Cascaded Non-Line-of-Sight Imaging

DIEGO ROYO*✉, Universidad de Zaragoza–I3A, Spain
MARÍA PEÑA*, Universidad de Zaragoza–I3A, Spain
FORREST B. PETERSON, University of Wisconsin–Madison, United States of America
ANDREAS VELTEN, University of Wisconsin–Madison, United States of America
JULIO MARCO, Universidad de Zaragoza–I3A, Spain
DIEGO GUTIÉRREZ, Universidad de Zaragoza–I3A, Spain

Time-of-flight non-line-of-sight (NLOS) imaging recovers information from hidden objects by analyzing the time of flight of indirect photons scattered on a visible (relay) wall. Most methods make the simplifying assumption that photons travel exclusively three-bounce paths, thus ignoring other useful information encoded in higher-order photons (with, e.g., four- or five-bounce paths). We present a novel *cascaded* NLOS imaging approach that leverages higher-order information and allows imaging a broader range of single- and multi-corner scenarios. We combine ultra-fast laser scanning with recent time-gated 2D sensor arrays to capture the scene's impulse response on a visible relay wall. From the captured impulse response, our method computes an analogous *virtual* impulse response at any other hidden wall. This effectively allows us to concatenate a second, virtual NLOS imaging system that leverages higher-order illumination. We validate our cascaded imaging method both in simulation and with a real prototype, demonstrating NLOS imaging with fourth- and fifth-bounce illumination of objects in challenging orientations and hidden around two corners. We also analyze how wave-based NLOS imaging interacts with rough hidden walls, which explains and helps overcome existing visibility limitations. We further illustrate how to image hidden objects from different perspectives, thus observing previously unseen features, by relying on multiple hidden walls.

CCS Concepts: • **Computing methodologies → Computational photography**; **3D imaging**.

Additional Key Words and Phrases: non-line-of-sight imaging, time-of-flight imaging, computational imaging

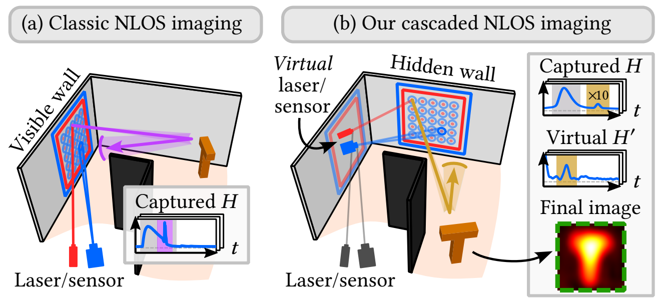


Fig. 1. (a) Classic time-of-flight NLOS imaging has shown how to image objects hidden around one corner, using time-resolved measurements of third-bounce indirect light (purple) from an ultra-fast laser (red) and sensor (blue). (b) Our *cascaded NLOS* imaging method allows imaging objects hidden around two corners by concatenating a second, *virtual* NLOS imaging system and without the need to capture any additional data, effectively using higher-order illumination (e.g., fifth-bounce light in yellow). Our work enables NLOS imaging over a wider, more general set of configurations, helping overcome visibility issues suffered by previous methods.

1 Introduction

Non-line-of-sight (NLOS) imaging methods are able to image objects hidden around a corner by capturing and analyzing indirect light on a visible relay wall. Such NLOS imaging methods take advantage of ultra-fast sensors, able to measure the time of flight of single photons at picosecond resolution (e.g., Liu et al. [2019b]; O'Toole et al. [2018]). This has led to a wide array of unprecedented imaging capabilities, with many potential applications in fields such as autonomous driving, remote sensing or medical imaging [Jarabo et al. 2017; Maeda et al. 2019].

However, most existing methods assume third-bounce-only illumination, where photons interact just once with the hidden scene before reaching the relay wall again (Figure 1a). Ignoring the broad range of higher-order illumination paths (i.e., fourth- and subsequent-bounce photons) that traverse the hidden scene limits NLOS imaging capabilities to single-corner setups with favorable surface orientations [Liu et al. 2019a]. Recent work that exploited higher-order illumination is also restricted to objects in particular locations and orientations, and to paths that follow precise specular directions [Royo et al. 2023]. Thus, in this work, we investigate the following question: can we fully exploit higher-order illumination to extract richer, useful information from hidden scenes?

We introduce a novel *cascaded* NLOS imaging method, showing how higher-order photons can lift some of the main limitations of existing methods. Recent wave-based NLOS systems [Liu et al. 2019b; Marco et al. 2021] introduced an imaging paradigm that interprets time-resolved photons captured on the relay wall as light arriving at a virtual aperture. The relay wall could potentially be transformed into a *virtual* line-of-sight (LOS) imaging system; this is the main intuition behind our cascaded approach.

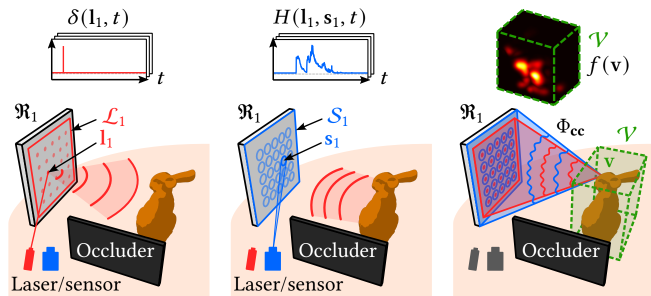
The key idea of cascaded NLOS imaging is to operate on the captured impulse response at the visible relay wall to compute an analogous, *virtual impulse response* at a hidden wall (Figure 1b, left). For this, we introduce two novel virtual LOS imaging tools: our *virtual projector* and *virtual camera* operators, analogous to the ultra-fast laser and sensor from classic NLOS captures. As a result, we concatenate a second virtual NLOS system *from a single capture*

*Both authors contributed equally to this research.

Authors' Contact Information: Diego Royo, droyo@unizar.es, Universidad de Zaragoza–I3A, Spain; María Peña, mpena@unizar.es, Universidad de Zaragoza–I3A, Spain; Forrest B. Peterson, fpeterson2@wisc.edu, University of Wisconsin–Madison, United States of America; Andreas Velten, velten@wisc.edu, University of Wisconsin–Madison, United States of America; Julio Marco, juliom@unizar.es, Universidad de Zaragoza–I3A, Spain; Diego Gutiérrez, diegog@unizar.es, Universidad de Zaragoza–I3A, Spain.

at the visible relay wall, which allows us, for instance, to image objects hidden around *two* corners (Figure 1b, right).

We demonstrate our cascaded imaging method both in simulation and with a real prototype. Our acquisition setup pairs an ultra-fast pulsed laser that scans a single, visible relay wall, with a time-gated, 16×16 2D sensor array [Riccardo et al. 2022]. Ultimately, our cascaded NLOS imaging methods unlock a much broader range of applications. We first employ classic third-bounce imaging to identify potential secondary relay walls within the hidden scene, which we then use to create secondary imaging systems. By leveraging fourth- and fifth-bounce illumination, we showcase NLOS imaging of objects hidden around two corners and objects with limited visibility under existing state-of-the-art methods due to challenging surface orientations. On top of this, we explore how wave-based NLOS imaging methods interact with rough walls, providing novel insights that explain and overcome object visibility limitations. Finally, we show how combining cascaded imaging systems at multiple secondary relay walls increases the visibility of hidden objects, enabling us to image them from different points of view. Code and data for this paper are available at https://graphics.unizar.es/projects/CascadedNLOS/.



