

# Acquiring and Characterizing Plane-to-Ray Indirect Light Transport

Hiroyuki Kubo<sup>1,2</sup>, Suren Jayasuriya<sup>2,3</sup>, Takafumi Iwaguchi<sup>1,2</sup>,

Takuya Funatomi<sup>1</sup>, Yasuhiro Mukaigawa<sup>1</sup>, Srinivasa G. Narasimhan<sup>2</sup>

<sup>1</sup>Nara Institute of Science and Technology, <sup>2</sup>Carnegie Mellon University, <sup>3</sup>Arizona State University

hkubo@is.naist.jp

## Abstract

Separation of light transport into direct and indirect paths has enabled new visualizations of light in everyday scenes. However, indirect light itself contains a variety of components from subsurface scattering to diffuse and specular interreflections, all of which contribute to complex visual appearance. In this paper, we present a new imaging technique that captures and analyzes these components of indirect light via light transport between epipolar planes of illumination and rays of received light. This plane-to-ray light transport is captured using a rectified projector-camera system where we vary the offset between projector and camera rows (implemented as synchronization delay) as well as the exposure of each camera row. The resulting delay-exposure stack of images can capture live short and long-range indirect light transport, disambiguate subsurface scattering, diffuse and specular interreflections, and distinguish materials according to their subsurface scattering properties.

## 1. Introduction

The light transport in a scene is the result of complex interactions between illumination and scene objects' geometry and materials. Light can bounce directly from illumination source to object to camera, or undergo multiple, indirect bounces amongst objects, or get scattered within particulate medium, before being captured by the camera. Light striking an object can be absorbed, reflected, or refracted.

Imaging different light transport paths improves algorithms in computer graphics and vision. Many active 3D scanning systems rely on a direct light/single bounce assumption to accurately estimate depth and/or surface normals [10]. Rendering the complex appearance of materials such as wax and human skin necessitates high fidelity modeling of subsurface scattering [12]. Finally, multi-bounce indirect light can yield information about objects outside the line-of-sight of the camera [32].

Capturing the full light transport for a scene involves

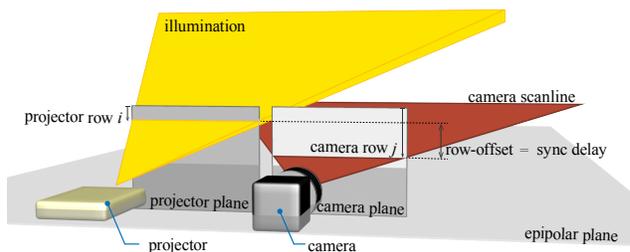


Figure 1: A rectified projector-camera system to implement plane-to-ray light transport. The laser projector illuminates the scene with a plane that is swept vertically. The camera synchronizes its rolling shutter to a fixed row-offset from the illumination plane. As each projector row advances, so does the exposed camera row. By varying both the delay between synced camera and projector rows as well as the camera row exposure, we can capture live various components of epipolar and non-epipolar light transport.

acquiring a large amount of measurements and memory for storage. To alleviate this, light transport components such as direct and indirect light have been captured directly. Direct-global separation [18] decomposed light transport into low (mostly global) and high-frequency (mostly direct) components. Indirect light has been further separated into short and long-range optical path lengths via homogeneous codes for primal-dual imaging [20]. Recent work has focused on the specific relationship between light transport and the epipolar geometry of projector-camera systems [20, 22]. *Epipolar light* is defined as light on the same epipolar plane aligned with the rows of a rectified projector-camera system. Note that epipolar light contains mostly direct light, which by definition must lie on the epipolar plane, but also may contain indirect light whose final bounce lies on the epipolar plane. Non-epipolar light necessarily contains only indirect light.

In this paper, we present a new imaging technique which exploits the light transport between planes of illumination and camera pixels for capturing and analyzing indirect light paths. In our scenario, a projector illuminates a plane of

light corresponding to a projector row as shown in Figure 1. Thus the full illumination is a set of vertically spaced planes defined by the projector rows. This light is received at each camera pixel for a given camera row. Since the camera and projector are rectified in an epipolar configuration, the difference between epipolar and non-epipolar light paths is determined by the vertical separation between projector and camera rows shown in Figure 1. By controlling light transport from illumination plane to camera ray, we can selectively image these indirect light paths.

Every camera ray is determined by the transport of 1D planar illumination from the projector to a 2D pixel. We observe that a rolling shutter camera, synchronized to the projector and in an epipolar configuration, performs illumination multiplexing of this transport using the parameters of (1) *synchronization delay or row offset* between the illuminated projector row and exposed camera row, and (2) *exposure* of the camera row itself. Throughout this paper, we refer to synchronization delay or row offset interchangeably<sup>1</sup>. In our new imaging technique, we capture image stacks of varying sync delay and exposure to analyze and capture different light transport paths in the scene.

In particular, we make the following contributions:

- Capture of short and long-range indirect light using varying delay-exposure image stacks.
- Analyzing delay profiles for subsurface scattering, diffuse and specular interreflections.
- Demultiplexing image stacks to recover 3D light transport between the projector and camera (1D light plane  $\times$  2D camera pixel).

We demonstrate applications including sharpened (coarse-to-fine) epipolar imaging, separation of epipolar-direct and epipolar-indirect light using high frequency illumination, and differentiating materials according to their subsurface scattering properties using their delay profiles.

We validate our ideas on real experimental data from the epipolar imaging system, Episcan3D (the system is described in Section 5 of [20]). This research exposes a new set of parametric control for these active illumination systems through the knobs of delay and exposure. This enables even more selectivity of the light transport paths we probe and capture.

## 2. Related work

**Light Transport and its Components:** There is a body of literature on the simulation and theory of light transport [13, 34]. Light transport matrices [4, 7, 19] have been used to describe the linear relationship between illumination sources and detectors. Further research to

efficiently acquire these matrices has focused on compressive sensing to acquire these large matrices at reduced measurements [24, 30], approximation via symmetry [6], and optical computing via matrix-vector products [21]. Methods have decomposed light transport into bounces of light [28] and its sub-components of light [2].

