Transient Light Transport Matrices for Non-Line-of-Sight Imaging

Anonymous Submission

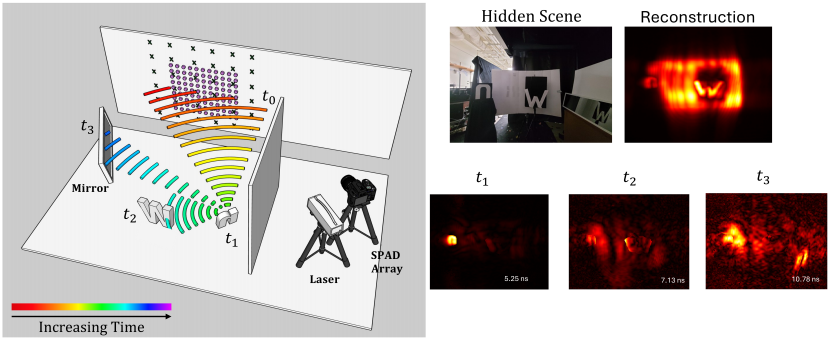

Figure 1. By employing the SPAD array to capture the complete laser and camera grid on the relay surface, we can create various virtual illumination patterns at the relay surface at time t_0 and project these patterns into the hidden scene. In this instance, we focus the illumination on the letter "n," which becomes illuminated at time t_1. Part of the light reflected from "n" then illuminates the left side of the letter "W" at time t_2, while the remaining light is reflected off a mirror at time t_3. By moving this focal spot around the hidden scene, we are able to construct the Transient Light Transport Matrix for the hidden scene.

Abstract

Transient Light Transport Matrices (LTM) leverage active imaging systems to measure the spatiotemporal light transport for a given scene. Time-resolved Non-line-of-sight (NLOS) imaging employs an active imaging system that uses transient measurements of indirect reflections of light to "see around corners", with applications in diverse areas such as disaster response, remote surveillance, and autonomous navigation. Existing NLOS systems usually measure a 3D subset of the full 5D transient measurement space. The development of fast-gated SPAD arrays [39] has made it feasible to probe the full measurement space. In this work, we demonstrate that the ***transient LTM for the hidden scene*** *can be computed from this measurement. We showcase three applications of Transient LTMs in NLOS imaging: Dual photography, Scene Relighting, and Separation of Indirect and Direct components. Additionally, we empirically demonstrate that SPAD arrays enable parallel acquisition of photons, effectively tackling the challenge of low SNR and long acquisition times.*

1. Introduction

Transient Light Transport Matrices (LTM) characterize spatiotemporal light transport from light sources to the detector for a given scene [28], and can be measured with an active imaging system using Time of flight detectors [31] (Fig 2). Once the LTM is measured, different aspects of the light transport can be manipulated in post-processing, such as separating direct from indirect light, specular from diffuse reflections, and even separating caustic from non-caustic effects [27, 30, 31].

Non-line-of-sight (NLOS) Imaging using time-of-flight (ToF) detectors is an example of an active imaging sys-

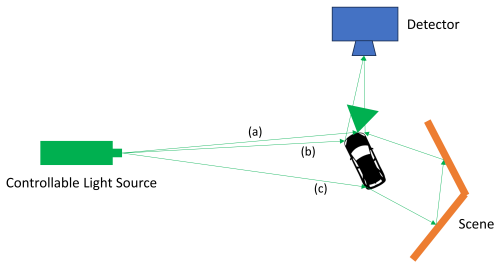

Figure 2. Active Imaging System for measuring Light Transport Matrix, which contains information about (a) Diffuse reflections (a), specular reflections (b), and (c) indirect components of Light Transport.

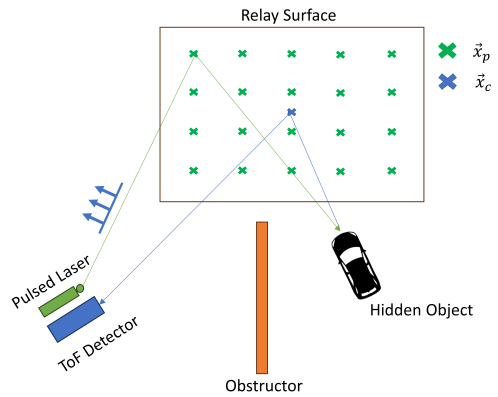

Figure 3. Non-line-of-sight Imaging System using Time of flight detectors is an active imaging system that measures the LTM of the relay surface.

tem that enables 2D and 3D visualizations of a surface hidden from the direct line of sight (LOS) of an observer [9, 46]. An intermediary relay surface in the LOS of both the observer and the hidden scene is scanned with a pulsed laser. Photons scatter from the surface, interact with the hidden scene, and reflect back to the relay surface where this third bounce optical signal consisting of photon time of flight (ToF) is measured using time-resolved photon detectors (Fig 3) such as streak cameras [46] or Single-Photon Avalanche Diodes (SPADs) [4].

This 5D transient measurement, $H(\vec{x_p}, \vec{x_c}, t)$, encodes temporal light transport from the laser to all illumination positions $\vec{x_p}$ and light detected by the sensor focused at all detection positions $\vec{x_c}$ at the visible relay surface. Consequently, the transient measurement is equivalent to measuring the transient Light Transport Matrix (LTM), but for the visible relay surface. Complex light transport, such as additional bounces or sub-surface scattering within the hidden scene, may be encoded in this transient measurement.

While various algorithms have been developed to extract information about the hidden scene from these transient measurements [1, 2, 16, 23, 32, 35], the phasor field framework enables us to build a virtual active imaging system that illuminates and images the hidden scene from the perspective of the relay surface. Using this framework, we can computationally focus and image our virtual illumination at 3D locations, $\vec{x_p}^{(2)}$ and $\vec{x_c}^{(2)}$ respectively, in the hidden scene. This ability allows us to capture a second order transient light transport matrix, $H^{(2)}(\vec{x}_p^{(2)}, \vec{x}_c^{(2)}, t)$, ***for the hidden scene*** from the transient measurement.

However, most existing systems for ToF-based NLOS imaging employ gated single-pixel SPAD detectors [3, 4] and collect the NLOS signal using one of two acquisition schemes. Confocal acquisition focuses the laser and the detector at the same spot at each discrete time interval i.e. $H(\vec{x_p} = \vec{x_c}, t)$. Non-confocal acquisition places the detector at a fixed location, $\vec{x_0}$, while the laser scans the relay surface independently $H(\vec{x_p}, \vec{x_c} = \vec{x_0}, t)$. As shown in Fig 4, both of these setups capture a 3D subset of the full 5D measurement space which limits computing the full LTM of the hidden scene. With the advent of new gated SPAD arrays [36, 39, 48], it has become possible to probe this 5D transient measurement space opening exciting avenues of research. However, the limited number of SPAD pixels poses a challenge to fully leveraging the full 5D measurement space.

Recent work used this LTM framework for NLOS imaging to separate the direct and indirect components of light transport [25]. However, their experimental setup employed linear SPAD arrays which limited the full demonstration of LTMs in NLOS imaging applications. Furthermore, the authors employ slow algorithms that become intractable for extracting indirect components of light that bounce more than one time in the hidden scene. Other work with SPAD arrays focuses on reducing acquisition times [26, 34] and consolidates the data to a single pixel 3D dataset to enable real-time reconstruction with existing algorithms.

Contributions In this work, we propose algorithms to fully utilize the 5 dimensions of the transient measurement using a 16x16 gated SPAD array [39]. We propose fast FFT - based algorithms that are able to compute the full 5D transient light transport matrix of the hidden scene from this 5D transient measurement [31]. We demonstrate three applications of TLTMs in NLOS imaging, namely dual photography, scene relighting, separation of direct and indirect light transport, and dual photography. Furthermore, we make the captured datasets publicly available to help push the field of NLOS imaging forward.

Limitations It is important to note that our data capture is currently *not* limited by capture noise or low light levels. We attenuate our laser to lower the photon count rate

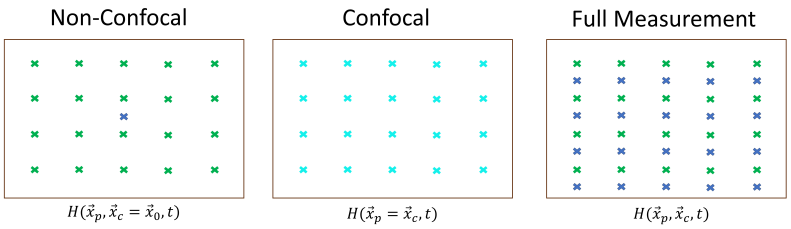

Figure 4. Existing acquisition schemes capture a subset of the full measurement

and avoid pileup in the sensor. Our datasets are captured with 16x16 pixels with 32 hot pixels that are turned off as they add excessive noise. This causes artifacts in our reconstruction. Furthermore, each SPAD pixel captures billions of photons each second. This calls for efficient implementations of specialized methods for data handling and processing. When scaling to larger resolutions, data processing and compression will become a major bottleneck. Additionally, we observe a phenomenon that looks similar to the formation of speckle (Fig S10). These synthetic speckle cause a grainy pattern in the reconstructions that eventually overwhelms the real signal. These grain patterns are stable over multiple measurements and therefore are not explained by capture noise.

2. Related Work

NLOS Imaging with ToF Detectors While many different approaches have emerged to tackle the problem of NLOS imaging [9, 10, 22], NLOS imaging using photon time of flight has demonstrated 3D reconstructions of a wide variety of complex room-sized scenes [16, 20] and relay surfaces [16, 21, 24] with high capture speeds [26, 34], and even reconstruct objects hidden around two corners [40].

Phasor Field Framework The development [5–8, 38, 43] and validation [37, 38, 42] of the phasor field framework has enabled the modeling of an NLOS imaging system as a virtual LOS imaging system at the relay surface. In this framework, light transmitted from and received at the relay surface is described as a virtual wavefront. This allows us to use familiar intuition and methodology from other LOS imaging sytems such as cameras, microwave phased arrays, or ultrasound imaging systems to understand the problem. Diffraction integrals from classical wave optics correctly describe the propagation of the virtual wavefront, and can be leveraged to backpropagate it and generate 3D reconstructions of the hidden scene [18]. It also correctly models the propagation of light inside the scene and explains phenomena like the NLOS missing cones problem [17, 40] and the formation of virtual mirror images in NLOS reconstructions [40].

Hardware for ToF NLOS The very first demonstration of ToF NLOS imaging utilized an expensive femtosecond pulsed laser and a Streak camera [46]. Cheaper systems that utilized modulated laser diodes for illumination and photonic mixer devices (PMDs) for detection [12, 13] suffered with poor image quality arising from low time resolution. To solve the problem of expensive hardware without sacrificing time resolution, a picosecond pulsed laser was paired with a single-pixel SPAD sensor to demonstrate NLOS with SPAD sensors for the first time [4].

The main drawback of such systems is that the 3rd bounce signal is very weak and the relay surface must be scanned for long exposure times or alternately with the laser power increased to unsafe levels. To circumvent this, a variety of different approaches have been proposed. One approach increases the signal at a specific location in the hidden scene by scanning or detecting along an ellipsoidal pattern on the relay surface [33] but this method requires sequential scanning of the laser or the single pixel detector with precise control.

Alternately, multi-pixel SPAD arrays are emerging as a viable technology that can reduce exposure time requirements [39]. A 32 x 32 pixel SPAD array was paired with a laser that was focused at a single illumination source [34] to reduce acquisition times (Fig 5), but this method was limited by a small field of view governed by the lens used to focus the array. Another approach paired two 16 x 1 pixel SPAD arrays [36] with a sparse laser grid to demonstrate live NLOS video reconstructions at 5 Hz [26]. To simultaneously detect the first bounce from the relay surface for calibration and the 3rd bounce from the hidden scene while mitigating pile-up, a recent approach applies a variable gate to individual SPAD pixels of a 6 x 6 SPAD array [48]. All these systems scan a lower dimensional subset of the 5D transient measurement (4, and have not been able to fully exploit the rich information available at the optical signal measured at the relay surface.

