

Slope Disparity Gating: System and Applications

Sreenithy Chandran, Hiroyuki Kubo, *Member, IEEE*, Tomoki Ueda, Takuya Funatomi, *Senior Member, IEEE*, Yasuhiro Mukaigawa, *Member, IEEE*, and Suren Jayasuriya, *Member, IEEE*

Abstract—Active illumination systems which have the ability to selectively image photons that arrive from a specified surface geometry some distance away have several applications in robotics, autonomous vehicles, and surveillance. In this paper, we present a new technique for sloped disparity gating utilizing a synchronized projector-camera system consisting of a raster scanning laser projector and a rolling shutter camera. This system enables disparity-based triangulation of light from specific sloped planar surfaces in the imaging volume. We demonstrate how to control the slope and thickness of these planar surfaces using hardware parameters of pixel clock, synchronization delay, and exposure. Using our system, we present a new decomposition of the light transport matrix in a scene as a function of disparity. We perform applications including real-time image masking and imaging and relighting in scattering media. We also demonstrate applications of the proposed light transport probing technique like novel disparity-dependent relighting, light transport-aware foreground/background separation and green screening, and improved visualization and relighting in scattering media.

I. INTRODUCTION

Active illumination systems, which comprise of a controllable light source and a detector (sensor) have been used for their numerous 3D imaging capabilities. These systems rely on the light source along with computational techniques for improved scene understanding and measurement. A robust active illumination system should handle ambient light and global light transport effects in the scene which typically cause errors in imaging and 3D scanning.

For active illumination systems to be deployable in the real world, energy efficiency is the key. Recently there has been a niche group of specialized optical setups that perform energy-efficient imaging of a scene by capturing light only from a certain desired distance, these works are broadly termed range or disparity gating [1]–[4]. Disparity gating has applications in imaging through scattering media such as fog since it can avoid the capture of backscattered photons and thus improve visibility. Such capabilities would be advantageous for computer vision in bad weather.

In this paper we introduce a disparity gating setup using a vertically aligned configuration of a synchronised projector-camera system. The rows of the raster scanning laser is synchronized with the rolling shutter of the camera. This system is a modification of the Episcan3D system [3]. The intersection

of the light rays and the camera paths will triangulate a sloped imaging plane. We demonstrate how by changing the hardware parameters such as the synchronization delay, camera exposure time and the pixel clock, we are able to control the properties of the imaging plane such as the distance from the setup, thickness of the imaging volume, and tilt of plane. Thus using this system, we can selectively capture desired regions of a scene while optically avoiding photons from other regions.

We utilize the system’s ability to control the imaging plane properties to probe a scene for light transport in a unique way. Capturing the full light transport matrix for a projector-camera system is expensive because of the acquisition time and the large size of the matrix. Therefore some works have tried to handle these limitations by probing the light transport matrix for various effects like direct/indirect separation [5] and epipolar/non-epipolar capture [3], [6], [7]. Our probing method enables capture of direct light transport paths based on the disparity between projector and camera, parameterized by the system’s hardware parameters. Since we can selectively discard the global component which includes light that is volumetrically scattered or undergoes multiple reflections, we can further improve disparity gating results as well as perform image-based relighting and visualization. We demonstrate various applications of this disparity-based light transport probing technique in this paper.

The main contributions of the paper include:

- A synchronized projector-camera system for capturing sloped planes in a scene,
- Formulation of the image formation model using the system and establishing the relationship between the various hardware parameters of the system to the properties of the triangulated imaging plane,
- Decomposition of direct light transport into components using disparity-based triangulation to probe light transport in a scene and utilizing the light transport probing for novel disparity-dependent relighting,
- Applications of sloped disparity gating including real-time image masking, imaging through scattering media, improved foreground/background separation and green screen compositing using both synthetic and real experiments from an experimental prototype.

A preliminary version of this paper appeared at the International Conference of Computational Photography [8]. This paper extends the conference version by interpreting disparity gating as light transport probing and introducing a new decomposition of the light transport matrix in a scene as a function of disparity. The capability of our disparity gating system to selectively image a specific area in the

S. Chandran is with the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ, 85281. Contact: schand56@asu.edu

H. Kubo is with the Department of Imaging Sciences, Chiba University, Chiba, Japan, 260–8670

T. Ueda, T. Funatomi, Y. Mukaigawa are with the Division of Information Science, Nara Institute of Science and Technology, Nara, Japan, 630–0192

S. Jayasuriya is with the School of Arts, Media and Engineering and the School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ, 85281.

scene coupled with our light transport probing technique enables new applications such as depth dependent relighting, improved foreground/background separation and green screen compositing and improved visualization in fog.

The rest of the paper is organized as follows. Section II outlines related work in range/disparity gating and light transport acquisition and probing. Section III introduces the main imaging model for slope disparity gating and discusses the effects of pixel clock, synchronization delay, and camera exposure on the triangulated imaging region. Section IV presents the system's light transport probing behavior in terms of probing matrices and includes a disparity-dependent relighting formula based on the decomposition of light transport (whose full derivation is presented in the Supplemental Appendix). Section V introduces both our simulation framework as well as the real hardware prototype we built in the lab (shown in Figure 2). Finally, experimental results are shown in Section VI for sloped disparity gating as well as light transport applications in Section VII before ending with a discussion in Section VIII.

II. RELATED WORK

Active Illumination Systems: A review of the recent advances in 3D surface imaging using structured light shows a variety of applications where 3D imaging is necessary [9]. Photometric stereo [10] uses independent light source directions to compute surface normals and thereby perform shape estimation. Recent technology using time-of-flight [11] and LIDAR [12] have enabled ubiquitous depth sensing. Active lighting has also helped to improve imaging through scattering media [13], [14], underwater imaging [15], and biomedical microscopy [16].

Range and Disparity Gating: Range gating involves using illumination systems to capture light that has travelled a certain distance before reaching the image sensor. Range gating was originally introduced for search and surveillance operations as well as automotive safety [17], [18]. Recently, it has been used for applications like imaging through scattering media [2], [19] due to its ability to block backscattered rays. The Episcan3D system in a rotated configuration was able to perform disparity gating and demonstrated depth mapping through smoke [3]. Future research introduced programmable triangulation light curtains [20]–[22] to perform gated imaging for arbitrary ruled surfaces in 3D using a line laser and line sensor in synchronisation. In this system the imaging plane and illumination plane are side-by-side to intersect in a line. This line can be then swept across the scene to capture objects that intersect this surface.

Full Light Transport Acquisition: Many systems have been proposed to capture full light transport in a scene. These include light stages for facial lighting [23], [24], moving actors in a scene using a LED dome [25], high-speed photometric stereo [26] and reflectance fields for people who are walking/running [27]. For specular objects, both scanning laser [28] and reflective [29] setups have been proposed. Beyond conventional light transport, researchers have included additional dimensions such as continuous-wave time-of-flight (represented as frequency) [30], as well as dynamic light

transport in time [31], [32]. In addition, several research methods have been proposed to reduce the acquisition cost of light transport including optical computing [33], Kernel method [34], and compressive sensing [35], [36].

Selective Imaging of Light Transport Components: Nayar et al. [37] first showed separation of direct/global illumination using high frequency patterns. This direct/global separation was then extended to situations involving defocused and global illumination [38], video via motion compensation [39], and multiple light sources [40]. Further decomposing global light into short and long-range indirect light has been made possible through spatially-varying illumination using multiple projections [41] as well as through primal-dual coding to capture these effects in a single exposure [5]. Other light transport probing methods have been introduced for capturing additional plenoptic dimensions using interferometry [42], [43] or polarization [44].

One useful representation of light transport paths is the decomposition into epipolar and nonepipolar light between the projector and camera. Exploiting the epipolar geometry between projector and camera has obtained epipolar and non-epipolar light ray separation at video rates using primal-dual coding techniques [45], [46]. The Episcan3D system contains of a rectified projector-camera system with aligned rows and synchronized laser scanning to the rolling shutter of the camera [3]. Since epipolar light paths contain mostly direct light (i.e. light that only bounces once in a scene), and non-epipolar light paths contain indirect light paths (i.e. light that bounces multiple times in the scene), this arrangement allows capture of the desired light paths through synchronization patterns. Extending the Episcan3D system to selectively capture non-epipolar light, researchers have shown applications including subsurface scattering material recognition based on subsurface scattering as well as improved vein visualization and remote photoplethysmography [6], [7]. Epipolar imaging was also extended by Achar et al. to time-of-flight [47] to obtain more accurate depth measurements and robustness to ambient light.

Comparison to O'Toole et al. [3]: Our work is an extension of the Episcan3D's disparity gating system [3]. However, we have advanced the system's capabilities by (1) changing synchronization and hardware parameters to control the slope and thickness of the gated region and (2) formulating the light transport probing matrices explicitly and deriving a relighting formula for disparity-dependent light transport. This work builds on the foundation set by O'Toole et al., and we show several new applications never shown by the original system including vein visualization, sloped disparity gating, improved foreground-background separation and green screen compositing, and relighting of objects in scattering media.

III. SLOPE DISPARITY GATING

In this section, we introduce the formulation for slope disparity gating for a synchronized projector-camera system. We proceed to discuss the importance of the hardware parameters camera exposure, pixel clock, and synchronization delay in controlling the geometry of the triangulated region.

Our hardware setup consists of a vertically aligned laser projector and rolling shutter camera. The laser projector il-

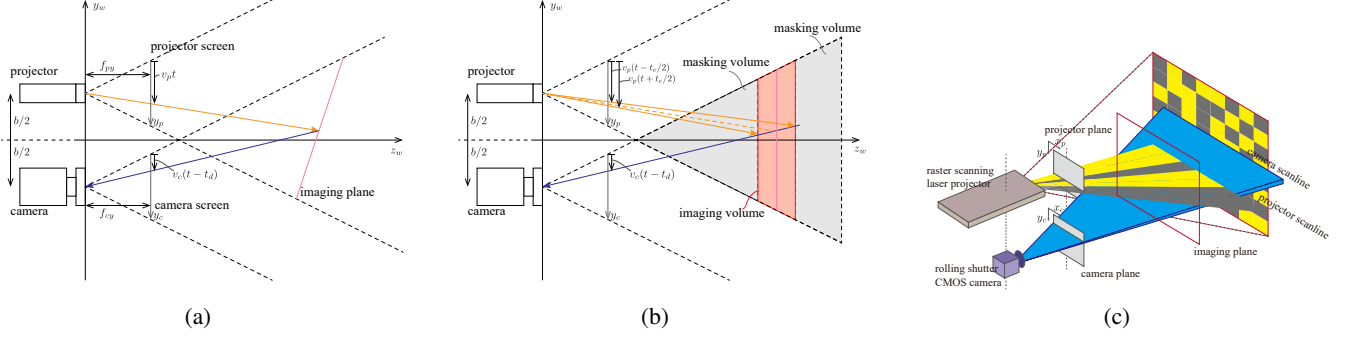


Fig. 1: (a) The ray geometry for our imaging system that determines a sloped plane to be imaged. (b) A schematic of the imaging volume in the scene where photons are captured based on the pixel clock, synchronization delay, and exposure time of the projector-camera system. (c) A visualization of the scanning illumination and the camera rolling shutter acquiring light transport by projecting coded patterns onto a selective image region determined via disparity gating.



Fig. 2: Our hardware prototype for the slope disparity gating system. This synchronized projector-camera system allows for optical gating of photons from selective regions in the scene. On the right is an example scene being captured in the lab.

lumines a single horizontal plane of the light at a time and the rolling shutter camera captures a single horizontal plane as well corresponding to each exposed pixel row. The illumination plane and camera plane are swept vertically from top row to bottom row. The ray geometry of our projector-camera system is shown in Figure 1(a). This hardware configuration was first introduced in [3] using a raster-scanning laser projector and a rolling shutter camera. Note while the laser is actually raster scanning pixel-by-pixel to illuminate a row, this is still faster than the shortest exposure time of a camera row and thus we view the projector as illuminating a single row at a time for our system for this paper.