(a) Illuminate $\mathbf{l}_1$ with a delta pulse (b) Capture $H(\mathbf{l}_1, \mathbf{s}_1, t)$ at points $\mathbf{s}_1$ (c) Image the hidden scene

Fig. 2. (a) A laser device emits delta light pulses $\delta(\mathbf{l}_1, t)$, illuminating each point $\mathbf{l}_1 \in \mathcal{L}_1$ on the relay wall $\mathfrak{R}_1$. (b) An ultra-fast sensing device captures the time-resolved impulse response $H(\mathbf{l}_1, \mathbf{s}_1, t)$ on multiple visible points $\mathbf{s}_1 \in \mathcal{S}_1$ of the imaging aperture $\mathcal{S}_1$ in $\mathfrak{R}_1$. (c) Imaging operators such as the confocal camera $\Phi_{cc}(\mathbf{v}; H)$ transform $H(\mathbf{l}_1, \mathbf{s}_1, t)$ to obtain an image $f(\mathbf{v})$ of the hidden bunny at points $\mathbf{v}$ on a bounding volume $\mathcal{V}$.

2 Related work

Time-of-flight (ToF) NLOS imaging: Kirmani et al. [2009] originally proposed an NLOS imaging method that emitted ultra-short light pulses towards a relay wall and analyzed the resulting indirect illumination using ToF sensors. This idea was later demonstrated by Velten et al. [2012] using ellipsoidal backprojection of signals captured by a streak camera at picosecond resolution. Single-photon avalanche diodes (SPADs) emerged as a cheaper alternative [Buttafava et al. 2015], leading to a wide variety of ToF NLOS imaging methods inverting the time of flight of indirect photons through physical [Liu et al. 2019b; Pediredla et al. 2019; Pueyo-Ciutad et al. 2024; Tsai et al. 2017; Xin et al. 2019] or deep learning models [Grau Chopite et al. 2020; Li et al. 2023; Shen et al. 2021; Ye et al. 2024]. Some works reduce computational costs using efficient algorithms constrained to confocal acquisition [Lindell et al. 2019; O'Toole et al. 2018], performing non-uniform sampling [Gu et al. 2023; Sultan et al. 2025], or providing efficient implementations [Luesia-Lahoz et al. 2023].

The phasor-field formulation [Liu et al. 2019b] poses NLOS imaging as a wave propagation problem by interpreting the real ToF measurements on the relay wall as *virtual* waves arriving at a wall-sized aperture of a virtual imaging device. This allows exploiting optics-based forward imaging operators to *virtually* illuminate and image the hidden scene [Choi et al. 2023; Dove and Shapiro 2019, 2020; Garcia-Pueyo and Muñoz 2025; Reza et al. 2019; Teichman 2019]. This formulation has fostered many efficient general NLOS imaging methods [Liu et al. 2020; Luesia-Lahoz et al. 2025; Marco et al. 2021; Nam et al. 2020]. Our work builds upon the phasor-field formulation, which we extend to reconstruct time-resolved illumination at hidden surfaces different from the visible relay wall, thus leveraging higher-order illumination instead.

NLOS imaging with higher-order illumination: Most NLOS imaging methods assume that light follows a three-bounce path: from the relay wall to the hidden scene and back. Some exceptions include the work by Yi et al. [2021], which used differentiable rendering to optimize an object's position around two corners using fifth-bounce illumination; Marco et al. [2021] separated fourth-bounce illumination by computing virtual light transport matrices (LTMs); and Sultan et al. [2026] extended the LTM computation to the time domain, showing time-dependent interreflections in the hidden scene. In our work, we go beyond visualization and parameter optimization tasks, enabling cascaded NLOS imaging in challenging scenes.

Wang et al. [2024] rely on polarization to separate third- and fifth-bounce photons. Our work does not require tracking polarization, as it computationally focuses the light to model different bounce paths as needed. Finally, Royo et al. [2023] leverage fourth- and fifth-bounce illumination to image objects with limited visibility and hidden around two corners, although their method requires planar hidden walls and is limited to hidden objects placed along a subset of specular paths. Our cascaded method mitigates these limitations by creating secondary virtual imaging systems. This allows us to leverage a broader collection of specular paths, thus imaging a wider range of scene configurations. Moreover, our method does not require these secondary hidden walls to be planar; on the contrary, rough walls allow us to exploit both specular *and non-specular* paths. Lastly, our cascaded method can combine third-, fourth-, and fifth-bounce information simultaneously.

3 Background on NLOS imaging

Figure 2 illustrates a conventional NLOS imaging setup. A laser individually illuminates points $\mathbf{l}_1 \in \mathcal{L}_1$ on a visible diffuse wall $\mathfrak{R}_1$ (the *relay wall*) using ultra-fast light pulses (Figure 2a). Light travels to the hidden scene and some is reflected back. An ultra-fast camera captures such indirect illumination from the hidden scene at points $\mathbf{s}_1 \in \mathcal{S}_1$ also on the relay wall (Figure 2b). This yields a time-resolved impulse response $H(\mathbf{l}_1, \mathbf{s}_1, t)$, where t denotes the total time of flight from $\mathbf{l}_1$ to $\mathbf{s}_1$. We refer the reader to Table 1 for a table of all symbols used in this paper.

Table 1. Symbols used throughout this paper. A hat $(\hat{\cdot})$ denotes the temporal Fourier transform, with the time t replaced by the frequency Ω.

Symbol	Description
$\mathfrak{R}_n,\ n=1$	The visible relay wall
$\mathfrak{R}_n,\ n>1$	A hidden, secondary relay wall
$\mathbf{l}_a, \mathbf{s}_b$	Illuminated and measured points on $\mathfrak{R}_a$ and $\mathfrak{R}_b$
$\mathcal{L}_a, \mathcal{S}_b$	Illumination and camera apertures on $\mathfrak{R}_a$ and $\mathfrak{R}_b$
$H(\mathbf{l}_1, \mathbf{s}_1, t)$	*Captured* impulse response
$H'_{a,b}(\mathbf{l}_a, \mathbf{s}_b, t)$	Our *virtual* impulse response
$K(t; \lambda_c, \sigma)$	Filter function with wavelength λ_c and pulse width σ
$\hat{R}(t, \Omega)$	RSD operator
$\Phi_{\text{bp}}(\mathbf{v}; H)$, $\Phi_{\text{cc}}(\mathbf{v}; H)$	Third-bounce imaging operators: backprojection and confocal camera
$\Phi_{\text{tp}}(\mathbf{l}_a, t; H)$, $\Phi_{\text{tc}}(\mathbf{s}_b, t; H)$	Our transient projector and transient camera operators
$\mathbf{v},\ \mathcal{V}$	Imaged point and volume in the hidden scene
$f_{a,b}(\mathbf{v})$, $f_{\text{all}}(\mathbf{v})$	Imaging result using $H_{a,b}$ or $H'_{a,b}$, or a combination of multiple imaging results