Nayar et al. showed that direct/global separation can be achieved using high spatial frequency illumination (e.g. a checkerboard pattern) [18]. This separation could be shown even for subsurface scattering materials with a high spatial frequency. Bounce decomposition can also yield direct/global separation [28], and the use of primal-dual coding allows for separation in live video [22, 23], and homogeneous codes decomposed indirect light into short and long-ranges [20]. In addition, diffuse and specular reflections have been separated using color and polarization cues [15, 16, 17, 33]. Similar to our paper but with time-resolved measurements, Wu et al. used temporal delay profiles to analyze indirect light transport and material properties [35]. In this work, we analyze components of indirect light through 3D light transport between projector and camera in a scene.

**Multiplexed Illumination and Exposure:** Several methods have been introduced to multiplex illumination for light transport acquisition. These include Hadamard coding for high SNR [27] and optimal multiplexing [26]. On the camera side, coded exposure methods enable motion deblurring [25], video from a single coded exposure [11], rolling shutter photography [8] and space-time voxels [9]. In this work, we utilize image stacks with varying parameter of delay and exposure to demultiplex illumination into 3D light transport.

**Imaging Systems for Light Transport Acquisition:** The light stage [3] uses a large installation of light sources and cameras to acquire light transport for human subjects. At a smaller scale, several smaller prototype systems have been created including regular projector-camera systems [29], coded exposure [9, 11], and primal-dual coding via digital micromirror devices [23]. For temporally synced projector-camera systems, researchers have modified DLP projectors for fast projection [14].

**Epipolar Imaging:** Episcan3D uses the syncing between a laser projector and the rolling shutter of a camera to perform real-time epipolar and non-epipolar imaging [20], and EpiToF extends this to a time-of-flight device [1]. We note that the synchronization of the camera’s rolling shutter to a raster-scanning projector to capture epipolar and non-epipolar light is the contribution of [20]. However, the insight that this mechanism can be interpreted as illumination multiplexing, and we can utilize delay-exposure

<sup>1</sup>This delay is unrelated to the temporal delay of transient imaging.

stacks to demultiplex and recover 3D light transport is novel to this paper. Further, we show new applications of epipolar imaging including epipolar sharpening, synthetic 1D relighting, enhanced direct/global separation, and material recognition for subsurface scattering.

### 3. Plane-to-ray Light Transport

In this section, we describe the principles of our imaging modality. We first expisit planar illumination and its light transport. We then show how light multiplexing occurs using the delay and exposure of a temporally-synced projector and rolling shutter camera system. We derive expressions for these parameters, and use them to model the illumination function and image stack formation.

#### 3.1. Epipolar Geometry and Light Transport

We follow the work of O’Toole et al. [20, 22] to specify the epipolar geometry of the projector-camera. In Figure 1, the projector and camera are rectified so that their rows are aligned on the same epipolar plane. Intuitively, this divides the scene into a stack of epipolar planes, each which contains the epipolar line connecting the projector and camera centers as well as a unique projector/camera row.

The relationship between light transport and epipolar geometry is determined by this alignment, and was first described in [20]. Direct light, which must lie on the same epipolar plane since it is defined by the intersection of the projector and camera rays, can only travel from projector row  $i$  to the same camera row  $i$ . Indirect light, which has undergone multiple bounces, can travel from projector row  $i$  to any camera row  $j$ . However, the amount of indirect light on the epipolar plane, light that has undergone multiple bounces but still travels from projector row  $i$  to camera row  $i$ , is a small percentage of the total indirect light (unless there are strong specular interreflections in the scene).

Thus, the planar light transport from illumination to camera can be parameterized by the relative offset between projector row  $i$  and camera row  $j$ . This row offset is controlled by the **synchronization delay**, the timing difference between the synchronized projector scanning and camera rolling shutter. In addition, the **exposure** of the camera row determines the amount of integrated light at the camera row.

In an ideal planar illumination system, one could capture the light transport by projecting one line at a time and taking an image. However, this impulse scanning suffers from low SNR due to low light levels [27]. This would particularly affect the capture of light paths such as subsurface scattering and long-range indirect light. In addition, in a real system, the laser itself has a temporal jitter, which may cause light to leak into neighboring rows as noticed in [20]. To solve these issues, we use light multiplexing as a way to increase the SNR for light transport acquisition.

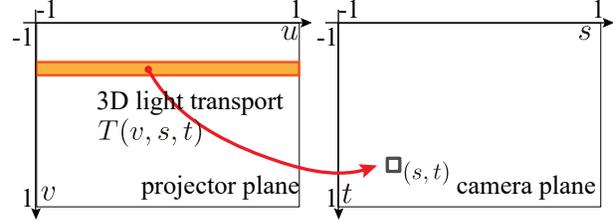


Figure 2: Illustration of 3D light transport  $T(v, s, t)$  from row  $v$  in projector plane to a camera pixel  $(s, t)$ .

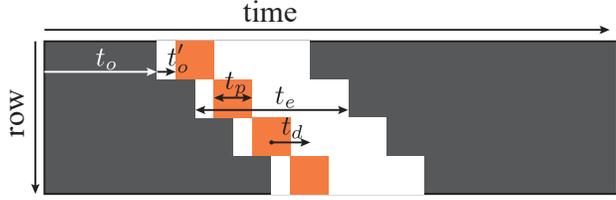
In our imaging system, we utilize a rolling shutter camera to capture these planes of light. Our key insight is that this rolling shutter, synchronized to the projector, performs light multiplexing for planar illumination. We now proceed to describe this light multiplexing using the parameters of delay and exposure in a rolling shutter system.

#### 3.2. Light Multiplexing using Delay and Exposure

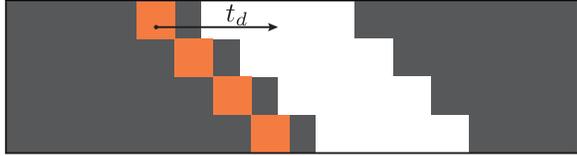
For a rolling shutter camera synchronized to the epipolar illumination of the projector, we can control the delay and exposure of this shutter to perform light multiplexing. The exposure determines the number of rows being exposed with larger exposures leading to larger sets of rows being exposed. The delay is the distance between the illuminated projector row and the center of the exposed rows.