Light Transport Matrices An active imaging system consists of a controlled light source paired with a detector. LTMs have been used extensively in computer graphics to characterize light transport from the source, usually a pro-

jector, to the detector for a given scene. Once the light transport matrix (LTM) is measured, the image can be manipulated in post-processing. For example, dual photography leverages the transpose of the LTM to swap the illumination and detection, and image the scene from the perspective of the light source [41]. Furthermore, the LTM can be used to separate the direct and indirect components of light transport [27, 30]. With the advent of fast Time-of-flight detectors, the transient LTM can be measured and manipulated using the temporal frequency domain [31]. The transient LTM can easily separate the direct and indirect components, while the fourier domain of the LTM can be used to separate caustics and non-caustic effects.

LTMs in NLOS Imaging In TOF NLOS, a relay surface is illuminated sequentially with a pulsed laser and the the temporal profile is collected with a transient detector. This basis scan is the slow way to probe the LTM [29, 41, 47], but is feasible for NLOS imaging since the number of illumination points is an order of magnitude smaller than what is commonly used in megapixel projectors used in line-of-sight imaging systems.

Julio et al [25] also exploit the phasor field framework to build Virtual Light Transport Matrices that can be used to separate the direct and indirect components of light transport in the hidden scene. However, their method is not able to extract full *Transient* Light Transport Matrix, requiring slow algorithms that become computationally intractable for distinguishing additional bounces within the indirect component of the light transport that occurs within the hidden scene. Furthermore, their experimental hardware scans a sparse laser grid and detects light with linear SPAD arrays which limits them from exploiting the full 5 dimensions of the transient measurement. Our method can generate the full *Transient* Light Transport Matrix with much lower complexity (see Section S3) using Fast Fourier Transform (FFT) based algorithms. Additionally, we utilize 2D SPAD arrays and demonstrate experimental results using the full 5D transient measurement.

3. Methods

3.1. Light Transport Matrix

The Light Transport Matrix (LTM) can be defined as

$$c(\vec{x}_c) = H(\vec{x}_p, \vec{x}_c)p(\vec{x}_p) \tag{1}$$

where H is the light transport matrix that describes how photons emitted by a controllable light source, p, at location $\vec{x}_p$ interact with a given scene and are detected by some sensor, c, at location $\vec{x}_c$ [28]. With fast ToF detectors, we can project and detect spatio-temporal patterns using the transient LTM in either the time domain:

$$c(\vec{x}_c, t) = H(\vec{x}_p, \vec{x}_c, t)p(\vec{x}_p, t) \tag{2}$$

or the frequency domain: [31]:

$$c(\vec{x}_c, \omega) = H(\vec{x}_p, \vec{x}_c, \omega)p(\vec{x}_p, \omega) \tag{3}$$

In dual photography [41], we synthesize photo-realistic images from the perspective of the light source using Helmholtz reciprocity [45]. Mathematically, this requires using the transpose of the LTM, H^T:

$$p' = H^T c' \tag{4}$$

So the dual image is the weighted combination of all light contributions from detector pixels at the source pixels:

$$\begin{aligned} p'(\vec{x}_p, t) &= H^T(\vec{x}_p, \vec{x}_c, t)c'(\vec{x}_c, t) \\ &= H(\vec{x}_c, \vec{x}_p, t)c'(\vec{x}_c, t) \end{aligned} \tag{5}$$

3.2. Phasor Field Imaging

To collect our transient measurement H, a picosecond laser sends laser pulses sequentially to laser positions $\vec{x}_p$ on the relay wall, and each SPAD pixel, $\vec{x}_c$, captures the number of photons that return from the hidden scene as a function of time t. Since these pulses can be treated as delta functions in time, each temporal measurement can be understood as the transient impulse response of the hidden scene $H(\vec{x}_p, \vec{x}_c, t)$.

Due to linearity and time-invariance of light transport within this measurement, we can generate the response $\mathcal{P}(\vec{x}_p, \vec{x}_c, t)$ to some virtual illumination function, $\mathcal{P}(t)$, by performing the following convolution:

$$\mathcal{P}(\vec{x}_p, \vec{x}_c, t) = \int \left[\mathcal{P}(t) * H(\vec{x}_p, \vec{x}_c, t)\right] \mathrm{d}t. \tag{6}$$

where each SPAD pixel $\vec{x}_p$ is treated as a virtual illumination source, and each laser position $\vec{x}_c$ is a point on the virtual aperture that captures the virtual wavefront. Helmholtz reciprocity can be used to show that the light transport, and therefore the measured impulse response, is identical if the illumination and detection positions are swapped [45]. While the virtual illumination function, $\mathcal{P}(t)$, can be arbitrary, we set it to the following virtual wave packet to generate the phasor field response:

$$\mathcal{P}(t) = e^{i\omega_0 t} e^{-\frac{t^2}{2\sigma^2}} \tag{7}$$

This packet is a modulated Gaussian in the time domain with some central wavelength of modulation given by $\omega_0 = \frac{2\pi c}{\lambda_C}$ and the FWHM of the gaussian given by $\sigma = 5\lambda_C/c$. In the frequency domain, we can show it to be a gaussian bandpass filter centered around ω_0:

$$\begin{aligned} \mathcal{P}_{\mathcal{F}}(\omega) &= \int \mathcal{P}(t)\, e^{-i\omega t} dt \\ &= e^{-\frac{(\omega-\omega_0)^2(\sigma_P)^2}{2}} \quad = \alpha(\omega) \end{aligned} \tag{8}$$

Then, we can show (details in supplement S1) that the hidden scene can be reconstructed at voxel x_v using the following operator:

$$\begin{aligned} I_{\mathcal{P}}(\vec{x}_v) = \iiint & \left[\mathcal{P}(\omega) H(\vec{x}_p, \vec{x}_c, \omega)\right] \\ & e^{-j\frac{\omega}{c}(|\vec{x}_p - \vec{x}_v| + |\vec{x}_c - \vec{x}_v|)} d\vec{x}_p d\vec{x}_c d\omega. \end{aligned} \tag{9}$$

where fast Fourier transforms can be used to accelerate the reconstruction [14, 19]. A spatiotemporal 4D video of the illumination moving through the hidden scene can also be generated (details in the supplement) from the transient impulse response in the time domain:

$$I_{\mathcal{F}}\left(\vec{x}_v, t' = t_0\right) = \iiint \left[\mathcal{P}(\omega) H\left(\vec{x}_p, \vec{x}_c, \omega\right)\right] e^{-j\omega t_0} e^{-j\frac{\omega}{c}\left(|\vec{x}_c - \vec{x}_v|\right)} d\omega d\vec{x}_c d\vec{x}_p. \tag{10}$$

where we add a voxel-dependent time shift, $t' = t + |\mathbf{x}_\mathrm{p} - \mathbf{x}_\mathrm{v}|/c$, to generate a global time index so that our illumination starts emitting light at $t' = 0$.

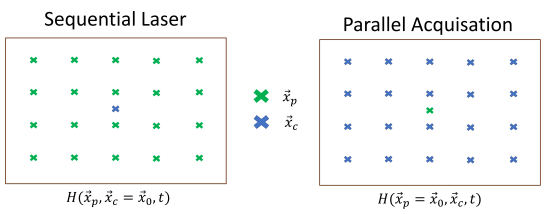

Figure 5. Equivalent Non-Confocal Acquisation schemes.

3.3. Connection to LTMs

It is clear from Eq 9 and Eq 10 that the reconstruction operators are linear operators that map the light transport encoded in the transient measurement, $H\left(\vec{x}_p, \vec{x}_c, t\right)$, to light transport to voxels x_v in the hidden scene. In Section 3.5.2 we demonstrate that if x_c and x_p are sampled adequately on the relay surface, we can project our illumination and focus our detection independently to voxels $\vec{x}_p^{(2)}$ and $\vec{x}_p^{(2)}$ respectively in the hidden scene. Therefore, we can build a second-order transient light transport matrix for the hidden scene:

$$H^{(2)}(\vec{x}_p^{(2)}, \vec{x}_c^{(2)}, t) = f_1(H\left(\vec{x}_p, \vec{x}_c, t\right)) \tag{11}$$

where f_1 denotes a series of linear operators, outlined in Eq 14. This provides a formalism for reconstructing scenes around multiple corners by iteratively reconstructing and using higher order matrix $H^{(n)}(\vec{x}_p^{(n)}, \vec{x}_c^{(n)}, t)$.

3.4. Dual Photography and SNR

The time of flight does not change whether a photon travels from location a to b or from b to a. Consequently, we can interchange illumination and detection in our reconstruction algorithm. In the phasor field framework, each SPAD pixel x_p is treated as a virtual illumination source, and each laser position x_c is a point on the virtual aperture that captures the virtual wavefront due to the higher density of the laser grid. Therefore, existing 3D reconstructions for non-confocal NLOS imaging systems with single-pixel ToF detectors are usually *dual images* [41] that image the hidden scene from the perspective of the laser grid/illumination positions.

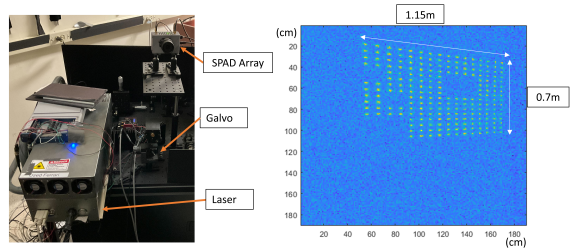

Figure 6. Our NLOS imaging system consisting of laser, galvanometer, and SPAD detectors. The right figure shows the SPAD pixels focused on the relay surface for large field of view

While fast reconstruction algorithms can be implemented [19] on chip [14, 15], the imaging pipeline is restricted by sequential scanning of the laser which often requires scanning a galvanometer mechanically, as shown on the left in Fig 5. The right side of Fig 5 shows an equivalent non-confocal acquisition scheme for the *primal* image, where we fix a single laser position and image from the perspective of detector pixels. While we don't see novel views since both grids image the hidden scene from the relay wall, this second scheme allows for parallel acquisition by leveraging multiple SPAD pixels to collect photons simultaneously.

In Supplement Section S4, we conduct an SNR analysis to demonstrate qualitatively and quantitatively that increasing the number of pixels with a fixed exposure time is equivalent to increasing the exposure time for a single pixel because both methods lead to capturing an increased number of photons. Consequently, we can reduce the acquisition time by increasing the number of pixels.

3.5. Scene Relighting

Once the LTM has been measured, different components of light transport can be manipulated in post-processing. Similarly, we can leverage the phasor field to manipulate our illumination and relight the scene in post-processing.

3.5.1 Novel Illumination

Under the Phasor field framework, each SPAD pixel illuminates the hidden scene from a different location and therefore our 3D reconstruction can be imaged under novel illumination from different locations of the relay wall:

$$I\left(\vec{x}_v, \vec{x}_p\right) = \iint \mathcal{P}(\omega) H\left(\vec{x}_p, \vec{x}_c, \omega\right) e^{-j\frac{\omega}{c}|\vec{x}_v - \vec{x}_c| + |\vec{x}_p - \vec{x}_v|} d\vec{x}_c d\omega \tag{12}$$

where we save reconstruction for each SPAD separately.

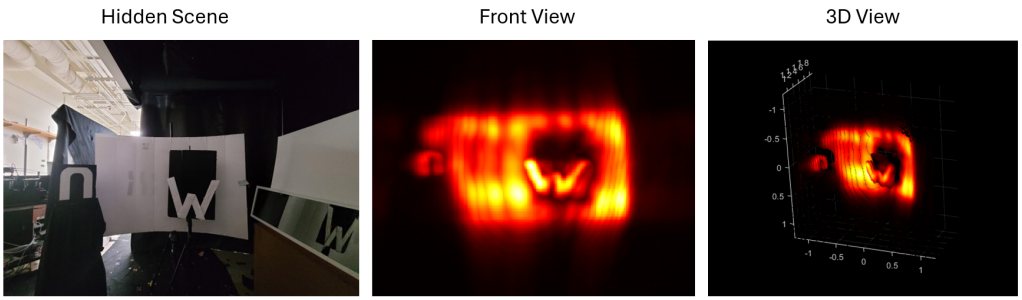

Figure 7. Hidden scene shown on the left. The steady-state reconstruction using Eq 9 is shown in the middle (front view) and on the left (3D view).