Imaging algorithms then define imaging operators $\Phi(\mathbf{v}; H)$ that transform $H(\mathbf{l}_1, \mathbf{s}_1, t)$ to obtain an image $f(\mathbf{v})$ of the hidden geometry at points $\mathbf{v}$ in a bounding volume $\mathcal{V}$ (Figure 2c) as:

$$f(\mathbf{v}) = \Phi(\mathbf{v}; H). \tag{1}$$

The commonly used backprojection method Φ_{bp} [Arellano et al. 2017; Velten et al. 2012] operates by temporally shifting and adding the impulse response:

$$\Phi_{\text{bp}}(\mathbf{v}; H) = \int_{\mathcal{S}_1} \int_{\mathcal{L}_1} H(\mathbf{l}_1, \mathbf{s}_1, t_{\mathbf{l}_1 \to \mathbf{v}} + t_{\mathbf{v} \to \mathbf{s}_1})\, \mathrm{d}\mathbf{l}_1 \mathrm{d}\mathbf{s}_1, \tag{2}$$

where $t_{\mathbf{a} \to \mathbf{b}} = |\mathbf{b} - \mathbf{a}|/c$ for any two points $\mathbf{a}$ and $\mathbf{b}$, with c the speed of light. The temporal shift term $t_{\mathbf{l}_1 \to \mathbf{v}} + t_{\mathbf{v} \to \mathbf{s}_1}$ corresponds to the time of flight of the third-bounce path $\mathbf{l}_1 \to \mathbf{v} \to \mathbf{s}_1$. Thus, $f(\mathbf{v}) = \Phi_{\text{bp}}(\mathbf{v}; H)$ provides a good estimation of the third-bounce light reflected by the hidden geometry, and thus of the hidden geometry itself. The majority of time-of-flight NLOS imaging algorithms are based on Equation 2, proposing different formulations and implementations to improve imaging quality and performance. A notable issue arises from multi-path interference (MPI) from higher-order illumination captured in H, e.g., fourth or fifth bounce. Methods that assume third-bounce-only illumination thus suffer from imaging artifacts in $f(\mathbf{v})$ due to MPI. A common strategy to improve imaging quality is to filter $f(\mathbf{v})$ [Velten et al. 2012] or $H(\mathbf{l}_1, \mathbf{s}_1, t)$ [Liu et al. 2019b; O'Toole et al. 2018] so as to enhance geometric features and mitigate imaging artifacts caused by noise.

Phasor fields. In particular, the phasor-field framework [Liu et al. 2019b] proposes using a band-pass filter with central wavelength λ_c and a standard deviation σ, denoted by $K(t; \lambda_c, \sigma)$:

$$K(t; \lambda_c, \sigma) = e^{i2\pi \frac{ct}{\lambda_c} - \frac{1}{2}\left(\frac{ct}{\sigma}\right)^2}, \tag{3}$$

which is applied to the impulse response function via temporal convolution $K(t; \lambda_c, \sigma) *_t H(\mathbf{l}_1, \mathbf{s}_1, t)$. In our work, we apply this filtering strategy to all impulse response functions $H(\mathbf{l}_1, \mathbf{s}_1, t)$. For notation clarity, we do not explicitly formulate the filtering step throughout the equations, and use H to denote already filtered impulse response functions. We will specify the parameters λ_c, σ used in each experiment to denote when and how filtering is done.

The phasor-field framework implements Equation 2 in the frequency domain using wave propagation operators, which allows bringing well-known wave-based LOS tools to the NLOS domain. Most relevant to this paper is their *confocal* camera operator Φ_{cc}; while Φ_{cc} is analogous to Φ_{bp}, its significance lies in its frequency-domain version $\widehat{\Phi}_{\text{cc}}$, defined for each imaging frequency Ω as:

$$\widehat{\Phi}_{\text{cc}}(\mathbf{v}, \Omega; \widehat{H}) = \int_{\mathcal{S}_1} \widehat{R}(t_{\mathbf{v} \to \mathbf{s}_1}, \Omega) \int_{\mathcal{L}_1} \widehat{R}(t_{\mathbf{l}_1 \to \mathbf{v}}, \Omega) \widehat{H}(\mathbf{l}_1, \mathbf{s}_1, \Omega)\, \mathrm{d}\mathbf{l}_1 \mathrm{d}\mathbf{s}_1, \tag{4}$$

where $\widehat{H}(\mathbf{l}_1, \mathbf{s}_1, \Omega) = \mathcal{F}\{H(\mathbf{l}_1, \mathbf{s}_1, t)\}$ is obtained by applying a Fourier transform over the time domain. Note that this confocal *operator* Φ_{cc} is unrelated to confocal *acquisitions* (where $\mathbf{l}_1 \equiv \mathbf{s}_1$).

The temporal shift in Equation 4 can be interpreted using Rayleigh-Sommerfeld Diffraction (RSD) operators $\widehat{R}(t, \Omega)$, which express the phase shift of a wave as a function of time t and frequency Ω:

$$\widehat{R}(t, \Omega) = e^{i2\pi\Omega t}. \tag{5}$$

Finally, to obtain the resulting image, one can convert back from the Fourier domain by integrating over all frequencies Ω:

$$\Phi_{\text{cc}}(\mathbf{v}; H) = \int_{-\infty}^{+\infty} \widehat{\Phi}_{\text{cc}}(\mathbf{v}, \Omega; \widehat{H})\, \mathrm{d}\Omega. \tag{6}$$

Using wave-based operators such as the RSD integral allows us to look at the same problem from a different perspective: we can interpret the relay wall $\mathfrak{R}_1$ as a computational lens that *focuses* light from points in the illumination aperture $\mathcal{L}_1$ (through $\widehat{R}(t_{\mathbf{l}_1 \to \mathbf{v}}, \Omega)$), and in the camera aperture $\mathcal{S}_1$ (through $\widehat{R}(t_{\mathbf{v} \to \mathbf{s}_1}, \Omega)$), both illustrated in Figure 2c. Notably, the confocal camera $\Phi_{\text{cc}}(\mathbf{v}; H)$ focuses light at both apertures $\mathcal{L}_1$ and $\mathcal{S}_1$ to each point $\mathbf{v}$ simultaneously, which works for third-bounce illumination. Throughout the rest of our work we will show how to leverage higher-order illumination in novel imaging algorithms, instead of filtering out this information.

4 Cascaded NLOS imaging

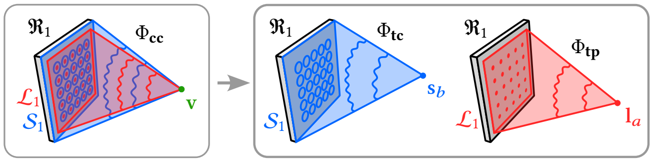
The key idea of our novel *cascaded NLOS imaging* method is that we can use the virtual imaging system at the visible relay wall to create and leverage *secondary* virtual imaging systems at different hidden walls in the scene (Figure 1b). These secondary systems allow us to take into account higher-order illumination, breaking the third-bounce-only assumption of most conventional methods.

Here we introduce the building blocks of our method. First, we introduce the transient camera Φ_{tc} and transient projector Φ_{tp} operators in Section 4.1; in Section 4.2, we show how to use them to compute the *virtual impulse response* H' of a hidden wall, which we then use in Section 4.3 to turn it into a secondary imaging system.

4.1 Transient camera and transient projector operators

We introduce the transient camera operator Φ_{tc} from work by Liu et al. [2019b], and our transient projector operator Φ_{tp} as a generalized virtual projector. We formalize these operators as the building blocks for phasor-based NLOS imaging methods.

