Looking Around Flatland: End-to-End 2D Real-Time NLOS Imaging

María Peña, Diego Gutierrez, and Julio Marco

***Abstract*—Time-gated non-line-of-sight (NLOS) imaging methods reconstruct scenes hidden around a corner by inverting the optical path of indirect photons measured at visible surfaces. These methods are, however, hindered by intricate, time-consuming calibration processes involving expensive capture hardware. Simulation of transient light transport in synthetic 3D scenes has become a powerful but computationally-intensive alternative for analysis and benchmarking of NLOS imaging methods. NLOS imaging methods also suffer from high computational complexity. In our work, we rely on dimensionality reduction to provide a real-time simulation framework for NLOS imaging performance analysis. We extend steady-state light transport in self-contained 2D worlds to take into account the propagation of time-resolved illumination by reformulating the transient path integral in 2D. We couple it with the recent phasor-field formulation of NLOS imaging to provide an end-to-end simulation and imaging pipeline that incorporates different NLOS imaging camera models. Our pipeline yields real-time NLOS images and progressive refinement of light transport simulations. We allow comprehensive control on a wide set of scene, rendering, and NLOS imaging parameters, providing effective real-time analysis of their impact on reconstruction quality. We illustrate the effectiveness of our pipeline by validating 2D counterparts of existing 3D NLOS imaging experiments, and provide an extensive analysis of imaging performance including a wider set of NLOS imaging conditions, such as filtering, reflectance, and geometric features in NLOS imaging setups.**

***Index Terms*—Transient rendering, NLOS imaging, computational imaging.**

I. Introduction

NON-LINE-OF-SIGHT (NLOS) imaging techniques aim to retrieve information about scenes that are not directly observable by a camera [1], [2], [3], including motion tracking [4], position detection [5], [6], and geometric reconstructions [7], [8], [9]. It has promising applications in diverse fields, such as cave exploration, rescue planning, medical imaging, and car navigation, to name a few. Time-gated NLOS imaging methods, in particular, rely on time-of-flight information to provide detailed reconstructions of a scene hidden either by triangulation [7], [8], [10], [11] or relying on wave-propagation principles [9], [12], [13], [14], [15], [16], [17], [18], [19].

However, the development of new time-gated NLOS imaging methods still suffers from three main problems. First, capture hardware is expensive, and some setups require hour-long captures (Liu et al. report 6.5 hours for some of their experiments [9]). Second, setting up scenes in a laboratory and calibrating the system can take hours or even days per experiment. Third, the high dimensionality and density of the captured data lead to severe memory constraints. Simulation-based pipelines, on the other hand, have become a fundamental tool for analyzing performance of computational imaging systems [20], [21] thanks to accurate physically-based rendering.

Existing transient rendering frameworks allow researchers to prototype and simulate transient light transport in synthetic 3D scene configurations [22], [23], [24]. Their flexibility provides additional insights by storing information that real capture devices cannot obtain, such as ground truth information of separate light bounces. However, simulated data in 3D scenes is still high-dimensional and dense, maintaining the severe memory constraints akin to captured data. Besides, accurate transient rendering in 3D scenes is computationally expensive, requiring render times of several minutes for a single scene even for a low number of simulation samples. As a consequence, NLOS imaging methods deal with high-dimensional captured and simulated data, leading to high computational costs, with high-resolution results taking several minutes or hours to compute.

Simulation of steady-state 2D light transport has proven to be a computationally-efficient approach to gain practical insights of light transport phenomena by relying on dimensionality reduction [25], [26], [27], [28], [29]. Inspired by this, we extend existing formulations of steady-state light transport in self-contained 2D worlds [28], [29] to account for *transient* light transport, based on the transient path integral formulation [22]. We then couple our 2D transient path integral formulation with our 2D re-formulation of phasor-based NLOS imaging models [9], [13], [14], [15], [16], [17]. We implement this combination in a WebGL-based end-to-end pipeline that simultaneously performs transient rendering and NLOS imaging at real-time rates, thanks to computational speedups of up to five orders of magnitude w.r.t. equivalent 3D counterparts. In our most complex 2D scene—a hidden Stanford bunny—our experiments perform transient light transport simulation and NLOS imaging (throughout the paper, we refer as imaging to the NLOS reconstruction step) in 17 ms to 19 ms for 16×10^3 simulation samples. In contrast, simulating and imaging a 3D Stanford bunny takes up to 47 min. In addition,

Received 1 July 2024; revised 21 October 2024 and 11 December 2024; accepted 7 January 2025. Date of publication 29 January 2025; date of current version 14 February 2025. This work was supported by the European Union's European Defense Fund Program through the ENLIGHTEN project under Grant Agreement number 101103242. The associate editor coordinating the review of this article and approving it for publication was Prof. Mohit Gupta. *(Corresponding author: María Peña.)*

The authors are with the Universidad de Zaragoza, 50008 Zaragoza, Aragón, Spain, and also with the I3A, 50018 Zaragoza, Aragón, Spain (e-mail: mpena@unizar.es; diegog@unizar.es; juliom@unizar.es).

This article has supplementary downloadable material available at https://doi.org/10.1109/TCI.2025.3536092, provided by the authors.

Digital Object Identifier 10.1109/TCI.2025.3536092

memory consumption is reduced by a factor of 280, from 3.4 GB in 3D to 12MB in 2D. We leverage our performance gains to progressively refine the simulated data by iteratively accumulating batches of simulation samples, showing the reconstruction results in real-time until a user-specified maximum of samples is reached. This provides reference reconstructions for analysis and benchmarking, and allows the user to modify the scene and imaging parameters and visualize their impact in the reconstruction in real-time as the simulation is refined. Our system allows much faster and more precise configuration of scenes than in 3D counterparts, facilitating exhaustive exploration of the parameter space, generating lighter data, and providing interactive and easier visualization of the imaging results.

We observe that NLOS imaging methods in 2D based on the phasor-field formulation retain the ability to image hidden geometry and the limited visibility caused by the missing cone problem [30]. We validate our system by adapting experiments from previous works to the 2D world, analyzing filtering functions, object visibility [30], and the mirror-like behavior of diffuse, planar curves (the 2D equivalent of surfaces) under NLOS imaging virtual illumination [31]. The speedup obtained from dimensionality reduction will allow researchers to do quick analysis of scene prototypes and help design the final 3D scene with the desired properties, e.g., avoiding known visibility problems to analyze different aspects, or forcing ill-posed orientations to test new methods that aim to improve visibility. Previous works usually assume diffuse materials, and it is common for hidden objects to present a locally-planar structure, e.g., when using planes and letters. We demonstrate the utility of our system by performing new studies on visibility with varying materials and geometric features.

Besides providing a tool for fast analysis and prototyping, our system can also become instrumental in introducing the topic of NLOS imaging to a larger public. It can also be generalized to other applications, such as visualization of light in motion.

In particular, we make the following contributions:

- Extending steady-state 2D light transport to transient light transport based on the transient path integral formulation (Section IV).
- Coupling transient 2D light transport with NLOS imaging algorithms to provide a real-time pipeline for NLOS imaging with support for the NLOS imaging camera models introduced with the phasor-field formulation. Our code and simulation tool are publicly available[1] (Section V).
- Leveraging our system to perform systematic and efficient analysis on varying NLOS imaging conditions, including filtering, reflectance, and geometric features (Section VI).

II. Related Work

a) NLOS Imaging: We focus on active NLOS imaging methods, which use controlled light sources to illuminate the hidden scene. Time-of-flight detectors at picosecond resolution enable the capture of light at high temporal resolutions comparable to the speed of light, which nurtured a wide range of NLOS imaging methods [8], [10], [11], [12], [18], [32], [33]. The first methods to obtain high-quality 3D reconstructions employed filtered backprojection [8], [10]. They are computationally expensive but many optimizations have been proposed, including GPU-based implementations [32]. The phasor-field formulation [9], [13], [14], [15], [16], [17] provides wave-based models to propagate virtual waves into the hidden scene, fostering the development of more efficient methods [18], [19], [34] and new analysis, such as the effect of hidden-object reflectance on the NLOS imaging results [35], [36]. The phasor-field formulation creates a virtual camera aperture at the relay wall (a visible surface) and derives different NLOS imaging camera models. We adapt three different NLOS imaging camera models to the 2D world. Naive backprojection-based solvers in this context are fast enough when implemented on a GPU. Still, the mentioned approaches can potentially be applied to improve efficiency for more complex methods or big, high-resolution scenes. Royo et al. [31] have recently shown imaging around two corners, exploiting the mirror-like properties of diffuse surfaces under wave-based NLOS imaging methods. Our work shows that these properties are maintained in the 2D world, so it can provide a faster prototyping tool for future work following this line of research.

b) Transient Light Transport: Smith et al. [37] proposed to extend the rendering equation by incorporating propagation time delays, which led to transient variations of Monte Carlo rendering [38], [39] and simulation of time-of-flight sensors [40]. Later, Jarabo et al. [22] introduced the transient path integral, that includes propagation and scattering delays. We rely on their formulation to develop a formal model of transient light transport in 2D. These ideas have also been applied to NLOS imaging to reduce the need for expensive data capture. However, the high dimensionality of the problem results in time-consuming and expensive algorithms. Previous works apply different strategies to alleviate this problem. Some authors account only for three-bounce paths, simplifying the transient path integral to perform efficient inverse rendering [41], [42]. Longer paths can provide additional information to tackle more challenging problems, such as seeing around two corners [31], so we do not limit our method in this way. Other works focus on noise reduction. Jarabo et al. [22] introduce a sampling technique to account for the temporal profile of light transport in participating media. In this work, we limit our analysis to media-free scenes. Pediredla et al. [23] introduce a sampling strategy for finding paths of a pre-determined length in media-free scenes through ellipsoidal connections. Their technique considerably improves the transient rendering of narrow temporal regions or dynamic scenes. It increases the cost of generating each path, so it does not offer a performance advantage over temporal path reuse when covering a larger temporal range [22]. Hidden objects in NLOS imaging setups are usually far from the relay wall and relatively small, while light sources cannot be sampled directly. This makes transient rendering especially challenging. Royo et al. [24] present geometry and laser sampling strategies specific for NLOS imaging. In this work, we do not implement these sampling strategies and leave them as future work.