The resulting projector-camera system performs light triangulation between the illumination and camera planes in the form of disparity gating. We formulate the parametric equations for the disparity gated surface as the cross-section between scanning projector and camera planes. The relation between a position in the world coordinate $\mathbf{x}_w$ and screen coordinate $\mathbf{x}_s$ is given by the following equation:

$$s\mathbf{x}_s = \mathbf{K}[\mathbf{R}|\mathbf{T}]\mathbf{x}_w, \quad (1)$$

where $\mathbf{x}_s = (x_s, y_s, 1)^\top$ and $\mathbf{x}_w = (x_w, y_w, z_w, 1)^\top$ are in homogeneous coordinates, and s , $\mathbf{R}$ and $\mathbf{T}$ represent the scaling factor, rotation matrix, and translation vector respectively. $\mathbf{K}$ is the camera matrix:

$$\mathbf{K} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}. \quad (2)$$

The camera is located at $(0, -b/2, 0)$ and the projector at $(0, b/2, 0)$. Both projector and camera are viewing in the $+z$ direction. The up direction vectors of both projector and camera are $(0, 1, 0)$, satisfying the right-hand rule for our coordinate system. In this configuration, the scale factor s is $(1, -1, 1)^\top$, the rotation matrix of the camera $\mathbf{R}_c$ and of the projector $\mathbf{R}_p$ are identity matrices, the camera translation vector $\mathbf{T}_c = (0, b/2, 0)^\top$, and the projector translation vector is $\mathbf{T}_p = (0, -b/2, 0)^\top$.

We can define both the camera screen coordinates $\mathbf{x}_c = (x_{cs}, y_{cs}, 1)^\top$ and projector screen coordinates $\mathbf{x}_p = (x_{ps}, y_{ps}, 1)^\top$ in terms of world coordinates:

$$\begin{bmatrix} -y_c \\ 1 \end{bmatrix} = \begin{bmatrix} f_{cy} & c_{cy} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & b/2 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_w \\ z_w \\ 1 \end{bmatrix}, \quad (3)$$

$$\begin{bmatrix} -y_p \\ 1 \end{bmatrix} = \begin{bmatrix} f_{py} & c_{py} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -b/2 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_w \\ z_w \\ 1 \end{bmatrix} \quad (4)$$

Since the equations are in homogeneous coordinates, we obtain

$$y_c = -\frac{f_{cy}(y_w + b/2)}{z_w} - c_{cy}, \quad (5)$$

$$y_p = -\frac{f_{py}(y_w - b/2)}{z_w} - c_{py}. \quad (6)$$

Note, we omit x coordinate in Equations (3)–(6) as the system cannot triangulate with respect to that coordinate axis. This is a limitation of our projector-camera geometry, while programmable light curtains [4], [21] can triangulate any arbitrary curved surface in 3D space.

For arbitrary time t , the illumination projector row for the scanning projector is given by

$$y_p = v_p t, \quad (7)$$

and the captured camera row in the rolling shutter is

$$y_c = v_c(t - t_d), \quad (8)$$

where v_p is the projector scanning velocity per row, v_c is the speed of the rolling shutter per row, and t_d is the synchronization delay. We convert the illumination row, Equation (7),

and a camera row, Equation (8) to world coordinates using Equation (6). The intersection of illumination and observation rows in world coordinate can be calculated as follows:

$$y_w = \frac{b(c_{cy}f_{py} + c_{py}f_{cy} + f_{py}v_c(t - t_d) + f_{cy}v_p t)}{2(c_{cy}f_{py} - c_{py}f_{cy} + f_{py}v_c(t - t_d) - f_{cy}v_p t)}, \quad (9)$$

$$z_w = \frac{-bf_{cy}f_{py}}{c_{cy}f_{py} - c_{py}f_{cy} + f_{py}v_c(t - t_d) - f_{cy}v_p t}. \quad (10)$$

These equations represent a single sloped plane in $y - z$ world coordinates parameterized by t . While these equations technically do not represent a plane but a very flat curve, we refer to it as a sloped plane for simplicity. For this reason, our imaging method is called **slope disparity gating**, since both the distance z and slope of the plane can be controlled by our system. In Figure 1(a), we visually depict this sloped plane that is triangulated by the projector and camera planes. Note that this surface is defined by parameters of the system including synchronized delay t_d , rolling shutter velocity of the camera v_c , and the projector scanning velocity v_p . This is the surface that we are performing gated imaging on.

Camera Exposure: In this system, the camera exposure t_e controls the thickness of the imaging volume being triangulated. For arbitrary camera row y_c which exposure starts and ends at $y_{cs} = v_c(t - t_d - t_e/2)$ and $y_{ce} = v_c(t - t_d + t_e/2)$, the projector rows starts from $y_{ps} = v_p(t - \frac{t_e}{2})$ to $y_{pe} = v_p(t + \frac{t_e}{2})$ during the exposure. The finite exposure causes the swept cross-section between the illumination and camera planes to become a 3D volume as shown in Figure 1(b). Thus, exposure enables control over the thickness of this volume. Previous works like the programmable light curtains [4], [21] require a physical change in either the baseline or optics of the system to control the thickness of the light curtains, thereby making this volume control by change in exposure, a unique feature of the slope disparity system.

Rolling Shutter Pixel Clock and Synchronization Delay:

The camera's rolling shutter pixel clock controls the scanning velocity of the camera. Changing the pixel clock, denoted v_c , controls the slope of the gated plane which is triangulated by the system. Similarly, the distance to the sloped surface can be controlled by the synchronization delay t_d where a shorter delay leads to a farther optical distance being triangulated.

IV. LIGHT TRANSPORT PROBING AND RELIGHTING VIA DISPARITY GATING

In this section we see how the imaging model of our slope disparity system described in Section III can be used for probing the light transport of selective regions in the scene by controlling the hardware parameters such as delay and exposure. In particular, we leverage the work of O'Toole et al. [46] who formulate light transport probing as the Hadamard product of the full light transport matrix with probing matrices dependent on the optical sensing mechanism of the system. This allows us to determine what components of light transport we capture in a scene, and has implications for relighting which we discuss later in this section.

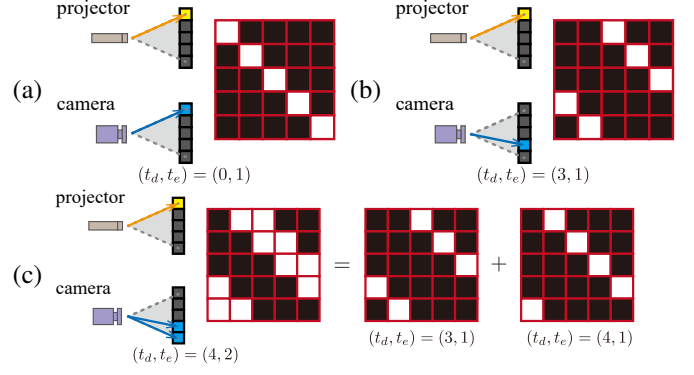


Fig. 3: We present a conceptual example for a five pixel 1D projector and camera and their corresponding probing matrix Π_{t_d, t_e} . Note in (a) with delay $t_d = 0$, that the system only triangulates objects at infinity, yielding a diagonal probing matrix. In (b), increasing the delay $t_d = 3$ yields to two components in the matrix with positive disparity in the lower triangular part, as well as three components in the upper triangular part that would be triangulated behind the camera (and thus must be indirect light). In (c), we show the matrix for $(t_d, t_e) = (4, 2)$ can be decomposed into the sum of $t_e = 1$ exposures with different delays. Note how the resulting matrix has widened its diagonal bands as a function of the exposure.

We begin by expressing the probing matrices for our system as a function of delay and exposure. Light transport probing can be written as the following as shown by [46]:

$$\mathbf{i} = (\Pi \odot \mathbf{T})\mathbf{p}, \quad (11)$$

where $\odot$ is the element-wise Hadamard product; $\mathbf{i}$ is the image represented as a column vector of N pixels; $\mathbf{p}$ is the P -pixel projected pattern, also represented as a column vector; Π is the matrix that determines what components of the light transport matrix are probed for the images captured and $\mathbf{T}$ is the scene's $N \times P$ instantaneous light transport matrix.

Our system performs selective imaging of a scene via the hardware parameters of synchronization delay and exposure time. Thus, given that we can project coded patterns (e.g. Hadamard patterns [48]), we can capture the light transport for a given delay and exposure time as light transport probing: $\mathbf{T}(t_d, t_e) := \Pi_{t_d, t_e} \odot \mathbf{T}$. In practical terms, this means we acquire light transport in the traditional way but while setting the delay t_d and exposure time t_e to particular values, and only imaging a selective region of the scene. Hadamard codes are used for multiplexed illumination as it has been shown that these are optimal in maximizing signal-to-noise ratio as compared to one illumination pixel at a time [48]. We demonstrate in Figure 13 using synthetic data.

Conceptual Example: In Figure 3, we show a conceptual example of what these probing matrices Π_{t_d, t_e} look like. Given 1D projector and camera with 5 pixels each, zero-indexed as 0–4 from top to bottom, let us take the case of $t_e = 1$ pixel, and $t_d = 0$ pixels. Then the probing matrix is

given as a diagonal matrix:

$$\Pi_{0,1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

which corresponds to probing only the disparity at $z = +\infty$. On the other hand, when the delay increases by 1 pixel ($t_d = 1$) or 2 pixels ($t_d = 2$), we get the following probing matrices:

$$\Pi_{1,1} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \Pi_{2,1} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

Extrapolating this pattern, we see that for $t_e = 1$,

$$\Pi_{t_d,1} = \sigma_{t_d}(\mathbf{I})$$

where $\sigma_n(\mathbf{I})$ is the permutation of the identity matrix $\mathbf{I}$ with the n -th row of the canonical identity matrix starting as the first (here defined as zeroth) row. Then we have that

$$\sum_{t_d=0}^4 \Pi_{t_d,1} \odot \mathbf{T} = \mathbf{T}$$

where the delay t_d sweeps from 0 pixels to 4 pixels given our camera only has five pixels. Thus we can recover back the full light transport if we add all of the light transports for different delays together.

General Formula: We can derive the general formula for our probing matrices for $t_e > 1$. The first observation to note is that the rolling shutter exposure is nothing but a superposition of multiple exposures of a pixel, delayed appropriately, as shown in Figure 3(c). Note that we use the convention that if delay is t_d and exposure is t_e , then rows $[t_d - (t_e - 1), t_d - (t_e - 2), \dots, t_d - 1, t_d]$ will be exposed at a time. Thus we can write the probing matrix for general t_d, t_e as follows:

$$\Pi_{t_d,t_e} = \sum_{k=0}^{t_e-1} \sigma_{t_d-k}(\mathbf{I}). \quad (12)$$

Exposure increasing widens the disparity off-diagonal elements in the probing matrix. A simple example of this can be seen in Figure 3(c). Summing over all t_d where N is the total number of camera rows, we obtain the following:

$$\begin{aligned} \sum_{t_d} \Pi_{t_d,t_e} \odot \mathbf{T} &= \left(\sum_{t_d=0}^{N-1} \sum_{k=0}^{t_e-1} \sigma_{t_d-k}(\mathbf{I}) \right) \odot \mathbf{T} \\ &= \left(\sum_{k=0}^{t_e-1} \sum_{t_d=0}^{N-1} \sigma_{t_d-k}(\mathbf{I}) \right) \odot \mathbf{T} = \left(\sum_{k=0}^{t_e-1} \mathbf{1} \right) \odot \mathbf{T} = t_e \cdot \mathbf{T} \end{aligned} \quad (13)$$

where $\mathbf{1}$ is a matrix of all ones. Thus the exposure acts as a global constant that multiplies the light transport matrix when all the delays are summed up. Note that this general formula also holds when we consider our 2D projector and camera, just that the probing matrices exhibit this permuted identity

matrix composition as blocks, similar to the block-structured light transport of [46].

Relationship with transient imaging: Our disparity-based light transport is similar to transient light transport decomposition [30], [49]–[51]. Transient imaging uses time-of-flight information of the light paths to decompose the scene and can capture “light-in-flight” as well as depth images of the scene. However, transient imaging systems require time-of-flight sensors and pulsed or modulated illumination which requires more complex optical hardware (such as spatial light modulators in addition to lasers and time-of-flight cameras as used in the setup of O’Toole et al. [30]) and can drive system costs up. In contrast, our disparity-based system requires an off-the-shelf projector and camera to operate. Further, although there has been some work on capturing transient images at interactive rates using single photon avalanche diodes [52], it is still difficult to perform real-time probing for applications such as selective image masking for transient imaging. It remains an area of future work to synergistically combine transient light transport techniques with disparity-based probing.