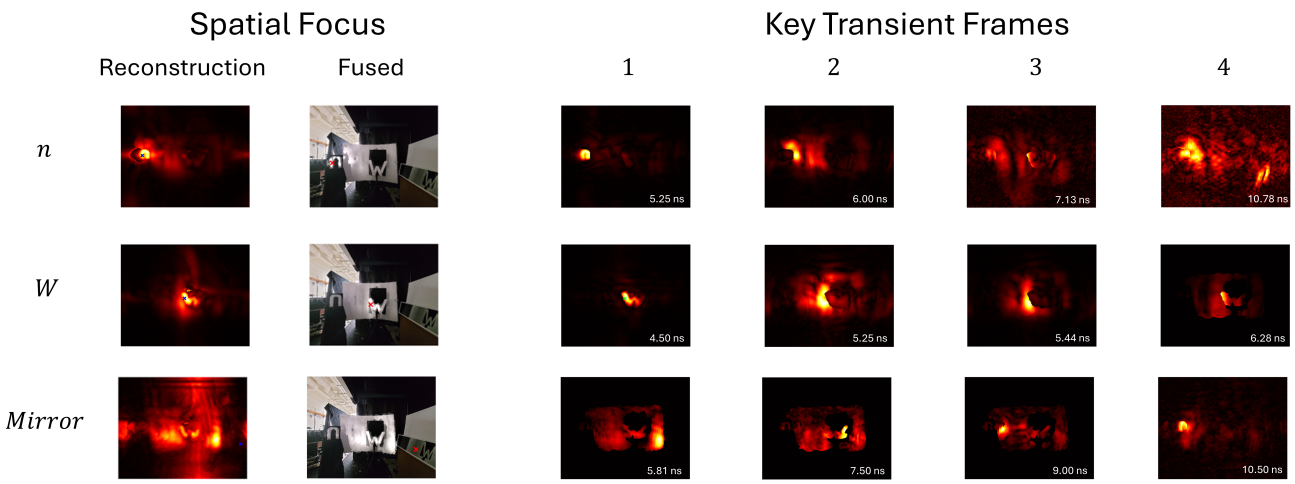

Figure 8. We demonstrate the TLTM on the hidden scene collected in Fig 7. Row 1, 2, and 3 show the impact of focusing the wavefront at different 3D locations corresponding to letter n, letter W, and the mirror in the hidden scene. Columns 1 and 2 are generated by integrating out the temporal component of the TLTM, showcasing the focal point using a blue and red cross respectively. Column 2 has the reconstruction overlaid on the ground truth image. Columns 3, 4, 5, 6 correspond to key frames within the video

3.5.2 Beamforming

Since the phasor field framework is defined by diffraction integrals from wave theory, we can leverage beamforming principles from electromagnetics and ultrasound imaging to focus and steer our virtual illumination. In supplement section S2, we show that these functions are simply spatio-temporal filters in the frequency domain and therefore can be used within a signal processing framework [44]:

$$I\left(\vec{x}_v, \vec{x}_p\right) = \iiint \mathcal{P}(\omega) H\left(\vec{x}_p, \vec{x}_c, \omega\right) e^{-j\omega\Delta t(x_p)} e^{-j\frac{\omega}{c}|\vec{x}_v - \vec{x}_c| + |\vec{x}_p - \vec{x}_v|} d\vec{x}_c d\omega d\vec{x}_p \tag{13}$$

A time-of-flight light propagation video can also be generated with these illumination coefficients:

$$I_{\mathcal{P}}\left(\vec{x}_{v_p}, \vec{x}_{v_c}, t\right) = \iiint \mathcal{P}(\omega) H\left(\vec{x}_p, \vec{x}_c, \omega\right) e^{-j\omega\Delta t(x_p)} e^{-j\frac{\omega}{c}|\vec{x}_v - \vec{x}_c|} d\vec{x}_c d\omega d\vec{x}_p \tag{14}$$

where $I_{\mathcal{P}}\left(\vec{x}_{v_p}, \vec{x}_{v_c}, t\right)$ is a 7D Transient Light Transport Matrix for the hidden scene. Therefore, phasor field beamforming enables us to extract a video while focusing on a 3D location denoted by $\vec{x}_{v_p}$ and imaging a 3D location denoted by $\vec{x}_{v_c}$

Beamforming Resolution and Requirements Previous work [5, 7, 37, 42] has shown the phasor field imaging resolution to be given by $\Delta x_c = 1.22\lambda_p z / D_c$ where $\lambda_p = c/2\pi\omega$ is the phasor field wavelength, z is the depth offset

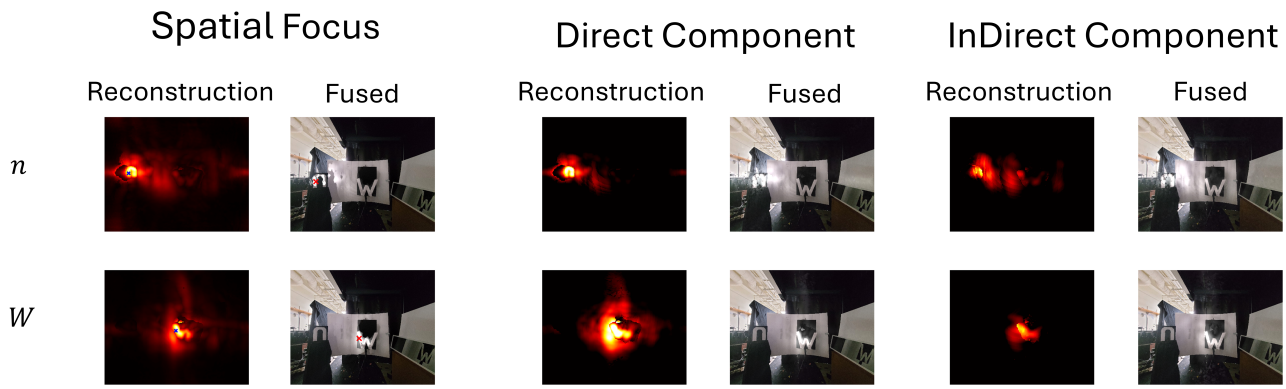

Figure 9. The TLTM can be used to separate the direct components (Columns 3,4) from the indirect component (Columns 5,6) of focusing the wavefront at location corresponding to n (Row 1) or W (Row 2). Columns 1 and 2 are generated by adding the direct and indirect components

from the relay wall, and D_c is the size of the laser grid, i.e. the virtual aperture. Similarly, the phasor field focusing resolution for our work is given by:

$$\Delta x_p = \frac{1.22\lambda_p z}{D_p} \tag{15}$$

where D_p is the diameter of the SPAD grid. Beam forming requires sampling the surface at the Nyquist rate which is difficult due to the limited amount of pixels in the array. Therefore, we can either use larger wavelengths or small apertures or wait for emerging SPAD arrays with more pixels. We note that increasing the number of SPAD pixels by an order of magnitude in each dimension does not significantly increase the computation time of generating the 7D TLTM since we can integrate the raw measurement with respect to SPAD pixels before computing the TLTM.

3.6. Seperation of Direct and Indirect Components

With the TLTM, the direct and indirect components can be separated efficiently with Time Gating on the temporal dimension. For each voxel, x_v, we can compute the time of flight, $t_i = |x_v - \bar{x}_p|/c + K$, for the illumination to reach x_v from the center of the SPAD array $\bar{x}_p$, i.e, the center of the illumination pulse. We add a constant K to account for the temporal width of the illumination pulse as it propagates in time. Then, the direct component is the sum of video frames that occur before this time t_i and the indirect component is the sum of video frames that occur after t_i i.e.

$$I_{\mathcal{D}}\left(\vec{x}_{v_p}, \vec{x}_{v_c}, t\right) = \int M_1(t) I_{\mathcal{P}}\left(\vec{x}_{v_p}, \vec{x}_{v_c}, t\right) dt \tag{16}$$

$$I_{\mathcal{I}}\left(\vec{x}_{v_p}, \vec{x}_{v_c}, t\right) = \int M_2(t) I_{\mathcal{P}}\left(\vec{x}_{v_p}, \vec{x}_{v_c}, t\right) dt \tag{17}$$

where $M_2(t)$ is the complement of $M_1(t)$, and $M_1(t)$ is given by:

$$M_1(t) = \begin{cases} 1, & t \leq t_i \\ 0, & t > t_i \end{cases} \tag{18}$$

We note that $M_1(t)$ and $M_2(t)$ needs to be computed only once per hidden scene with $O(N^3)$ complexity, where N is the side length of reconstruction volume. In Section S3, we discuss how our algorithm has significantly lower complexity than previous work [25].

4. Hardware

As shown in Fig. 3, our NLOS system is an active imaging system that consists of the illumination source and a ToF detector. Our illumination source is a PM-1.03-25^{TM} laser from Polar Laser Laboratories coupled with a frequency doubler that generates 515 nm pulses with a pulse width smaller than 35 ps. The laser operates in burst mode where the burst rep rate is 357.14 kHz, and each burst emits 14 pulses per burst at 10MHz. This yields an average rep rate of 5Mhz and an average output power of 375 mW. We use ND filters to adjust the output power for different lens configurations to avoid pile-up, so the average power for the laser varies between 100 - 375 mW.

A two-mirror Thorlabs galvanometer (Thorlabs GVS012) is used to scan our relay surface. While the total scan time varies by experiment, the scan area is fixed to a 1.9m x 1.9m planar relay surface. We use a customized gated 16x16 SPAD array [39] for our measurements. The deadtime of the SPAD array is smaller than 100ns, and the temporal resolution is characterized by a Full-Width at Half Maximum (FWHM) of approximately 60 ps. The detector is roughly 2.7m away from the relay surface, and we use different lens configurations so we can span different fields of view on the relay surface.

For SNR experiments shown in the Supplement Section

S4, we use a Canon EF 85 mm f/1.8 USM Lens to focus to a small 7.1 cm by 4.7 cm area of the relay surface. To perform beamforming and simultaneously image our hidden scene, we need the spacing between the SPAD pixels to be roughly the same as the spacing between laser positions. For this, we use an Edmund Optics 6X Manual Zoom Video lens with the focal length set to 30mm, which focuses the SPAD array to a 20 cm x 12.5 cm area on the relay wall. To image the hidden scene under novel illumination and demonstrate dual photography, we use a Fujinon 3 MP Varifocal Lens (3.8-13mm, 3.4x Zoom) which focuses on a 1.1m x 0.7m area on the relay wall (Figure 6). This figure also shows the "holes" arising from the non-functional SPAD pixels, as well as the distortion in spacing arising from projecting the pixels from one side.

5. Results

In this section, we demonstrate novel applications that are made possible from probing the full 5D measurement space from with our SPAD array.

5.1. Beamforming

In Figure 8, we show the effect of focusing our illumination. We can utilize Eq 14 on the transient measurement to compute $H^{(2)}(\vec{x}_p^{(2)}, \vec{x}_c^{(2)}, t)$, which is a 7D spatiotemporal video of our virtual illumination propagating through a complex scene, conisting of a mirror, letters n and W, and a back wall. Figure 7 shows the reconstruction - note that the mirror is in the missing cone, and does not show up in the reconstruction.