The rolling shutter of the camera can be synchronized to the projector illumination, as described in [20]. In particular, this means the pixel clock is fixed and focal length of the lens adjusted so that the projector rows and camera rows change with the same vertical velocity. In epipolar imaging mode, the delay is zero, so that the band of exposed camera rows is on the same epipolar plane as the light being projected, while in non-epipolar mode, the band of exposed camera rows does not include the epipolar plane where the light is. Light multiplexing occurs since each row gets light from multiple projector lines due to the width of the exposure and the value of the delay. To describe the demultiplexing algorithm in Section 4 necessary to estimate 3D light transport, we first must derive the relationship between delay and exposure, and use it to model the illumination.

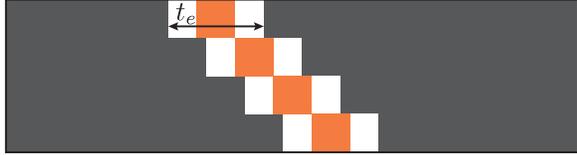
**Relationship between delay and exposure:** We use the same notation as [20] to parametrize delay and exposure in a rolling shutter system. Let  $t_p$  denote the amount of time for which the projector illuminates a single scanline (with some finite band width),  $t_e$  be the exposure time which corresponds to a contiguous block of rows being exposed, and  $t_o$  denote the time offset of synchronization between the projector and camera. Additionally, we denote  $t'_o$  as the time difference between the start of exposure and when the projector illuminates that row of pixels. Please see Figure 3 for a visual description of these parameters.



(a) The projector illuminates a single row for a time  $t_p$  in orange. At the same time, the rolling shutter exposes a single row for length  $t_e$ . Light from a single row (orange) will be captured not just by the same camera row, but rows above and below it that are being exposed (white). Delay  $t_d$  is the distance from center of exposure to the center of illumination.  $t'_o$  is the time between start of exposure and illumination in a row, and  $t_o$  is the synchronization offset from projector to camera.



(b) As the delay is increased, the illuminated projector row sends light to camera rows that are at least one row above it. This corresponds to short range non-epipolar light paths in the scene.



(c) As the exposure is decreased, the illuminated projector row leaks less light into neighboring rows, resulting in a majority of epipolar light paths captured.

Figure 3: Timing diagram of projector illumination and camera rolling shutter for epipolar imaging.

As we change  $t_o$ , this changes  $t'_o$ , and thus we express delay:  $t_d$  as the difference between the center times of exposure and illumination, as following:

$$t_d = \frac{1}{2}t_e - \frac{1}{2}t'_o. \quad (1)$$

Positive  $t_d > t_e/2$  means the camera row receives light from a vertically lower epipolar plane. Similarly, negative  $t_d < -t_e/2$  means the light arrives from a vertically higher epipolar plane. If  $0 \leq |t_d| \leq t_e/2$ , then the exposed row receives a majority of illuminated light from the same epipolar plane. Typically, epipolar imaging operates with  $t_d = 0$  and  $t_e$  as short as possible (as shown in Fig. 3(c)).

**Illumination Model:** We formulate a model for the illumination as a function of delay and exposure. Using calibration, we obtain the speed of the projector scanline  $v_p$  sec/line in the scene. Given this, we express illumination band width  $I_w$  and its center location  $I_d$  by the following

equations:

$$I_w(t_e) = v_p t_e, \quad I_d(t_d) = v_p t_d. \quad (2)$$

Let  $v$  denote a row of the projector plane. We then define the illumination function  $L(v, t_d, t_e)$ :

$$L(v, t_d, t_e) = \begin{cases} 1, & \text{if } \|v - I_d(t_d)\| < \frac{1}{2}I_w(t_e); \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Note that we define the maximum intensity of the projector as 1. We will use this illumination function to estimate 3D light transport amongst rows of the projector/camera in Section 4.

### 3.3. Delay-Exposure Image Stacks

We thus capture a series of images while varying delay  $t_d$  and exposure  $t_e$ . We typically use uniformly sampled points between minimum and maximum values for both delay and exposure as part of our sweep.

By controlling the delay and exposure, we have the ability to capture short and long-range non-epipolar light. As the delay increases, light from the illumination plane has to travel a longer vertical distance to reach the camera row. This gives a minimum bound of the optical path length traveled by the indirect light. By controlling the exposure, we can allow more or less amount of light that has traveled this minimum bound, thus creating a *band* of non-epipolar light. This corresponds to banded diagonals off the main diagonal in a light transport matrix [20, 22]. In Section 6.1 we demonstrate results of this banded imaging.

## 4. 3D Light Transport Estimation via Demultiplexing

In this section, we use delay-exposure image stacks to perform illumination demultiplexing and estimate 3D light transport in the scene. For given  $t_d$  and  $t_e$ , the observation  $I$  at pixel  $(s, t)$  is given by a convolution of the illumination with the light transport operator:

$$I(s, t) = L(v, t_d, t_e) * T(v, s, t). \quad (4)$$

Note that,  $T(v, s, t)$  is 3D light transport from row  $v$  to a pixel  $(s, t)$ . This relationship is illustrated in Fig. 2. We note that this equation can be discretized to the standard matrix-vector product of light transport.