A. Disparity-dependent Relighting

One of the main applications of our light transport acquisition technique is image-based relighting of scenes in post-capture [23], [27], [53]. In traditional relighting, we can form a new image $\mathbf{i} = \mathbf{T}\mathbf{p}$ using a new illumination vector $\mathbf{p}$. In other words, novel images can be synthesized as linear combinations of the light transport matrix columns weighted by illumination pixels. However our slope disparity system performs matrix probing (Π_{t_d,t_e}) to capture only slices of light transport, and thus we cannot immediately acquire the full light transport to perform relighting.

However, we can capture light transport for direct light at individual disparity slices. To do so, we acquire light transport via coded acquisition by projecting Hadamard codes [48], but while setting delay and exposure to isolate an individual disparity slice in the scene. We denote this acquired transport matrix as $\mathbf{T}_{direct}(z(t_d))$ and can perform relighting of the direct components in the scene by summing over t_d :

$$\mathbf{i} = \sum_{t_d} \mathbf{T}_{direct}(z(t_d)) \mathbf{p}(t_d), \quad (14)$$

where $\mathbf{p}(t_d)$ is illumination that changes with delay/disparity. Note we can only perform relighting for direct light because of the assumptions involving triangulation and our image formation model in Section III.

While this relighting is not physically realistic as illumination does not typically vary with disparity, it can be useful for certain visualization applications including green screen compositing and relighting in scattering media as shown in Section VI. It does also increase the number of patterns projected to acquire light transport for every disparity slice, but the light captured has higher signal-to-noise (SNR) and contrast since other photons outside the region of interest are being optically rejected by the system.

We can also recover a floodlit direct image from these individual disparity slices:

$$\mathbf{i}_{\text{floodlit}} = \sum_z \mathbf{T}_{\text{direct}}(z) \mathbf{1} = \sum_{t_d} \mathbf{T}_{\text{direct}}(z(t_d)) \frac{dz}{dt_d} \mathbf{1}, \quad (15)$$

where $\mathbf{1}$ is an illumination vector of all ones and $\frac{dz}{dt_d}$ is a change of variables formula. In the Supplemental Appendix, we derive this equation and the change of variables factor which is given by the final expression:

$$\frac{dz}{dt_d} = \frac{-bf}{vt_d^2} \quad (16)$$

under the assumptions that $v_c = v_p = v$ is scanning velocity, $f_{cy} = f_{py} = f$ is focal length, and b is the baseline between camera and projector. Please see the Appendix for the full derivation of this formula. In our experimental results, we utilize $\left| \frac{dz}{dt_d} \right| = \frac{bf}{t_d^2}$ assuming $v = 1$ in the system.

V. IMPLEMENTATION

Simulator: To simulate our disparity gating system, we utilize the PBRT-v3 rendering software [54]. Since there is no rolling shutter mechanism in the simulator, we render multiple images with a different projected horizontal line corresponding to one row of the projector. We then extract a row from the rendered image corresponding to the exposed camera row, and digitally stitch and merge them into a final image.

For our light transport experiments, we primarily use 512×512 camera images with projector resolution varying from 4×4 up to 32×32 depending on the spatial frequency of illumination desired. Total rendering time varies depending on the PBRT-v3 integrator used. Simulating the rolling shutter results takes approximately 30 minutes to a couple of hours depending on the complexity of the scene. For each scene, we save in EXR format and then perform white balancing and tone mapping for displaying the images.

Hardware Setup: Our experimental prototype uses a synchronised laser projector and a rolling shutter camera similar to the configuration in [8]. We use a SONY MP-CL1A laser projector with 1280×720 resolution and a 60Hz refresh rate, and an IDS UI-3250CP-C-HQ color camera with both global and rolling shutter capabilities. Figure 2 (a) shows the system configuration we actually implemented. Using a model shaped by 3D printing, we aligned vertically the optical axes of the laser projector and a rolling shutter camera. The main camera triggers the rolling shutter using the vertical scanning signal of the projector as the trigger signal. We refer the readers to the supplemental material of [3] for further information about the synchronization circuitry.

Acquisition: We used 12 bit capture for each captured image, the range of pixel clock frequencies allowable by our hardware is 10–50MHz with increments of 1MHz, 50–100MHz with increments of 2MHz, and 100–128MHz with increments of 4MHz. We are able to capture a single image using our hardware prototype at approximately 20FPS. Thus it is possible to perform experiments in Section VI such as selective image masking, fog removal, vein visualisation, and automatic parameter optimization in real time.

For the acquisition time for light transport, this depends on the illumination resolution desired or equivalently the number of Hadamard code projections used. For performing small light transport acquisition with a single delay value, image resolution 512×512 , projector resolution of 4×2 , the acquisition time is approximately 50 seconds. The acquisition time scales up to 15 minutes approximately when we capture light transport at around 5 different depth slices in the scene with image resolution 1200×1600 and projector resolution 32×32 . If we want to perform very high resolution spatial projections, we can increase the projector resolution to 512×512 , in which case the acquisition time can go up to about 4 hours.

We typically capture between 2 – 5 depth slices in real experiments for our light transport acquisition. To obtain the projector and camera intrinsic parameters, we perform a standard projector-camera calibration method [55]. This involves using a checkerboard and projecting a structured light pattern onto it. The local homography around the checkerboard is used to calculate the corresponding projector pixels.

VI. EXPERIMENTAL RESULTS

In this section, we present real results demonstrating the applications of our slope disparity gating system. This includes applications such as selective image masking and imaging through fog, and vein visualisation.

A. Real-time Image Masking

In Figure 4, we show a demonstration of slope disparity gating to selectively image portions of the scene. In the scene shown in Figure 4(a), we display three different planar targets, tilted at $+15, 0, -15$ degrees and spaced 15cm apart. Using pixel clocks of 76, 78, 80MHz and delays 12630, 13200, 13920 μ s, we can selectively image the three planar targets in Figure 4(b). Please refer to the supplemental material for a video imaging this scene. Experimentally, we even can image the planar targets of $+75$ degrees using a pixel clock of 50MHz, and -75 degrees using a pixel clock of 108MHz.

In addition to selectively gating targets, we can also mask regions in the scene in Figure 4(c) where all the targets are tilted at $+15$ degrees. We show that we can selectively mask the planar targets with different delay and exposure values in Figure 4(d–f).

In Figure 5, we show the effect of using pixel clock to control the slope of our disparity gating region in order to perform either imaging or masking. Here, in Figure 5(b), the slope is set in such a way that only the photons that arrive from the table are captured, thereby masking the cup and the mouse. The inverse of this happens in (c), where the table with the document is selectively masked and the cup and mouse is imaged. This is achieved by taking the complement of the exposure from (b), in other words, exposing all other rows except the rows that correspond to the sloped plane at the table.

This image masking is useful for security applications since it is a hardware solution that discards photons before it even reaches the sensor, similar to other privacy-preserving optics

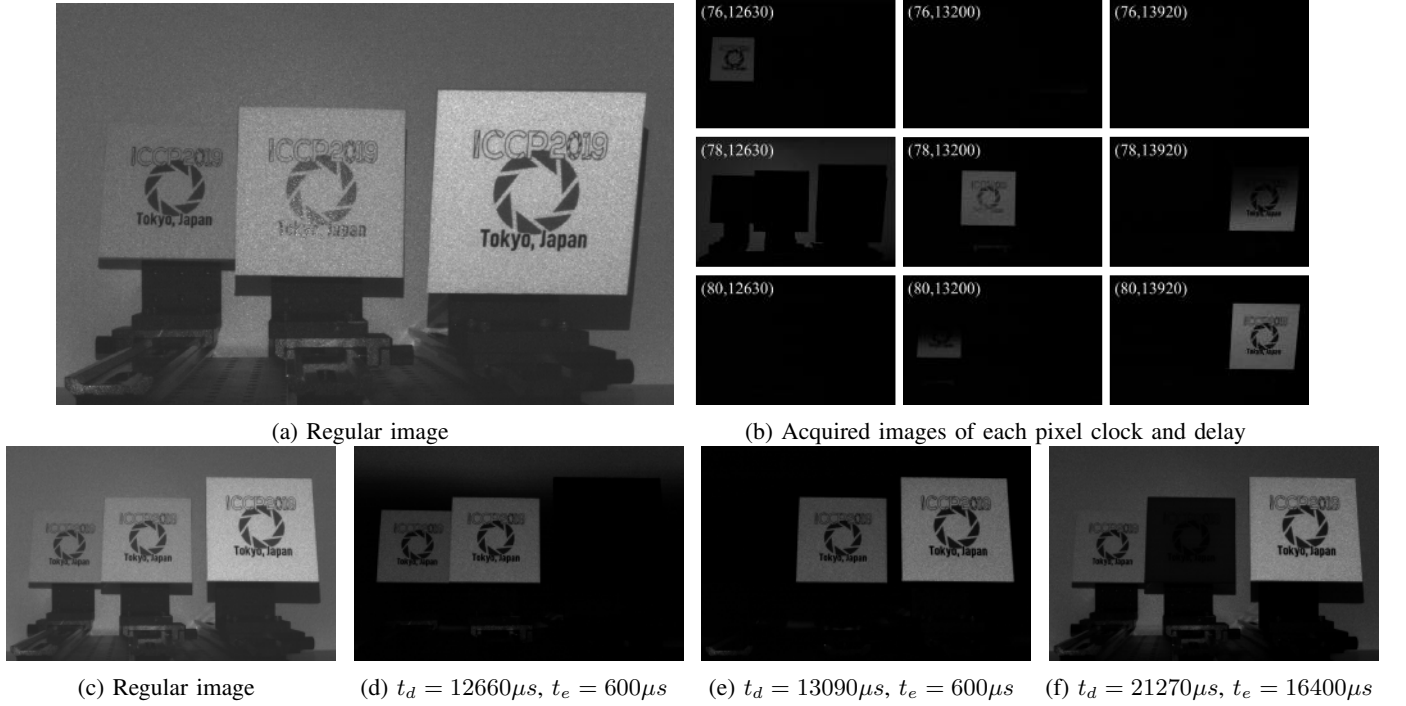


Fig. 4: Slope disparity gating results: (a) Planar targets oriented at $+15, 0, -15$ degrees and 15cm apart. (b) Selective imaging of these targets using different pixel clock and delay, denoted in the upper left as (v_c, t_d) . (c) Planar targets $+15$ degrees oriented. (d)–(f) Selectively masking each planar target by changing delay and exposure to control the thickness and placement of the volumetric gated region. Each image was an average of 10 frames for visualization enhancement.

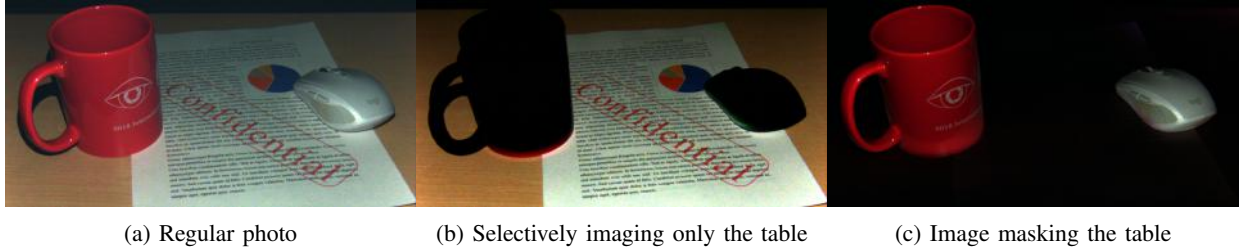
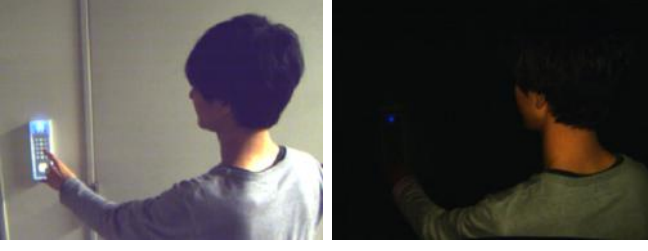


Fig. 5: Image (a) shows a regular image of a scene captured with a normal camera. Using the slope disparity gating system we can selectively image certain regions of a scene. By controlling the geometry of the sloped imaging plane we can capture only the paper on the table and selectively mask out all the other objects as shown in (b). We can also perform the inverse of this masking, where we do not capture the light from the table as shown in (c). Note all these images were captured in real-time.

solutions [56]. In Figure 6 we demonstrate a potential use case for image masking. Figure 6(a) shows scene capture from a surveillance camera of a person trying to enter a room by entering a pin on a keypad door lock. In Figure 6(b) we image the same scene using our system under the same illumination conditions, i.e. without changing the ambient lighting of the room. The hardware parameters are set in such a way that all regions besides the keypad is captured. This enables us to monitor who is entering the room without capturing sensitive information such as people’s pins. This can be easily installed in automated teller machines (ATMs) as a hardware solution to privacy preservation. Please refer to the supplemental material for a video imaging this scene. Both disparity gating and masking can be performed in real-time by setting the appropriate pixel clock, exposure, and synchronization delay

and also by automatic hardware parameter tuning which we describe below.