[1] https://github.com/mpenalm/nlos_flatland, https://mpenalm.github.io/nlos_flatland

c) Light Transport in 2D: Some previous works on light transport analysis reduce the problem to two spatial dimensions to perform prior analyses to the full 3D problem [26], [27], [28], while various interactive 2D light transport simulators are available online [29], [43], [44], [45]. We follow this approach for NLOS imaging, extending Jarosz et al.'s formal definition of steady-state light transport in 2D worlds [28] by incorporating the temporal dimension. Bitterli's Tantalum renderer [29] uses light tracing to show light transport in 2D scenes in steady-state. Most renderers compute exiting radiance from curve (analogous to a surface in 3D) points or media points in the scene, projected into the image line of the camera (analogous to the typical image plane in 3D). Instead, Tantalum shows light transport in the full scene by computing fluence. Fluence is the average amount of light passing through a point in space, not limited to curves and media. Bitterli also developed a 2D transient light transport visualizer based on Tantalum [46]. However, NLOS imaging requires a device inside the scene to capture incoming light from a set of points in its geometry. Therefore, his approach is not directly applicable. We extend the Tantalum renderer to support propagation time and a transient capture device placed inside the scene. Fluence shows which parts of the scene are properly illuminated, which can help understand some NLOS imaging results, so we keep the original steady-state fluence visualization.

III. NLOS Imaging Background

This section presents an overview of the capture and computational principles of time-gated NLOS imaging on which our work is based. We build upon the phasor-field formulation [9] to summarize the different NLOS imaging camera models implemented in our pipeline. We include a table defining the most relevant symbols in the supplementary material, Table III.

A. Data Acquisition

Fig. 1 illustrates a conventional time-gated NLOS imaging setup. During the capture process, an ultra-fast laser (red) located at $\mathbf{x}_{l0}$ illuminates a set of points $\mathbf{x}_l \in \mathcal{L}$ on a relay wall visible to the camera using short illumination pulses. Light is then scattered by the relay wall towards points $\mathbf{x}_v \in \mathcal{V}$ in the hidden scene and back to the relay wall. For every illuminated point, an ultra-fast sensor (blue) located at $\mathbf{x}_{s0}$ captures indirect photons arriving at a set of points $\mathbf{x}_s \in \mathcal{S}$ on the relay wall. The resulting time-resolved radiance measure (commonly called a transient) is typically a tensor $\mathbf{H}$ discretized in $N_l \times N_s \times N_t$ spatio-temporal bins on the dimensions of laser targets $\mathbf{x}_l$, sensor targets $\mathbf{x}_s$, and time t, respectively. Transient light transport simulation in 3D synthetic NLOS setups mimics the capture process to obtain $\mathbf{H}$ through different transient rendering methods [22], [23], [24], [41], [42], [47], [48]. $\mathbf{H}$ approximates the continuous *impulse response function* $H(\mathbf{x}_l, \mathbf{x}_s, t)$ of the hidden scene, which represents global illumination at the relay wall measured by a sensor with a delta response function after illuminating the wall with delta illumination pulses. NLOS imaging methods are commonly formulated in terms of the impulse response function H, while in practice its discrete counterpart $\mathbf{H}$ is used throughout

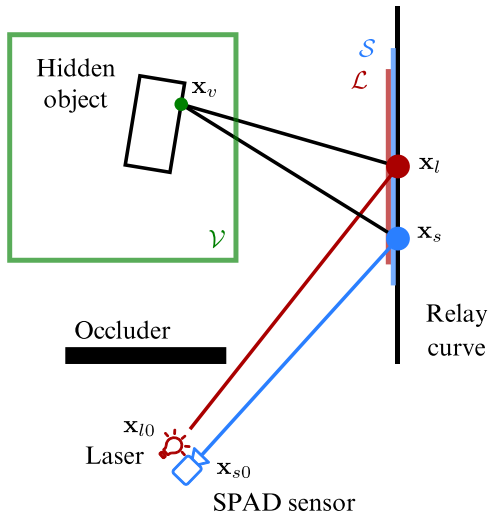

Fig. 1. NLOS imaging in the 2D world. A laser at $\mathbf{x}_{l0}$ illuminates points $\mathbf{x}_l$ in the relay curve, which corresponds to the usual 3D relay wall. A sensor at $\mathbf{x}_{s0}$ observes all points $\mathbf{x}_s$.

the reconstruction process. In 2D, light interacts with 1D boundaries called curves, equivalent to surfaces in 3D. Therefore, the 3D relay wall becomes a *relay segment*, and the integration domain of the laser and sensor targets $\mathbf{x}_l \in \mathcal{L}, \mathbf{x}_s \in \mathcal{S}$ become curves, too, while the domain of the hidden scene $\mathbf{x}_v \in \mathcal{V}$ is a plane instead of a volume.

B. NLOS Forward Transport

To provide tractable reconstruction algorithms, most NLOS imaging methods assume the impulse response function $H(\mathbf{x}_l, \mathbf{x}_s, t)$ represents occlusion-free light transport from third-bounce-only illumination paths $\overline{\mathbf{x}} = \mathbf{x}_{l0} \rightarrow \mathbf{x}_l \rightarrow \mathbf{x}_v \rightarrow \mathbf{x}_s \rightarrow \mathbf{x}_{s0}$ where $\mathbf{x}_v$ is an arbitrary point in the hidden scene (see Fig. 1), and all objects exhibit diffuse reflectance. Under these assumptions, H can be defined as:

$$H(\mathbf{x}_l, \mathbf{x}_s, t) \approx \int_{\mathcal{V}} f(\mathbf{x}_v) G(\overline{\mathbf{x}}) \delta(t - \Delta t_v) \mathrm{d}\mathbf{x}_v, \qquad (1)$$

where $f(\mathbf{x}_v)$ represents the albedo of diffuse surfaces in the hidden scene; $G(\overline{\mathbf{x}})$ is the geometric attenuation of three-bounce paths $\overline{\mathbf{x}}$; Δt_v is the time of flight of path $\overline{\mathbf{x}}$; and the delta term constrains light transport to three-bounce paths. The laser and sensor devices are fixed at the same location during the entire capture process, and therefore $\mathbf{x}_{l0}$ and $\mathbf{x}_{s0}$ are constant values.

C. Phasor-Based NLOS Imaging

The phasor-field formulation [9] builds upon wave-optics principles to formulate different NLOS imaging camera models based on virtual-wave propagation operators. Such formulation encompasses classic NLOS reconstruction methods, which aim to invert (1) to estimate f, as well as novel NLOS imaging camera models [31], [49]. Under this formulation, the illuminated points $\mathbf{x}_l \in \mathcal{L}$ define a virtual illumination aperture, and the sensor points $\mathbf{x}_s \in \mathcal{S}$ define a virtual camera aperture. These two virtual apertures can be used to

1) implement virtual illumination functions to computationally emit and/or focus light at specific points of the hidden scene; and
2) generate images of the hidden scene that result from such illumination by computationally focusing the virtual camera aperture at specific points in the hidden scene.

Given any virtual time-resolved illumination function $\mathcal{P}(\mathbf{x}_l, t)$, the response $\mathcal{P}(\mathbf{x}_s, t)$ of the hidden scene to such illumination at $\mathbf{x}_s \in \mathcal{S}$ is defined by

$$\mathcal{P}(\mathbf{x}_s, t) = \int_{\mathcal{L}} \mathcal{P}(\mathbf{x}_l, t) *_t H(\mathbf{x}_l, \mathbf{x}_s, t) \mathrm{d}\mathbf{x}_l, \tag{2}$$

where $*_t$ denotes a convolution in time, and $\mathcal{P}(\mathbf{x}_l, t)$ and $\mathcal{P}(\mathbf{x}_s, t)$ can be complex-valued functions. Their frequency-domain counterparts $\widehat{\mathcal{P}}(\mathbf{x}_., \Omega) = \mathcal{F}_t\{\mathcal{P}(\mathbf{x}_., t)\}$—with $\mathcal{F}_t$ the Fourier transform—represent virtual illumination *phasors* at frequency Ω, and the collection of such phasors at apertures $\mathbf{x}_l \in \mathcal{L}, \mathbf{x}_s \in \mathcal{S}$ is known as a *phasor field*. By applying well-known wave-based lens operators to such phasor fields, the phasor-field formulation allows to define different NLOS imaging camera models and efficient solvers [9], [18], [31], [49]. This imaging step is compactly defined as

$$\widehat{f}(\mathbf{x}_v, \Omega) = \Phi(\mathbf{x}_v, \widehat{\mathcal{P}}(\mathbf{x}_s, \Omega)), \tag{3}$$

where Φ represents an imaging operator for the camera aperture based on Rayleigh-Sommerfeld (RSD) propagation, and $\widehat{f}(\mathbf{x}_v, \Omega)$ is a set of phasors that represents the resulting image of the hidden scene at imaging frequencies Ω.

The entire phasor-based NLOS imaging process can be expressed as

$$\widehat{f}(\mathbf{x}_v, \Omega) = \int_{\mathcal{S}} \frac{e^{ikd_{sv}}}{d_{sv}} \int_{\mathcal{L}} \widehat{\mathcal{P}}(\mathbf{x}_l, \Omega) \widehat{H}(\mathbf{x}_l, \mathbf{x}_s, \Omega) \mathrm{d}\mathbf{x}_l \mathrm{d}\mathbf{x}_s, \tag{4}$$

where $\widehat{H}(\mathbf{x}_l, \mathbf{x}_s, \Omega) = \mathcal{F}_t\{H(\mathbf{x}_l, \mathbf{x}_s, t)\}$, $d_{sv} = \|\mathbf{x}_s - \mathbf{x}_v\|$, and $k = 2\pi\Omega/c$ is the wavenumber, with c the speed of light. The frequency-domain virtual illumination function $\widehat{\mathcal{P}}(\mathbf{x}_l, \Omega)$ can be defined to e.g. constrain the set of imaging frequencies, introduce temporal offsets to virtual emitters, or implement virtual illumination lenses in the domain of laser targets $\mathbf{x}_l \in \mathcal{L}$. Different capture setups lead to specific trade-offs in both capture and computational efficiency. Confocal setups co-locate the laser and sensor targets $\mathbf{x}_l \equiv \mathbf{x}_s$ and lead to closed-form formulations that can be efficiently solved [11], [12]. Non-confocal setups do not impose such constraint, enabling the use of multi-pixel SPAD sensors.

In the remainder of this manuscript, we couple the simulation of 2D transient light transport with the phasor-field formulation to define and implement different NLOS imaging camera models in an efficient 2D end-to-end pipeline. For this, in Section IV we describe transient global illumination in 2D scenes under the transient path integral formulation and show how to leverage it to simulate impulse response functions of synthetic 2D NLOS scenes. In Section V, we describe how we efficiently simulate $\mathbf{H}$, how to couple it with the phasor-based NLOS imaging camera models presented, and the solvers we implement for such camera models.