However, the transient frames demonstrate that the light transport is captured, including higher-order bounces from the mirror. Row 1 shows the effect of focusing our illumination to the letter n, lighting up the n (column 3), then hitting the backwall (column 4). The letter n reflects light and illuminates the left side of the letter W (column 5) before illuminating the mirror (column 6). We see an image of n form on the mirror (bottom right of row 1, column 6). We can also image the reciprocal path of focusing on the mirror (Row 3). While we don't see the mirror light up since it is in the missing cone when illuminated from the relay wall, we see the indirect 4th bounce of the illumination from the mirror to the right side of the back wall (column 3), the right side of letter W (column 4), the left side of the back wall (column 5), and finally the letter n (column 6). Row 2 demonstrates the effect of focusing on letter W so W lights up (column 3), then the back wall (columns 4) and finally the reflection from the back wall lights up the left side of the letter W (column 6). Column 5 shows that the letter n does not light up when focusing on W. These effects are easier to visualize in a video format, and in supplementary videos, we showcase the effect of moving our spatial focus as well as transient videos for focusing on at key spatial locations.

In supplement section S5 and supplementary videos, we demonstrate how focusing on the virtual illumination helps reveal shadows in higher-order bounces as well as difference in scattering properties of water and milk, demonstrating that the full light transport can be extracted from the transient measurement.

5.2. Separating Direct and Indirect Light Transport

We utilize the time dimension to gate out the direct illumination (*3rd bounce*) from indirect (*higher order bounces*). We focus on n (Row 1) and W (Row 2), and see that the n and W light up in the direct component (columns 3 and 4). The indirect components (columns 5 and 6) display additional bounces. For example, light scatters from the n and hits left side of W, and various places for the back wall (Row 1, Columns 5 and 6). When focusing on W, the light propagates to W, then lights up the back wall, before finally lighting up the edge of W again (Row 2, Columns 5 and 6). Since the mirror is in the missing cone, there is no direct component and so it does not make sense to run the algorithm for this scenario.

5.3. Novel Illumination

The diversity of information collected is greatly enhanced as different parts of the hidden scene reflect light to different parts of the relay wall. This is equivalent to illuminating the hidden scene from different angles from the relay surface, and then generating a 3D image for each illumination position. We can either add all the images together as shown in Fig S12 or we can stack individual 3D images to build a 3D light field as shown in Fig S14 where each illumination location is shown on the left image.

6. Discussion

Due to limited lab space, we can not place the SPAD array more than 2.67m away from the relay surface. This creates two issues that prevent us from fully exploiting the array. First, the field of view is less than a quarter than that of the laser grid. Second, the distortion of projecting the SPAD pixels from one side of the relay surface is enhanced (Fig 6) since the difference in distances to the nearest SPAD pixel on the relay wall relative to the furthest SPAD pixel is large. Both these issues can be mitigated by moving the system further from the relay wall. While 16x16 SPAD array is a big step towards probing the full 5D transient measurement, in practice we are limited by the limited number of SPAD pixels. This issue is further exacerbated by the non-functional hot pixels, limiting how well we can probe the hidden scene, and consequently, the resolution with which we can extract the second order light transport matrix, $H^{(2)}(\vec{x}_p^{(2)}, \vec{x}_c^{(2)}, t)$.

7. Conclusion

In this work, we have shown that the 5D transient measurement, $H(\vec{x}_p, \vec{x}_c, t)$ contains information about the Light Transport Matrix of the hidden scene $H^{(2)}(\vec{x}_p^{(2)}, \vec{x}_c^{(2)}, t)$ which can be computed using the phasor field framework. Once we have the TLTM, we can generate the 7D spatiotemporal light transport within the hidden scene. We can use this formalism to build higher order TLTM, $H^{(n)}(\vec{x}_p^{(n)}, \vec{x}_c^{(n)}, t)$, and show reconstructions around multiple corners, which will be possible as the number of pixels increases in future SPAD arrays. On the other hand, the increased amount of SPAD pixels comes at the cost of increased data rates. Future work will explore how we can compress or reduce our measurement without losing our ability to extract the transient light transport matrix.

References

[1] Byeongjoo Ahn, Akshat Dave, Ashok Veeraraghavan, Ioannis Gkioulekas, and Aswin Sankaranarayanan. Convolutional Approximations to the General Non-Line-of-Sight Imaging Operator. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 7888–7898, Seoul, Korea (South), 2019. IEEE. 2

[2] Victor Arellano, Diego Gutierrez, and Adrian Jarabo. Fast back-projection for non-line of sight reconstruction. *Optics Express*, 25(10):11574–11583, 2017. 2

[3] Mauro Buttafava, Gianluca Boso, Alessandro Ruggeri, Alberto Dalla Mora, and Alberto Tosi. Time-gated single-photon detection module with 110 ps transition time and up to 80 MHz repetition rate. *Review of Scientific Instruments*, 85(8):083114, 2014. Publisher: American Institute of Physics. 2

[4] Mauro Buttafava, Jessica Zeman, Alberto Tosi, Kevin Eliceiri, and Andreas Velten. Non-line-of-sight imaging using a time-gated single photon avalanche diode. *Optics Express*, 23(16):20997–21011, 2015. 2, 3

[5] Justin Dove and Jeffrey H. Shapiro. Paraxial theory of phasor-field imaging. *Optics Express*, 27(13):18016–18037, 2019. 3, 6

[6] Justin Dove and Jeffrey H. Shapiro. Nonparaxial phasor-field propagation. *Optics Express*, 28(20):29212–29229, 2020. Publisher: Optical Society of America.

[7] Justin Dove and Jeffrey H. Shapiro. Paraxial phasor-field physical optics. *Optics Express*, 28(14):21095–21109, 2020. Publisher: Optical Society of America. 6

[8] Justin Dove and Jeffrey H. Shapiro. Speckled speckled speckle. *Optics Express*, 28(15):22105–22120, 2020. Publisher: Optical Society of America. 3

[9] Daniele Faccio, Andreas Velten, and Gordon Wetzstein. Non-line-of-sight imaging. *Nature Reviews Physics*, 2(6): 318–327, 2020. Number: 6 Publisher: Nature Publishing Group. 2, 3

[10] Ruixu Geng, Yang Hu, and Yan Chen. Recent Advances on Non-Line-of-Sight Imaging: Conventional Physical Models, Deep Learning, and New Scenes. *APSIPA Transactions on Signal and Information Processing*, 11(1), 2022. 3

[11] Chaoying Gu, Talha Sultan, Khadijeh Masumnia-Bisheh, Laura Waller, and Andreas Velten. Fast Non-line-of-sight Imaging with Non-planar Relay Surfaces. In *2023 IEEE International Conference on Computational Photography (ICCP)*, pages 1–12, 2023. ISSN: 2472-7636. 2

[12] Felix Heide, Matthias B. Hullin, James Gregson, and Wolfgang Heidrich. Low-budget transient imaging using photonic mixer devices. *ACM Transactions on Graphics*, 32(4):1–10, 2013. 3

[13] Felix Heide, Lei Xiao, Wolfgang Heidrich, and Matthias B. Hullin. Diffuse Mirrors: 3D Reconstruction from Diffuse Indirect Illumination Using Inexpensive Time-of-Flight Sensors. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pages 3222–3229, Columbus, OH, USA, 2014. IEEE. 3

[14] Deyang Jiang, Xiaochun Liu, Jianwen Luo, Zhengpeng Liao, Andreas Velten, and Xin Lou. Ring and Radius Sampling Based Phasor Field Diffraction Algorithm for Non-Line-of-Sight Reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11):7841–7853, 2022. Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence. 5

[15] Zhengpeng Liao, Deyang Jiang, Xiaochun Liu, Andreas Velten, Yajun Ha, and Xin Lou. FPGA Accelerator for Real-Time Non-Line-of-Sight Imaging. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 69(2):721–734, 2022. Conference Name: IEEE Transactions on Circuits and Systems I: Regular Papers. 5

[16] David B. Lindell, Gordon Wetzstein, and Matthew O'Toole. Wave-based non-line-of-sight imaging using fast *f-k* migration. *ACM Transactions on Graphics*, 38(4):1–13, 2019. 2, 3

[17] Xiaochun Liu, Sebastian Bauer, and Andreas Velten. Analysis of Feature Visibility in Non-Line-Of-Sight Measurements. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10132–10140, Long Beach, CA, USA, 2019. IEEE. 3, 7

[18] Xiaochun Liu, Ibón Guillén, Marco La Manna, Ji Hyun Nam, Syed Azer Reza, Toan Huu Le, Adrian Jarabo, Diego Gutierrez, and Andreas Velten. Non-line-of-sight imaging using phasor-field virtual wave optics. *Nature*, 572(7771): 620–623, 2019. Number: 7771 Publisher: Nature Publishing Group. 3, 1

[19] Xiaochun Liu, Sebastian Bauer, and Andreas Velten. Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems. *Nature Communications*, 11(1): 1–13, 2020. Number: 1 Publisher: Nature Publishing Group. 5, 1, 2, 4

[20] Xintong Liu, Jianyu Wang, Zhupeng Li, Zuoqiang Shi, Xing Fu, and Lingyun Qiu. Non-line-of-sight reconstruction with signal–object collaborative regularization. *Light: Science & Applications*, 10(1):198, 2021. 3

[21] Xintong Liu, Jianyu Wang, Leping Xiao, Zuoqiang Shi, Xing Fu, and Lingyun Qiu. Non-line-of-sight imaging with arbitrary illumination and detection pattern. *Nature Communications*, 14(1):3230, 2023. 3

[22] Tomohiro Maeda, Guy Satat, Tristan Swedish, Lagnojita Sinha, and Ramesh Raskar. Recent Advances in Imaging Around Corners, 2019. arXiv:1910.05613 [cs, eess]. 3

[23] M. La Manna, F. Kine, E. Breitbach, J. Jackson, T. Sultan, and A. Velten. Error Backprojection Algorithms for Non-Line-of-Sight Imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–1, 2018. 2

[24] Marco La Manna, Ji-Hyun Nam, Syed Azer Reza, Andreas Velten, and Andreas Velten. Non-line-of-sight-imaging using dynamic relay surfaces. *Optics Express*, 28(4):5331–5339, 2020. Publisher: Optica Publishing Group. 3

[25] Julio Marco, Adrian Jarabo, Ji Hyun Nam, Xiaochun Liu, Miguel Ángel Cosculluela, Andreas Velten, and Diego Gutierrez. Virtual Light Transport Matrices for Non-Line-of-Sight Imaging. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 2420–2429, 2021. ISSN: 2380-7504. 2, 4, 7

[26] Ji Hyun Nam, Eric Brandt, Sebastian Bauer, Xiaochun Liu, Marco Renna, Alberto Tosi, Eftychios Sifakis, and Andreas

Velten. Low-latency time-of-flight non-line-of-sight imaging at 5 frames per second. *Nature Communications*, 12(1):6526, 2021. 2, 3, 4

[27] Shree K. Nayar, Gurunandan Krishnan, Michael D. Grossberg, and Ramesh Raskar. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Transactions on Graphics*, 25(3):935–944, 2006. 1, 4

[28] Matthew O'Toole. *Optical Linear Algebra for Computational Light Transport*. PhD thesis, 2016. Publication Title: Ph.D. Thesis ADS Bibcode: 2016PhDT.....18O. 1, 4

[29] Matthew O'Toole and Kiriakos N. Kutulakos. Optical computing for fast light transport analysis. *ACM Transactions on Graphics*, 29(6):164:1–164:12, 2010. 4

[30] Matthew O'Toole, Ramesh Raskar, and Kiriakos N. Kutulakos. Primal-dual coding to probe light transport. *ACM Transactions on Graphics*, 31(4):39:1–39:11, 2012. 1, 4

[31] Matthew O'Toole, Felix Heide, Lei Xiao, Matthias B. Hullin, Wolfgang Heidrich, and Kiriakos N. Kutulakos. Temporal Frequency Probing for 5D Transient Analysis of Global Light Transport. *ACM Trans. Graph.*, 33(4):87:1–87:11, 2014. 1, 2, 4