Another application of image masking is for a proximity detector where the plane can be set in front of a moving platform to detect if anything crosses within a certain distance from the platform. This could have uses for robotics or autonomous vehicles. Finally, a novel application of slope disparity gating was recently presented by Tsuji et al. who created an interactive touch screen out of ordinary surfaces by capturing only the region slightly above the projected screen necessary for touch sensing [57]. The authors show that disparity gating is a compelling option for interactive touch surfaces as compared to other methods shown in Table 1 of their paper [57]. All of these examples speak to the potential of selective imaging using disparity gating.



(a) Regular photo from a surveillance camera (b) Selectively imaging only the person entering the room while masking the keypad

Fig. 6: (a) shows a scene of a person trying to enter a room with a pin-pad lock as captured by a surveillance camera; (b) shows the capture using the slope disparity gating system, where we set the imaging volume to not include the wall with the pin-pad. Hence, we do not capture the password/pin a person is entering to access the room. This is a potential practical application of the real-time image masking that can be accomplished using our system.

B. Automatic Hardware Parameter Optimization:

In the above subsection, the hardware parameters of pixel clock, sync delay and exposure had to be manually tuned for a particular scene. This is inconvenient and tedious for users, and requires experience to set parameters correctly. We describe a simple method to automatically determine the pixel clock, synchronization delay and exposure time to focus on a particular object or region in the scene. We assume we have a known target object captured by the system with full exposure. We first perform ORB feature extraction [58] and store keypoints in a list. We then set the delay and exposure to minimum values, $t_d = 1\mu s$, $t_e = 0.1ms$. Then we do the following steps: (1) Increase pixel clock v_c until a keypoint of the object is visible in the scene, (2) change the delay t_d to maximize new keypoints to be captured or until already visible keypoints are lost, (3) increase the exposure t_e to maximize the number of keypoints. By doing these three steps, we are able to automatically find an object in a scene, and the resulting parameters give us a triangulated volume that has the entire image in focus. In Figure 7, we show the results of this process for a target object, and later in Section VI we utilize this technique to help identify the target in which we want to selectively capture its light transport.

C. Imaging through Scattering Media

Fog Visualization: The slope disparity gating system can also be used for imaging through scattering media such as imaging through fog or smoke [8]. The main advantage is the ability to capture only photons from the depth of interest while rejecting all other back-scattered photons due to the environment. This results in better visibility and contrast.

We first show simulated results imaging through fog for the Cornell Box scene as shown in Figure 8. Note that regular imaging is not able to see the full box due to the degraded visibility from the scattered photons. Further, another type of light transport probing based on epipolar imaging (we leverage

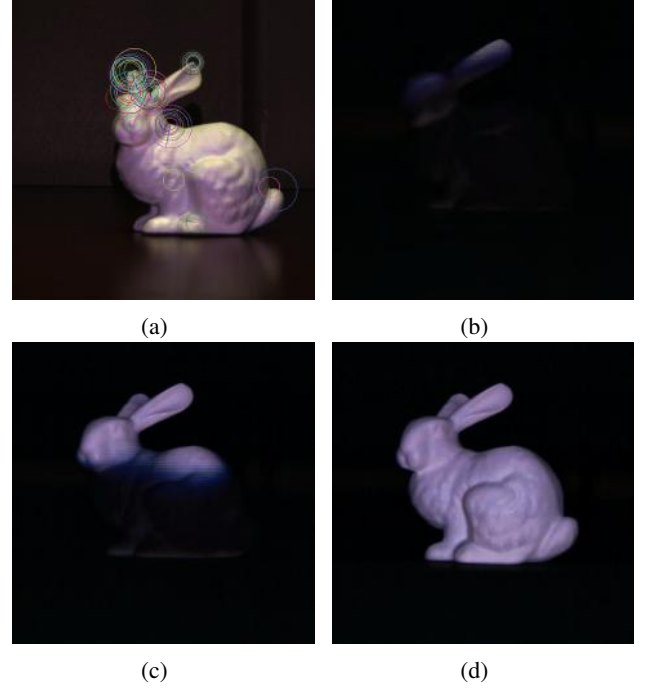


Fig. 7: Automatic parameter optimization: (a) shows the ORB keypoints detection which are then stored; (b) is the initial plane intersection. The pixel clock is swept until at least one keypoint is identified in the scene; (c) delay t_d is increased to maximize the number of detected keypoints without losing any; and (d) finally the exposure time t_e is increased so that the volume captures the entire object.

the Mitsuba extensions provided by Tian and Gkioulekas to perform these simulations [59]) can remove most indirect light paths, but still cannot remove scattering on the epipolar (mostly direct) light paths. This is shown in Figure 8(b). Our method is able to clearly visualize the box without scattering, showing the promise for vision in bad weather. This is shown in Figure 8(c). Note that transient imaging techniques have shown similar capabilities but with time-of-flight sensors and more computational post-processing [2], [60]–[62].

For our real hardware results, in Figure 9(a), we image a small scene of cars driving on a road. A fog machine is used to generate fog and reduce the visibility of the camera (Figure 9(b)). In Figure 9(c) and (d) is a capture of the scene with and without fog using a regular camera. In Figure 9(e), we capture the car in the front and in Figure (f), we capture a portion of the turn symbol on the road using conventional disparity gating [3]. Because of the vertical geometry of the gated regions, conventional disparity gating reduces the flexibility of the capture and controls how much scene is captured. In contrast, our method of slope disparity gating shown in Figure 9(g) can capture the full turn symbol by using a sloped planar region to capture the arrow. In Figure 9 (h), even though conventional disparity gating can resolve the stop sign, there is still issue with less visibility on upper region because the sign is placed on the sloped road. In contrast, our slope disparity gating can resolve the stop sign in Figure 9(i), and we can also expand the volumetric region to see neighbouring entities on the road in Figure 9(j). For video results, please view the

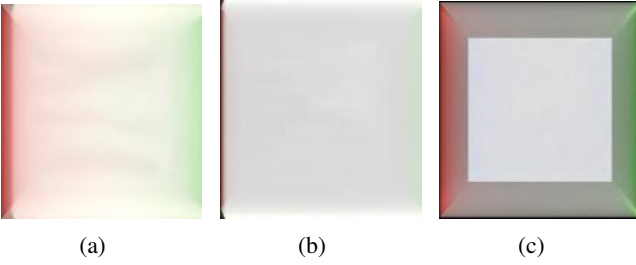


Fig. 8: (a) is a simulation of the Cornell box scene in a room with fog; (b) is the result of epipolar imaging as a form of light transport probing; (c) shows our slope disparity result. Note how our method is able to resolve the entire Cornell box without scattering.

supplementary material for imaging through fog.

Our method’s imaging capabilities in the presence of scattering media is governed by the contrast preserved while rejecting backscattering photons from the media. To validate the effectiveness of our contrast improvement, we graph the effect of increasing the camera’s exposure time. In Figure 10, we empirically show the reduced normalized cross-correlation (NCC) between a reference image without fog and a foggy image as a function of exposure time, on sweeping exposure time from $300\mu\text{s}$ – $14,000\mu\text{s}$. As one can see, there is an optimal value where the sign is most exposed without much fog and then it decreases in contrast back to a regular image.

In addition, we plotted the modulation transfer function (MTF) of our system’s performance in fog, this is shown in Figure 11. The MTF describes how different objects and their textures would be sharply captured using the system as a function of the texture’s spatial frequency. We obtain the MTF by performing the well-known slanted edge measurement described previously in the literature [63]. Figure 11(b) is a regular camera capture of the slanted edge target in a foggy scene and Figure 11(c) is a capture using our slope disparity gating system of the slanted edge present in a foggy scene. Note that we utilize the same camera and optics for our system for both regular and disparity gating capture, and thus any differences in the MTF are due solely to changes in contrast due to scattering. As one can see from the graph, our disparity gating preserves higher spatial frequencies with an MTF50 of 0.201 cycles per pixel as compared to MTF50 of 0.14 cycles/pixel for conventional imaging, and captures a sharper, higher contrast image.

Vein Visualization: We can image the veins under the skin in visible light using our system. By controlling the hardware parameter, delay, to capture the short-range indirect light scattered by the vein, we image veins, similar to the approach of [7] using non-epipolar light captured via Episcan3D. Using slope disparity gating, we can also change the pixel clock and thus the slope of capture for these vein images. In Figure 12, we see two arms with different vertical slopes in (a) and (d). By changing the pixel clock, we are able to selectively image the veins in the arm at different orientations in (c) and (e). This shows the importance of pixel clock to adjust the disparity gating for off-axis limbs. A limitation of this

experiment is that the imaging can be done only in the sloped vertical direction, since, horizontal disparity gating is not possible with our hardware configuration. Thus a method such as non-epipolar imaging [7] may be more practical for robust vein visualization, but this example still shows the ability of disparity gating to also visualize these structures.

VII. LIGHT TRANSPORT APPLICATIONS

In this section, we present several applications of light transport with our system. We show a mixture of rendered synthetic scenes as well as real experimental data captured from our prototype hardware setup. We begin with results for our novel disparity-dependent relighting. We show how this method can be used for projecting high frequency illumination patterns in a scene and improving the fog visualization results obtained in Section VI-C. Following this we present results for light transport based green screen compositing in VII-B, which combines both the ideas of image masking and the disparity-based relighting.

A. Disparity-based Relighting

As described in Section IV-A, the light transport probing technique can be used to visualize the disparity in the scene by color coding the various slices captured by the slope disparity system.

Synthetic Data: We first confirm in simulation that Hadamard codes give us improved light transport acquisition at higher SNR as compared to turning on one projector pixel at a time. In Figure 13, we show the floodlit results of a Cornell Box scene with fog. Note that our method is able to recover the scene while the naive impulse scanning acquisition suffers from SNR issues on the walls and back of the box.

In Figure 14, we show the effects of capturing disparity-based light transport information for two different synthetic scenes. We utilized a projector resolution of 4×4 for the illumination dimension, thus only showing low spatial frequency relightings in this example. Note that, projector resolution here refers to the order of the Hadamard matrix used for probing the scene. How the light transport probing is expressed as a function of the time delay t_d and exposure t_e has been described in depth in Section IV. In Figure 14(a), we show the floodlit image from using the full light transport captured in the scene. In (b), we show the results of reconstructing the floodlit using $\sum_{t_d} \mathbf{T}_{direct}(t_d)$ using 30 delays and no weighting factor. In (c) we show the results of reconstructing the floodlit when we include the weighting factor $\frac{dz}{dt_d}$ as in Equation 15. We see that the weighting factor makes a difference to reconstruct back to the original light transport. Note that for the bunny PSNR was 32.4dB for the light transport reconstruction and the SSIM was 0.981 for comparing the floodlit images. For the winged horse, PSNR was 24.8dB, and SSIM was 0.974. These results seem to confirm with our theory and equations in Section IV.