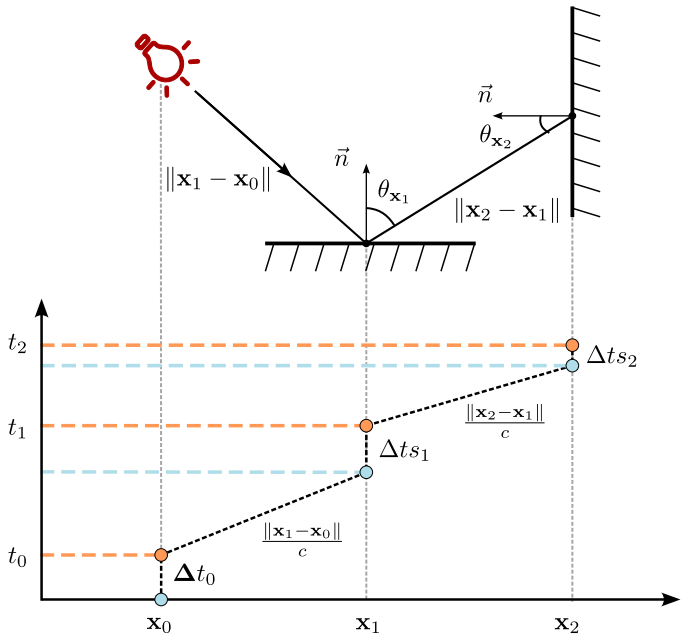

Fig. 2. Spatio-temporal diagram of light propagation for a path of length $m = 2$. Light is emitted from $\mathbf{x}_0$ at time t_0, and reaches $\mathbf{x}_1$ at $t_0 + \|\mathbf{x}_1 - \mathbf{x}_0\|/c$. After a scattering delay Δts_1, light emerges from $\mathbf{x}_1$ at t_1 and takes $\|\mathbf{x}_2 - \mathbf{x}_1\|/c$ time to reach $\mathbf{x}_2$, where another scattering delay Δts_2 may occur before light is reflected. The angles between the normal to the curve and outgoing $\theta_{\mathbf{x}_1}$ and incoming directions $\theta_{\mathbf{x}_2}$ affect light attenuation. Figure after Jarabo et al. [22].

IV. Transient Light Transport in 2D NLOS Scenes

In the following, we formulate transient light transport in self-contained 2D worlds based on the transient path integral [22] and 2D radiometry [28], and define the impulse response function of 2D NLOS imaging setups based on this formulation. We include a table defining the most relevant symbols in the supplementary material, Table II.

A. 2D Transient Path Integral

The original transient path integral formulation [22] defines any spatio-temporal measurement I_j as the integral over the domain of light transport paths between 3D surfaces and media, and over the domain of temporal delays resulting from scattering and optical distances of such paths. In Fig. 2, we illustrate said temporal delays on a path of length $m = 2$, where Δts_i corresponds to scattering delays. Considering an intrinsic self-contained 2D world, where all light is generated, scattered, absorbed, and measured inside it, this integral can be expressed as

$$I_j = \int_{\Psi_{2D}} \int_{\Delta T} \rho_j(\overline{\mathbf{x}}, \overline{\mathbf{\Delta t}}) \mathrm{d}\mu(\overline{\mathbf{\Delta t}}) \mathrm{d}\mu_{2D}(\overline{\mathbf{x}}), \tag{5}$$

where Ψ_{2D} is the space of 2D light transport paths, ΔT is the space of temporal delays of all paths, $\overline{\mathbf{x}} = \mathbf{x}_0 \ldots \mathbf{x}_m, m \geq 1$ represents the 2D spatial coordinates of $m+1$ vertices of a light path $\overline{\mathbf{x}} \in \Psi_{2D}$ between a light source at $\mathbf{x}_0$ and a camera sensor at $\mathbf{x}_m$, with $\mathbf{x}_1 \ldots \mathbf{x}_{m-1}$ intermediate scattering vertices at curves or media; $\overline{\mathbf{\Delta t}} = \mathbf{\Delta} t_0 \ldots \mathbf{\Delta} t_m$ represents a sequence

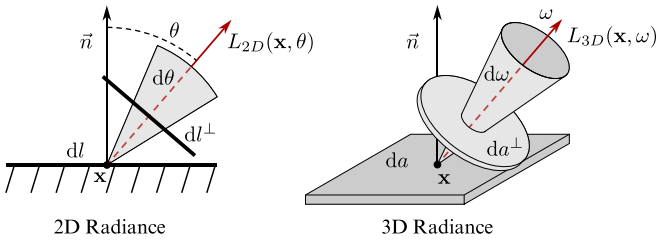

Fig. 3. Comparison of radiance in 2D (left) and 3D (right). It expresses the amount of light arriving at or leaving a single point $\mathbf{x}$ from a differential plane angle $\mathrm{d}\theta$ (2D) or solid angle $\mathrm{d}\omega$ (3D). Figure adapted from Jarosz et al. [28].

of time delays resulting from the optical distance and scattering events of the path $\overline{\mathbf{x}}$; the differential measure $\mathrm{d}\mu_{2D}(\overline{\mathbf{x}})$ denotes *length* integration for vertices at curves and *area* integration for vertices at media; and $\mathrm{d}\mu(\overline{\mathbf{\Delta t}})$ denotes temporal integration at each path vertex.

The measurement contribution function $\rho_j(\overline{\mathbf{x}}, \overline{\mathbf{\Delta t}})$ represents the time-dependent contribution of a path $\overline{\mathbf{x}}$ as

$$\rho_j\left(\overline{\mathbf{x}}, \overline{\mathbf{\Delta t}}\right) = L_e(\mathbf{x}_0, \mathbf{x}_1, \mathbf{\Delta} t_0) G_{2D}(\mathbf{x}_0, \mathbf{x}_1) \cdot \prod_{i=1}^{m-1} f(\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1}, \mathbf{\Delta} t_i) G_{2D}(\mathbf{x}_i, \mathbf{x}_{i+1}) \cdot W_e^{(j)}(\mathbf{x}_{m-1}, \mathbf{x}_m, \mathbf{\Delta} t_m), \tag{6}$$

where $L_e(\mathbf{x}_0, \mathbf{x}_1, \mathbf{\Delta} t_0)$ is emitted radiance towards $\mathbf{x}_1$ at time $\mathbf{\Delta} t_0$, $G_{2D}(\mathbf{x}_i, \mathbf{x}_{i+1})$ is the geometric term modeling attenuation between any pair of points on a curve or in media, $f(\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1}, \mathbf{\Delta} t_i)$ is the scattering coefficient at $\mathbf{x}_i$ after a time delay $\mathbf{\Delta} t_i$, and $W_e^{(j)}(\mathbf{x}_{m-1}, \mathbf{x}_m, \mathbf{\Delta} t_m)$ is the temporal sensor importance at pixel j for a light path ending at $\mathbf{x}_m$ with time of flight $\mathbf{\Delta} t_m$. In this work, we choose scattering coefficients according to three BRDFs: perfectly diffuse ($f = 1/2$), perfectly specular ($f = 1$ for the specular direction, 0 otherwise), and a logistic-based microfacet model with microfacet distribution $D(\theta) = \frac{1}{4s}\coth(\frac{\pi}{4s})\mathrm{sech}^2(\frac{\theta}{2s})$, where s is a roughness parameter [29]. An important consideration in 2D is that light interacts with 1D boundaries (curves), and scattering at curves $f(\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1}, \mathbf{\Delta} t_i)$ takes place in a semicircular domain (Fig. 3, left), in contrast to hemispherical scattering at surfaces in 3D scenes (Fig. 3, right). As a consequence, the 2D geometric term $G_{2D}(\mathbf{x}_i, \mathbf{x}_{i+1})$ represents the transformation between differential planar angle $\mathrm{d}\theta$ and differential length $\mathrm{d}l$, involving inverse falloff instead of the usual inverse-squared falloff in 3D [28], as

$$G_{2D}(\mathbf{x}_i, \mathbf{x}_{i+1}) = \frac{\cos\theta_{\mathbf{x}_i} \cos\theta_{\mathbf{x}_{i+1}}}{\|\mathbf{x}_i - \mathbf{x}_{i+1}\|}, \tag{7}$$

where $\theta_{\mathbf{x}_i}$ represents the angle between the normal at $\mathbf{x}_i$ and light exiting that point, and $\theta_{\mathbf{x}_{i+1}}$ represents the angle between the normal at $\mathbf{x}_{i+1}$ and incoming light from $\mathbf{x}_i$, as illustrated in Fig. 2.

B. 2D Impulse Response Function

The spatio-temporal delta emission of an emitter at $\mathbf{x}_0$ and the spatio-temporal delta response function of a sensor at $\mathbf{x}_m$ can be formally expressed as $L_e(\mathbf{x}_0, \mathbf{x}_1, \mathbf{\Delta} t_0) = \delta(\mathbf{x}_0 - \mathbf{x}_1, t - \mathbf{\Delta} t_0)$ and $W_e(\mathbf{x}_{m-1}, \mathbf{x}_m, \mathbf{\Delta} t_m) = \delta(\mathbf{x}_m - \mathbf{x}_{m-1}, t - \mathbf{\Delta} t_m)$, respectively. Substituting these two terms in the measurement contribution function ((6)) and restricting possible paths $\overline{\mathbf{x}} = \mathbf{x}_0\mathbf{x}_1 \ldots \mathbf{x}_{m-1}\mathbf{x}_m$ to a subset $\Psi_{2D}^{(l,s)}$ of the entire path space Ψ_{2D}, where $\mathbf{x}_0 \equiv \mathbf{x}_{l0}, \mathbf{x}_1 \equiv \mathbf{x}_l, \mathbf{x}_{m-1} \equiv \mathbf{x}_s, \mathbf{x}_m \equiv \mathbf{x}_{s0}$, results in the following expression for the 2D impulse response function:

$$H(\mathbf{x}_l, \mathbf{x}_s, t) = \int_{\Psi_{2D}^{(l,s)}} f(\overline{\mathbf{x}}) G_{2D}(\overline{\mathbf{x}}) \delta(\mathbf{x}_0 - \mathbf{x}_1, t - \Delta t_0) \, \delta(\mathbf{x}_m - \mathbf{x}_{m-1}, t - \Delta t_m) \, \mathrm{d}\mu_{2D}(\overline{\mathbf{x}}), \tag{8}$$

where $f(\overline{\mathbf{x}}) = \prod_{i=1}^{m-1} f(\mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1}, \mathbf{\Delta} t_i)$ and $G_{2D}(\overline{\mathbf{x}}) = \prod_{i=0}^{m-1} G_{2D}(\mathbf{x}_i, \mathbf{x}_{i+1})$. In our pipeline, we approximate the impulse response function under a discrete set of laser and sensor targets, as well as temporal bins, and use it as input for different NLOS imaging camera models.