We can write the epipolar and non-epipolar images by the following convolutional equations:

$$I_e(s, t) = \delta(t - v) * T(v, s, t), \quad (5)$$

$$I_n(s, t) = (1 - \delta(t - v)) * T(v, s, t). \quad (6)$$

Hence, if we can estimate  $T$  from the image stack of varying  $t_d$  and  $t_e$ , we can synthesize epipolar and non-epipolar images. We denote the  $i$ -th image with delay  $t_{d,i}$

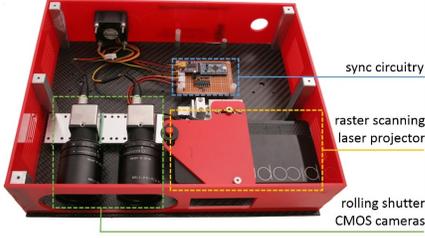


Figure 4: The Episcan3D prototype [20] which we use for our epipolar imaging system.

and exposure  $t_{e,i}$ . Thus we estimate the light transport  $T^*$  as the solution to the following optimization problem:

$$\min_{T(v,s,t)} \sum_i^N \|I_i - \{L(v, t_{d,i}, t_{e,i}) * T(v, s, t)\}\|_2^2 + \alpha E_c + \beta E_s$$

$$\text{subject to } T \geq 0, \forall v, E_c = \|\frac{\partial}{\partial v} T\|_2^2, E_s = \|T\|_1.$$

We use additional regularization for smoothness and sparsity in the light transport:  $\alpha$  and  $\beta$  are coefficients of smoothness and sparsity respectively. This helps with the optimization to reduce noise and other image artifacts. The total number of images in the stack is  $N$ . In practice, we utilize  $\alpha = 0.01$ ,  $\beta = 0.01$ , and  $N = 75$ .

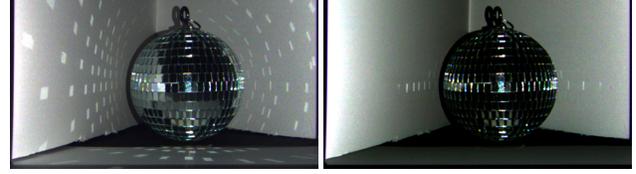
Since the formulation is per-pixel, the optimization is easily parallelizable. We use the CVXPY framework for convex optimization to solve this [5]. We feed most delay-exposure images to the solver except for those delay images which lie on the boundary between epipolar and non-epipolar imaging ( $t_d \approx t_e/2$ ), which have significant horizontal artifacts due to synchronization problems. One limitation of our algorithm is the sparsity condition prevents recovering dense light transport effects.

## 5. Hardware Implementation

To implement our ideas, we utilize the Episcan3D system described in [20]. We also discuss calibration and acquisition process for capturing delay-exposure image stacks.

**Prototype:** We utilize a similar prototype to [20] and shown in Figure 4. We use a Celluon PicoPro as a projector with display resolution  $1280 \times 720$ , and an IDS UI-3250CP-C-HQ color camera with both global and rolling shutter capabilities. The shutter is triggered by the VSYNC signal generated by the projector, and also can be delayed by the camera using  $t_o$ . We refer the reader to the supplemental material of [20] about the support circuitry and physical alignment required for the prototype.

**Calibration:** For our calibration procedure, it is necessary to determine the illumination bandwidth  $I_w$  and center location  $I_d$  with respect to  $t_e$ . From Equation 2, we see that



(a) regular

(b)  $t_d = 0\mu s, t_e = 300\mu s$



(c)  $t_d = 1200\mu s, t_e = 450\mu s$ , (scaled  $\times 3$  for visualization)

(d)  $t_d = 1200\mu s, t_e = 2000\mu s$ , (scaled  $\times 3$  for visualization)

Figure 5: Using delay and exposure, we can capture bands of indirect light. In this scene, we show the (a) a regular image of a disco ball, (b) the epipolar image, (c) a band of indirect light at a fixed delay, (d) the band of indirect light with increased exposure at the same delay.

we need to estimate  $v_p$ . To do this, we project a single-pixel horizontal white line on a black background, and sweep the line vertically from top to bottom. We image this sweep in global shutter mode with  $t_e = 500\mu s$ . By counting the number of visible lines  $n_v$ , we obtain the projector scanline velocity  $v_p = n_v/t_e$ .

**Acquisition:** To acquire images, we use 12 bit capture and average 8 images for each delay and exposure. We performed no gamma correction ( $\gamma = 1.0$ ) to ensure our image measurements were linear. For a typical sweep, we used delay  $t_d = -1500\mu s$  to  $t_d = 1500\mu s$ , and exposure  $t_e = 600\mu s$  to  $t_e = 1200\mu s$ .

Total acquisition time is as follows: Calibration takes approximately 4 minutes (one-time only process), acquisition for 75 images (a typical delay-exposure stack) takes approximately 2 minutes where each image is an average of 8 frames, and it takes 9 minutes to demultiplex these images for  $128 \times 128$  resolution.

## 6. Experimental Results

### 6.1. Short and Long-range Indirect Light

As described earlier, controlling the delay and exposure can selectively image bands of indirect light in a scene. In Figure 5, a mirrored disco ball is illuminated in the scene. With  $t_d = 0\mu s$ , mostly direct light is captured, but more indirect light leaks in as the exposure is increased from  $450\mu s$  to  $1200\mu s$  (Figure 5(b-c)). If the delay is increased to  $800\mu s$ , we can visualize the vertical distance separating the projector and camera rows in Figure 5(d). We note that this effect is particularly noticeable due to the brightness of