Real data: We also performed our disparity-dependent relighting on real scenes. In Figure 15, we show the results from our hardware prototype. We utilized a projector resolution of 4×2 with 4 disparity slices. In this scene, each object of

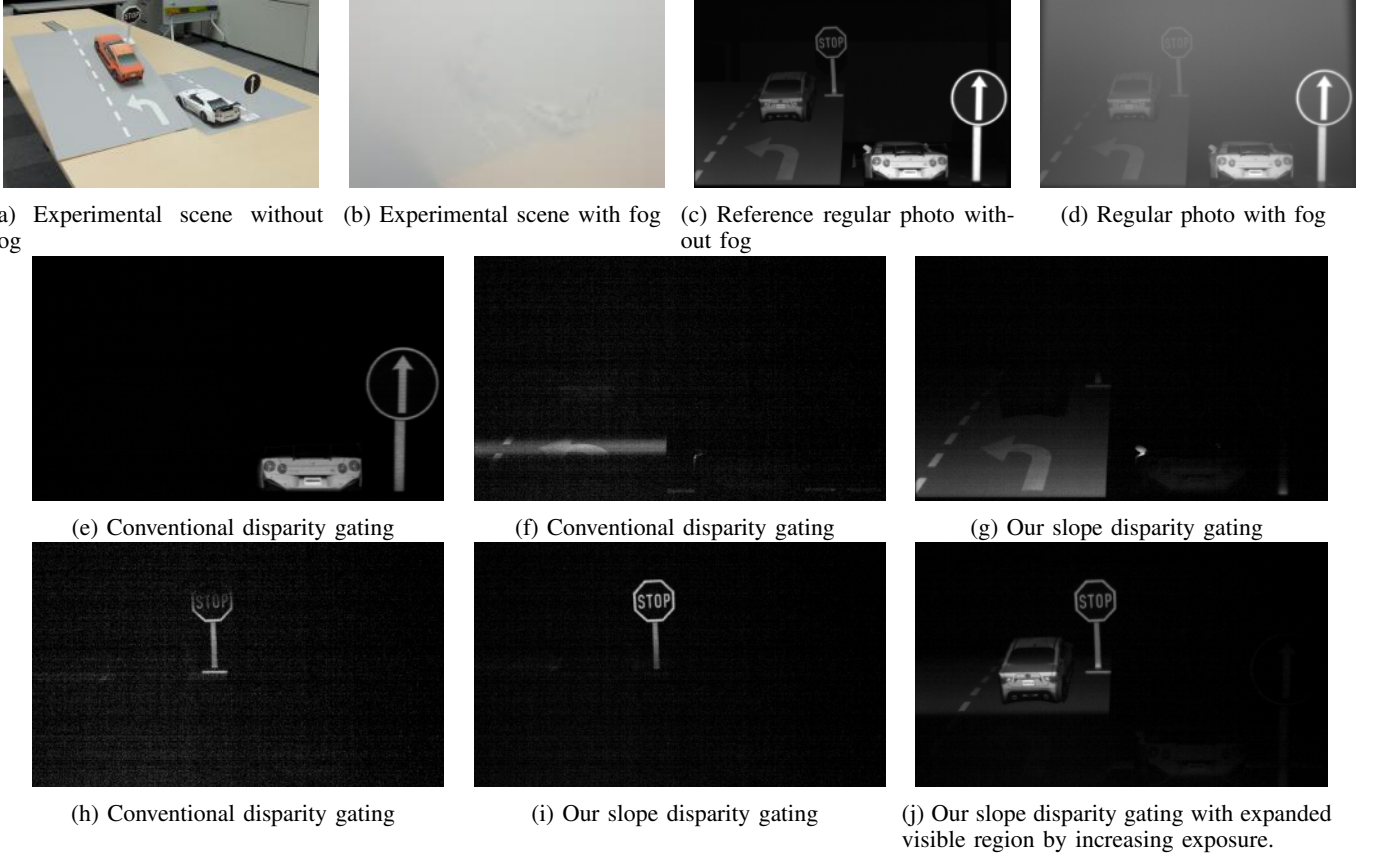


Fig. 9: Imaging through scattering media(fog). A sample test traffic scene that was built in the lab (a) after being filled with fog is shown in (b). Compared to a regular photo without fog (c), the captured image in fog (d) decreases the contrast and visibility for regular cameras. Conventional disparity gating [3] can significantly enhance the visibility of the objects on the same vertical plane (e), but can not visualize tilted surfaces (f). By controlling the synchronization delay and pixel clock, our device can visualize the entire area of the tilted surface and resolve the arrow sign on the road (g). Similarly, while conventional disparity gating can partially capture the sign on the tilted surface (h). Our method can capture the sign entirely (i). Moreover, our device can expand the visible volume by increasing exposure time as shown in (j). Each image in this figure was an average of 5 frames to improve the visualization.

various material types are placed at different disparities, and each is relit with a different solid color. This shows the proof of concept that this relighting works on our actual system.

High frequency illumination: In the previous examples, we showed relatively modest spatial frequency illuminations. To show that our method can scale up to very large light transport matrices, particularly high frequency illumination, we captured a real scene with projector resolution of 512×512 . In Figure 16, we show the results of disparity-dependent relighting on the scene in (a). Here the backwall is illuminated with a high resolution ACM Siggraph logo (b) while the bottle is masked, and then alternatively only the bottle is relit with green specularities in (c) while the wall is masked. Finally the resulting composite image showcase the combination of the two effects in (d). Note how the green specularities are not cast upon the backwall as the light transport is disparity-dependent, and thus yielding an interesting, non-physically realistic effect.

We compare our results against two state-of-the-art deep learning based matting methods. The first method of Forte et al. [64] estimates the foreground, background and the

alpha map from a single natural image and a good trimap estimate [66]. We also compare against Sengupta et al. [65], which requires an additional photo of the background without the subject at the time of capture. As shown in Figure 16(f), the trimap based method is unable to perform foreground separation of the glass bottle. Also shown in Figure 16(g) is the trimap free method. The caustics associated with the bottle which appears on the background is estimated to belong to the foreground. Also, only the top half of the bottle is estimated. Clearly our technique can perform better background-foreground separation in scenes and thus better relighting of required depth slices, containing caustic effects, and strong reflections.

We also attempt to compare our disparity based relighting with a depth based relighting method. Thus we obtain a depth map of our scenes using the Intel RealSense Depth Camera 82635. The result of the depth map captured is shown in Figure 16(h). The depth camera is unable to handle the specularity in the scene. The goal is to perform relighting on the various depth slices to obtain an equivalent comparison,

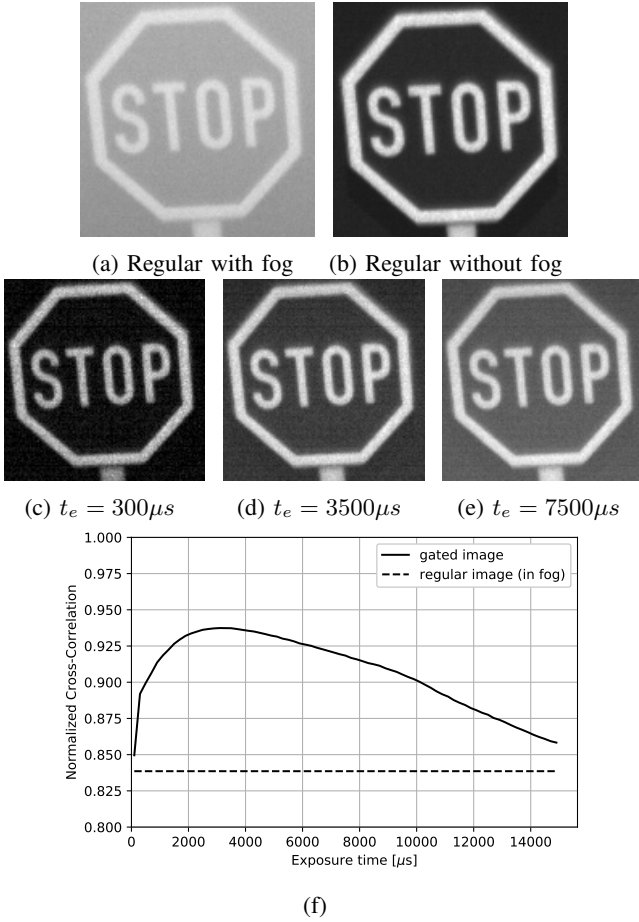


Fig. 10: Exposure time vs. Normalized Cross-Correlation (NCC). (a) – (e) The displayed signs are cropped from the same scene in Figure 9. (a) Regular image in fog. (b) Regular image without fog as a reference. (c) – (e) The gated images acquired for varying exposure time t_e while fixing delay $t_d = 12980$ to image the stop sign. (f) The NCC of each image compared to (b), sweeping $t_e = 300 \sim 14900 \mu s$. Each image was averaged over 5 frames to enhance visualization.

however, since the depth estimation has errors, we were unable to relight the depth slices.

Improving fog visualization: Using our light transport acquisition techniques, we can improve on the results presented in Section VI-C. Since we can obtain the full light transport, we can perform selective visualization of the slices, and thus focus illumination only on certain objects in the scene.

In Figure 17, we conducted a real fog experiment to demonstrate the improved visualization in fog. We utilized a projector resolution of 512×512 with 2 disparity slices. In Figure 17(a) we show the room with the generated fog from an observer camera. In (b) and (c) we show that we can relight the scenes for better post-capture visualization, and focus the light on just the targets of interest. Figure 17(d) is the depth map of the scene once again the depth sensor fails to handle scattering media correctly. This experiment helps validate our approach as offering a way to focus viewer’s attentions to salient parts of the image using programmable illumination.

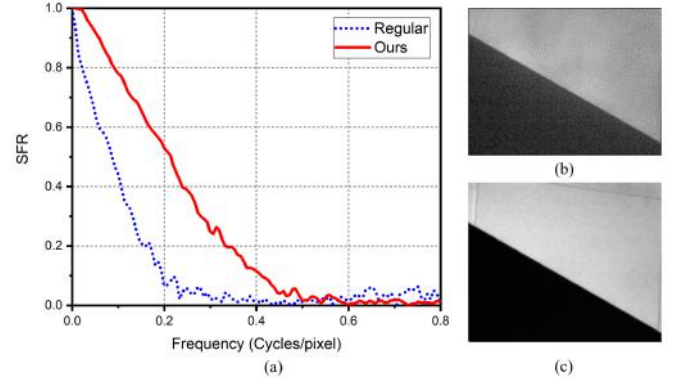


Fig. 11: (a) shows the MTF of both our system and a regular camera capture of a target in fog; (b) a slant edge pattern that is projected in a foggy scene under regular camera capture, and has MTF50 of 0.140 cycles per pixel; (c) is the capture using our disparity gating system with an MTF50 of 0.201 cycles per pixel.

B. Light Transport-based Green Screen Compositing

One important application of light transport relighting is in conjunction with image compositing for visual effects. In image compositing, a matte of per-pixel color and alpha values is utilized to composite images from multiple domains together. Typically this is achieved using a green screen or labeled trimaps as described in Section VII-A. However, scenes with small structures such as hair or those which exhibit complex lighting effects such as specularities can pose real challenges for image matting. Please see [67] for a comprehensive survey of techniques for image and video matting.

With our method, we obtain an easy way to perform image compositing based on the depth difference between the foreground objects and background. We compare our method against the previous two deep learning methods for comparison [64], [65].

In Figure 18, we show an example from our real hardware prototype performing light transport-aware green screening via disparity gating. We utilize a projector resolution of 32×32 with 2 slices in the scene. In (a) the original scene as shot on a green screen background is shown. In (b) and (c) we acquire the light transport based on foreground and background disparity slices and show a separation. We can then composite a desired background image correctly in the scene without errors as shown in (d). We even utilize the light transport of the foreground to realistically illuminate the foreground objects with physically-realistic lighting to match the background as shown. Notice how the pepper’s specularities take on the color of the orange background. In (f) we can observe severe errors when we attempt a HSV Matting on (a), due to the green objects in the foreground. We also compare our result against (g) and (h) which show background and foreground separation using deep learning based matting method of Forte et al. [64]. And (i) shows the result of foreground separation using trimap free method of Sengupta et al. [65] which also produces several errors. Thus our method provides a light-transport

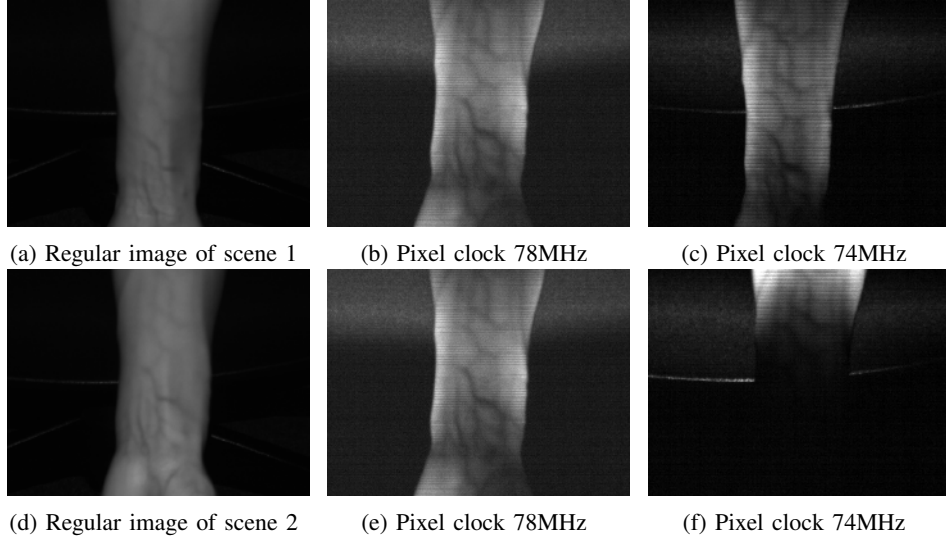


Fig. 12: We visualize human veins by capturing short range indirect light. This requires setting an appropriate value of pixel clock. Here we tested two scenes of tilted arms with different angles. Each image in this figure is an average of 10 frames for enhanced visualization. For the tilted arm shown in (a), the veins of the arm are more visible with a pixel clock of 74 MHz as shown in (c) as compared to (b) with pixel clock of 78 MHz. For another arm oriented differently in (d), image (e) with pixel clock of 78 MHz is more visible as compared to the image with pixel clock of 74 MHz in (f). Note that using conventional disparity gating one cannot change the pixel clock, thus, only our system can perform this imaging for different vertically sloped arms.