V. 2D NLOS Imaging Pipeline

Here we describe how to couple 2D transient light transport simulation to obtain $\mathbf{H}$ with phasor-based NLOS imaging camera models and solvers. In the supplementary material, Section S.II, we provide additional details on our implementation.

A. Non-Line-of-Sight Light Tracing

We build upon our 2D transient path integral formulation (Section IV) to add support for transient light transport simulation to the Monte-Carlo-based 2D steady-state light tracer Tantalum [29]. We approximate the impulse response function (8) as the discretized tensor $\mathbf{H}$ by aiming the laser source and sensor pixel at different targets $\mathbf{x}_l, \mathbf{x}_s$ on a visible relay segment. We model the ultra-fast laser as a collimated beam of light perfectly targeted at $\mathbf{x}_l$. In 2D, we model the sensor to integrate outgoing radiance on a finite length surrounding $\mathbf{x}_s$, mimicking the area measured by a real sensor pixel focused at $\mathbf{x}_s$ in 3D. We point the laser at a fixed location $\mathbf{x}_l$ while measuring a set of sensor targets $\mathbf{x}_s$ on the relay segment, reusing sampled light paths for all sensor pixels; this mimics the use of SPAD arrays for efficient capture of indirect photons in real non-confocal setups [31], [34], [49], [50]. For confocal setups, we co-locate $\mathbf{x}_l$ and $\mathbf{x}_s$, which is analogous to the use of single-pixel SPADs [11], [12]. Real confocal setups usually mitigate sensor saturation and low SNR due to first-bounce photons by introducing a slight offset between $\mathbf{x}_l$ and $\mathbf{x}_s$ [2]. Instead of implementing this offset, our pipeline allows to selectively discard illumination samples when computing $\mathbf{H}$ based on the number of bounces traversed by the corresponding light paths.

We incorporate support for several capture parameters relevant to NLOS imaging analysis: The laser and sensor setup greatly affects the ability of NLOS imaging methods to reconstruct specific scene features due to the missing cone problem [30], [31] and due to resolution constraints determined

by the extent and density of captured sensor points [9], [11]. Therefore, the location of both devices $\mathbf{x}_{l0}, \mathbf{x}_{s0}$, the laser target $\mathbf{x}_l$, and the spatial extent and number N_s of target sensor points $\mathbf{x}_s$ on the relay segment is customizable by the user. Exhaustive capture setups [49] can be potentially implemented by simulating non-confocal captures at $N_l > 1$ laser targets $\mathbf{x}_l$. Ultra-fast SPAD sensors have limited temporal resolution Δt_e and range T_{max}. The former introduces a lower bound for reconstruction resolution, while the latter constrains $H(\mathbf{x}_l, \mathbf{x}_s, t)$ to light paths with a maximum optical distance $T_{\text{max}}c$. To mimic those limitations, our system permits modifying the width of temporal bins (Δt_e) and the maximum temporal range measured (T_{max}). Our system contains a set of predefined scenes and a semi-automated tool to create or load new ones.

B. NLOS Imaging Camera Models and Solvers

We adapt three different NLOS imaging camera models derived from the phasor-field formulation [9] (4) to the 2D domain. We implement solvers for such models based on backprojection algorithms. Since $\mathbf{H}$ is discretized across the sensor and laser domains, the derived NLOS imaging camera models can be numerically solved through discrete summations. We represent the discretized versions of $H(\mathbf{x}_l, \mathbf{x}_s, t)$ and $f(\mathbf{x}_v, t)$ as $\mathbf{H}[m_l, m_s, m_t]$ and $\mathbf{f}[j_v, j_t]$, respectively, where $m_l = 1..N_l, m_s = 1..N_s$ and $j_v = 1..N_v$ index the discretized spatial domains $\mathbf{x}_l[m_l] \in \mathcal{L}_{N_l}, \mathbf{x}_s[m_s] \in \mathcal{S}_{N_s}, \mathbf{x}_v[j_v] \in \mathcal{V}_{N_v}$, and m_t, j_t index the discretized temporal domain.

Time-gated camera: This NLOS imaging model estimates transient transport of ultra-short pulses of light emitted at $t = 0$ by a single virtual light source at $\tilde{\mathbf{x}}_l \in \mathcal{L}$ and propagating through the hidden scene at $t > 0$ [9]. It is the result of defining virtual illumination as

$$\mathcal{P}(\mathbf{x}_l, t) = \delta(\tilde{\mathbf{x}}_l - \mathbf{x}_l)K(t),$$

where $K(t)$ represents a pulsed function in the temporal domain, such as a delta, a short Gaussian pulse, or a carrier frequency with a Gaussian envelope. Under this virtual illumination, the phasor-based imaging operator from (4) in the temporal domain yields the virtual time-gated camera model f_{tc} as

$$f_{tc}(\mathbf{x}_v, t) = \int_{\mathcal{S}} \frac{1}{d_{sv}} H_l(\tilde{\mathbf{x}}_l, \mathbf{x}_s, t + t_{sv}) \mathrm{d}\mathbf{x}_s, \tag{9}$$

where $t_{sv} = d_{sv}/c$ and

$$H_l(\tilde{\mathbf{x}}_l, \mathbf{x}_s, t) = \mathcal{P}(\mathbf{x}_l, t) *_t H(\mathbf{x}_l, \mathbf{x}_s, t). \tag{10}$$

We implement a backprojection solver for (9) under the discretized $\mathbf{H}_l$ and $\mathbf{f}_{\text{tc}}$ as

$$\mathbf{f}_{\text{tc}}[j_v, j_t] = \sum_{m_s=1}^{N_s} \frac{1}{d_{sv}} \mathbf{H}_l \left[\tilde{m}_l, m_s, m_t^{\text{tc}} + j_t\right], \tag{11}$$

where $\tilde{m}_l$ indexes the laser target $\tilde{\mathbf{x}}_l \in \mathcal{L}$ used as a virtual emitter, and

$$m_t^{\text{tc}} = \left\lceil \frac{\|\mathbf{x}_s[m_s] - \mathbf{x}_v[j_v]\|}{c\Delta t_e} \right\rceil. \tag{12}$$

Confocal camera: This NLOS imaging model estimates transient light transport in the hidden scene produced by virtual emitters $\mathbf{x}_l \in \mathcal{L}$ *focused* at voxels $\mathbf{x}_v$ [9]. It is the result of defining virtual illumination as

$$\widehat{\mathcal{P}}(\mathbf{x}_l, t) = K(t),$$

with a pulsed function $K(t)$. Under this virtual illumination, the phasor-based imaging operator from (4) in the temporal domain yields the virtual confocal camera model f_{cc} as

$$f_{cc}(\mathbf{x}_v, t) = \int_{\mathcal{S}} \frac{1}{d_{sv}} \int_{\mathcal{L}} \frac{1}{d_{lv}} H_K(\mathbf{x}_l, \mathbf{x}_s, t + t_{lv} + t_{sv}) \mathrm{d}\mathbf{x}_l \mathrm{d}\mathbf{x}_s, \tag{13}$$

where $d_{lv} = \|\mathbf{x}_l - \mathbf{x}_v\|$, $t_{lv} = d_{lv}/c$, and

$$H_K(\mathbf{x}_l, \mathbf{x}_s, t) = \mathcal{P}(\mathbf{x}_l, t) *_t H(\mathbf{x}_l, \mathbf{x}_s, t). \tag{14}$$

We implement a backprojection solver for (13) as

$$\mathbf{f}_{\text{cc}}[j_v, j_t] = \sum_{m_s=1}^{N_s} \frac{1}{d_{sv}} \sum_{m_l=1}^{N_l} \frac{1}{d_{lv}} \mathbf{H}_K[m_l, m_s, m_t^{\text{cc}} + j_t], \tag{15}$$

where

$$m_t^{\text{cc}} = \left\lceil \frac{\|\mathbf{x}_l[m_l] - \mathbf{x}_v[j_v]\| + \|\mathbf{x}_s[m_s] - \mathbf{x}_v[j_v]\|}{c\Delta t_e} \right\rceil \tag{16}$$

Evaluating the confocal camera at $t = 0$ ($j_t = 1$) shows an approximation of the hidden geometry based on third-bounce illumination [9], which is the basis of the majority of NLOS imaging methods.

Steady-state camera: We model the behavior of a conventional steady-state camera $f_{sc}(\mathbf{x}_v)$ with an exposure time Δt_{sc} much larger than the ultra-fast exposure time of time-gated cameras $\Delta t_e \ll \Delta t_{sc}$. In practice, $f_{sc}(\mathbf{x}_v)$ is the result of integrating the time-gated camera ((9)) along the temporal domain

$$f_{sc}(\mathbf{x}_v) = \int f_{tc}(\mathbf{x}_v, t) dt,$$

which we solve numerically as a discrete summation

$$\mathbf{f}_{sc}[j_v] = \sum_{j_t=1}^{N_t} \mathbf{f}_{tc}[j_v, j_t].$$

Liu et al. [9] showcased a specific implementation of a steady-state camera under monochromatic virtual illumination, which they called a photography camera. Our pipeline supports a more general steady-state camera not constrained to a single frequency. Liu et al.'s photography camera can be implemented in our system by filtering H by a delta function in the frequency domain.

Applying a virtual illumination function $\mathcal{P}(\mathbf{x}_l, t)$ is equivalent to filtering the impulse response function H of the scene. A filtering step is necessary to remove artifacts and see the actual shape of hidden objects. We implement temporal filtering over $\mathbf{H}$ ((10), (14)) following pulsed virtual illumination functions implemented in the phasor-field formulation [9], which are defined by a carrier frequency Ω_c—corresponding to a central wavelength $\lambda_c = 1/\Omega_c$—with a Gaussian envelope with standard deviation σ:

$$K(t) = e^{2\pi i \Omega_c t} e^{-t^2/(2\sigma^2)}. \tag{17}$$

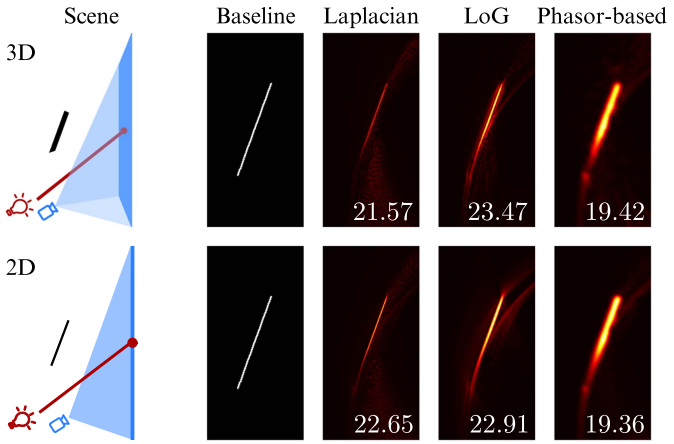

Fig. 4. We create the 3D scene ROTATED PATCH (top row) and the 2D scene ROTATED SEGMENT (bottom row), with equivalent composition. In the bottom-right corner of each image, we show PSNR with respect to the shown baseline. We compare the effect of different filtering strategies, showing matching results and similar PSNR values. Spatial filters based on Laplacian and Laplacian of Gaussian (LoG) functions remove most artifacts but can create new ones, while temporal filtering with phasor-based pulsed virtual illumination provides cleaner images.