As illustrated in Figure 3a, the confocal camera operator Φ_{cc} (Section 3) focuses all points $\mathbf{s}_1$ of the camera aperture $\mathcal{S}_1$ and all



(a) Confocal camera ($\widehat{\Phi}_{cc}$, Equation 4) (b) Transient camera ($\widehat{\Phi}_{tc}$, Equation 7) (c) Transient projector ($\widehat{\Phi}_{tp}$, Equation 9)

Fig. 3. (a) The confocal camera operator Φ_{cc} computationally focuses light at the illumination aperture $\mathcal{L}_1$ and at the camera aperture $\mathcal{S}_1$ of $\mathfrak{R}_1$ on the same point $\mathbf{v}$. We separate this procedure into (b) the transient camera Φ_{tc} and (c) our transient projector Φ_{tp} operators. Each focuses the light at $\mathcal{S}_1$ or $\mathcal{L}_1$ to individual points $\mathbf{s}_b$ and $\mathbf{l}_a$ respectively.

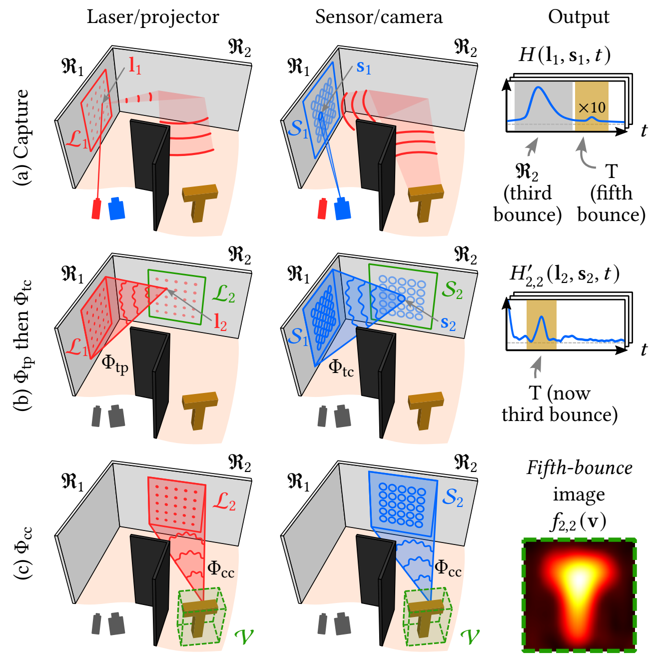


Fig. 4. Cascading NLOS imaging systems. (a) A laser emits pulses to each point $\mathbf{l}_1 \in \mathcal{L}_1$ on $\mathfrak{R}_1$ and an ultra-fast camera captures $H(\mathbf{l}_1, \mathbf{s}_1, t)$ at points $\mathbf{s}_1 \in \mathcal{S}_1$ also on $\mathfrak{R}_1$. Here $H(\mathbf{l}_1, \mathbf{s}_1, t)$ contains third- and fifth-bounce illumination from $\mathfrak{R}_2$ and T, respectively. (b) We compute our virtual impulse response $H'_{a,b}(\mathbf{l}_a, \mathbf{s}_b, t)$ on $\mathfrak{R}_2$ from H, i.e., $H'_{2,2}(\mathbf{l}_2, \mathbf{s}_2, t)$, using the transient projector Φ_{tp} and camera Φ_{tc} operators to focus light at $\mathcal{L}_1$ or $\mathcal{S}_1$ to each point $\mathbf{l}_2 \in \mathcal{L}_2$ or $\mathbf{s}_2 \in \mathcal{S}_2$ in $\mathfrak{R}_2$ respectively. Here $H'_{2,2}$ shows light reflected by the T as third-bounce illumination. (c) Finally we concatenate a confocal camera operator Φ_{cc} with $H'_{2,2}$, resulting in an image $f_{2,2}(\mathbf{v})$ for $\mathbf{v} \in \mathcal{V}$ of the T effectively computed using fifth-bounce illumination.

points $\mathbf{l}_1$ of the illumination aperture $\mathcal{L}_1$ simultaneously on the same point $\mathbf{v}$ of the hidden scene (Equation 4). Our general paradigm independently formulates the focusing operations of $\mathcal{S}_1$ and $\mathcal{L}_1$ using the transient camera operator Φ_{tc} (Figure 3b) and transient projector operator Φ_{tp} (Figure 3c) respectively; each focuses light to individual points $\mathbf{s}_b$ and $\mathbf{l}_a$ in the hidden scene. Note that $\mathbf{s}_b$ and $\mathbf{l}_a$ are analogous to $\mathbf{v}$ in the general case; we use a different notation to highlight that these operations have a different purpose. This formulation allows us to model light paths from the relay wall to the hidden scene and vice versa, and to reconstruct the time-resolved light transport through the hidden scene.

The *transient camera operator* Φ_{tc} (Figure 3b) focuses light from all $\mathbf{s}_1 \in \mathcal{S}_1$ to each $\mathbf{s}_b$. We first define its frequency-domain counterpart $\widehat{\Phi}_{tc}$ for each imaging frequency Ω:

$$\widehat{\Phi}_{tc}(\mathbf{s}_b, \Omega; \widehat{H}) = \int_{\mathcal{S}_1} \widehat{R}(t_{\mathbf{s}_b \to \mathbf{s}_1}, \Omega)\, \widehat{H}(\mathbf{l}_1, \mathbf{s}_1, \Omega)\, \mathrm{d}\mathbf{s}_1, \tag{7}$$

with $\widehat{H}(\mathbf{l}_1, \mathbf{s}_1, \Omega) = \mathcal{F}\{H(\mathbf{l}_1, \mathbf{s}_1, t)\}$ the Fourier transform in time, and only using one of the two RSD operators $\widehat{R}$ (Equation 5) from the confocal camera operator $\widehat{\Phi}_{cc}$ (Equation 4), since the illuminated points $\mathbf{l}_1 \in \mathcal{L}_1$ do not vary. We obtain the time-resolved signal via an inverse Fourier transform over all imaging frequencies Ω:

$$\Phi_{tc}(\mathbf{s}_b, t; H) = \mathcal{F}^{-1}\left\{ \widehat{\Phi}_{tc}\left(\mathbf{s}_b, \Omega; \widehat{H}\right) \right\}. \tag{8}$$

Our *transient projector operator* Φ_{tp} (Figure 3c) is similar to Φ_{tc}, except that Φ_{tp} focuses light from $\mathcal{L}_1$ to $\mathbf{l}_a$, leaving $\mathcal{S}_1$ unchanged. We define its frequency-domain counterpart $\widehat{\Phi}_{tp}$ as:

$$\widehat{\Phi}_{tp}(\mathbf{l}_a, \Omega; \widehat{H}) = \int_{\mathcal{L}_1} \widehat{R}(t_{\mathbf{l}_1 \to \mathbf{l}_a}, \Omega)\, \widehat{H}(\mathbf{l}_1, \mathbf{s}_1, \Omega)\, \mathrm{d}\mathbf{l}_1, \tag{9}$$

using the other RSD operator $\widehat{R}$ from Equation 4. Its time-domain version becomes:

$$\Phi_{tp}(\mathbf{l}_a, t; H) = \mathcal{F}^{-1}\left\{ \widehat{\Phi}_{tp}\left(\mathbf{l}_a, \Omega; \widehat{H}\right) \right\}. \tag{10}$$

Capture procedure. Our transient projector Φ_{tp} and transient camera Φ_{tc} operators enable focusing light onto the hidden scene, analogous to beamforming operators. Effectively focusing light onto a single point requires illuminating and capturing multiple points $\mathbf{l}_1$ and $\mathbf{s}_1$ in the impulse response $H(\mathbf{l}_1, \mathbf{s}_1, t)$ over finite apertures $\mathcal{L}_1$ and $\mathcal{S}_1$, respectively. To enable single-point focusing using both operators, we illuminate a 2D grid of points $\mathbf{l}_1$ on the relay wall with an ultra-fast laser. For each illuminated point $\mathbf{l}_1$ we exhaustively capture the resulting indirect photons at a 2D grid of points $\mathbf{s}_1$ with a time-gated sensor array. This results in a five-dimensional $H(\mathbf{l}_1, \mathbf{s}_1, t)$, which can be reduced to 3D in applications that only require single-point focusing using either Φ_{tc} or Φ_{tp}.