[32] Matthew O'Toole, David B. Lindell, and Gordon Wetzstein. Confocal non-line-of-sight imaging based on the light-cone transform. *Nature*, 555(7696):338–341, 2018. 2

[33] Adithya Pediredla, Akshat Dave, and Ashok Veeraraghavan. SNLOS: Non-line-of-sight Scanning through Temporal Focusing. In *2019 IEEE International Conference on Computational Photography (ICCP)*, pages 1–13, 2019. ISSN: 2472-7636. 3

[34] Chengquan Pei, Anke Zhang, Yue Deng, Feihu Xu, Jiamin Wu, David U.-Lei Li, Hui Qiao, Lu Fang, and Qionghai Dai. Dynamic non-line-of-sight imaging system based on the optimization of point spread functions. *Optics Express*, 29(20): 32349–32364, 2021. Publisher: Optica Publishing Group. 2, 3

[35] Joshua Rapp, Charles Saunders, Julián Tachella, John Murray-Bruce, Yoann Altmann, Jean-Yves Tourneret, Stephen McLaughlin, Robin M. A. Dawson, Franco N. C. Wong, and Vivek K. Goyal. Seeing around corners with edge-resolved transient imaging. *Nature Communications*, 11(1):5929, 2020. Number: 1 Publisher: Nature Publishing Group. 2

[36] Marco Renna, Ji Hyun Nam, Mauro Buttafava, Federica Villa, Andreas Velten, and Alberto Tosi. Fast-Gated 16 × 1 SPAD Array for Non-Line-of-Sight Imaging Applications. *Instruments*, 4(2):14, 2020. Number: 2 Publisher: Multidisciplinary Digital Publishing Institute. 2, 3

[37] Syed Azer Reza, Marco La Manna, Sebastian Bauer, and Andreas Velten. Phasor field waves: experimental demonstrations of wave-like properties. *Optics Express*, 27(22):32587, 2019. 3, 6

[38] Syed Azer Reza, Marco La Manna, Sebastian Bauer, Andreas Velten, and Andreas Velten. Phasor field waves: A Huygens-like light transport model for non-line-of-sight imaging applications. *Optics Express*, 27(20):29380–29400, 2019. Publisher: Optical Society of America. 3

[39] S. Riccardo, E, Conca, V. Sesta, A. Velten, and A. Tosi. Fast-Gated 16 × 16 SPAD Array with 16 on-chip 6 ps Time-to-Digital Converters for Non-Line-of-Sight Imaging. *IEEE Sensors Journal*, pages 1–1, 2022. Conference Name: IEEE Sensors Journal. 1, 2, 3, 7

[40] Diego Royo, Talha Sultan, Adolfo Muñoz, Khadijeh Masumnia-Bisheh, Eric Brandt, Diego Gutierrez, Andreas Velten, and Julio Marco. Virtual Mirrors: Non-Line-of-Sight Imaging Beyond the Third Bounce. *ACM Transactions on Graphics*, 42(4):140:1–140:15, 2023. 3, 7

[41] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik P. A. Lensch. Dual Photography. In *ACM SIGGRAPH 2005 Papers*, pages 745–755, New York, NY, USA, 2005. ACM. 4, 5

[42] Talha Sultan, Syed Azer Reza, and Andreas Velten. Towards a more accurate light transport model for non-line-of-sight imaging. *Optics Express*, 32(5):7731–7761, 2024. Publisher: Optica Publishing Group. 3, 6

[43] Jeremy A. Teichman. Phasor field waves: a mathematical treatment. *Optics Express*, 27(20):27500–27506, 2019. Publisher: Optical Society of America. 3

[44] B.D. Van Veen and K.M. Buckley. Beamforming: a versatile approach to spatial filtering. *IEEE ASSP Magazine*, 5(2): 4–24, 1988. Conference Name: IEEE ASSP Magazine. 6

[45] Eric Veach. Robust Monte Carlo Methods For Light Transport Simulation. *PhD thesis, Stanford University*, 1997. 4

[46] Andreas Velten, Thomas Willwacher, Otkrist Gupta, Ashok Veeraraghavan, Moungi G. Bawendi, and Ramesh Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Communications*, 3: 745, 2012. 2, 3, 1

[47] Jiaping Wang, Yue Dong, Xin Tong, Zhouchen Lin, and Baining Guo. Kernel Nyström method for light transport. In *ACM SIGGRAPH 2009 papers*, pages 1–10, New York, NY, USA, 2009. Association for Computing Machinery. 4

[48] Jiuxuan Zhao, Francesco Gramuglia, Pouyan Keshavarzian, Eng-Huat Toh, Michelle Tng, Louis Lim, Vinit Dhulla, Elgin Quek, Myung-Jae Lee, and Edoardo Charbon. A Gradient-Gated SPAD Array for Non-Line-of-Sight Imaging. *IEEE Journal of Selected Topics in Quantum Electronics*, pages 1–10, 2023. Conference Name: IEEE Journal of Selected Topics in Quantum Electronics. 2, 3

Transient Light Transport Matrices for Non-Line-of-Sight Imaging

Supplementary Material

S1. Phasor Field Imaging

Here, we show our illumination function for Phasor Field imaging is a bandpass filter in the Fourier domain. This filters our signal using a frequency-dependent complex number:

$$\begin{aligned}\mathcal{P}_{\mathcal{F}}(\omega) &= \int \mathcal{P}(t)\, e^{-i\omega t} dt \\ &= \int e^{-\frac{t^2}{2\sigma^2}} e^{-i(\omega-\omega_0)t} dt \\ &= \delta(\omega-\omega_0) \circledast e^{-\frac{(\omega\sigma_P)^2}{2}} \\ &= e^{-\frac{(\omega-\omega_0)^2(\sigma_P)^2}{2}} \qquad\qquad = \alpha(\omega)\end{aligned} \tag{19}$$

S1.1. Volumetric 3D Reconstruction

The back projection algorithm [46] reconstructs each voxel, $\mathbf{x}_\mathrm{v}$, at the time it is illuminated by the light source. Therefore, it accounts for the time shifts to each voxel in our hidden scene from each illumination $\mathbf{x}_\mathrm{p}$ and detector position $\mathbf{x}_\mathrm{c}$ at the relay surface:

$$I(\mathbf{x}_\mathrm{v}) = \iiint H(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{c}, t)\, \delta\left(t - \frac{|\mathbf{x}_\mathrm{p}-\mathbf{x}_\mathrm{v}| + |\mathbf{x}_\mathrm{c}-\mathbf{x}_\mathrm{v}|}{c}\right) d\mathbf{x}_\mathrm{p} d\mathbf{x}_\mathrm{c} dt. \tag{20}$$

where the measurement, $H(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{c}, t)$ is the temporal impulse response of the hidden scene. Therefore, we can compute the virtual response to an arbitrary illumination function, $P(t)$, by performing a temporal convolution:

$$\begin{aligned}\mathcal{P}(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{c}, t) &= [\mathcal{P}(t) * H(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{c}, t)] \\ &= \int_{-\infty}^{\infty} P(t-\tau) H(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{c}, \tau)\, d\tau\end{aligned} \tag{21}$$

The back projection algorithm can be applied on this virtual response to generate high quality reconstructions of the hidden scene [18]:

$$I_{\mathcal{P}}(\mathbf{x}_\mathrm{v}) = \iiint \mathcal{P}(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{c}, t)\, \delta\left(t - \frac{|\mathbf{x}_\mathrm{p}-\mathbf{x}_\mathrm{v}| + |\mathbf{x}_\mathrm{c}-\mathbf{x}_\mathrm{v}|}{c}\right) d\mathbf{x}_\mathrm{p} d\mathbf{x}_\mathrm{c} dt. \tag{22}$$

This operator can be implemented on the temporal fourier transform of the captured measurement to reduce computational complexity and accerelate reconstruction [19]:

$$\begin{aligned}I_{\mathcal{P}}(\mathbf{x}_\mathrm{v}) &= \iiint [\mathcal{P}(t) * H(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{c}, t)]\, \delta\left(t - \frac{|\mathbf{x}_\mathrm{p}-\mathbf{x}_\mathrm{v}| + |\mathbf{x}_\mathrm{c}-\mathbf{x}_\mathrm{v}|}{c}\right) d\mathbf{x}_\mathrm{p} d\mathbf{x}_\mathrm{c} dt. \\ &= \iiint [\mathcal{P}(\omega) H(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{c}, \omega)]\, e^{-j\frac{\omega}{c}(|\mathbf{x}_\mathrm{p}-\mathbf{x}_\mathrm{v}|+|\mathbf{x}_\mathrm{c}-\mathbf{x}_\mathrm{v}|)} d\mathbf{x}_\mathrm{p} d\mathbf{x}_\mathrm{c} d\omega. \\ &= \iiint [\mathcal{P}(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{c}, \omega)]\, e^{-j\frac{\omega}{c}(|\mathbf{x}_\mathrm{p}-\mathbf{x}_\mathrm{v}|+|\mathbf{x}_\mathrm{c}-\mathbf{x}_\mathrm{v}|)} d\mathbf{x}_\mathrm{p} d\mathbf{x}_\mathrm{c} d\omega.\end{aligned} \tag{23}$$

where the following relation follows from taking the temporal Fourier transform on both sides of Eq 21:

$$\mathcal{P}(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{c}, \omega) = [\mathcal{P}(\omega) H(\mathbf{x}_\mathrm{p}, \mathbf{x}_\mathrm{c}, \omega)] \tag{24}$$

Rearranging Eq 23 as:

$$I_{\mathcal{P}}\left(\mathbf{x}_{\mathrm{v}}\right)=\iint e^{-j\frac{\omega}{c}\left|\mathbf{x}_{\mathrm{p}}-\mathbf{x}_{\mathrm{v}}\right|}\underbrace{\left\{\int \mathcal{P}\left(\mathbf{x}_{\mathrm{p}},\mathbf{x}_{\mathrm{c}},\omega\right)e^{-j\frac{\omega}{c}\left|\mathbf{x}_{\mathrm{v}}-\mathbf{x}_{\mathrm{c}}\right|}d\mathbf{x}_{\mathrm{c}}\right\}}_{\mathcal{R}_{\mathbf{x}_{\mathrm{v}}}}d\omega d\mathbf{x}_{\mathrm{p}} \tag{25}$$

Provided all the detection points, x_c, lie on a regular grid on a planar surface, then the the inner integral, denoted by $\mathcal{R}_{\mathbf{x}_{\mathrm{v}}}$, can be written as a convolution:

$$\mathcal{R}_{\mathbf{x}_{\mathrm{v}}}\left(x_c,\omega\right)=\mathcal{P}_{\mathcal{F}}\left(x_c,y_c,0,\omega\right)\underset{\text{2D}}{*}G\left(x_c,y_c,z_v,\omega\right) \tag{26}$$

where $G\left(x_c,y_c,z_v,\omega\right)=e^{i\frac{\omega}{c}\sqrt{{x_c}^2+y_c^2+z_v^2}}/\sqrt{x_c^2+y_c^2+z_v^2}$. The reconstruction for all x_v at a fixed depth z_v and frequency ω can plane can be computed simultaneously by using a single 2D FFT which significantly reduces the computational complexity [11, 19].