We also implement filtering strategies commonly used in NLOS imaging spatially over f, based on Laplacian and Laplacian of Gaussian filters [8], [10]. We define $K(t) = \delta(t)$ and apply discrete 2D convolutions over the spatial domain of $\mathbf{f}_{tc}, \mathbf{f}_{cc}, \mathbf{f}_{sc}$ when evaluated at specific time instants.

The confocal and time-gated camera models ((11), (15)) allow us to obtain multiple images of the hidden scene at different instants. Our system computes a single image at the j_t-th instant, chosen by the user, and allows the user to visualize a sequence of frames by sequentially increasing j_t. The resulting image is updated every time a batch of paths ends its simulation.

VI. RESULTS

2D fast prototyping helps to design scenes and test different configurations (material, geometry, localization, etc), before launching their significantly more costly 3D and real counterparts. In the following, we validate our system by comparing its 2D results with an equivalent 3D experiment, and analyze its performance by exploring hidden object visibility. Unless stated otherwise, all data is simulated with non-confocal capture setups with one illumination point at the center of the relay segment, and images of the hidden scene are obtained with the camera model described by (13), (15), evaluated at $t = 0$ (i.e., $j_t = 1$).

A. Validation

We present two equivalent scenes showing that results in 2D match 3D observations, allowing the study and analysis of visibility issues. Laplacian and Laplacian-of-Gaussian functions are commonly used to spatially filter the result, which removes most artifacts but can create new ones. Temporal filtering of H with a phasor-based pulsed virtual illumination function provides cleaner images. These effects are shown in Fig. 4, with matching results for a pair of equivalent 2D and 3D scenes.

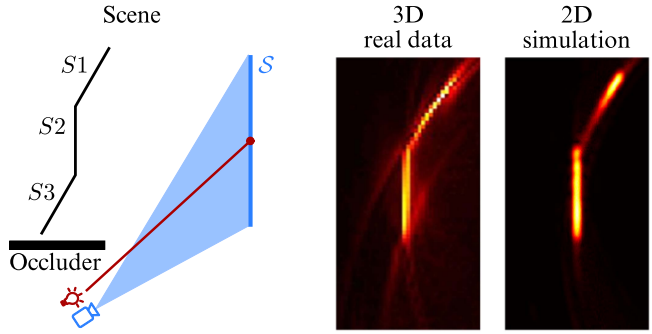

Fig. 5. Scenes THREE PATCHES (3D, top view) and THREE SEGMENTS (2D), with equivalent composition. Liu et al. [30] illustrate the missing cone problem in the THREE PATCHES scene, where it prevents $S3$ from being imaged. We create the equivalent 2D scene THREE SEGMENTS, with matching results.

The 3D scene ROTATED PATCH is composed of a 0.5×0.5 m patch at 0.75 m from the relay wall and rotated $20°$ in the vertical axis. The 2D scene ROTATED SEGMENT is composed of a 0.5 m segment rotated $20°$, facing away from the relay segment. For equivalent comparison, in the 3D scene we image a horizontal cross-section of the scene at the center of the patch. For quantitative analysis, we compute the PSNR for each image (displayed at the bottom of each result) with respect to a baseline consisting of a binary image with 1's at pixels that intersect the geometry, and 0's everywhere else. All imaging results are normalized to the range [0,1]. We can observe 3D and 2D yield equivalent results quantitatively and qualitatively, especially when using pulsed virtual illumination. Minor differences are caused by dimensionality changes in light transport propagation. In subsequent experiments, we always use pulsed virtual illumination.

B. Visibility Analysis

Third-bounce NLOS imaging methods have limited visibility due to the position and orientation of certain objects with respect to the laser and SPAD baseline on the relay segment [30]. This is a pathological issue known as the missing cone problem and it is inherent to several imaging techniques beyond NLOS imaging [51], [52]. We illustrate this behavior in Fig. 5, where we show and replicate Liu et al.'s experiment [30]. Liu et al. capture real data from a THREE PATCHES scene containing three patches $S1, S2, S3$ in different locations and orientations, where $S2$ is parallel to the relay segment and $S1, S3$ are not. $S1$ and $S2$ can be imaged, while $S3$ remains invisible to the imaging process. Our equivalent 2D THREE SEGMENTS scene represents a horizontal slice of THREE PATCHES that replicates this behavior. Note that some artifacts can appear in the image due to triangulation errors of backprojection, such as the vertical edge of energy below $S2$.

Confocal setups increase angular coverage at the cost of longer capture and simulation times. This partially alleviates the effect of the missing cone problem but does not delete it. For instance, when placing the silhouette of a Stanford bunny in front of the relay segment, its ear faces away from the relay segment and will not be imaged with data from a non-confocal

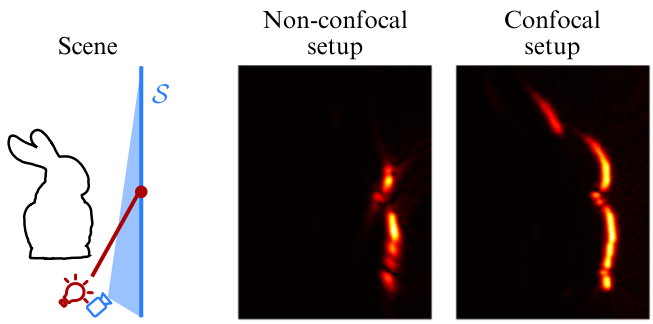

Fig. 6. BUNNY scene using different capture setups. We place the silhouette of a Stanford bunny at 0.5m from the relay segment. Data simulated with a non-confocal setup results in an image with very few details because of its limited angular coverage. A confocal setup improves visibility, showing the front and ear of the bunny in great detail. Some features are still affected by the missing cone problem and remain invisible, such as the bottom of the bunny. In both images, virtual illumination has $\lambda_c = 6cm$ and $\sigma = 3cm$.

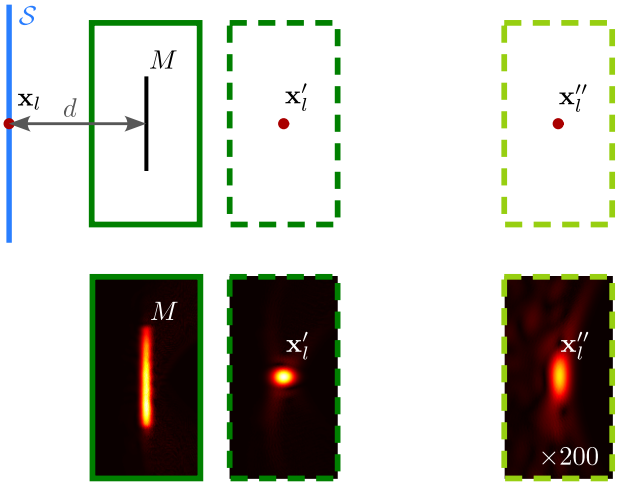

Fig. 7. 2D equivalent of the INFINITY MIRROR experiment by Royo et al. [31], with a hidden segment M placed parallel to $\mathcal{S}$ at a distance d. We apply virtual illumination with $\lambda_c = 4cm$ and $\sigma = 3cm$. The confocal camera model images the geometry of M at a distance d. The mirror-like behavior of $\mathcal{S}$ and M under the chosen virtual illumination produces multiple virtual reflections of $\mathbf{x}_l$. The time-gated camera model images the first virtual reflection $\mathbf{x}'_l$ at distance $2d$ (third-bounce illumination), and the second one $\mathbf{x}''_l$ at distance $4d$ (fifth-bounce illumination).

setup. The increased coverage of a confocal setup obtains more information and the ear appears in the image. Some positions are ill-posed for both capture setups, e.g., the bottom of the bunny remains invisible in both cases. We illustrate this behavior in Fig. 6.

Recent work by Royo et al. [31] demonstrated that diffuse surfaces may exhibit mirror-like behavior under specific virtual illumination functions. Therefore, objects can produce virtual reflections of other elements in the scene. This intuition is shown in Fig. 7, where a hidden segment M at a distance d from the relay segment reflects $\mathbf{x}_l$ producing a virtual reflection $\mathbf{x}'_l$ at a distance $2d$. Fifth-bounce paths of the form $\mathbf{x}_l \rightarrow M \rightarrow \mathcal{S} \rightarrow M \rightarrow \mathcal{S}$ produce a second virtual reflection $\mathbf{x}''_l$ at a distance $4d$. Under specific scene configurations, specular reflections can be used to address the missing cone or image around a second

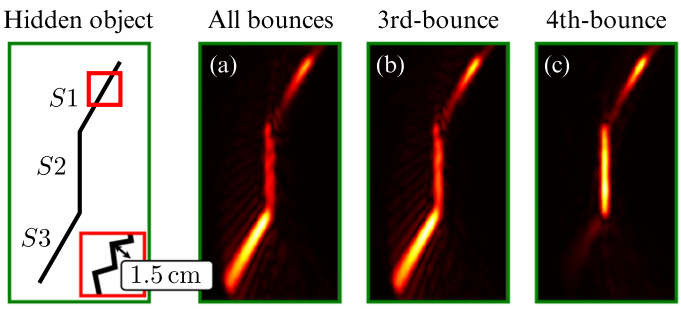

Fig. 8. THREE SEGMENTS scene with facets of 1.5cm. We apply virtual illumination with $\lambda_c = 4cm$ and $\sigma = 3cm$. (a) The presence of facets changes the virtual behavior of the object to more diffuse-like and we can image $S3$, no longer affected by the missing cone problem. To further analyze this effect, we limit simulation to three-bounce paths only (b) and four-bounce paths only (c). Three-bounce paths allow us to image $S3$ because the virtual specular reflection of $\mathbf{x}_l$ with respect to most facets reaches $\mathcal{S}$. Four-bounce paths are produced by interreflection between facets. Since facets are close, delays are small in $S1$ and $S2$, but these paths do not reach $\mathcal{S}$ from $S3$.

corner. For further detail, we redirect the reader to the original publication.