4.2 Virtual impulse response function

We next introduce our *virtual impulse response* H' of the hidden scene, computed from the impulse response H captured at the relay wall. For reference, $H(\mathbf{l}_1, \mathbf{s}_1, t)$ contains the temporal responses at $\mathbf{s}_1 \in \mathcal{S}_1$ of light pulses emitted towards each $\mathbf{l}_1 \in \mathcal{L}_1$ of the relay wall $\mathfrak{R}_1$. Throughout our work, we will use $H'_{a,b}(\mathbf{l}_a, \mathbf{s}_b, t)$ to denote our virtual impulse response, which analogously to H contains the *reconstructed* temporal response at $\mathbf{s}_b$ on $\mathfrak{R}_b$ of light focused at $\mathbf{l}_a$ on $\mathfrak{R}_a$ (note that $\mathfrak{R}_a$ and $\mathfrak{R}_b$ can be different). This will allow us in Section 4.3 to turn walls $\mathfrak{R}_a$ and $\mathfrak{R}_b$ into secondary virtual imaging systems. We use the apostrophe in H' to highlight that it is *computed* from H and not captured, with the special case $a = b = 1$ corresponding to $H_{1,1} \equiv H$ in $\mathfrak{R}_1$.

We formulate our virtual impulse response $H'_{a,b}$ as:

$$H'_{a,b}(\mathbf{l}_a, \mathbf{s}_b, t) = \Phi_{tc}\left(\mathbf{s}_b, t; \Phi_{tp}\left(\mathbf{l}_a, t; H\right)\right), \tag{11}$$

using Φ_{tc} (Equation 8) to compute at $\mathbf{s}_b$ the light that Φ_{tp} focuses at $\mathbf{l}_a$ (Equation 10). In Appendix A of our Supplemental Material we detail an alternative convolution-based formulation, optimized for computational efficiency.

Note that $H'_{a,b}(\mathbf{l}_a, \mathbf{s}_b, t)$ is effectively computed by temporally shifting and adding the filtered impulse response $H(\mathbf{l}_1, \mathbf{s}_1, t)$ using the distances of paths $\mathbf{l}_1 \rightarrow \mathbf{l}_a$ and $\mathbf{s}_b \rightarrow \mathbf{s}_1$. Given that light follows higher-order paths of the form $\mathbf{l}_1 \rightarrow \mathbf{l}_a \rightarrow \cdots \rightarrow \mathbf{s}_b \rightarrow \mathbf{s}_1$, the time t of $H'_{a,b}(\mathbf{l}_a, \mathbf{s}_b, t)$ represents the time of flight of light from $\mathbf{l}_a$ to $\mathbf{s}_b$, i.e., light subpaths of the form $\mathbf{l}_a \rightarrow \cdots \rightarrow \mathbf{s}_b$.

In the special case of $b = 1$, we only focus the illumination at $\mathbf{l}_a$ while keeping the camera aperture $\mathcal{S}_1$ untouched, which produces a time of flight t associated with subpaths $\mathbf{l}_a \rightarrow \cdots \rightarrow \mathbf{s}_1$ as:

$$H'_{a,1}(\mathbf{l}_a, \mathbf{s}_1, t) = \Phi_{tp}\left(\mathbf{l}_a, t; H\right). \tag{12}$$

Similarly, when $a = 1$, we keep the illumination aperture $\mathcal{L}_1$ untouched and measure other points $\mathbf{s}_b$ of the hidden scene, with the time of flight t corresponding to subpaths $\mathbf{l}_1 \rightarrow \cdots \rightarrow \mathbf{s}_b$, as:

$$H'_{1,b}(\mathbf{l}_1, \mathbf{s}_b, t) = \Phi_{tc}\left(\mathbf{s}_b, t; H\right). \tag{13}$$

Figure 4 shows an illustrative example with a letter T hidden around two corners, thus invisible to third-bounce methods. To image it, we leverage $\mathfrak{R}_2$ as a secondary relay wall. We illustrate the impulse response $H(\mathbf{l}_1, \mathbf{s}_1, t)$ at a grid of points for both $\mathbf{l}_1$ and $\mathbf{s}_1$ on $\mathfrak{R}_1$ (Figure 4a). As shown in the rightmost graphic, H contains third-bounce illumination from $\mathfrak{R}_2$ and, importantly, fifth-bounce illumination from the hidden T. From this, we compute the virtual impulse response $H'_{2,2}(\mathbf{l}_2, \mathbf{s}_2, t)$ (Equation 11), effectively shifting the illumination and capture positions to $\mathbf{l}_2$ and $\mathbf{s}_2$ in $\mathfrak{R}_2$ (Figure 4b). We show how the fifth-bounce illumination in the captured H now corresponds to third-bounce illumination in the reconstructed $H'_{2,2}$.

4.3 Cascading virtual imaging systems

We can now use our virtual impulse response $H'_{a,b}(\mathbf{l}_a, \mathbf{s}_b, t)$ with the same imaging operators as the conventional captured impulse response $H(\mathbf{l}_1, \mathbf{s}_1, t)$. In particular, we can use conventional imaging operators such as the confocal camera Φ_{cc} (Equation 6). As this will be common throughout our work, we use $f_{a,b}$ notation (adapted from Equation 1) to refer to images computed using Φ_{cc} with $H'_{a,b}$ as input:

$$f_{a,b}(\mathbf{v}) \equiv \Phi_{cc}(\mathbf{v}; H'_{a,b}). \tag{14}$$

In Figure 4c, we use Φ_{cc} with $H'_{2,2}(\mathbf{l}_2, \mathbf{s}_2, t)$, effectively using fifth-bounce illumination to compute an image $f_{2,2}(\mathbf{v})$ of the hidden T. Alternatively, we could also use our Φ_{tc} and Φ_{tp} operators in Equations 11, 12 and 13, which would yield an additional cascaded virtual impulse response H'' from H', further discussed in Section 9.

Cascaded imaging resolution. Compared to third-bounce imaging, higher-order imaging is a harder problem with more challenging conditions. Intuitively, the effective imaging resolution obtained from concatenated NLOS setups primarily depends on two factors: it increases with the area and resolution of *all* apertures $\mathcal{L}$ and $\mathcal{S}$, and decreases with the average distance along the paths $\mathcal{L} \rightarrow \mathcal{V} \rightarrow \mathcal{S}$ (i.e., between the apertures and the imaged volume). We refer the reader to Appendix B in our Supplemental Material for more details, where we reason about the theoretical imaging resolution and provide experimental validation using our hardware prototype.