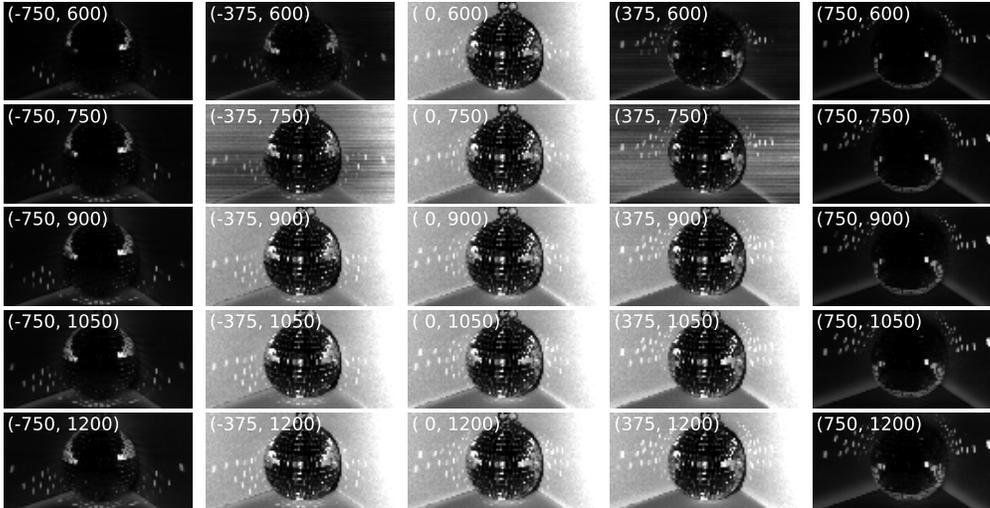


Figure 6: An image stack of varying delay on the x-axis and exposure on the y-axis, with values  $(t_d, t_e)$  in microseconds. Epipolar images occur when  $|t_d| < t_e/2$ , and non-epipolar or indirect light otherwise. Notice how the specular interreflections of the disco ball move vertically as delay increases in one row corresponding to indirect light paths jumping from illumination to received plane. The size of the interreflections band is controlled by exposure. All images are brightened for visualization.

specular interreflections of the disco ball, and other diffuse interreflections cannot have tight indirect bands due to low exposure. Please view a video of sweeping the band in the supplemental material.

## 6.2. Acquired Image Stacks

We captured several image stacks sweeping exposure and delay with uniform increments. In Figure 6, we visualize the specular interreflections of a disco ball shifting vertically as the delay changes.

Noise is primarily determined by the amount of light reaching the pixels (although there are synchronization artifacts at very short exposure times due to jitter in the laser raster scan). For indirect imaging, specular interreflections (such as the disco-ball reflections) are brighter and thus less noisy than diffuse interreflection or subsurface scattering effects. Since exposure is coupled to the band of light received by the rolling shutter, there is a tradeoff between integrating more light and the tightness of the band of indirect light (i.e. the resolution of the illumination function).

## 6.3. Delay Profiles

For each pixel, we can plot the pixel intensity as a function of the delay  $t_d$ . We call this a *delay profile*, and it yields information about the scattering of light with respect to the planar illumination of the projector. Delay profiles look qualitatively different for subsurface scattering and diffuse interreflections, which are short-range indirect light effects, versus specular interreflection that has long range. We note that Wu et al. performed a similar analysis using temporal

delay for time-of-flight imaging [35].

In Figure 7, we image a scene with a variety of these effects and show their delay profiles. Note how specular interreflections from the mirror ball (blue) have two peaks in their delay profile. This is due to a diffuse reflection from the page at  $t_d = 0$  coupled with a peak from the specular reflection of the mirror ball. For the near corner of the book (red) and candle (yellow), their broadened delay profiles are due to subsurface scattering. The more translucent the object, the more broader its delay profile (see also milk results in Section 6.7). Note that the delay profiles are not symmetric around zero as one would expect, but are affected by the surface geometry/surface normal at those points. This relationship between symmetry and surface normal is a subject of further investigation.

In addition to being qualitatively different, delay profiles can also help identify materials based on their scattering properties as shown in Section 6.7.

## 6.4. Sharpening epipolar imaging

In epipolar imaging, there is an inherent trade off between the amount of non-epipolar light that leaks into the signal and exposure  $t_e$ . Thus it is difficult to capture epipolar images with large exposure as the amount of non-epipolar light inside the epipolar image scales with  $t_e - t_p$ .

However, as noted in Equation 6, if we can resolve light transport to a fine resolution in projector rows  $v$ , we can synthesize a “sharper” epipolar image. We can computationally render an epipolar image to the limit of the light’s illumination width  $I_w$ .

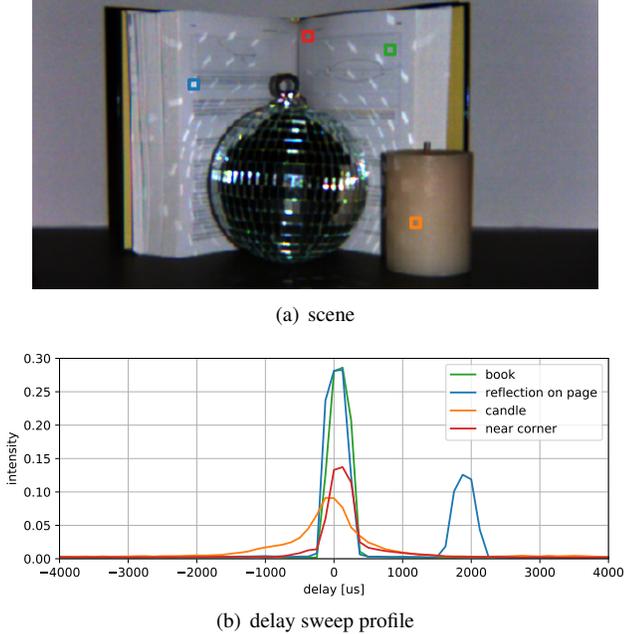


Figure 7: We perform a delay sweep of the scene shown in (a), and plot the delay profiles for selected pixels in the image (b). Note how subsurface scattering material like the candle has a wide broadening profile (orange), while diffuse interreflection in the near corner has a steeper profile (red). The diffuse reflection from the book page itself has an unimodal peak (green), but the specular interreflection has a bimodal peak (blue).

In Figure 8, we image a rose candle made of translucent wax. We synthesize in Figure 8(c) a tighter epipolar image than a regular epipolar image with exposure  $t_e = 600\mu s$  shown in Figure 8(b). Note how the regular epipolar image cannot remove the subsurface scattering of the candle, but the sharpened epipolar image removes all these effects. Looking at the cross-section pixel values in Figure 8(d), the sharper epipolar image has more contrast amongst its rose petals. This sharpening has applications for when the system has a large exposure, and thus needs computation to generate a tighter epipolar image.

### 6.5. Relighting

In addition, 1D light transport allows us to synthesize novel images. For instance, we can render a new image with a novel illumination pattern of any linear combination of projector rows using the  $T$  operator. In Figure 9, we synthesize relighting from a single line illumination for the imaged rose. Please see the supplemental material for a video of this relighting effect.