This mirror-like behavior is related to the missing cone problem: objects cannot be imaged when the specular reflection of $\mathbf{x}_l$ does not reach $\mathcal{S}$. According to microfacet theory of surface reflectance [53], [54], [55], objects with larger variations of micrometric surface facets exhibit a more diffuse appearance. Similarly, objects with variations of larger-size curve facets should exhibit diffuse-like behavior under virtual illumination used for NLOS imaging, allowing us to image them. We show and analyze this behavior in Fig. 8, where we modify the THREE SEGMENTS scene so that $S1$, $S2$ and $S3$ have facets of 1.5 cm. Fig. 8 a shows success in imaging the three segments under virtual illumination with $\lambda_c = 4$cm and $\sigma = 3$cm. Each facet acts as a small virtual mirror, and the combination of their orientations produces the diffuse-like behavior of the object. Many facets in $S3$ now face $\mathcal{S}$, making it visible for third-bounce NLOS imaging methods. On the other hand, the rest of the facets in $S3$ face away from $\mathcal{S}$, so interreflection between facets does not reach it. We illustrate this behavior in Fig. 8(b) and (c) by limiting captured illumination to 3rd- and 4th-bounce only, respectively. Using 3rd-bounce illumination, we image all segments, while interreflection in 4th-bounce illumination allows us to see $S1$ and $S2$ but not $S3$.

Object visibility depends on the relationship between facet size and the virtual illumination wavelength. In the previous experiment, these values are close and we obtain a detailed image. Shorter wavelengths can provide finer detail, up to the point of showing each facet individually. However, they tend to enhance high-frequency noise and artifacts. On the other hand, light does not interact with features much smaller than its wavelength. Therefore, when using a virtual illumination wavelength much larger than the facet size, the object will behave analogous to a non-faceted one. We illustrate this behavior in Fig. 9 under different virtual illumination functions and facet sizes. We create the PERPENDICULAR scene, where a hidden segment is placed in a particularly ill-posed orientation: at 90° with respect to the

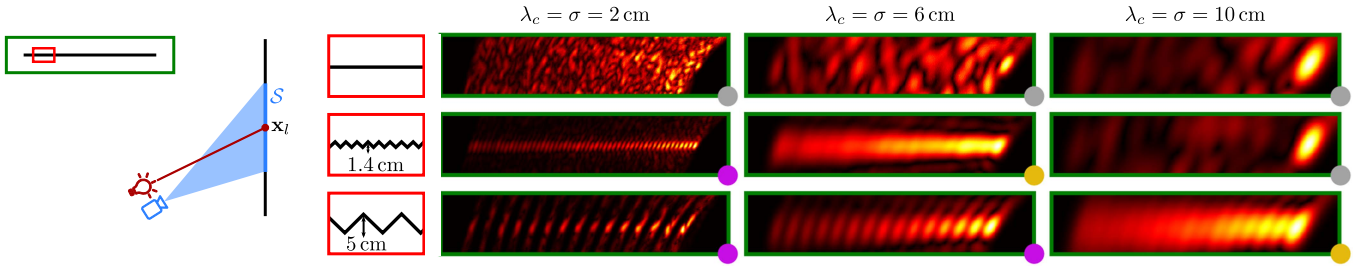

Fig. 9. We analyze object visibility on the PERPENDICULAR scene with facets of different sizes and different virtual illumination functions. A non-faceted object (first row) exhibits mirror-like behavior and cannot be imaged. We can image an object with centimeter-scale facets (second and third row) at different levels of detail depending on the chosen virtual illumination. If its wavelength is much larger than the facet size, the mirror-like behavior holds (gray-marked insets). With virtual illumination wavelengths close to but slightly larger than the facet size, the facets blur together producing a diffuse-like behavior and we can image the global shape of the object (yellow-marked insets). Shorter wavelengths make facets distinguishable as small virtual mirrors but tend to enhance noise (purple-marked insets).

TABLE I
TIME AND MEMORY REQUIREMENTS COMPARISON BETWEEN EQUIVALENT 2D AND 3D SCENES WITH AN EQUIVALENT NUMBER OF SIMULATED SAMPLES

		Simulation			Imaging			Memory		
Scene	Δx	3D	2D (Ours)	Speedup	3D	2D (Ours)	Speedup	3D	2D (Ours)	Saving
BUNNY	256	11.12 s	16 ms	694.9	4.8 min	1.6 ms	1.8×10^5	871 MB	6.56 MB	132.77
BUNNY	512	76.27 s	16 ms	4867.5	46.4 min	2.8 ms	9.9×10^5	3.4 GB	12 MB	283.33

We achieve imaging speedups of up to five orders of magnitude, allowing us to compute and visualize the imaging output while we progressively accumulate simulation samples.

relay segment $\mathcal{S}$, which is a pathological case of the missing cone problem. We use three versions of the scene, where

1) the hidden segment is planar,
2) it has facets of 1.4cm, and
3) it has facets of 5cm.

Gray-marked insets represent cases where the segment remains invisible because of its mirror-like behavior. Yellow-marked insets represent cases where it behaves diffuse-like and all facets blur to show the global shape of the object. Purple-marked insets represent cases where the virtual illumination wavelength is shorter than the feature size or very close to it and allows us to see each facet.

C. Performance

We demonstrate the benefits of our pipeline in terms of computational speedups and total memory requirements by evaluating two pairs of equivalent 3D and 2D NLOS setups simulating non-confocal captures and the NLOS imaging step. The 3D scene consists of a Stanford bunny facing the relay wall, and the 2D scene is the flattened version of the bunny shown in Fig. 6. We evaluate the scenes under non-confocal captures with two different lateral resolutions: 256×256 and 512×512 sensor points in 3D, which we compare to 256 and 512 sensor points in 2D, respectively. We image an equivalent number of voxels in 3D and pixels in 2D, resulting in an output resolution of 512×910 in 2D, and $98 \times 98 \times 49$ in 3D. In Table I we provide a quantitative comparison of the time- and memory-saving benefits of 2D prototyping in our most complex scene, composed of a hidden Stanford bunny. We use $16{\times}10^3$ simulation samples in 2D and an equivalent number of samples in 3D. Combining the simulation step to perform transient rendering, and the NLOS imaging step to obtain a reconstruction of the hidden scene, a 3D hidden scene with a Stanford bunny takes a total of 5 minutes at a coarse resolution ($\Delta x = 256$, first row), and 47 minutes at the highest resolution tested ($\Delta x = 512$, second row). In contrast, our 2D pipeline yields results in real-time (17 ms to 19 ms) regardless of the tested resolution, achieving speedups from four to five orders of magnitude w.r.t. the 3D counterparts. The memory requirements for generated data can reach several gigabytes for 3D scenes, while removing one spatial dimension reduces the size to a dozen megabytes, i.e. two orders of magnitude smaller. NLOS imaging in 2D benefits from this memory-size reduction and the lower value of N_s, as for the same lateral resolution of the capture grid, N_s grows quadratically in 3D, while N_s grows linearly in 2D. Our implementation obtains overwhelming NLOS imaging speedups of up to five to six orders of magnitude compared to 3D, obtaining full images in a few milliseconds. This allows us to couple the NLOS imaging process with progressive simulation of light transport samples, showing imaging results in real-time. This allows the user to modify any parameter and simultaneously visualize changes in the output while the simulation converges to the total sample upper bound defined by the user. Note we use mitransient [56] for transient light transport simulation in 3D, which performs highly optimized vectorized operations, yet our 2D simulator takes shorter times, with improved speedups at higher lateral resolutions.

3D confocal setups require longer simulation times of up to several hours due to the inability to reuse sampled paths for all measured points, as both laser and SPAD positions change synchronously; this time is reduced to just a few minutes in our 2D system.

D. Variations of Curve Reflectance

Conventional NLOS imaging methods assume a diffuse relay segment and diffuse hidden objects. We analyze their performance when one of these assumptions breaks. Our experiments show that the ability to image hidden objects strongly depends

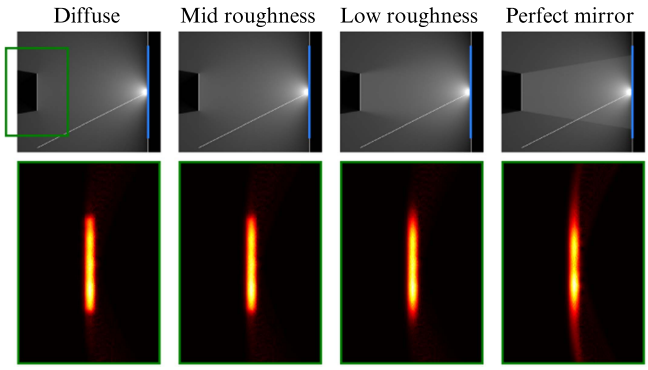

Fig. 10. LINE scene: a hidden segment is placed parallel to the relay segment. We decrease the roughness of the hidden segment (left to right). Breaking the diffuse assumption results in coarser images because each point reflects emitted light to a narrower region of the relay segment. Top row shows steady-state fluence in the scene. Bottom row shows the obtained image under virtual illumination with $\lambda_c = \sigma = 3cm$.