S1.2. Spatio-Temporal 4D Reconstruction

We can compute the reconstruction as a function of time if we refrain from integrating over the time dimension in Eq 20 and Eq 22. We generate 4D reconstructions by calculating the reconstruction at different times on the raw measurement:

$$I\left(\mathbf{x}_{\mathrm{v}},t\right)\triangleq\iint H\left(\mathbf{x}_{\mathrm{p}},\mathbf{x}_{\mathrm{c}},t\right)\delta\left(t-\frac{\left|\mathbf{x}_{\mathrm{p}}-\mathbf{x}_{\mathrm{v}}\right|+\left|\mathbf{x}_{\mathrm{c}}-\mathbf{x}_{\mathrm{v}}\right|}{c}\right)d\mathbf{x}_{\mathrm{c}}d\mathbf{x}_{\mathrm{p}} \tag{27}$$

Or the filtered Phasor field measurement:

$$I_{\mathcal{P}}\left(\mathbf{x}_{\mathrm{v}},t\right)\triangleq\iint\left[\mathcal{P}(t)*H\left(\mathbf{x}_{\mathrm{p}},\mathbf{x}_{\mathrm{c}},t\right)\right]\delta\left(t-\frac{\left|\mathbf{x}_{\mathrm{p}}-\mathbf{x}_{\mathrm{v}}\right|+\left|\mathbf{x}_{\mathrm{c}}-\mathbf{x}_{\mathrm{v}}\right|}{c}\right)d\mathbf{x}_{\mathrm{c}}d\mathbf{x}_{\mathrm{p}} \tag{28}$$

This can be efficiently implemented in the fourier domain as:

$$\begin{aligned}I_{\mathcal{F}}\left(\mathbf{x}_{\mathrm{v}},\Omega\right)&=\iint\mathcal{P}(\omega)H\left(\mathbf{x}_{\mathrm{p}},\mathbf{x}_{\mathrm{c}},\Omega\right)*e^{j\frac{\Omega}{c}\left(\left|\mathbf{x}_{\mathrm{p}}-\mathbf{x}_{\mathrm{v}}\right|+\left|\mathbf{x}_{\mathrm{c}}-\mathbf{x}_{\mathrm{v}}\right|\right)}d\mathbf{x}_{\mathrm{c}}d\mathbf{x}_{\mathrm{p}}\\&=\iiint\mathcal{P}(\omega)H\left(\mathbf{x}_{\mathrm{p}},\mathbf{x}_{\mathrm{c}},\omega\right)e^{-j\frac{\Omega-\omega}{c}\left(\left|\mathbf{x}_{\mathrm{p}}-\mathbf{x}_{\mathrm{v}}\right|+\left|\mathbf{x}_{\mathrm{c}}-\mathbf{x}_{\mathrm{v}}\right|\right)}d\omega d\mathbf{x}_{\mathrm{c}}d\mathbf{x}_{\mathrm{p}}\\&=\iiint\mathcal{P}\left(\mathbf{x}_{\mathrm{p}},\mathbf{x}_{\mathrm{c}},\omega\right)e^{-j\frac{\Omega-\omega}{c}\left(\left|\mathbf{x}_{\mathrm{p}}-\mathbf{x}_{\mathrm{v}}\right|+\left|\mathbf{x}_{\mathrm{c}}-\mathbf{x}_{\mathrm{v}}\right|\right)}d\omega d\mathbf{x}_{\mathrm{c}}d\mathbf{x}_{\mathrm{p}}.\end{aligned} \tag{29}$$

The time index for this video is voxel-dependent. After reconstruction, each voxel in the reconstructions lights up at $t=0$ since we have shifted the time index to account for the photon time of flight from the illumination to the voxel and back to the sensor. So the 3D reconstruction can be rewritten as the DC component of the 4D video, which can be generated using the temporal Fourier transform ($\mathcal{F}$):

$$I\left(\mathbf{x}_{\mathrm{v}}\right)=\int\mathcal{F}^{-1}\left\{I_{\mathcal{F}}\left(\mathbf{x}_{\mathrm{v}},\Omega\right)\right\}dt=I_{\mathcal{F}}\left(\mathbf{x}_{\mathrm{v}},0\right). \tag{30}$$

To generate a transient video of the illumination moving through the hidden scene, we can add a voxel-dependent time shift to generate a global time index $t'=t+\left|\mathbf{x}_{\mathrm{p}}-\mathbf{x}_{\mathrm{v}}\right|/c$. For this shifted index, $t'=0$ is when our illumination source starts emitting light and light arrives at each voxel at a time index determined by the speed of light.

$$\begin{aligned}I_{\mathcal{P}}\left(\mathbf{x}_{\mathrm{v}},t'\right)&\triangleq\iint\left[\mathcal{P}(t)*H\left(\mathbf{x}_{\mathrm{p}},\mathbf{x}_{\mathrm{c}},t\right)\right]\delta\left(t'-\frac{\left|\mathbf{x}_{\mathrm{p}}-\mathbf{x}_{\mathrm{v}}\right|+\left|\mathbf{x}_{\mathrm{c}}-\mathbf{x}_{\mathrm{v}}\right|}{c}\right)d\mathbf{x}_{\mathrm{c}}d\mathbf{x}_{\mathrm{p}}\\&=\iint\mathcal{P}\left(\mathbf{x}_{\mathrm{p}},\mathbf{x}_{\mathrm{c}},t\right)\delta\left(t-\frac{\left|\mathbf{x}_{\mathrm{c}}-\mathbf{x}_{\mathrm{v}}\right|}{c}\right)d\mathbf{x}_{\mathrm{c}}d\mathbf{x}_{\mathrm{p}}\end{aligned} \tag{31}$$

In the Fourier domain, this is equivalent to adding a frequency-dependent phase shift, $e^{-j\frac{\Omega}{c}}=e^{j\omega\frac{\left|\mathbf{x}_{\mathrm{p}}-\mathbf{x}_{\mathrm{v}}\right|}{c}}$:

$$\begin{aligned}I_{\mathcal{F}}\left(\mathbf{x}_{\mathrm{v}},\Omega\right)&=\iiint\mathcal{P}(\omega)H\left(\mathbf{x}_{\mathrm{p}},\mathbf{x}_{\mathrm{c}},\omega\right)e^{-j\frac{\Omega-\omega}{c}\left(\left|\mathbf{x}_{\mathrm{p}}-\mathbf{x}_{\mathrm{v}}\right|+\left|\mathbf{x}_{\mathrm{c}}-\mathbf{x}_{\mathrm{v}}\right|\right)}d\omega d\mathbf{x}_{\mathrm{c}}d\mathbf{x}_{\mathrm{p}}.\\&=\iiint\mathcal{P}\left(\mathbf{x}_{\mathrm{p}},\mathbf{x}_{\mathrm{c}},\omega\right)e^{-j\frac{\omega}{c}\left(\left|\mathbf{x}_{\mathrm{c}}-\mathbf{x}_{\mathrm{v}}\right|\right)}d\omega d\mathbf{x}_{\mathrm{c}}d\mathbf{x}_{\mathrm{p}}.\end{aligned} \tag{32}$$

This algorithm can be accelerated if we integrate the raw measurement with respect to virtual illumination (x_p) before applying the RSD reconstruction algorithm.

$$I_{\mathcal{F}}\left(\mathbf{x}_{\mathrm{v}}, \Omega\right) = \iint \left[\int \mathcal{P}\left(\mathbf{x}_{\mathrm{p}}, \mathbf{x}_{\mathrm{c}}, \omega\right) d\mathbf{x}_{\mathrm{p}}\right] e^{-j\frac{\omega}{c}\left(\left|\mathbf{x}_{\mathrm{c}}-\mathbf{x}_{\mathrm{v}}\right|\right)} d\omega d\mathbf{x}_{\mathrm{c}}. \tag{33}$$

S2. Beamforming

S2.1. Delay and Sum Beamformer

We can take our impulse response, and add a time delay that depends on the position of our virtual illimunation souce, $\Delta t(x_p)$:

$$\begin{aligned} I\left(\mathbf{x}_{\mathrm{v}}, \mathbf{x}_{\mathrm{p}}\right) &= \int \left\{ e^{-j\frac{\omega}{c}\left|\mathbf{x}_{\mathrm{p}}-\mathbf{x}_{\mathrm{v}}\right|} \int \left[\int \mathcal{P}(\omega)\mathcal{F}\left\{H\left(\mathbf{x}_{\mathrm{p}}, \mathbf{x}_{\mathrm{c}}, t-\Delta t(x_p)\right)\right\} e^{-j\frac{\omega}{c}\left|\mathbf{x}_{\mathrm{v}}-\mathbf{x}_{\mathrm{c}}\right|} d\mathbf{x}_{\mathrm{c}}\right] d\omega \right\} d\mathbf{x}_{\mathrm{p}} \\ &= \int \left\{ e^{-j\frac{\omega}{c}\left|\mathbf{x}_{\mathrm{p}}-\mathbf{x}_{\mathrm{v}}\right|} \int \left[\int \mathcal{P}(\omega) H\left(\mathbf{x}_{\mathrm{p}}, \mathbf{x}_{\mathrm{c}}, \omega\right) e^{-j\omega\Delta t(x_p)} e^{-j\frac{\omega}{c}\left|\mathbf{x}_{\mathrm{v}}-\mathbf{x}_{\mathrm{c}}\right|} d\mathbf{x}_{\mathrm{c}}\right] d\omega \right\} d\mathbf{x}_{\mathrm{p}} \end{aligned} \tag{34}$$

We can generate a transient video of the modified virtual illumination:

$$\begin{aligned} I\left(\mathbf{x}_{\mathrm{v}}, \mathbf{x}_{\mathrm{p}}, t\right) &= \iiint \mathcal{P}(\omega) H\left(\mathbf{x}_{\mathrm{p}}, \mathbf{x}_{\mathrm{c}}, \omega\right) e^{-j\omega\Delta t(x_p)} e^{-j\frac{\omega}{c}\left(\left|\mathbf{x}_{\mathrm{c}}-\mathbf{x}_{\mathrm{v}}\right|\right)} d\omega d\mathbf{x}_{\mathrm{c}} d\mathbf{x}_{\mathrm{p}}. \\ &= \iint \left[\int \mathcal{P}(\omega) H\left(\mathbf{x}_{\mathrm{p}}, \mathbf{x}_{\mathrm{c}}, \omega\right) e^{-j\omega\Delta t(x_p)} d\mathbf{x}_{\mathrm{p}}\right] e^{-j\frac{\omega}{c}\left(\left|\mathbf{x}_{\mathrm{c}}-\mathbf{x}_{\mathrm{v}}\right|\right)} d\omega d\mathbf{x}_{\mathrm{c}}. \end{aligned} \tag{35}$$

By linking our time delay to the spatial location of our virtual illumination source on the relay wall, we can define appropriate functions that can steer or focus our virtual beam.

S2.1.1 BeamSteering

Our virtual illumination position x_p is a 3D vector with an x-component, x_{px}, y-component x_{py}, and z-component x_{pz}. The mean illumination location, $\bar{x}_p$, can be considered to be the beam origin and a spherical coordinate system consisting of (r, Θ, Φ) can be defined to steer the beam to any voxel location given by x_v, along a radial vector $\vec{r}$. For this coordinate system $r = |\vec{r}| = |x_v - \bar{x}_p|$ is the radial distance, Θ is the zenith angle between the z-axis and the radial vector $\vec{r}$, and Φ is the azimuthal angle between the x-axis and the orthogonal projection of the radial vector, $\vec{r}$, onto the relay surface given by $z = 0$ plane. Assuming our relay wall is planar, we can steer our beam in 3D by introducing time delays that depend only on x_{px} and x_{py}. Then, using beamforming principles from ultrasound phased arrays, we can show that the beam can be steered to a voxel x_v using the following formulas:

$$\begin{aligned} \Theta &= \arccos\left(\frac{x_{vz} - \bar{x}_{pz}}{r}\right) \\ \Phi &= \arctan\left(\frac{x_{vy} - \bar{x}_{py}}{x_{vx} - \bar{x}_{px}}\right) \\ \Delta t(x_p) &= \left(\frac{(x_{px} - \bar{x}_{px})\sin\Theta\cos\Phi + (x_{py} - \bar{x}_{py})\sin\Theta\sin\Phi}{c}\right) \end{aligned} \tag{36}$$

S2.1.2 BeamFocusing and BeamSteering

To steer a focused beam to a voxel x_v, we can define time delays that impart an elliptical curvature to each virtual illumination x_p and focuses the beam originating from $\bar{x_p}$ at a focal distance $r = |\vec{r}| = |x_v - \bar{x}_p|$:

$$\Delta t(x_p) = \frac{1}{c}\sqrt{\left(r\sin\Theta\cos\Phi - (x_{px} - \bar{x}_{px})\right)^2 + \left(r\sin\Theta\sin\Phi - (x_{py} - \bar{x}_{py})\right)^2 + \left(r\cos(\Phi)\right)^2} \tag{37}$$

S3. Complexity

Let $L \times L$ denote the number of laser positions, $M \times M$ denote the number of spad pixels focused on the wall, T denote the number of time bins, and F the number of frequency components. Assuming we use the RSD algorithm to reconstruct a volumetric cube of side length N, we can show that the complexity of this algorithm is $O(N^3 \log N)$ [19] for a single SPAD pixel. To get the steady reconstruction for each SPAD pixel, we simply run this algorithm M^2 times and add the results together. This yields a complexity of $O(M^2 N^3 \log N)$. Sequential reconstruction can be used to preserve the memory complexity relative to reconstructing an individual SPAD pixel, while the process can be easily parallelized to speed up reconstruction performance.