For this single line relighting of the scene, we compare our method in Figure 9(f) versus conventional imaging techniques. We show the comparison against a single projected

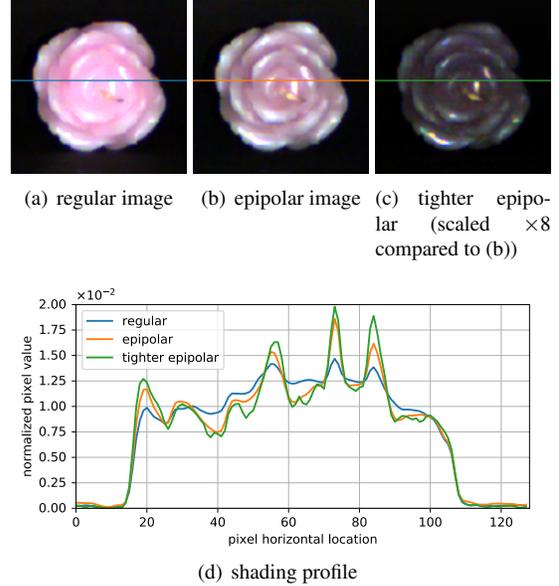


Figure 8: Imaging a wax rose candle in epipolar mode (b) with an exposure of  $600\mu s$  does not remove the subsurface scattering. Demultiplexing the image stack to recover 3D light transport, we synthesize a tighter epipolar image in (c) which preserves sharp features and highlights while removing the subsurface scattering from the epipolar image. In (d), we plot pixel values for a single scan line for comparison, note how the tighter epipolar image has larger contrast.

line with exposure of 16ms in Figure 9(g), and a single line projection with exposure of 800ms in Figure 9(h). Note that our method achieves better noise performance than Figure 9(g) since we utilize multiplexed illumination to capture our delay-exposure stack. Our method achieves similar performance to Figure 9(h) in terms of noise, but requires multiple images and does not capture the long range light transport effects for far away rows due to the sparsity assumption in our optimization.

### 6.6. Epipolar Direct/Global Separation

One of the disadvantages of epipolar imaging in this configuration is that it is difficult to separate epipolar indirect light from the image, and thus achieve true direct/global separation. To solve this, we thus apply the method from Nayar et al. [18] on the epipolar images alone. We used 128 shifting patterns of  $24 \times 24$  pixel checkerboard for our implementation.

In Figure 10, we perform direct/global separation on a scene consisting of a wax bowl and disco ball. We note that the method from Nayar et al. fails on removing the specular interreflections of the disco ball as shown in Figure 10(b). Epipolar imaging thus improves upon Nayar et al. by removing these highlights in Figure 10(c). Combining the two methods results in an epipolar-direct image in

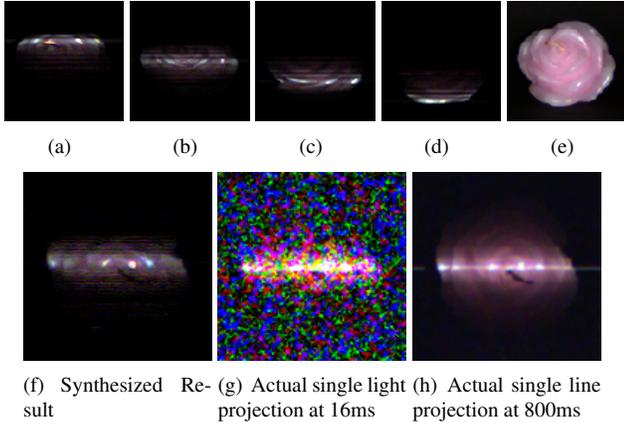


Figure 9: Relighting under virtual single line illumination sweeping from up (a) to bottom (d) with regular imaging (e). Note that the synthesized result (f) has much better noise properties than an actual single line projection at exposure 16ms (f). Our method achieves similar noise performance to (h) a 800ms exposure, but is limited to short light transport effects due to our assumption of sparsity in optimization.

Figure 10(d) and an epipolar-global image in Figure 10(e). Note that the epipolar-direct image is improved over each method alone, but still cannot completely remove all the specular interreflections on the epipolar plane. This is still an open problem for direct/global separation and warrants further study.

### 6.7. Material Recognition of Subsurface Scattering

Finally, the use of delay and exposure can yield fundamental new information about light scattering in materials, particularly subsurface scattering. Previous researchers have used time-of-flight measurements to achieve a similar result [31, 35]. Consider the delay profile for a given material. We expect the maximum of this plot to be at  $t_d = 0$ . However, our intuition is that the more subsurface scattering present in the material, the more spread out the delay profile will be.

In Figure 11, we tested this hypothesis and its usefulness for material recognition of subsurface scattering in common household items. We imaged hand soap, fat free, 2%, and whole milk, and toothpaste. All of these items were white in color, and difficult to identify with RGB information alone. We plotted their average delay profiles for a set of their pixels shown in Figure 11(b). We normalized these delay profiles using the area under the curve to cancel out the effects of albedo.

Using training and test images, we trained a support vector machine (SVM) with nonlinear kernel (radial Gaussian basis function) to get a per-pixel semantic segmentation of the materials (Figure 11(c)) and a confusion matrix (Fig-

ure 11(d)). We achieved over 90% recognition for all the materials. We note that the only errors occurred for pixels near the edge of the container, where possibly the scattering profile changes for the materials due to the asymmetry of a boundary condition. This is an interesting avenue of future research to use delay profiles to better model or inverse render subsurface scattering. This application is not meant for robust instance-level material recognition, but highlights the usefulness of delay profiles for understanding subsurface scattering in materials.

## 7. Discussion

We have presented a new imaging modality based on the light transport between planar illumination and camera pixels. We showed how the synchronized rolling shutter of a camera multiplexes this planar illumination using its parameters of sync delay and exposure. Using delay-exposure image stacks, we perform demultiplexing with a convex optimization algorithm to estimate 3D light transport. We show applications including analyzing the delay profiles of various light transport paths, and enhanced light separation algorithms and material recognition of subsurface scattering. We hope that such parametric analysis of epipolar and non-epipolar light can lead to future insights into the physical nature of these visual effects.