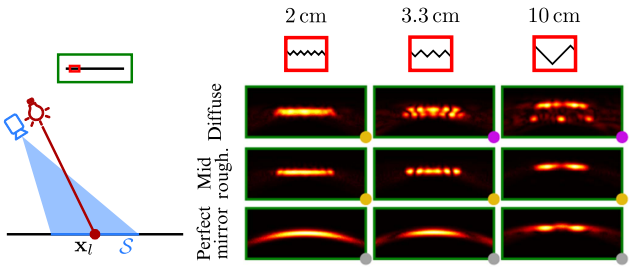

Fig. 11. We modify the roughness (rows) and faceted-structure (columns) of the hidden segment in the LINE scene. We analyze the impact on imaging performance under virtual illumination with $\lambda_c = \sigma = 3cm$. The strong interreflection between facets prevents them from being precisely imaged, but we can see their endpoints when they are diffuse (purple-marked insets). As roughness decreases, the impact of interreflection increases and the closest endpoints become invisible. The object is imaged as a planar segment, similar to the one with facets smaller than the wavelength (yellow-marked insets). Perfectly-specular facets in this scene focus light on a narrow region of $\mathcal{S}$, producing very coarse images (gray-marked insets).

on the directionality of light scattered by both the hidden objects and relay segment. Imaging precision drops when very specular (low-roughness) materials are involved, but it holds well for intermediate ones. Curve points with higher specularity (i.e. lower roughness) reflect light to a narrower region of the relay segment, which reduces the baseline measuring light from such points. This will result in coarser images, especially when the hidden object is very close to a perfect mirror. We illustrate this behavior in Fig. 10 with the LINE scene, where we place a segment parallel to $\mathcal{S}$ with decreasing roughness (left to right) from perfectly diffuse to a perfect mirror. We also show the steady-state fluence visualization of the scene to illustrate the directionality of reflected light. Next, we analyze NLOS imaging performance under joint changes in curve reflectance and faceted structure by modifying the LINE scene and fixing the virtual illumination function to $\lambda_c = \sigma = 3cm$. Note that hidden objects have different properties under emitted light—whether the facets are diffuse, rough or a mirror—and under virtual illumination used for NLOS imaging—whether the object exhibits *virtual* mirror-like or diffuse-like behavior. We show our analysis in Fig. 11. With virtual illumination wavelength close to facet

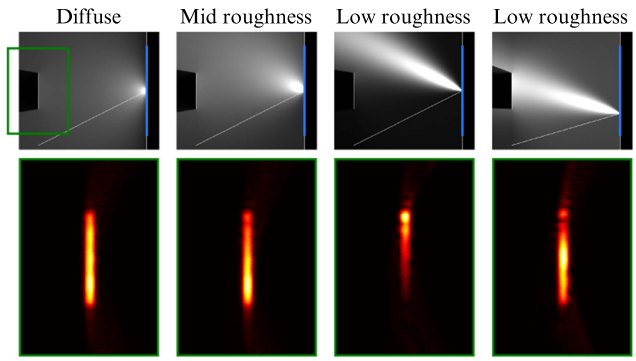

Fig. 12. We modify the roughness of the relay segment in the LINE scene. Imaging performance barely changes between using a diffuse and a close-to-diffuse relay segment (first and second columns). As roughness decreases, specularity increases and most light is reflected towards the specular direction (third and fourth columns), and only geometries close to it can be properly imaged. Top row shows steady-state fluence in the scene. Bottom row shows the obtained image under virtual illumination with $\lambda_c = \sigma = 3cm$.

size (first column), the virtual diffuse-like behavior observed in Figs. 8 and 9 is maintained (yellow-marked insets) except for perfectly or almost-perfectly specular facets (gray-marked inset), where interreflection between the facets focuses all received light on a narrow relay segment baseline, producing a very coarse image. Interreflection in this scene has a stronger influence than in previous scenes. It prevents larger facets from being precisely imaged with third-bounce methods (second and third columns) but diffuse facet endpoints remain visible (purple-marked insets). Imaging resolution decreases with roughness, yielding images of planar-like objects even for large facet sizes (yellow-marked insets). Perfectly specular reflectance yields coarse images in all cases because of the reduced baseline. The reflectance of the relay segment determines how the hidden objects are illuminated. Rough segments do not have a strong specular lobe, and light can reach most points in the hidden scene. If the relay segment has very low roughness, most light is reflected towards or around the specular reflection direction. It will only be possible to image objects placed near this direction since they are the only ones that receive and reflect light. We illustrate this behavior in Fig. 12 using the LINE scene with a non-faceted, diffuse hidden segment and varying the roughness of the relay segment. Imaging performance is very similar for a diffuse relay segment and one with intermediate roughness (first and second columns). As roughness decreases, performance strongly depends on the relative position and orientation of the laser, sensor, and hidden objects (third and fourth columns).

E. Additional Experiments

SPAD sensors often suffer from time jitter and Poisson noise. This impacts imaging quality, making it difficult to match captured and simulated results. In Fig. 13, we show the degradation of the signal and imaging result when incorporating these factors into the simulation, which may help researchers estimate optimal capture times to obtain a reasonable SNR. Exhaustive capture increases angular coverage, improving imaging quality as shown in Fig. 14. It leads to a dramatic increase of the

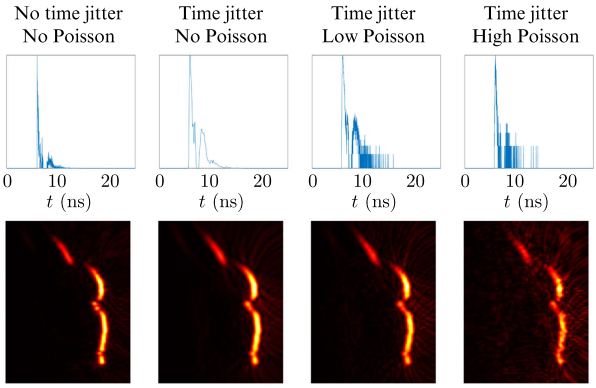

Fig. 13. We apply Gaussian time jitter with FWHM $= 160ps, \sigma = 35ps$, and Poisson noise to a confocal capture on the BUNNY scene. Top row shows the degradation of the signal at one captured point, and bottom row shows the result of the NLOS imaging process. Time jitter barely degrades the result, but Poisson noise significantly hinders NLOS imaging, emphasizing the importance of long captures to reduce noise.

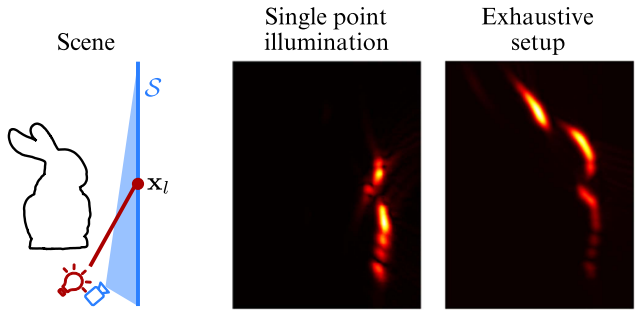

Fig. 14. Comparison of the BUNNY scene illuminating a single point $\mathbf{x}_l$ and using an exhaustive capture setup, where $\mathcal{L} = \mathcal{S}$ and $N_l = N_s = 128$. From the exhaustive data, we can see the shape of the object with greater detail than when illuminating a single point $\mathbf{x}_l$.

memory requirements, so our 2D pipeline becomes fundamental to analyze performance with high lateral resolutions. We expand the discussion on these topics in the supplementary material, Sections S.III and S.IV.

VII. DISCUSSION

We proposed an efficient framework for coupled NLOS light transport simulation and imaging in self-contained 2D worlds that allows us to perform systematic real-time analysis of NLOS imaging performance in simulated scenes considering radiometric changes in a lower-dimensional domain. Such a tool is very useful for quickly prototyping laboratory experiments. Researchers can see the result of imaging a 2D version of the experiment in real-time, and see interactively the result of adjusting the shape and orientation of the objects and the imaging parameters. This prevents the costly 3D rendering or data capture of scenes with undesired visibility problems. Simple 3D scenes with no self-occlusions could be decomposed into a series of parallel slices conforming 2D scenes to approximate the final 3D NLOS image as the stack of all 2D NLOS images. However, this method would lose all interreflection between slices, so the result on complex objects may differ greatly. This dimensionality reduction allows us to implement our formulation in an end-to-end WebGL-based pipeline that couples physically-based transient rendering in 2D with efficient backprojection solvers for different phasor-based NLOS imaging camera models. Our system provides flexible control over multiple scene and NLOS imaging parameters, which we leverage to validate and analyze NLOS imaging performance in various scenes thoroughly. We showed that our results match existing 3D NLOS imaging experiments [8], [9], [30], including NLOS imaging of higher-order illumination bounces [31], while providing dramatic computational speedups. Our analysis provides novel insights on NLOS imaging performance in new scene configurations that diverge from classic Lambertian scenes with locally-planar objects. We show the impact of faceted objects to mitigate the missing cone problem, how decreasing roughness of hidden objects and the relay segment hinders the imaging process, and how imaging wavelengths and baselines are strong determiners of imaging quality.

We constrain our analysis to opaque objects with planar and regular faceted structures within media-free scenes. An interesting avenue of future work would be extending our formulation to incorporate simulation of 2D transient radiative transfer [47], [57], [58] for efficient analysis of NLOS imaging in participating media, as well as analyzing more diverse geometries such as objects with non-regular faceted structures, or transparent objects with varying refractive indices. Additional light properties, such as spectral and polarization information, could be computed by using more intermediate textures. The limitations of WebGL make it difficult to incorporate if the number of wavelengths is large. In particular, it is interesting to explore adding the polarization of light, as recent work has shown its benefits for active and passive NLOS imaging [59], [60], [61]. Our light transport simulation is based on Monte Carlo light tracing. The incorporation of participating media and translucent objects could benefit from rendering techniques based on density estimation accounting for radiometric considerations in 2D [28]. Our pipeline is based on backprojection solvers for the phasor-field formulation for NLOS imaging, which yield real-time imaging rates in 2D. Recent wave-based NLOS imaging works provide significant computational speedups by leveraging efficient operators from classic optics and seismology [9], [12], which would be interesting to migrate and incorporate into our 2D pipeline.

ACKNOWLEDGMENT

We would like to thank Diego Royo for his help generating the 2D exhaustive capture and Daniel Subias for his help with figures.

REFERENCES

[1] A. Jarabo, B. Masia, J. Marco, and D. Gutierrez, "Recent advances in transient imaging: A computer graphics and vision perspective," *Vis. Inform.*, vol. 1, no. 1, pp. 65–79, 2017.
[2] T. Maeda, G. Satat, T. Swedish, L. Sinha, and R. Raskar, "Recent advances in imaging around corners," 2019, *arXiv:1910.05613*.
[3] D. Faccio, A. Velten, and G. Wetzstein, "Non-line-of-sight imaging," *Nature Rev. Phys.*, vol. 2, no. 6, pp. 318–327, 2020.

[4] J. Klein, C. Peters, J. Martín, M. Laurenzis, and M. B. Hullin, "Tracking objects outside the line of sight using 2D intensity images," *Sci. Rep.*, vol. 6, 2016, Art. no. 32491.

[5] K. L. Bouman et al., "Turning corners into cameras: Principles and methods," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2270–2278.

[6] C. Saunders, J. Murray-Bruce, and V. K. Goyal, "Computational periscopy with an ordinary digital camera," *Nature*, vol. 565, no. 7740, pp. 472–475, 2019.