Secondary relay wall. Our method uses a hidden wall as a secondary relay wall. However, note that we do *not* require prior knowledge of the room layout. Instead, we can leverage existing third-bounce methods to reveal such secondary walls with just 1–3 cm of error both in simulated and real captures, which does not affect our cascaded imaging results. We provide several examples in Sections 7.1, 7.3 and 8.2 of our results. In Appendix C of our Supplemental Material, we show how using third-bounce methods to find unknown secondary relay walls leads to cascaded imaging results that are practically identical to using exact prior knowledge about the location of such secondary relay walls. This highlights the feasibility of our cascaded approach, without relying on any prior scene information. Finally, to sample points $\mathbf{l}_a$ and $\mathbf{s}_b$ in secondary walls, we use the same sampling spacing as in the primary wall unless stated otherwise.

5 Imaging methods beyond the third bounce

We use our building blocks from Section 4 to design new NLOS imaging methods that leverage illumination beyond the third bounce. Here, we present two methods that use fourth- and fifth-bounce information, respectively, and then discuss how to obtain a combined image that merges third-, fourth-, and fifth-bounce imaging results.

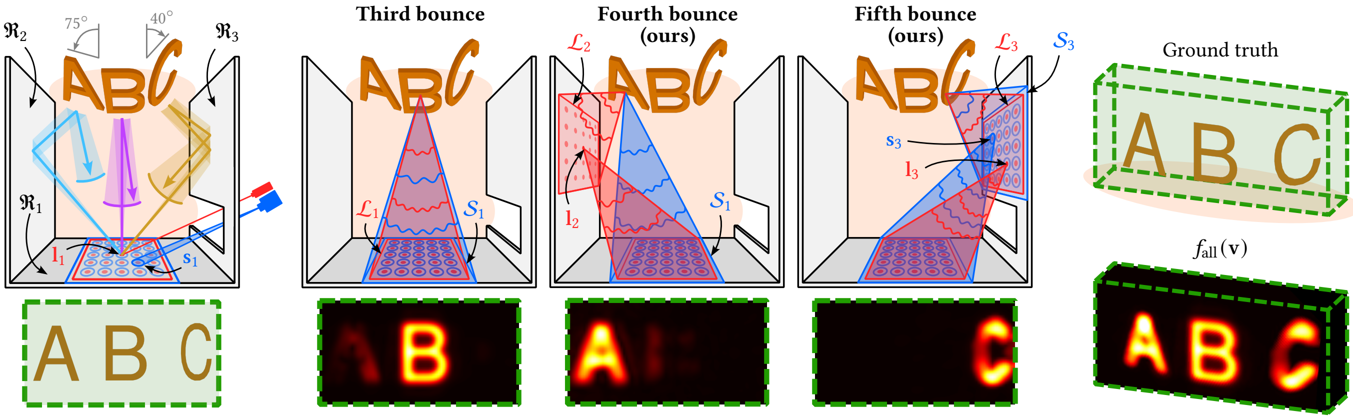
To illustrate our methods, we use the scene in Figure 5a, which contains three letters hidden from direct view: A, B, and C, each one with a different orientation. The captured impulse response H contains indirect light from all three letters. Conventional NLOS imaging methods (Figure 5b) use only third-bounce illumination (purple arrows in Figure 5a), and can only image the letter B. Due to their unfavorable orientation, the letters A and C have limited visibility. This is related to the finite size of the relay wall, which prevents certain surfaces from being fully imaged depending on their position and orientation [Liu et al. 2019a; Peña et al. 2025]. We discuss this problem and our approach to it in detail in Section 6.

To mitigate this, we leverage light coming from the left $\mathfrak{R}_2$ or right $\mathfrak{R}_3$ walls inside the hidden scene (which will act as secondary relay walls). This light then interacts with letters A and C, and finally is captured in the impulse response H as fourth- and fifth-bounce illumination, respectively (cyan and yellow arrows in Figure 5a).

In the following, we formulate our methods using this scene as an example, illustrated in Figure 5c and 5d. Note that all methods use the same input data: the impulse response H, only changing the computations. In Sections 7 and 8 we show and discuss the results of our methods in simulated and real scenarios.

5.1 Fourth-bounce imaging method

We are interested in leveraging fourth-bounce illumination (cyan arrow in Figure 5a). Generally, we can model either $\mathbf{l}_1 \rightarrow \mathbf{l}_a \rightarrow \mathbf{v} \rightarrow \mathbf{s}_1$ or $\mathbf{l}_1 \rightarrow \mathbf{v} \rightarrow \mathbf{s}_b \rightarrow \mathbf{s}_1$, with $\mathbf{l}_a$ or $\mathbf{s}_b$ on a secondary wall $\mathfrak{R}_a$ or $\mathfrak{R}_b$, and $\mathbf{v}$ on the target geometry. For the first case, we use our transient projector operator Φ_{tp} to compute $H'_{a,1}(\mathbf{l}_a, \mathbf{s}_1, t)$ as per Equation 12. Similarly, for the second case, we use our transient camera operator Φ_{tc} to compute $H'_{1,b}(\mathbf{l}_1, \mathbf{s}_b, t)$ as per Equation 13.



(a) Scene setup and relevant light paths

(b) Third bounce $f_{1,1}(\mathbf{v})$ Projector $\mathcal{L}_1$, camera $\mathcal{S}_1$

(c) Our $f_{2,1}(\mathbf{v})$ (Section 5.1) Projector $\mathcal{L}_2$, camera $\mathcal{S}_1$

(d) Our $f_{3,3}(\mathbf{v})$ (Section 5.2) Projector $\mathcal{L}_3$, camera $\mathcal{S}_3$

(e) Our combined image $f_{\text{all}}(\mathbf{v})$ (Section 5.3)

Fig. 5. Cascaded NLOS imaging methods. (a) The simulated impulse response $H(\mathbf{l}_1, \mathbf{s}_1, t)$ at $\mathfrak{R}_1$ contains light that has scattered on the relay wall $\mathfrak{R}_1$, the three letters A, B, and C (oriented differently), and on the left $\mathfrak{R}_2$ and right $\mathfrak{R}_3$ walls. Hence, H contains third- (purple), fourth- (cyan) and fifth-bounce (yellow) illumination. (b) Third-bounce imaging methods use apertures $\mathcal{L}_1$ and $\mathcal{S}_1$ and thus can only use part of H. (c) In this example, our fourth-bounce imaging method concatenates our transient projector operator Φ_{tp} focused at $\mathbf{l}_2 \in \mathcal{L}_2$ on the left wall $\mathfrak{R}_2$ along with the confocal camera operator Φ_{cc}. (d) Our fifth-bounce method concatenates our transient projector Φ_{tp} and camera Φ_{tc} operators focused at $\mathbf{l}_3 \in \mathcal{L}_3$ and $\mathbf{s}_3 \in \mathcal{S}_3$ respectively on the right wall $\mathfrak{R}_3$ along with the confocal camera operator Φ_{cc}. (e) Combined result.

Note that both cases would give the same results in our ABC scene, as the time of flight is the same for both paths.

Our fourth-bounce method consists of computing one of these two virtual impulse responses, then using it as input for the confocal camera operator Φ_{cc} as per Equation 14. Figure 5c illustrates an example using Equations 12 and 14 with $a = 2$ and $b = 1$.