Some limitations of our method include that it is not real-time, requiring about 80-100 images in the stack for demultiplexing. We are limited to 3D light transport estimation since our imaging model assumes projector illumination as impulse rows rather than the full 4D light transport. Further, our optimization algorithm uses no spatial cues from neighboring pixels that could potentially aid the light transport estimation.

Future work includes extending our method to imaging systems where exposure is decoupled from delay/row-offset as in ROI systems like EpiTOF [1]. In addition, it would be useful to have control over the projector’s raster scanning behavior, and estimate 4D light transport from this illumination. We hope that this research shows a path forward for temporal synchronization between projectors and cameras to selectively capture many different components of light transport.

## Acknowledgement

The authors would like to thank Supreeth Achar and Joe Bartels for help with Episcan3D prototype development, and Vishwanath Saragadam for his helpful comments. This work was sponsored by the JSPS program for advancing strategic international networks to accelerate the circulation of talented researchers (G2802), KAKENHI grant number JP15K16027, the NAIST global collaboration program,

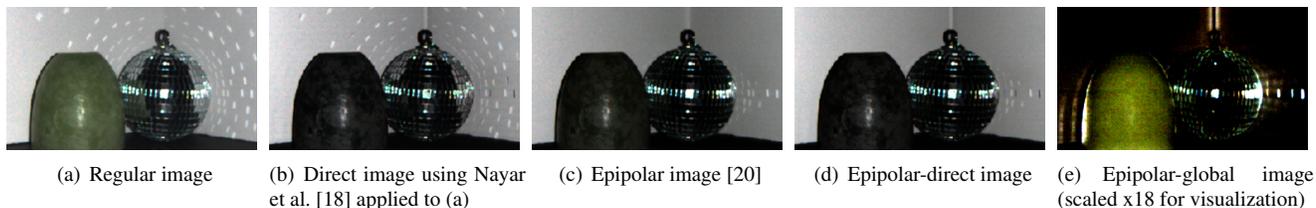


Figure 10: We show the results of two direct/global separation methods on our scene (a): Nayar et al. [18] in (b), and epipolar imaging [20] in (c). By combining the two methods, we are able to visualize (d) epipolar-direct only light, and (e) epipolar-global light. Notice how the specular interreflections are only removed in epipolar imaging, and the remaining subsurface scattering light in the epipolar image is separated by combining both algorithms.

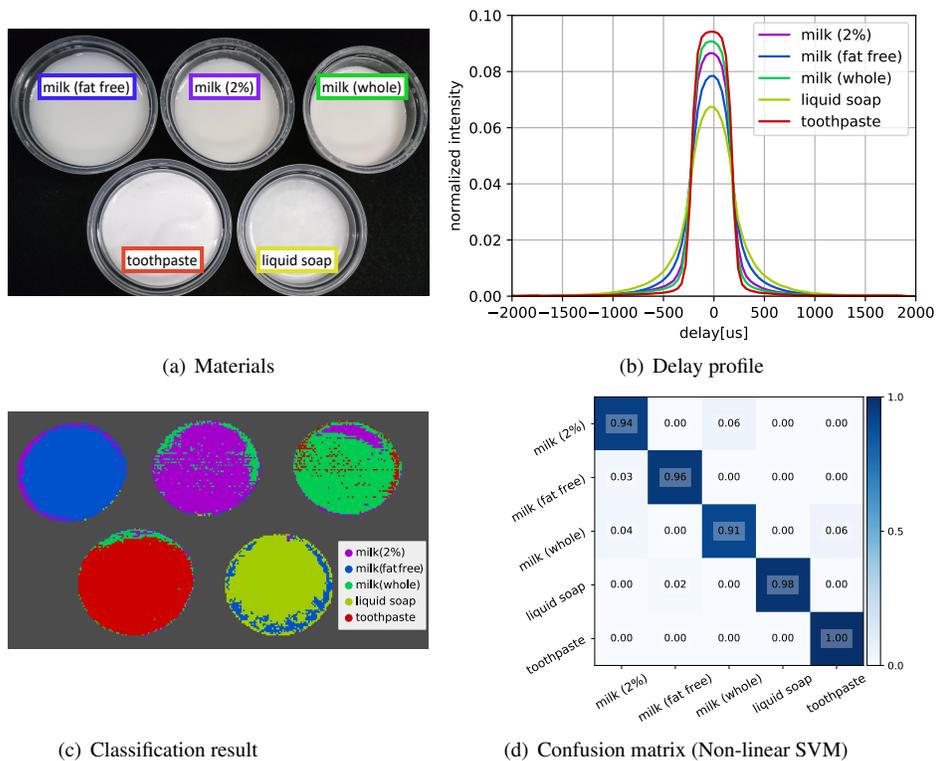


Figure 11: Using varying delay, we can perform recognition and analysis of subsurface scattering materials. In (a), fat free, 2%, whole milk, toothpaste, and liquid soap were imaged. Note that all materials are white in color, and difficult to distinguish visually. Their average pixel delay profiles are shown in (b). We use a nonlinear SVM to help perform semantic classification (c) and show the results of the confusion matrix (d). Note that all materials were correctly identified with most pixels semantically segmented correctly.

and the Defense Advanced Research Projects Agency (RE-VEAL Grant HR00111620021).