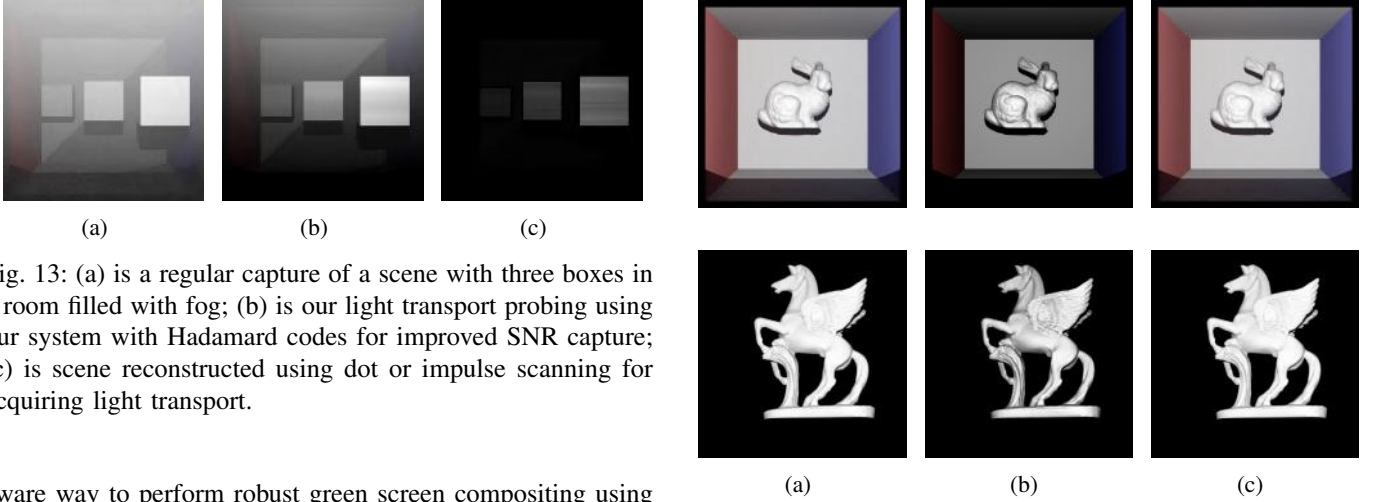


Fig. 13: (a) is a regular capture of a scene with three boxes in a room filled with fog; (b) is our light transport probing using our system with Hadamard codes for improved SNR capture; (c) is scene reconstructed using dot or impulse scanning for acquiring light transport.

aware way to perform robust green screen compositing using photometric and geometric information.

VIII. DISCUSSION

In this paper, we present an active illumination system that is able to perform gated imaging controllable for planes sloped in the y-z plane. Using a vertically-aligned configuration of the Episcan3D system [3], the system can be controlled using the hardware parameters such as the synchronization delay, pixel clock and exposure. The key application of the system includes selective region-masking in real time and improving imaging through scattering media such as fog which are demonstrated using real results in this paper.

We also present a new probing of light transport based on our disparity gating system. We formulate the general framework for this light transport probing and discuss how the

Fig. 14: Light transport acquisition using disparity gating. Here we show comparisons of a floodlit image reconstructed from original light transport (a), summation of disparity components without (b) and with (c) the weighting factor. Note for the bunny in (c), the PSNR was 32.4dB and the SSIM was 0.981 compared the floodlit images in (a). For the winged horse, the PSNR was 24.8dB, and the SSIM was 0.974.

hardware parameters of the system determine these probing matrices Π . Finally, we also show applications using both rendered and real experimental scenes that included disparity-dependent relighting. We showed how the visualization in fog can be improved by our disparity-based relighting technique. This technique also enables relighting for the image

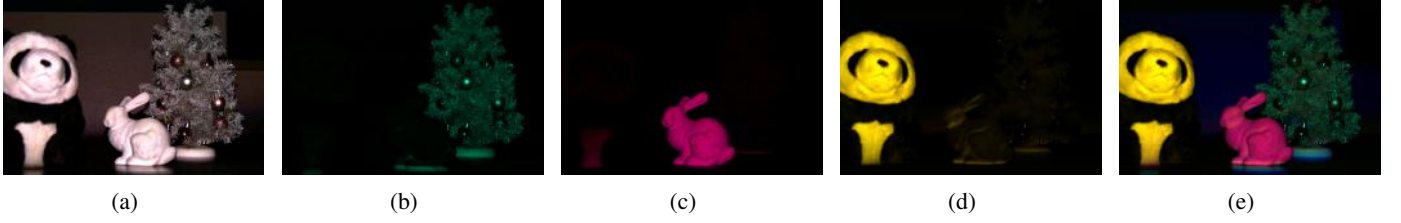


Fig. 15: Real Experimental Scene. Here we show results captured from our prototype hardware setup: (a) Floodlit image; (b)-(d) colored disparity slices; (e) Relit scene from summing these slices.

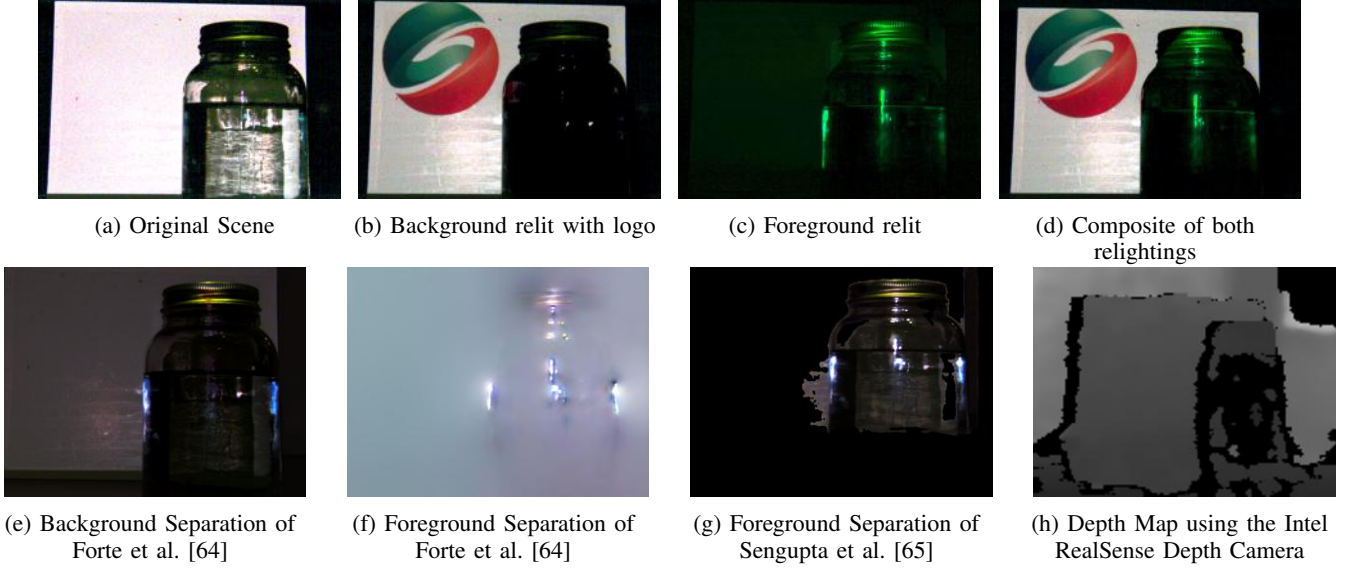


Fig. 16: Real experimental scene with caustic effects relighting using high frequency illumination of 512×512 . Here we show results captured from our prototype hardware setup: (a) Floodlit image; (b)-(c) colored disparity slices; (e) Relit scene from summing these slices. (f) Foreground separation by [64] (g) Foreground separation by [65] (h) Depth map of scene

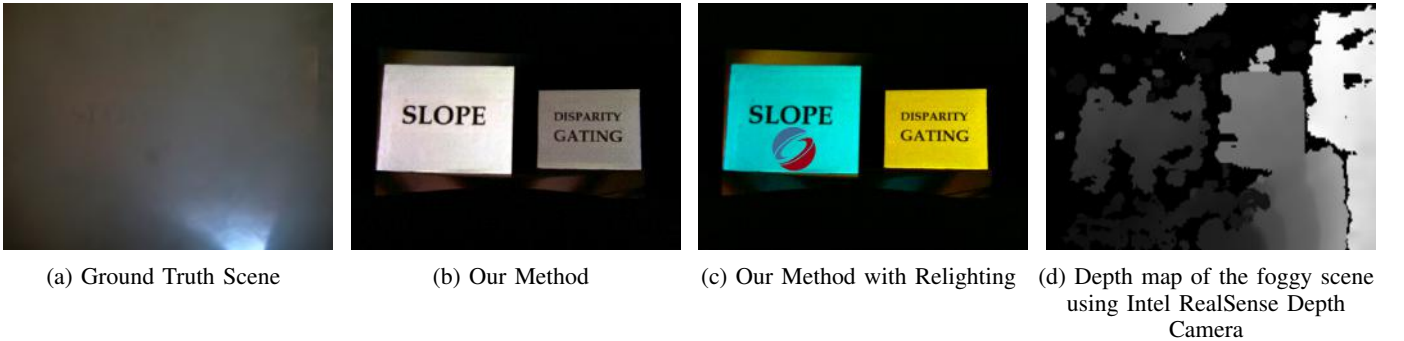


Fig. 17: Real Experimental Scene: Fog. We show the results from a real fog machine using our hardware prototype. Given the scene in (a), a regular camera with normal exposure yields image (b) In contrast, our disparity slices have high visibility and SNR (c) Utilizing the light transport decomposition, we can relight these images to remove spurious captured regions such as the sidewalls in the background of slice 1, and color the targets or project patterns like arrows to enhance visualization even further (d) Depth map of scene. Please note the images has been brightened for visualization here

matting example for only the foreground in a conceptually straightforward way. When compared to state-of-the-art image matting methods [64], [65], which rely on a good estimation of the trimap and training datasets, our method performs better foreground-background separation with the ability to relight the scene.