[7] A. Kirmani, T. Hutchison, J. Davis, and R. Raskar, "Looking around the corner using transient imaging," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, 2009, pp. 159–166.

[8] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," *Nature Commun.*, vol. 3, no. 1, Mar. 2012, Art. no. 745.

[9] X. Liu et al., "Non-line-of-sight imaging using phasor-field virtual wave optics," *Nature*, vol. 572, no. 7771, pp. 620–623, Aug. 2019, doi: 10.1038/s41586-019-1461-3.

[10] M. Buttafava, J. Zeman, A. Tosi, K. Eliceiri, and A. Velten, "Non-line-of-sight imaging using a time-gated single photon avalanche diode," *Opt. Exp.*, vol. 23, no. 16, pp. 20997–21011, 2015.

[11] M. O'Toole, D. B. Lindell, and G. Wetzstein, "Confocal non-line-of-sight imaging based on the light-cone transform," *Nature*, vol. 555, no. 7696, pp. 338–341, Mar. 2018.

[12] D. B. Lindell, G. Wetzstein, and M. O'Toole, "Wave-based non-line-of-sight imaging using fast f-k migration," *ACM Trans. Graph.*, vol. 38, no. 4, pp. 1–13, Jul. 2019.

[13] S. A. Reza, M. L. Manna, S. Bauer, and A. Velten, "Phasor field waves: Experimental demonstrations of wave-like properties," *Opt. Exp.*, vol. 27, no. 22, pp. 32587–32608, Oct. 2019.

[14] S. A. Reza, M. L. Manna, S. Bauer, and A. Velten, "Phasor field waves: A huygens-like light transport model for non-line-of-sight imaging applications," *Opt. Exp.*, vol. 27, no. 20, pp. 29380–29400, Sep. 2019.

[15] J. A. Teichman, "Phasor field waves: A mathematical treatment," *Opt. Exp.*, vol. 27, no. 20, pp. 27500–27506, Sep. 2019.

[16] J. Dove and J. H. Shapiro, "Paraxial theory of phasor-field imaging," *Opt. Exp.*, vol. 27, no. 13, pp. 18016–18037, Jun. 2019.

[17] J. Dove and J. H. Shapiro, "Paraxial phasor-field physical optics," *Opt. Exp.*, vol. 28, no. 14, pp. 21095–21109, Jul. 2020.

[18] X. Liu, S. Bauer, and A. Velten, "Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems," *Nature Commun.*, vol. 11, no. 1, Apr. 2020, Art. no. 1645.

[19] Z. Liao, D. Jiang, X. Liu, A. Velten, Y. Ha, and X. Lou, "FPGA accelerator for real-time non-line-of-sight imaging," *IEEE Trans. Circuits Syst. I: Reg. Papers*, vol. 69, no. 2, pp. 721–734, Feb. 2022.

[20] J. E. Farrell, F. Xiao, P. B. Catrysse, and B. A. Wandell, "A simulation tool for evaluating digital camera image quality," *Proc. SPIE*, vol. 5294, pp. 124–131, 2003.

[21] J. E. Farrell and B. A. Wandell, "Image systems simulation," *Handbook Digit. Imag.*, vol. 1, pp. 373–400, 2015.

[22] A. Jarabo, J. Marco, A. Muñoz, R. Buisan, W. Jarosz, and D. Gutierrez, "A framework for transient rendering," *ACM Trans. Graph.*, vol. 33, no. 6, pp. 1–10, Nov. 2014.

[23] A. Pediredla, A. Veeraraghavan, and I. Gkioulekas, "Ellipsoidal path connections for time-gated rendering," *ACM Trans. Graph.*, vol. 38, no. 4, pp. 1–12, 2019.

[24] D. Royo, J. García, A. Muñoz, and A. Jarabo, "Non-line-of-sight transient rendering," *Comput. Graph.*, vol. 107, pp. 84–92, 2022.

[25] P. S. Heckhert, "Radiosity in flatland," *Comput. Graph. Forum*, vol. 11, pp. 181–192, 1992.

[26] F. Durand, N. Holzschuch, C. Soler, E. Chan, and F. X. Sillion, "A frequency analysis of light transport," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 1115–1126, 2005.

[27] R. Ramamoorthi, D. Mahajan, and P. Belhumeur, "A first-order analysis of lighting, shading, and shadows," *ACM Trans. Graph.*, vol. 26, no. 1, Jan. 2007, Art. no. 2–es.

[28] W. Jarosz, V. Schönefeld, L. Kobbelt, and H. W. Jensen, "Theory, analysis and applications of 2d global illumination," *ACM Trans. Graph.*, vol. 31, no. 5, Aug. 2012, Art. no. 125.

[29] B. Bitterli, "The secret life of photons," 2015. [Online]. Available: https://benedikt-bitterli.me/tantalum/

[30] X. Liu, S. Bauer, and A. Velten, "Analysis of feature visibility in non-line-of-sight measurements," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 10140–10148.

[31] D. Royo et al., "Virtual mirrors: Non-line-of-sight imaging beyond the third bounce," *ACM Trans. Graph.*, vol. 42, no. 4, pp. 1–15, 2023.

[32] V. Arellano, D. Gutierrez, and A. Jarabo, "Fast back-projection for non-line of sight reconstruction," *Opt. Exp.*, vol. 25, no. 10, pp. 11574–11583, 2017.

[33] J. Ye, Y. Hong, X. Su, X. Yuan, and F. Xu, "Plug-and-play algorithms for dynamic non-line-of-sight imaging," *ACM Trans. Graph.*, vol. 43, no. 5, Jun. 2024, Art. no. 155.

[34] J. H. Nam et al., "Low-latency time-of-flight non-line-of-sight imaging at 5 frames per second," *Nature Commun.*, vol. 12, no. 1, 2021, Art. no. 6526.

[35] I. Guillén, X. Liu, A. Velten, D. Gutierrez, and A. Jarabo, "On the effect of reflectance on phasor field non-line-of-sight imaging," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2020, pp. 9269–9273.

[36] Y. Yang, K. Yang, and A. Zhang, "Influence of target surface BRDF on non-line-of-sight imaging," *J. Imag.*, vol. 10, no. 11, 2024, Art. no. 273.

[37] A. Smith, J. Skorupski, and J. Davis, "Transient rendering," School of Engineering, University of California, Santa Cruz, CA, USA, Tech. Rep. UCSC-SOE-08-26, 2008.

[38] A. Jarabo, "Femto-photography: Visualizing light in motion," master's thesis, Universidad de Zaragoza, Zaragoza, Spain, 2012.

[39] A. Adam, C. Dann, O. Yair, S. Mazor, and S. Nowozin, "Bayesian time-of-flight for realtime shape, illumination and albedo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 5, pp. 851–864, May 2017.

[40] M. Keller and A. Kolb, "Real-time simulation of time-of-flight sensors," *Simul. Modelling Pract. Theory*, vol. 17, no. 5, pp. 967–978, 2009.

[41] J. Iseringhausen and M. B. Hullin, "Non-line-of-sight reconstruction using efficient transient rendering," *ACM Trans. Graph.*, vol. 39, no. 1, pp. 1–14, Jan. 2020.

[42] C.-Y. Tsai, A. C. Sankaranarayanan, and I. Gkioulekas, "Beyond volumetric albedo a surface optimization framework for non-line-of-sight imaging," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2019, pp. 1545–1555.

[43] M. E. Scott, "Zen photon garden," 2013. [Online]. Available: https://zenphoton.com/

[44] Y.-T. Tu and W.-F. Sun, "Ray optics simulation," 2016. [Online]. Available: https://phydemo.app/ray-optics/

[45] S. Hilton, "Refractorium," 2017. [Online]. Available: https://samreay.github.io/Refractorium/

[46] B. Bitterli, "Virtual femto photography," 2016. [Online]. Available: https://benedikt-bitterli.me/femto.html

[47] J. Marco, "Transient light transport in participating media," M.S. thesis, Universidad de Zaragoza, Zaragoza, Spain, 2013.

[48] Y. Liu, S. Jiao, and W. Jarosz, "Temporally sliced photon primitives for time-of-flight rendering," *Comput. Graph. Forum*, vol. 41, pp. 29–40, 2022.

[49] J. Marco et al., "Virtual light transport matrices for non-line-of-sight imaging," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, Oct. 2021, pp. 2440–2449.

[50] M. Renna, J. H. Nam, M. Buttafava, F. Villa, A. Velten, and A. Tosi, "Fast-gated 16×1 SPAD array for non-line-of-sight imaging applications," *Instruments*, vol. 4, no. 2, 2020, Art. no. 14.

[51] A. H. Delaney and Y. Bresler, "Globally convergent edge-preserving regularized reconstruction: An application to limited-angle tomography," *IEEE Trans. Image Process.*, vol. 7, no. 2, pp. 204–221, Feb. 1998.

[52] J.W. Lim et al., "Comparative study of iterative reconstruction algorithms for missing cone problems in optical diffraction tomography," *Opt. Exp.*, vol. 23, no. 13, pp. 16933–16948, Jun. 2015.

[53] K. E. Torrance and E. M. Sparrow, "Theory for off-specular reflection from roughened surfaces," *Josa*, vol. 57, no. 9, pp. 1105–1114, 1967.

[54] P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces*. Norwood, MA, USA: Artech House, Inc., 1987.

[55] M. Oren and S. K. Nayar, "Generalization of Lambert's reflectance model," in *Proc. 21st Annu. Conf. Comput. Graph. Interactive Techn.*, 1994, pp. 239–246.

[56] D. Royo, M. Crespo, and J. Garcia-Pueyo, "Mitransient," 2023. [Online]. Available: https://github.com/diegoroyo/mitransient

[57] A. Liemert and A. Kienle, "Radiative transfer in two-dimensional infinitely extended scattering media," *J. Phys. A: Math. Theor.*, vol. 44, no. 50, 2011, Art. no. 505206.

[58] M. Machida, "The radiative transport equation in flatland with separation of variables," *J. Math. Phys.*, vol. 57, no. 7, 2016, Art. no. 073301.

[59] O. Pueyo-Ciutad et al., "Time-gated polarization for active non-line-of-sight imaging," in *Proc. ACM SIGGRAPH Asia Conf.*, 2024, pp. 1–11.

[60] K. Tanaka, Y. Mukaigawa, and A. Kadambi, "Polarized non-line-of-sight imaging," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 2136–2145.

[61] Z. Wang et al., "Vectorial-optics-enabled multi-view non-line-of-sight imaging with high signal-to-noise ratio," *Laser Photon. Rev.*, vol. 18, 2024, Art. no. 2300909.