Fast separation of direct and global components of a scene using high frequency illumination. *ACM SIGGRAPH*, 935–944.

NG, R., RAMAMOORTHI, R., AND HANRAHAN, P. 2003. All-frequency shadows using non-linear wavelet lighting approximation. *ACM SIGGRAPH*, 376–381.

O'TOOLE, M., AND KUTULAKOS, K. N. 2010. Optical computing for fast light transport analysis. *ACM SIGGRAPH Asia*.

PEERS, P., MAHAJAN, D. K., LAMOND, B., GHOSH, A., MATUSIK, W., RAMAMOORTHI, R., AND DEBEVEC, P. 2009. Compressive light transport sensing. *ACM Trans. on Graphics 28*, 1.

POPOFF, S. M., LEROSEY, G., CARMINATI, R., FINK, M., BOCCARA, A. C., AND GIGAN, S. 2010. Measuring the transmission matrix in optics. *Phys. Rev. Lett. 104*, 10.

SCHECHNER, Y. Y., NAYAR, S. K., AND BELHUMEUR, P. N. 2007. Multiplexing for optimal lighting. *IEEE T-PAMI 29*, 8, 1339–1354.

SEITZ, S. M., MATSUSHITA, Y., AND KUTULAKOS, K. N. 2005. A theory of inverse light transport. In *Proc. ICCV*, 1440–1447.

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

SEN, P., CHEN, B., GARG, G., MARSCHNER, S., HOROWITZ, M., LEVOY, M., AND LENSCH, H. P. A. 2005. Dual photography. *ACM SIGGRAPH*, 745–755.

SLOAN, P.-P., KAUTZ, J., AND SNYDER, J. 2002. Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. *ACM SIGGRAPH*, 527–536.

TANG, J. M., AND SAAD, Y. 2012. A probing method for computing the diagonal of a matrix inverse. *Numer. Linear Algebra Appl. 19*, 3, 485–501.

VEERARAGHAVAN, A., REDDY, D., AND RASKAR, R. 2011. Coded strobing photography: compressive sensing of high speed periodic videos. *IEEE T-PAMI 33*, 4, 671–686.

WANG, J., DONG, Y., TONG, X., LIN, Z., AND GUO, B. 2009. Kernel Nyström method for light transport. *ACM SIGGRAPH*.

WETZSTEIN, G., HEIDRICH, W., AND LUEBKE, D. 2010. Optical image processing using light modulation displays. *Computer Graphics Forum 29*, 6, 1934–1944.

WILSON, T., JUŠKAITIS, R., AND NEIL, M. 1996. Confocal microscopy by aperture correlation. *Optics Letters 21*, 23.

ZHANG, L., AND NAYAR, S. K. 2006. Projection defocus analysis for scene capture and image display. *ACM SIGGRAPH*, 907–915.

| scene under white light | global component | direct + back-scattering + retro-reflection |

Figure 12: *Left to right: Photo of scene under white light, global component of scene captured by a single RAW photo, direct component computed by subtracting the global image from the scene under white light. No processing was done on the photos, other than a resize, flip, white balance, gamma correction and crop. All scenes use 512 primal-dual codes, with the exception of Row 4 which uses 1024 codes.*
Row 1: *A disco ball produces high frequency light transport across the scene, with diffuse inter-reflections occurring near the edge of the floor. Primal-dual coding cleanly separates the photo into its direct component and its global component, which contains inter-reflections and specularities.* ***Row 2:*** *A scene composed of two styrofoam balls inside a glass container generates specular reflections, sub-surface scattering and caustics. We used small aperture to photograph this scene in order to avoid pixel saturation due to caustics. The photo is slightly noisier as a result.* ***Row 3:*** *A coke bottle filled with milky water strongly scatters light that enters the bottle.* ***Row 4:*** *Observe that the direct component contains none of the light that is reflected by the mirror (left) onto the vase (right) and the wall. However, the mirror does not appear black in the direct photo even though it has little to no direct component. This is because of back-scattering and retro-reflection (Figure 4c): light leaves a projector pixel, hits the mirror at a point, diffusely reflects off the wall, then hits the mirror at the* same *point, and returns to the same pixel. This light path will always contribute to the diagonal of the transport matrix and cannot be distinguished from direct illumination using methods that rely purely on intensity, and without regard to arrival time or polarization. Note that according to Equation 7, the variance of stochastic diagonal estimation may be higher for indirect specular transport than for diffuse transport. To compensate for this, we used 1024 primal-dual codes for this scene.*