Comparison to Programmable Light Curtains: Pro-

grammable light curtains (PLCs) [20], [21] can be viewed as a generalization of disparity gating systems where they are able to triangulate an arbitrary 3D surface. As such, our system which can only triangulate a plane in the y-z plane does not have this level of programmability mainly due to the fact that we use a fixed raster-scanning laser projector and do not

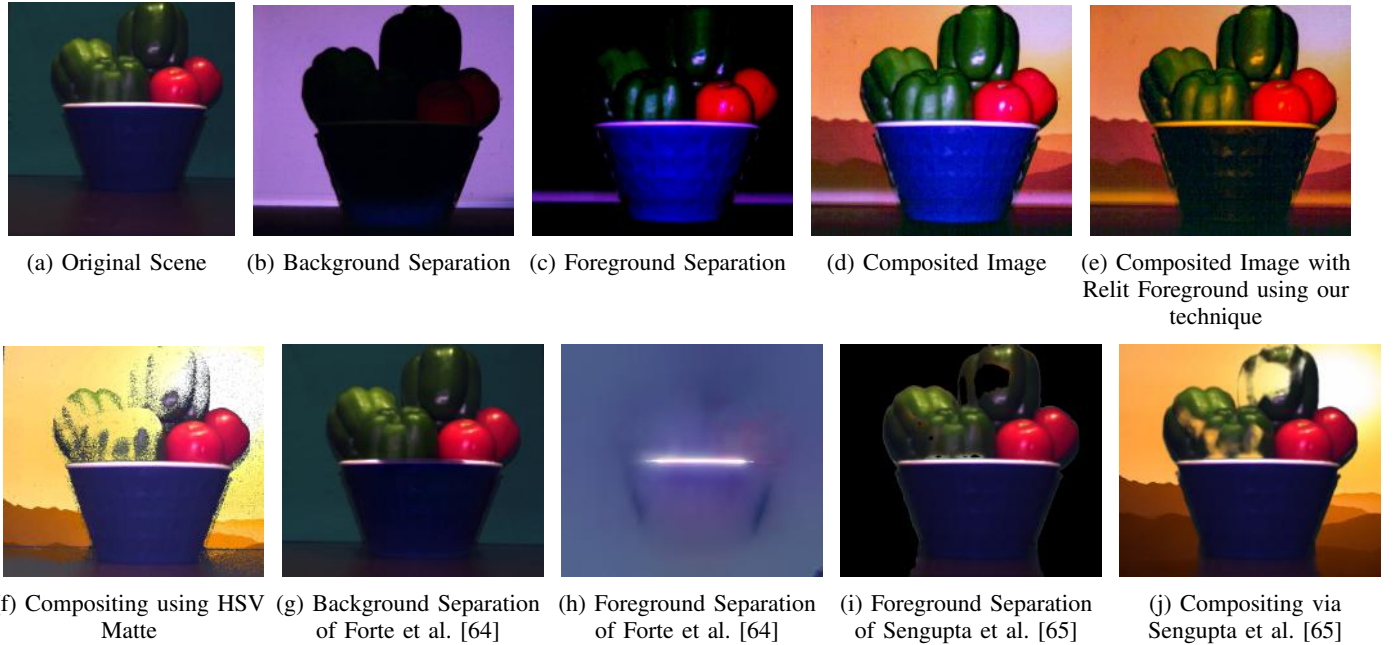


Fig. 18: We show an application to image matting and compositing. (a) The original scene as shot on a green screen background (b) and (c) shows the background and foreground separation using our disparity gating based light transport acquisition method (d) We can then composite a desired background image correctly in the scene without errors in (e) We even utilize the light transport of the foreground to realistically illuminate the foreground objects with physically-realistic lighting (h) Severe errors are obtained when we attempt a HSV Matting on (a), due to the green objects in the foreground (g) and (h) shows background and foreground separation using state-of-the-art deep learning based matting method [64], (i) shows the result of foreground separation using [65] and (j) shows the result of compositing using the same method. While this method manages to handle the basket region it clearly fails to handle the peppers in the scene.

have control over the light source direction as PLCs do. We acknowledge that this is a main advantage of the PLCs over our system for sensing applications. However, it is possible for our system to increase the thickness of the intersection plane with the camera exposure digitally, which is not possible with the PLC without modifying the physical baseline between the light source and camera (see Section IV of [20]).

The main comparative advantage of slope disparity gating is the ability to perform light transport acquisition. While it may be theoretically possible for a PLC to acquire light transport (although this has not been demonstrated to the best of our knowledge in the literature), our projector allows us to perform light multiplexing and projecting content into the scene while performing disparity gating. For a PLC to perform light transport acquisition, it must capture images of a flying dot or line scanning over the scene as the columns of the light transport matrix. This type of impulse scanning would suffer from low SNR (as discussed in other flying dot setups [31]) while a projector can perform Hadamard multiplexing.

Limitations: Due to the arrangement of the projector camera system, our method is limited to imaging only in the $y - z$ plane. With regard to the proposed light transport decomposition method, we capture direct light only and do not take into account the indirect light. Our probing method requires more acquisition time than traditional light transport (minutes to hours depending on illumination resolution), which scales with the number of disparity slices that are captured in

the scene. Also, due to the 1D nature of our triangulation using a line projector, we cannot perform arbitrary ray-based relighting along a 3D surface as one could theoretically do with a programmable light curtain system [20], [21].

Future Work: There are many avenues of future work for exploring light transport for such triangulation-based systems. In particular, using more fine-grained triangulation such as programmable light curtains [20], [21] might allow ray-dependent light transport decompositions, or probing individual entries of the light transport matrix. This could allow for very interesting post-capture control of scenes in terms of illumination. Further, we are also interested in fast capture systems such as in [31], [32] which could enable such probing to occur for dynamic light transport in moving scenes.

ACKNOWLEDGEMENTS

This work was supported by NSF grant IIS-1909192, JSPS Kakenhi(JP19H04138, JP21K19799), JST CREST (JPMJCR1764), JST FOREST (JPMJFR2061), and JST SICORP (JPMJSC2003).

REFERENCES

- [1] O. David, N. S. Kopeika, and B. Weizer, "Range gated active night vision system for automobiles," *Applied Optics*, vol. 45, no. 28, pp. 7248–7254, 2006.
- [2] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar, "Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles," *ACM Transactions on Graphics*, vol. 32, no. 6, p. 167, 2013.

- [3] M. O'Toole, S. Achar, S. G. Narasimhan, and K. N. Kutulakos, "Homogeneous codes for energy-efficient illumination and imaging," *ACM Transactions on Graphics*, vol. 34, no. 4, pp. 35:1–35:13, Jul. 2015. [Online]. Available: <http://doi.acm.org/10.1145/2766897>
- [4] J. Wang, J. Bartels, W. Whittaker, A. C. Sankaranarayanan, and S. G. Narasimhan, "Programmable triangulation light curtains," in *The European Conference on Computer Vision (ECCV)*, September 2018.
- [5] M. O'Toole, R. Raskar, and K. N. Kutulakos, "Primal-dual coding to probe light transport," *ACM Trans. Graph.*, vol. 31, no. 4, pp. 39–1, 2012.
- [6] H. Kubo, S. Jayasuriya, T. Iwaguchi, T. Funatomi, Y. Mukaigawa, and S. G. Narasimhan, "Acquiring and characterizing plane-to-ray indirect light transport," in *Computational Photography (ICCP), 2018 IEEE International Conference on*. IEEE, 2018, pp. 1–10.
- [7] —, "Programmable non-epipolar indirect light transport: Capture and analysis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 4, pp. 2421–2436, 2019.
- [8] T. Ueda, H. Kubo, S. Jayasuriya, T. Funatomi, and Y. Mukaigawa, "Slope disparity gating using a synchronized projector-camera system," in *2019 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2019, pp. 1–9.
- [9] J. Geng, "Structured-light 3d surface imaging: a tutorial," *Advances in Optics and Photonics*, vol. 3, no. 2, pp. 128–160, Jun 2011. [Online]. Available: <http://aop.osa.org/abstract.cfm?URI=aop-3-2-128>
- [10] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical Engineering*, vol. 19, no. 1, p. 191139, 1980.
- [11] S. B. Gokturk, H. Yalcin, and C. Bamji, "A time-of-flight depth sensor-system description, issues and solutions," in *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW'04. Conference on*. IEEE, 2004, pp. 35–35.
- [12] B. Schwarz, "Lidar: Mapping the world in 3d," *Nature Photonics*, vol. 4, no. 7, p. 429, 2010.
- [13] M. Gupta, S. G. Narasimhan, and Y. Y. Schechner, "On controlling light transport in poor visibility environments," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2008, pp. 1–8.
- [14] S. G. Narasimhan, S. K. Nayar, B. Sun, and S. J. Koppal, "Structured light in scattering media," in *IEEE International Conference on Computer Vision (ICCV)*, vol. 1. IEEE, 2005, pp. 420–427.
- [15] T. Treibitz and Y. Y. Schechner, "Active polarization descattering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 3, pp. 385–399, 2008.
- [16] L. Tian, X. Li, K. Ramchandran, and L. Waller, "Multiplexed coded illumination for fourier ptychography with an led array microscope," *Biomedical Optics Express*, vol. 5, no. 7, pp. 2376–2389, 2014.
- [17] D. Bonnier and V. Laroche, "Range-gated active-imaging system for search-and-rescue and surveillance operations," in *Infrared Technology and Applications XXII*, vol. 2744. International Society for Optics and Photonics, 1996, pp. 134–146.
- [18] Y. Grauer and E. Sonn, "Active gated imaging for automotive safety applications," in *Video Surveillance and Transportation Imaging Applications 2015*, vol. 9407. International Society for Optics and Photonics, 2015, p. 9407OF.
- [19] R. Tadano, A. Kumar Pediredla, and A. Veeraraghavan, "Depth selective camera: A direct, on-chip, programmable technique for depth selectivity in photography," in *IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 3595–3603.
- [20] J. Wang, J. Bartels, W. Whittaker, A. C. Sankaranarayanan, and S. G. Narasimhan, "Programmable triangulation light curtains," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 19–34.
- [21] J. R. Bartels, J. Wang, W. Whittaker, S. G. Narasimhan *et al.*, "Agile depth sensing using triangulation light curtains," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 7900–7908.
- [22] Y. Raaj, S. Ancha, R. Tamburo, D. Held, and S. G. Narasimhan, "Exploiting & refining depth distributions with triangulation light curtains," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2021, pp. 7434–7442.
- [23] 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*, 2000, pp. 145–156.
- [24] T. Hawkins, J. Cohen, C. Tchou, and P. Debevec, "Light Stage 2.0," in *SIGGRAPH Technical Sketches*, 2001, p. 217.
- [25] P. Debevec, A. Wenger, C. Tchou, A. Gardner, J. Waese, and T. Hawkins, "A Lighting Reproduction Approach to Live-Action Compositing," in *SIGGRAPH 2002*, San Antonio, TX, Jul. 2002, pp. 547–556.
- [26] A. Wenger, A. Gardner, C. Tchou, J. Unger, T. Hawkins, and P. Debevec, "Performance relighting and reflectance transformation with time-multiplexed illumination," in *ACM Transactions on Graphics (TOG)*, vol. 24, no. 3. ACM, 2005, pp. 756–764.
- [27] P. Einarsson, C.-F. Chabert, A. Jones, W.-C. Ma, B. Lamond, T. Hawkins, M. Bolas, S. Sylwan, and P. Debevec, "Relighting human locomotion with flowed reflectance fields," in *Proceedings of the 17th Eurographics conference on Rendering Techniques*. Eurographics Association, 2006, pp. 183–194.
- [28] T. Hawkins, P. Einarsson, and P. E. Debevec, "A dual light stage," *Rendering Techniques*, vol. 5, pp. 91–98, 2005.
- [29] P. Peers, T. Hawkins, and P. Debevec, "A Reflective Light Stage," University of Southern California Institute for Creative Technologies, ICT Technical Report ICT TR 04 2006, 2006.
- [30] M. O'Toole, F. Heide, L. Xiao, M. B. Hullin, W. Heidrich, and K. N. Kutulakos, "Temporal frequency probing for 5d transient analysis of global light transport," *ACM Transactions on Graphics (ToG)*, vol. 33, no. 4, p. 87, 2014.
- [31] K. Henderson, X. Liu, J. Folden, B. Tilmon, S. Jayasuriya, and S. Koppal, "Design and calibration of a fast flying-dot projector for dynamic light transport acquisition," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 529–543, 2020.
- [32] X. Liu, K. Henderson, J. Rego, S. Jayasuriya, and S. Koppal, "Dense lissajous sampling and interpolation for dynamic light-transport," *Optics Express*, vol. 29, no. 12, pp. 18 362–18 381, 2021.
- [33] M. O'Toole and K. N. Kutulakos, "Optical computing for fast light transport analysis," in *ACM Transactions on Graphics (TOG)*, vol. 29, no. 6. ACM, 2010, p. 164.
- [34] J. Wang, Y. Dong, X. Tong, Z. Lin, and B. Guo, "Kernel nyström method for light transport," *ACM Trans. Graph.*, vol. 28, no. 3, Jul. 2009. [Online]. Available: <https://doi.org/10.1145/1531326.1531335>
- [35] P. Sen and S. Darabi, "Compressive dual photography," in *Computer Graphics Forum*, vol. 28, no. 2. Wiley Online Library, 2009, pp. 609–618.
- [36] 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)*, vol. 28, no. 1, p. 3, 2009.
- [37] 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)*, vol. 25, no. 3. ACM, 2006, pp. 935–944.
- [38] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan, "A practical approach to 3d scanning in the presence of interreflections, subsurface scattering and defocus," *International Journal of Computer Vision*, vol. 102, no. 1-3, pp. 33–55, 2013.
- [39] S. Achar, S. T. Nuske, and S. G. Narasimhan, "Compensating for motion during direct-global separation," in *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 1481–1488.
- [40] J. Gu, T. Kobayashi, M. Gupta, and S. K. Nayar, "Multiplexed illumination for scene recovery in the presence of global illumination," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 691–698.
- [41] D. Reddy, R. Ramamoorthi, and B. Curless, "Frequency-space decomposition and acquisition of light transport under spatially varying illumination," in *European Conference on Computer Vision*. Springer, 2012, pp. 596–610.
- [42] I. Gkioulekas, A. Levin, F. Durand, and T. Zickler, "Micron-scale light transport decomposition using interferometry," *ACM Transactions on Graphics (ToG)*, vol. 34, no. 4, pp. 1–14, 2015.
- [43] A. Kotwal, A. Levin, and I. Gkioulekas, "Interferometric transmission probing with coded mutual intensity," *ACM Transactions on Graphics (TOG)*, vol. 39, no. 4, pp. 74–1, 2020.
- [44] S.-H. Baek and F. Heide, "Polarimetric spatio-temporal light transport probing," *ACM Transactions on Graphics (TOG)*, vol. 40, no. 6, pp. 1–18, 2021.
- [45] 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)*, 2014, pp. 3246–3253.
- [46] M. O'Toole, J. Mather, and K. N. Kutulakos, "3d shape and indirect appearance by structured light transport," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 7, pp. 1298–1312, 2016.
- [47] S. Achar, J. R. Bartels, W. L. Whittaker, K. N. Kutulakos, and S. G. Narasimhan, "Epipolar time-of-flight imaging," *ACM Transactions on Graphics*, vol. 36, no. 4, p. 37, 2017.