5.2 Fifth-bounce imaging method

To leverage fifth-bounce illumination (yellow arrow in Figure 5a), we model paths of the form $\mathbf{l}_1 \to \mathbf{l}_a \to \mathbf{v} \to \mathbf{s}_b \to \mathbf{s}_1$, with $\mathbf{l}_a$ and $\mathbf{s}_b$ lying on secondary walls $\mathfrak{R}_a$ and $\mathfrak{R}_b$, respectively, and then use both Φ_{tp} and Φ_{tc} to compute $H'_{a,b}$ as per Equation 11. Our fifth-bounce method uses our virtual impulse response $H'_{a,b}$ as input for the confocal camera operator Φ_{cc} as per Equation 14, which in this case will yield a fifth-bounce image. Figure 5d illustrates an example using Equations 11 and 14 with $a = b = 3$.

5.3 Combining multiple images

Last, we can add up the third-, fourth- and fifth-bounce images to obtain the combined image $f_{\text{all}}(\mathbf{v})$, which in our example would be equal to $f_{1,1}(\mathbf{v}) + f_{2,1}(\mathbf{v}) + f_{3,3}(\mathbf{v})$. Note that our method is phasor-based: specifically, we add the amplitudes of each image.

6 Virtual surface reflectance and surface visibility

Here, we provide a theoretical and practical analysis of visibility limits of previous NLOS methods, and how our cascaded imaging overcomes such limits. We first introduce the concept of *virtual reflectance* of a surface; although all surfaces in our setups are diffuse under visible light, they may exhibit different reflectance properties when illuminated by virtual phasor waves. We explore here how virtual reflectance affects both the hidden objects' visibility and the imaging results when using a secondary virtual imaging system, and provide an extended analysis in Appendix D of our Supplemental Material. The insights developed here underpin the NLOS imaging results presented later in Sections 7 and 8.

6.1 Virtual specular and diffuse behavior

Figure 6a shows two NLOS setups, each with a single hidden wall: planar (top) and rough with 3 cm facets (bottom). In simulation, we illuminate one point $\mathbf{l}_1$ and capture the impulse response H at all points in $\mathcal{S}_1$. Although this capture uses visible light, the phasor-field formulation convolves H with a virtual illumination function $K(t; \lambda_c, \sigma)$ (Equation 3), producing phasor waves whose wavelength is set by λ_c. Practical λ_c values lie in the centimeter range [Liu et al. 2020, 2019b; Royo et al. 2023] (see Appendix D).

To better understand how each hidden wall reflects these phasor waves, we simulate how they reflect a spherical wavefront of $\lambda = 3$ cm emitted from $\mathbf{l}_1$ (Figure 6b). The top wall is planar with respect to λ and exhibits virtual *specular* behavior, as observed by Royo et al. [2023]. The wall reflects light from $\mathbf{l}_1$ into the specular direction (cyan lines). Only an area A_p redirects light towards the camera aperture $\mathcal{S}_1$, which corresponds to the area recovered in the NLOS imaging result (Figure 6c, computed with the confocal camera operator Φ_{cc} and filtering $\lambda_c = \sigma = 3$ cm). The rough wall instead exhibits virtual *diffuse* behavior: reflected light reaches $\mathcal{S}_1$ from all of its points, so the resulting image recovers a larger area A_r, but also introduces speckle-like artifacts from the wall roughness.

This relates to the missing-cone problem, and illustrates why rough walls have a greater coverage compared to planar walls, since an NLOS imaging system can only image parts of the surface that reflect phasor waves towards its sensor aperture [Liu et al. 2019a; Peña et al. 2025; Royo et al. 2023].

Effect of the phasor-field parameters. The virtual reflectance of a surface depends on the parameters λ_c and σ of $K(t; \lambda_c, \sigma)$. In Figure 7,

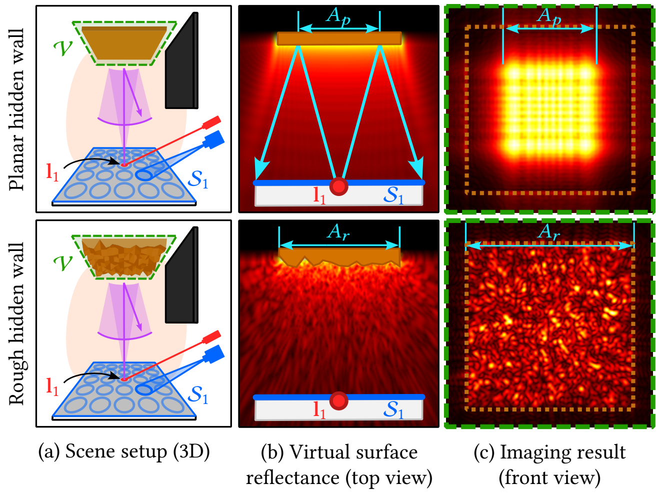


Fig. 6. Virtual surface reflectance. Top: planar hidden wall. Bottom: rough hidden wall with 3 cm × 3 cm facets. (a) Scene setup: we illuminate one point $\mathbf{l}_1$, and capture the impulse response in $\mathcal{S}_1$. We mark the imaged region as $\mathcal{V}$. (b) Simulated propagation of a phasor wave of wavelength $\lambda = 3$ cm emitted from $\mathbf{l}_1$ and reflected by the hidden wall. The planar wall reflects light specularly (cyan lines); only the area A_p returns light to the aperture $\mathcal{S}_1$. In contrast, the rough wall reflects diffusely, returning light to the aperture from a larger area A_r. (c) Resulting image $f_{1,1}(\mathbf{v}) = \Phi_{cc}(\mathbf{v}; H)$ over $\mathcal{V}$. The orange dotted line marks the expected wall location.

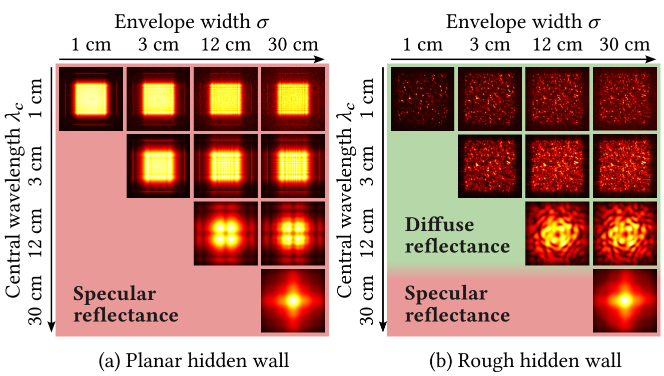


Fig. 7. Effect of the filter parameters λ_c and σ of $K(t; \lambda_c, \sigma)$ (Equation 3) for the setups in Figure 6. (a) The hidden wall is planar with respect to every λ_c and is thus always virtually specular; its edges have limited visibility. (b) The rough wall is virtually diffuse for λ_c around 3 cm (the size of its facets), and its full surface is imaged. For larger λ_c, it becomes planar with respect to the wavelength and converges to the same virtually specular behavior. Larger σ, corresponding to a narrower bandwidth of K, makes the NLOS imaging result more uniform, suggesting increased diffuseness.

we explore how λ_c and σ affect the NLOS imaging result. When λ_c is close to the rough wall's facet size, it behaves as virtually diffuse, enhancing its visibility. As λ_c grows, the phasor waves become planar with respect to both walls, which converge to virtual specular behavior. In this regime, parts of the wall disappear at the edges, where phasor waves no longer reach $\mathcal{S}_1$. Also, increasing σ narrows the filter bandwidth, which further enhances diffuseness when λ_c is near the facet size.