## References

- [1] S. Achar, J. R. Bartels, W. L. Whittaker, K. N. Kutulakos, and S. G. Narasimhan. Epipolar time-of-flight imaging. *ACM Transactions on Graphics (TOG)*, 36(4):37, 2017.
- [2] J. Bai, M. Chandraker, T.-T. Ng, and R. Ramamoorthi. A dual theory of inverse and forward light transport. *European Conference on Computer Vision (ECCV)*, pages 294–307, 2010.
- [3] P. Debevec. The light stages and their applications to photoreal digital actors. *SIGGRAPH Asia*, 2(4), 2012.
- [4] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. In *ACM SIGGRAPH*, pages 145–156, 2000.
- [5] S. Diamond and S. Boyd. Cvxpy: A python-embedded modeling language for convex optimization. *The Journal of Machine Learning Research*, 17(1):2909–2913, 2016.
- [6] G. Garg, E.-V. Talvala, M. Levoy, and H. P. Lensch. Symmetric photography: exploiting data-sparseness in reflectance fields. In *Proceedings of the 17th Eurographics Conference*

- on *Rendering Techniques*, pages 251–262. Eurographics Association, 2006.
- [7] C. M. Goral, K. E. Torrance, D. P. Greenberg, and B. Bataille. Modeling the interaction of light between diffuse surfaces. In *ACM SIGGRAPH Computer Graphics*, volume 18, pages 213–222. ACM, 1984.
- [8] J. Gu, Y. Hitomi, T. Mitsunaga, and S. Nayar. Coded rolling shutter photography: Flexible space-time sampling. In *IEEE International Conference on Computational Photography (ICCP)*, pages 1–8. IEEE, 2010.
- [9] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan. Flexible voxels for motion-aware videography. In *European Conference on Computer Vision (ECCV)*, pages 100–114. Springer, 2010.
- [10] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan. Structured light 3d scanning in the presence of global illumination. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 713–720. IEEE, 2011.
- [11] Y. Hitomi, J. Gu, M. Gupta, T. Mitsunaga, and S. K. Nayar. Video from a single coded exposure photograph using a learned over-complete dictionary. In *IEEE International Conference on Computer Vision (ICCV)*, pages 287–294. IEEE, 2011.
- [12] H. W. Jensen, S. R. Marschner, M. Levoy, and P. Hanrahan. A practical model for subsurface light transport. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, pages 511–518. ACM, 2001.
- [13] J. T. Kajiya. The rendering equation. In *ACM Siggraph Computer Graphics*, volume 20, pages 143–150. ACM, 1986.
- [14] S. J. Koppal, S. Yamazaki, and S. G. Narasimhan. Exploiting dlp illumination dithering for reconstruction and photography of high-speed scenes. *International Journal of Computer Vision*, 96(1):125–144, 2012.
- [15] S. Lin, Y. Li, S. B. Kang, X. Tong, and H.-Y. Shum. Diffuse-specular separation and depth recovery from image sequences. In *European Conference on Computer Vision (ECCV)*, pages 210–224. Springer, 2002.
- [16] W.-C. Ma, T. Hawkins, P. Peers, C.-F. Chabert, M. Weiss, and P. Debevec. Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination. In *Proceedings of the 18th Eurographics Conference on Rendering Techniques*, pages 183–194. Eurographics Association, 2007.
- [17] S. K. Nayar, X.-S. Fang, and T. Boult. Separation of reflection components using color and polarization. *International Journal of Computer Vision*, 21(3):163–186, 1997.
- [18] S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar. Fast separation of direct and global components of a scene using high frequency illumination. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 935–944. ACM, 2006.
- [19] R. Ng, R. Ramamoorthi, and P. Hanrahan. All-frequency shadows using non-linear wavelet lighting approximation. In *ACM Transactions on Graphics (TOG)*, volume 22, pages 376–381. ACM, 2003.
- [20] M. O’Toole, S. Achar, S. G. Narasimhan, and K. N. Kutulakos. Homogeneous codes for energy-efficient illumination and imaging. *ACM Transactions on Graphics (TOG)*, 34(4):35, 2015.
- [21] M. O’Toole and K. N. Kutulakos. Optical computing for fast light transport analysis. In *ACM Transactions on Graphics (TOG)*, volume 29, page 164. ACM, 2010.
- [22] M. O’Toole, J. Mather, and K. N. Kutulakos. 3d shape and indirect appearance by structured light transport. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3246–3253, 2014.
- [23] M. O’Toole, R. Raskar, and K. N. Kutulakos. Primal-dual coding to probe light transport. *ACM Transactions on Graphics (TOG)*, 31(4):39–1, 2012.
- [24] P. Peers, D. K. Mahajan, B. Lamond, A. Ghosh, W. Matusik, R. Ramamoorthi, and P. Debevec. Compressive light transport sensing. *ACM Transactions on Graphics (TOG)*, 28(1):3, 2009.
- [25] R. Raskar, A. Agrawal, J. Tumblin, R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography. *ACM Transactions on Graphics*, 25(3):795, 7 2006.
- [26] N. Ratner and Y. Y. Schechner. Illumination Multiplexing within Fundamental Limits. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 6 2007.
- [27] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur. Multiplexing for optimal lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(8):1339–1354, 2007.
- [28] S. M. Seitz, Y. Matsushita, and K. N. Kutulakos. A theory of inverse light transport. In *IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 1440–1447. IEEE, 2005.
- [29] P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. Lensch. Dual photography. In *ACM Transactions on Graphics (TOG)*, volume 24, pages 745–755. ACM, 2005.
- [30] P. Sen and S. Darabi. Compressive dual photography. In *Computer Graphics Forum*, volume 28, pages 609–618. Wiley Online Library, 2009.
- [31] S. Su, F. Heide, R. Swanson, J. Klein, C. Callenberg, M. Hullin, and W. Heidrich. Material classification using raw time-of-flight measurements. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3503–3511, 2016.
- [32] C.-Y. Tsai, K. N. Kutulakos, S. G. Narasimhan, and A. C. Sankaranarayanan. The geometry of first-returning photons for non-line-of-sight imaging. In *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [33] S. Umeyama and G. Godin. Separation of diffuse and specular components of surface reflection by use of polarization and statistical analysis of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(5):639–647, 2004.
- [34] E. Veach. *Robust monte carlo methods for light transport simulation*. Stanford University Stanford, 1998.
- [35] D. Wu, A. Velten, M. Otoole, B. Masia, A. Agrawal, Q. Dai, and R. Raskar. Decomposing global light transport using time of flight imaging. *International Journal of Computer Vision*, 107(2):123–138, 2014.