To generate a transient video for k total frames, we first compute the inner integral in Eq 33 with respect to x_p. Adding the transient measurements together preserves the memory complexity. Then each frame has the same reconstruction complexity as RSD algorithm for a single SPAD, and running this k times gives the complexity $O(kN^3 \log N)$. This is the complexity for generating the LTM for a single spatial focus for the illumination.

This complexity is order of magnitudes faster than previous work [25], where the corresponding complexity for a single spatial focus was given by $O(N^3 * L^2 * M^2)$ and can be approximated as $O(N^7)$ when $N \approx L \approx M$. Additionally, the output was a Light transport matrix without any timing information. To generate the full TLTM, we need to repeat the previous computation for each spatial location in the hidden volume, yielding an overall complexity of $O(SkN^3 log(N))$, where S is the number of spatial focus locations.

S3.1. Memory Complexity

For sequential reconstruction, the memory complexity remains the same as that of previous algorithms reported for single pixel SPAD [19]. For steady-state reconstruction, we load each histogram of dimension L^2T into memory and compute the reconstruction for an N^3 volume, which yields a complexity of $O(L^2T + N^3)$. This can be improved further if we directly load the temporal Fourier transform of the histogram, yielding a memory complexity of $O(L^2F + N^3)$ where $F < T$, usually by an order of magnitude.

Reconstructing the video for a single spatial focus has a similar memory complexity, since we load the histogram for each SPAD and add them together after applying a SPAD dependent temporal shift to focus the wavefront before reconstructing a 3D volume for each frame. To avoid increasing the memory complexity, we only save an image of a 2D surface for each frame (the location with maximum intensity along depth) yielding a complexity of $O(L^2F + kN^2)$ and avoiding storing a 4D cube.

GPU-based reconstructions can increase reconstruction speed at the expense of requiring additional memory. To provide context, our GPU-based implementation can output a video for a single focus within 1 - 2 minutes while matlab-based sequential reconstruction takes around 30 minutes for 228 SPAD pixels and 141 video frames.

S4. SNR

A practical benefit of using a multi-pixel SPAD array is the ability to capture significantly more photons per unit of time than when using a single-pixel photon detector. While this may seem intuitively obvious, we believe the consequences of this benefit are foundational to advancing NLOS research and furthering the development of practical NLOS applications. Therefore, this fact warrants closer inspection.

A researcher accustomed to imaging with almost any transient camera method is familiar with the unavoidable trade-off: exposure time versus image quality. SNR is largely a function of the number of photons collected. Short exposure times collect fewer photons, therefore have higher Poisson noise, and consequently, lower SNR. Longer exposure times ameliorate the low-SNR problem because Poisson noise grows at a rate proportional to the square root of the number of photons detected. However, long exposure times are not a universal panacea because many compelling practical applications of NLOS imaging either involve dynamic scenes or require near real-time analysis, e.g. [26].

The expectation is that for a given single-pixel exposure time, t_1, an equivalent SNR can be achieved by instead using n pixels with an exposure time of $t_n = t_1/n$. We show for the first time experimentally that this expectation does indeed hold true. To achieve this trade-off, it is necessary to project the SPAD array onto a small area on the relay wall, ideally having dimensions of the same order as the phasor-field wavelength. By doing so, all pixels observe nearly the same perspective of the hidden scene and can be easily combined without introducing any additional blurring in the combined reconstruction. In our experimental data, all SPAD pixels were projected onto a rectangular area of 7.1 $\times$ 4.7 cm.

In our experiment, we constructed a simple scene containing a Lambertian number '2' cut from expanded polystyrene foam. We then captured 125 exposures of 7.25s each using all available SPAD pixels on our array. Analysis was performed

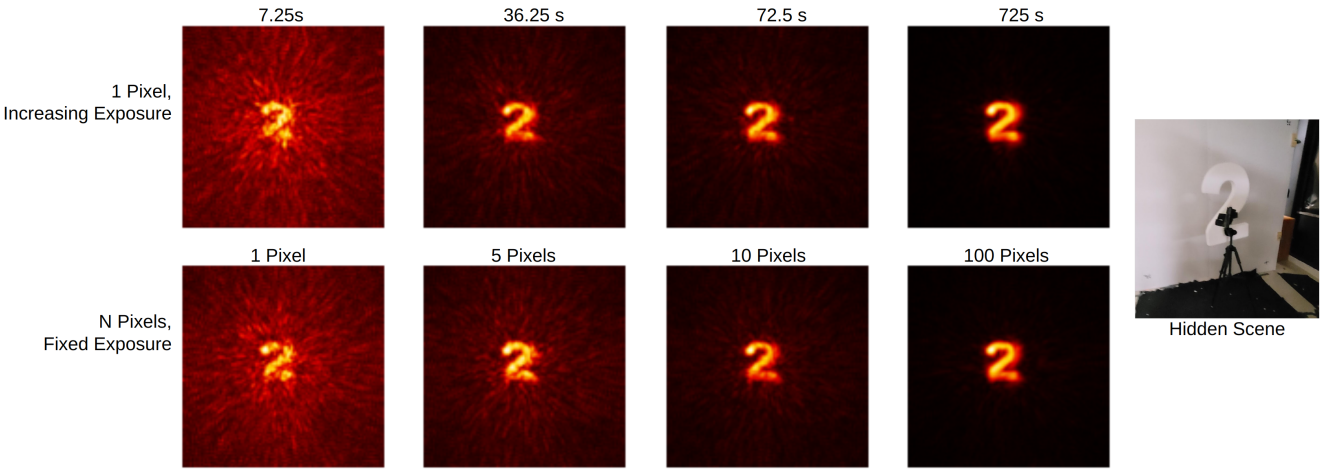

Figure S10. Qualitative comparison of SNR. **Left, Top row:** Single-pixel reconstructions at progressively longer exposure times. **Left, Bottom row:** Reconstructions with a fixed 7.25s exposure time using an increasing number of SPAD pixels. **Right:** Hidden object

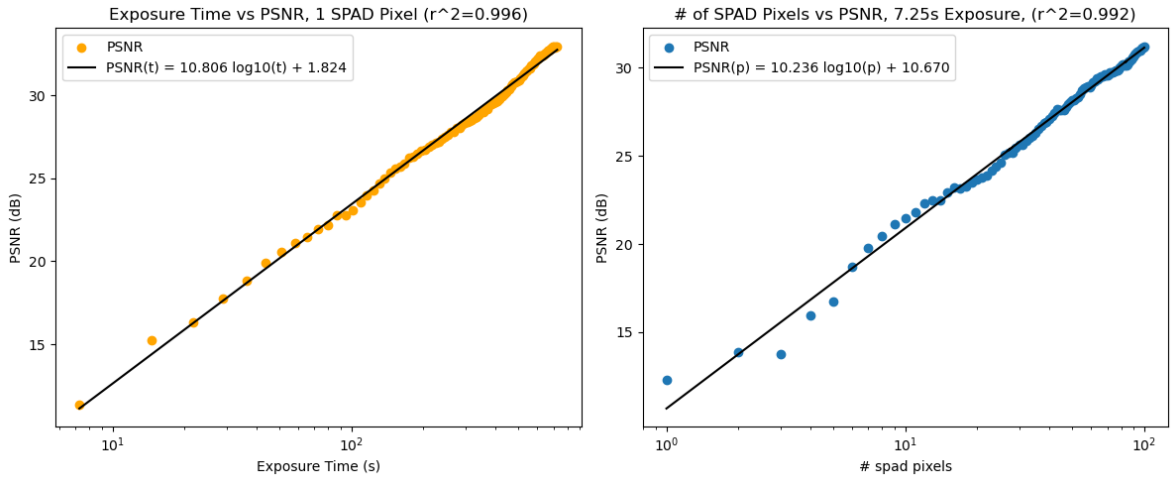

Figure S11. Quantitative comparison of SNR, showcasing that the growth factor for PSNR is similar when increasing the exposure time (10.806) or the number of SPAD pixels (10.236).

in the following manner: First, we selected a single pixel at random from the array. Using only photons collected by this pixel, we performed 100 phasor-field reconstructions having exposure times $7.25 \times n$ seconds for $n = 1 \ldots 100$. Second, we selected a single 7.25-second exposure, and performed 100 phasor-field reconstructions, using n arbitrarily selected pixels for $n = 1 \ldots 100$. As expected, we observed a linear increase in total detected photon count with both the increasing exposure and increasing pixels analyses. A sampling of the reconstructions for $n = \{1, 5, 10, 100\}$ are shown in figure S10.

For quantitative analysis, we used the widely accepted Peak Signal-to-Noise Ratio (PSNR) calculation as a metric to assess improvement in image quality with the addition of more photons, whether by increased exposure or increased number of pixels. The PSNR formula requires calculating the mean square error with respect to a noise-free image. For the noise-free image in our calculations, we used a reconstruction from 224 pixels of a 906.25 second exposure. The PSNRs of each of the 200 reconstructions were calculated and are shown in figure S11. A least-squares, best-fit exponential curve overlays each data set. Note that both methods of adding additional photons (longer exposure vs additional pixels) have nearly the same growth rate for the reconstructions' PSNR (0.186 vs 0.184). From this, we can conclude that when a tight spacing of SPAD pixels on the relay wall is used, capturing with multiple SPAD pixels is experimentally equivalent to capturing over a longer exposure time.

S4.1. Primal-Dual Equivalence

The photon time of flight is identical whether the photon travels from the illumination to the detection or from the detection to the illumination. Additionally, fast reconstruction utilizes Fourier transforms which require sampling the relay wall using a uniform grid. While the illumination and detection grids can be interchanged in principle, it is easier to generate a uniform laser grid since visual feedback can be used to focus the laser and the galvo can be used to scan the uniform grid sequentially. Conversely, the SPAD pixels must be focused simultaneously and the uniformity of the detection grid is ultimately limited by the focusing optics and the presence of hot pixels on the sensor, leading to distortions seen in Figure 6. Since we don't have a uniform SPAD grid, we use a slow phasor field back projection algorithm to demonstrate the equivalence between the primal and dual images.

Figure S12 shows the hidden scene (left) and the reconstruction (right) after adding up the reconstructions from all the SPAD pixels for the full laser grid. Next, we reconstruct the primal (left) and dual (right) images for a single laser and camera position respectively, and display them in Fig S13. To get the SPAD and laser grids to match (Row 1, Figure S13), the laser grid is downsampled leading to a loss of reconstruction resolution. Additionally, the laser positions outside the corners of the SPAD grid are discarded (Figure 6), which is the bottom of the 2 and the T is missing in the reconstruction. The 2 in the laser grid reconstruction looks cleaner, but that may be due to the prescence of hot pixels within the SPAD grid.