- [48] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur, "Multiplexing for optimal lighting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 8, pp. 1339–1354, 2007.
- [49] D. Wu, G. Wetzstein, C. Barsi, T. Willwacher, M. O'Toole, N. Naik, Q. Dai, K. Kutulakos, and R. Raskar, "Frequency analysis of transient light transport with applications in bare sensor imaging," in *European Conference on Computer Vision*. Springer, 2012, pp. 542–555.
- [50] A. Jarabo, J. Marco, A. Munoz, R. Buisan, W. Jarosz, and D. Gutierrez, "A framework for transient rendering," *ACM Transactions on Graphics (ToG)*, vol. 33, no. 6, pp. 1–10, 2014.
- [51] M. O'Toole, F. Heide, D. B. Lindell, K. Zang, S. Diamond, and G. Wetzstein, "Reconstructing transient images from single-photon sensors," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 1539–1547.
- [52] D. B. Lindell, M. O'Toole, and G. Wetzstein, "Towards transient imaging at interactive rates with single-photon detectors," in *2018 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2018, pp. 1–8.
- [53] P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. Lensch, "Dual photography," in *ACM Transactions on Graphics (TOG)*, vol. 24, no. 3. ACM, 2005, pp. 745–755.
- [54] M. Pharr, W. Jakob, and G. Humphreys, *Physically based rendering: From theory to implementation*. Morgan Kaufmann, 2016.
- [55] D. Moreno and G. Taubin, "Simple, accurate, and robust projector-camera calibration," in *2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*. IEEE, 2012, pp. 464–471.
- [56] F. Pittaluga and S. J. Koppal, "Pre-capture privacy for small vision sensors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 11, pp. 2215–2226, 2017.
- [57] M. Tsuji, H. Kubo, S. Jayasuriya, T. Funatomi, and Y. Mukaigawa, "Touch sensing for a projected screen using slope disparity gating," *IEEE Access*, vol. 9, pp. 106 005–106 013, 2021.
- [58] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: An efficient alternative to sift or surf," in *2011 International Conference on Computer Vision*. IEEE, 2011, pp. 2564–2571.
- [59] J. Sun and I. Gkioulekas, "Mitsuba renderer extensions for computational light transport." [Online]. Available: https://github.com/cmu-ci-lab/mitsuba_clt
- [60] F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich, "Low-budget transient imaging using photonic mixer devices," *ACM Transactions on Graphics*, vol. 32, no. 4, p. 45, 2013.
- [61] S. Jayasuriya, A. Pediredla, S. Sivaramakrishnan, A. Molnar, and A. Veeraraghavan, "Depth fields: Extending light field techniques to time-of-flight imaging," in *2015 International Conference on 3D Vision*. IEEE, 2015, pp. 1–9.
- [62] A. Kadambi, H. Zhao, B. Shi, and R. Raskar, "Occluded imaging with time-of-flight sensors," *ACM Transactions on Graphics (ToG)*, vol. 35, no. 2, pp. 1–12, 2016.
- [63] P. D. Burns *et al.*, "Slanted-edge mtf for digital camera and scanner analysis," in *Is and Ts Pics Conference*. Society for Imaging Science & Technology, 2000, pp. 135–138.
- [64] M. Forte and F. Pitié, "F,b.alpha matting," *CoRR*, vol. abs/2003.07711, 2020.
- [65] S. Sengupta, V. Jayaram, B. Curless, S. M. Seitz, and I. Kemelmacher-Shlizerman, "Background matting: The world is your green screen," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 2291–2300.
- [66] Y.-Y. Chuang, B. Curless, D. H. Salesin, and R. Szeliski, "A bayesian approach to digital matting," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, vol. 2. IEEE, 2001, pp. II–II.
- [67] J. Wang and M. F. Cohen, *Image and video matting: a survey*. Now Publishers Inc, 2008.



Sreenithy Chandran is currently a Ph.D. student at the School of Electrical, Computer and Energy Engineering (ECEE) at Arizona State University. She received her MS in ECEE from ASU in 2019, and a BEng in Electronics and Communications from Anna University, India in 2017. Her research interests are in computer vision and machine learning, particularly light transport and non-line-of-sight (NLOS) imaging.



Hiroyuki Kubo is an associate professor at Chiba university, Japan since 2022. He received his M.S. and Ph.D. degrees from Waseda University, in 2008 and 2012, respectively. His research interests include computer graphics, computer vision and computational photography.



Tomoki Ueda graduated with his MS from Nara Institute of Science and Technology (NAIST), Japan in 2019. His research interests are in computational photography and computer vision.



Takuya Funatomi is currently an associate professor at the Division of Information Science, Nara Institute of Science and Technology (NAIST), Japan since 2015 and concurrently a PRESTO Researcher of Japan Science and Technology Agency (JST) since 2020. He was an assistant professor at Kyoto University from 2007 to 2015, and a visiting assistant professor at Stanford University in 2014. He received B.E., M.S. and PhD in Informatics from Kyoto University in 2002, 2004, and 2007, respectively. His research interests include computational photography, 3D model processing, computer vision, computer graphics, and pattern recognition. He is a senior member of IEEE.



member of IEEE.

Yasuhiro Mukaigawa is currently a professor at division of Information Science, Nara Institute of Science and Technology (NAIST), Japan since 2014. He was a research associate at Okayama University, Japan from 1997 to 2003, an assistant professor at University of Tsukuba, Japan from 2003 to 2004, an associate professor at Osaka University, Japan from 2004 to 2014. He received his ME and Ph.D. degrees from University of Tsukuba in 1994 and 1997, respectively. His research interests include photometric analysis and computational photography. He is a



of IEEE.

Suren Jayasuriya is an assistant professor at Arizona State University (ASU), USA since 2018. He was a postdoctoral fellow at the Robotics Institute at Carnegie Mellon University, USA in 2017. He received his B.S. degree in mathematics and a B.A. degree in philosophy from the University of Pittsburgh, PA in 2012, and received his Ph.D. degree in Electrical and Computer Engineering from Cornell University, USA in 2017. His research interests are in computational photography and imaging, computer vision, and image sensors. He is a member

Appendix: Slope Disparity Gating: System and Applications

Sreenithy Chandran, Hiroyuki Kubo, *Member, IEEE*, Tomoki Ueda, Takuya Funatomi, *Member, IEEE*, Yasuhiro Mukaigawa, *Member, IEEE*, and Suren Jayasuriya, *Member, IEEE*

DERIVATION OF THE CHANGE OF VARIABLES FORMULA FOR DISPARITY-DEPENDENT RELIGHTING

In the main paper, we described light transport probing in terms of matrices. Now we utilize this probing technique to perform decomposition of the light transport matrix based on disparity. This decomposition is used to perform a novel disparity dependent relighting discussed in the main paper. We begin by establishing the relationship between the light transport matrix and the world coordinate system, then we discuss this decomposition for both the vertical and sloped case of the imaging plane. Note that we assume dominant direct light $\mathbf{T}_{direct}$ for the following discussion, since, in Section 3, we assume direct (one bounce) light while triangulating the image formation model.

In Section 3, we showed that if we consider the effect of only delay t_d without exposure time, the system triangulates a 2D plane in the world for direct light (not indirect light). We assume $v_p = v_c$, thus, this plane is vertically-oriented with its normal in the $+z$ direction. Hence, every t_d corresponds to a unique xy -plane in the world coordinate system, i.e. $z_w(t_d)$ is a 1-to-1 function. Note for direct light, $\mathbf{T}_{direct}[i, p]$ only occurs at one unique z -plane, as light only bounces once in the scene before arriving at the camera. The summation along z will thus yield the following equation:

$$\mathbf{T}_{direct} = \int_z \mathbf{T}_{direct}(z) dz, \quad (1)$$

where z represents the depth at which the plane is along the z -axis.

As already stated, our system cannot capture $\mathbf{T}_{direct}(z)$ directly, so we must parametrize the equation in terms of delay t_d using a change in variables:

$$\mathbf{T}_{direct} = \int_z \mathbf{T}_{direct}(z, t_e) dz = \int_{t_d} \mathbf{T}_{direct}(z(t_d), t_e) \left| \frac{dz}{dt_d} \right| dt_d. \quad (2)$$

Note, we have fixed the value of t_e here. For a normal projector-camera system, $\mathbf{T}_{direct}$ is actually a function of exposure $\mathbf{T}_{direct}(t_e)$, but all this exposure does is uniformly scale the matrix by t_e (assuming linear measurements).

The change of variable term $\frac{dz}{dt_d}$ can be derived from Equation 10 as:

$$\frac{dz}{dt_d} = \frac{-bf_{cy}f_{py}^2v_c}{(c_{cy}f_{py} - c_{py}f_{cy} + f_{py}v_c(t - t_d) - f_{cy}v_pt)^2}. \quad (3)$$

Recall that f_{cy} and f_{py} are the camera and projector focal lengths, b is the baseline, and c_{cy} and c_{py} are the camera

and projector translation y -coordinates from their intrinsic matrices.

We can simplify this weighting factor, assuming $v_c = v_p = v$, $f_{cy} = f_{py} = f$, and $c_{cy} = c_{py} = c$ to get the following relationship:

$$\frac{dz}{dt_d} = \frac{-bf^3v}{(cf - cf + fv(t - t_d) - fvt)^2} = \frac{-bfv}{(vt_d)^2} \propto \frac{1}{t_d^2}. \quad (4)$$

This implies that the weighting factor is inversely proportional to t_d . In practice, we utilize $\left| \frac{dz}{dt_d} \right| = \frac{bf}{t_d^2}$ assuming $v_c = v_p = 1$.

General case: For all of our experimental results involving light transport in the main paper, we leverage the assumption that the imaging plane is vertical to simplify the acquisition. However, we can also consider the case where $v_p \neq v_c$, which results in the imaging plane to no longer be vertical but sloped. According to the Equations 9 and 10 of y_w and z_w , the slope of the equations can be obtained by using the following formula:

$$m = \frac{y_w(t=1) - y_w(t=0)}{z_w(t=1) - z_w(t=0)} \quad (5)$$

From this, we can compute $\theta = \tan^{-1}(m)$. Thus we can think of this as sloped planes traveling along some axis at an angle θ . Let's call this new axis r_w . Note that r_w is just z_w rotated clockwise around x -axis by $\pi/2 - \theta$. We can parameterize this axis $r = \frac{z_w}{\cos(\pi/2 - \theta)}$. Then we can still do the decomposition as follows:

$$\mathbf{T}_{direct} = \int_r \mathbf{T}_{direct}(r) dr = \int_{t_d} \mathbf{T}_{direct}(r(t_d)) \left| \frac{dr}{dt_d} \right| dt_d. \quad (6)$$

Thus, our methodology can also be used for handling sloped cases.