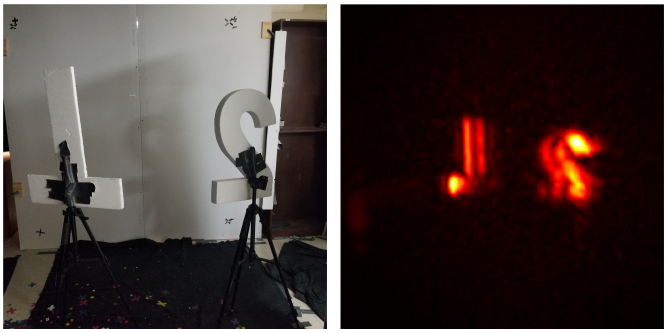

Figure S12. Hidden scene shown on the left, and reconstruction from all SPAD pixels for the large field of view camera grid is shown on the right. Adding up all the contributions yields 3D shading and lighting effects for the hidden scene.

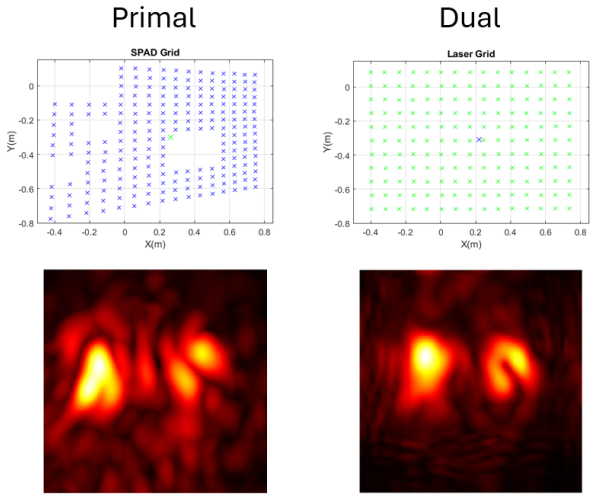

Figure S13. Qualitative comparison of primal (column 1) and dual (column 2) images. Row 1 displays the two grids used, with camera positions marked in blue and laser positions marked in green. Row 2 displays the corresponding reconstruction.

S5. Supplementary Results

S5.1. Novel Illumination

Our current pixels utilize each SPAD pixel as a virtual illumination source. Therefore, with a SPAD array focused to a large area on the relay surface, we can demonstrate the impact of illuminating the hidden scene from Figure S12 from different parts of the relay wall.

Each reconstruction (and illumination) is averaged over 4x4 subgroup of neighboring SPAD pixels to improve SNR and is shown on the right. These pixels share a single Time-to-Digital Converter (TDC), and are labelled using their TDC index. We note that the 2 is seen most clearly by the bottom left pixels while the T is seen the most clearly by top right pixels on the relay surface. This is because the opposing tilts of the hidden objects mean that T is in the missing cone [17, 40] of the pixels located on the right-hand side of the relay surface while 2 is in the missing cones of the pixels on the left.

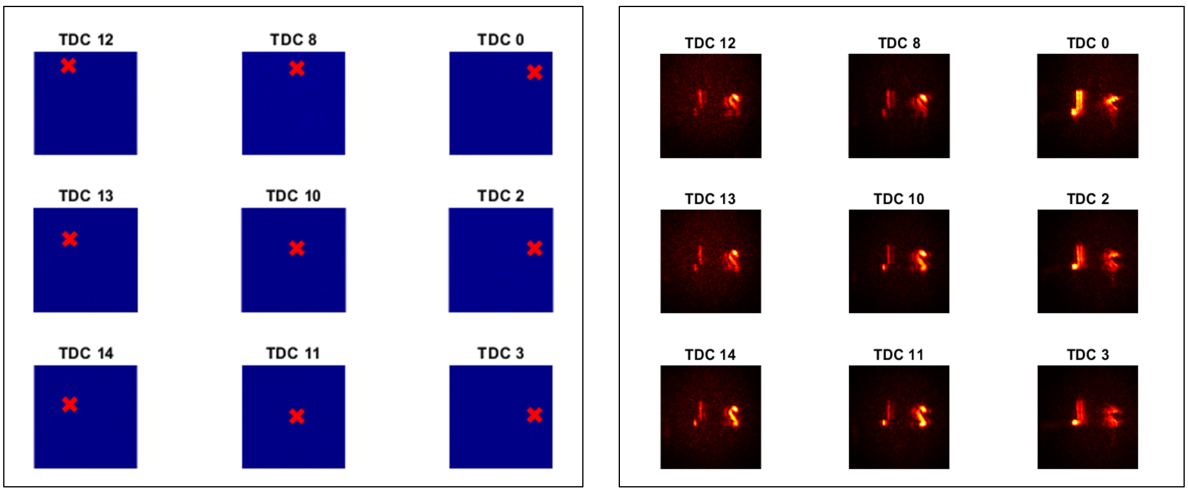

Figure S14. The left image shows rough location for the illumination on the relay surface. The image on the right shows the front view of the 3D light field. Each single image is generated by averaging reconstructions over a 4x4 subgroup of neighboring SPAD pixels that share a TDC, and thus each image is labelled using corresponding TDC index.

S5.2. Beamforming

We virtually focus on an additional scene, illuminating the chair, mannequin, and shelf at the correct times (Row 1 of Fig 8). By leveraging principles of beamforming, we can add a curvature to the virtual illumination wavefront so that it focuses on the mannequin (Row 2 of Fig 8) and the illumination does not encounter the shelf or the chair anymore.

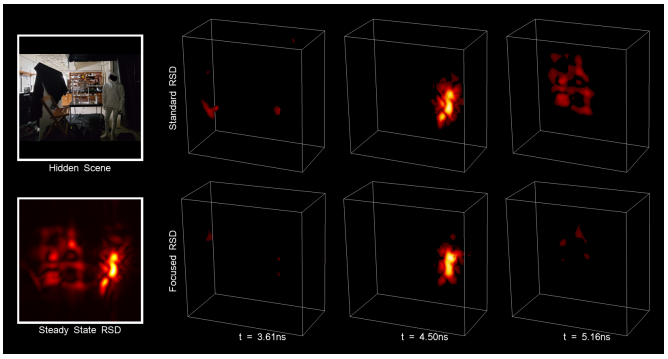

Figure S15. Hidden scene shown on the left. Still-images from spatial-temporal video show light propagating through the scene. Top row: Standard RSD transient reconstruction. Bottom row: Beamforming reconstruction, focusing the virtual wavefront on the mannequin. Notice the chair and shelf are not illuminated in this reconstruction. See supplemental materials for video files of these reconstructions.

S5.3. Complex Light Transport: Subsurface Scattering

S5.3.1 Subsurface Scattering

In Figure S16, we place an acquarium in the hidden scene and play a video of light propagating through the tank. Row 1 shows how the illumination interacts with water, showcasing caustic direct reflections from the front wall of the tank (column 2), the back wall of the tank (column 3), as well as the indirect caustic reflection from the reflection of the back wall hitting the front wall (column 4). We then add milk, and demonstrate how the illumination interacts differently (Row 2) with the milk-water mixture. We encounter diffuse scattering from within the tank (0.14 ns) after the reflection from glass front (column 2),, and observe subsurface scattering from the milky solution (column 3). Finally, we don't see any reflection from the back wall of the tank at the corresponding frame (column 3).

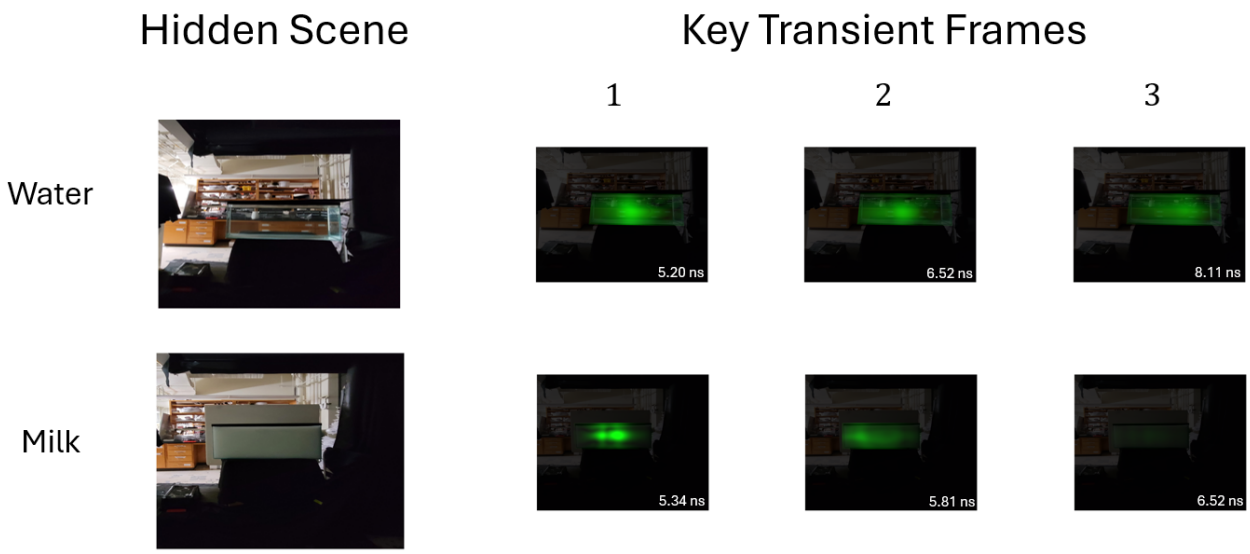

Figure S16. Hidden scene shown on the left. Still-images from spatial-temporal video showing light propagating through the glass tank containing water (Row 1) and milk (Row 2). Top row: Illuminate the tank with a plane wave, lighting up the tank (column 2). Some light propagates through the clear tank and illuminates the back wall (column 3). Finally, we see the reflection from the back wall illuminating the front wall (column 4). Bottom row: The light coming back is from the milky solution within the tank (column 2), with the illumination persisting within the tank due to subsurface scattering (column 3). No light is reflected back from the back wall of the tank due to the opaque milky mixture (column 4). See supplemental materials for video files of these reconstructions.

S5.4. Complex Light Transport: 4th Bounce Shadow

In Figure S17, we showcase the light transport effects within the indirect component of the TLTM. Here, we place a vase in front of a back wall, along with a mirror on the side. Focusing our virtual illumination on the mirror reveals a 4th bounce of light hitting the mirror followed by hitting the back wall (Row 2). When the vase is present, a shadow corresponding to the indirect illumination hitting the vase (and the vase stand) can be seen on the back wall (Row 2, Column 1). We verify this by removing the vase and the vase stand, leading to the disappearance of the shadow (Row 2, Column 2)

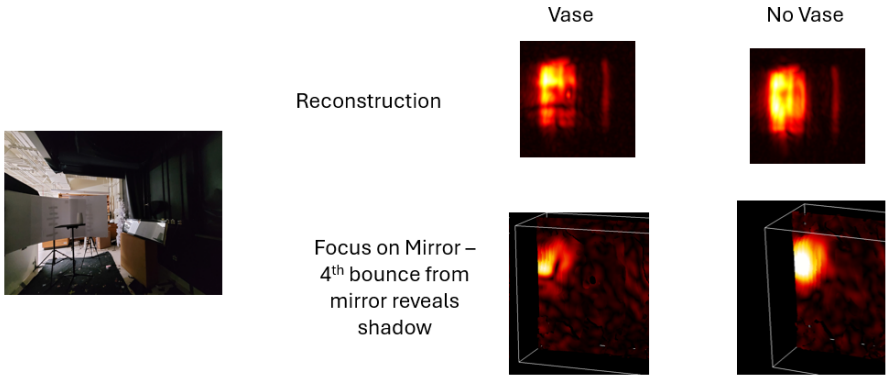

Figure S17. Hidden scene shows on the left, consisting of a vase, back wall, and a mirror. When we focus our virtual illumintion on the mirror, we see the backwall light up due to the 4th bounce of light from light hitting the mirror and then the back wall (Row 2). We also see a shadow of the vase blocking some of this light (Column 2, Row 2). When we remove the vase, we see that this shadow disappears showcasing that light transport is preserved.