6.2 Surface visibility under higher-order illumination

In the third-bounce case, a hidden surface is visible only if it virtually reflects the illumination from $\mathcal{L}_1$ towards the aperture $\mathcal{S}_1$. Our insights on virtual reflectance and surface visibility extend to higher-order illumination: along every step of the path, each surface must virtually reflect light towards the next one, and the last one back towards $\mathcal{S}_1$. Visibility therefore depends on the virtual reflectance of every surface along the path, not just on the surface being imaged.

Leveraging the virtual diffuse behavior of rough walls allows us to turn hidden surfaces into secondary virtual imaging systems, whose angular visibility is much wider than the specular directions available to previous works [Royo et al. 2023]. This improvement entails a trade-off, since the speckle-like artifacts from virtual diffuse reflection limit the resolution of secondary virtual imaging systems.

Note that (virtually diffuse) rough walls are *not* a requirement of our method. On virtually specular surfaces our transient projector Φ_{tp} and camera Φ_{tc} operators form new apertures on secondary walls. This allows us to leverage more specular paths than previous works, which place apertures $\mathcal{L}_1$ and $\mathcal{S}_1$ only on the primary relay wall. We provide further analysis and illustrations in Appendix D.1.

7 Results in simulation

We show the versatility of our cascaded NLOS imaging in three different scenarios that represent novel imaging possibilities: overcoming visibility limits from third-bounce methods (Section 7.1), enhanced multi-view imaging of complex geometry (Section 7.2), and imaging and tracking around two corners in non-specular directions (Section 7.3). The results in this section are obtained from simulations using publicly available software [Peña et al. 2025; Royo et al. 2025], and include realistic noise as detailed in Appendix E of our Supplemental Material. We match SNR, calibration errors, and temporal uncertainty to values reported in previous works and measured in our own real-world captures. Later, Section 8 discusses additional results obtained with our real prototype. We reiterate that, both for simulated and real-world results, the impulse response H needs to be captured only once.

7.1 Overcoming third-bounce visibility limits

The bottom row of Figure 5 shows results of the scene previously introduced in Section 5. In our simulation, we exhaustively capture the impulse response $H(\mathbf{l}_1, \mathbf{s}_1, t)$ by scanning a 32×32 grid of illumination points $\mathbf{l}_1 \in \mathcal{L}_1$ over a $2\,\text{m} \times 2\,\text{m}$ area, while measuring the reflected signal at an equivalent, co-located grid of sensing points $\mathbf{s}_1 \in \mathcal{S}_1$. The impulse response H captures interreflections from the ABC letters and the left $\mathfrak{R}_2$ and right $\mathfrak{R}_3$ walls. Third-bounce photons account for 90% of the total, fourth-bounce for 7.5%, and fifth-bounce for just 2.5%. We show a classic, third-bounce image $f_{1,1}(\mathbf{v})$ in Figure 5b from the confocal camera operator Φ_{cc} (Equation 6). The image only shows the B letter, since A and C exhibit limited visibility due to their orientation. To image these, we leverage higher-order illumination from $\mathfrak{R}_2$ and $\mathfrak{R}_3$.

First, we estimate the position of $\mathfrak{R}_2$ and $\mathfrak{R}_3$. Inspired by work by Royo et al. [2023], we image the *reflection* $\mathbf{l}'_1$ of illuminated points $\mathbf{l}_1$ produced by $\mathfrak{R}_2$ and $\mathfrak{R}_3$, and estimate each wall as the perpendicular bisector of $\mathbf{l}'_1$ and $\mathbf{l}_1$. This estimate is accurate to about 1 cm.

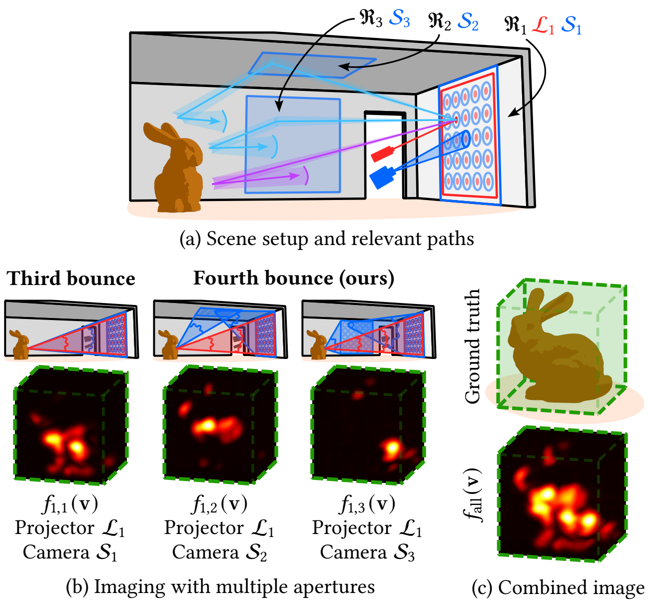


Fig. 8. Enhanced multi-view imaging of complex geometry. (a) The simulated capture $H(\mathbf{l}_1, \mathbf{s}_1, t)$ contains third-bounce illumination from the target object (Bunny) and also higher-order interreflections that arise from interactions with the top $\mathfrak{R}_2$ and right $\mathfrak{R}_3$ walls. (b) We show that fourth-bounce illumination extends the coverage of conventional NLOS imaging methods. Here we focus the camera aperture $\mathcal{S}_1$ of the first relay wall $\mathfrak{R}_1$ on the top $\mathfrak{R}_2$ and right $\mathfrak{R}_3$ walls (creating virtual imaging devices with apertures $\mathcal{S}_2$ and $\mathcal{S}_3$, respectively). Each yields a different image $f_{1,1}(\mathbf{v}), f_{1,2}(\mathbf{v}), f_{1,3}(\mathbf{v})$ of the target object, and (c) we combine all of them in $f_{\mathrm{all}}(\mathbf{v})$ to show the enhanced visibility of our cascaded NLOS imaging paradigm.

Then, we create secondary imaging systems on $\mathfrak{R}_2$ and $\mathfrak{R}_3$ to image the A and C letters. Figure 5c shows the results using fourth-bounce illumination (Section 5.1) from the left wall $\mathfrak{R}_2$. We apply Φ_{tp} from Equation 12 to compute $H'_{2,1}$ from H, then compute $f_{2,1}(\mathbf{v})$ as per Equation 14. Figure 5d shows results using fifth-bounce illumination (Section 5.2) using the right wall $\mathfrak{R}_3$, applying both Φ_{tp} and Φ_{tc} (Equation 11) to compute $H'_{3,3}$ from H, and then computing $f_{3,3}(\mathbf{v})$ as per Equation 14. The combined result $f_{\mathrm{all}}(\mathbf{v}) = f_{1,1}(\mathbf{v}) + f_{2,1}(\mathbf{v}) + f_{3,3}(\mathbf{v})$ in Figure 5e shows all three letters. We use $\lambda_c = \sigma = 10$ cm for imaging with third and fourth bounces, and $\lambda_c = \sigma = 14$ cm for the fifth bounce, due to increased noise in fifth-bounce photons.

7.2 Enhanced multi-view imaging of complex geometry

The bunny in Figure 8a presents self-occlusions and surface features oriented in a wide range of directions. Again, we simulate the capture of H at the $2\,\mathrm{m} \times 2\,\mathrm{m}$ relay wall $\mathfrak{R}_1$ using a grid of 8×8 laser $\mathcal{L}_1$ and 128×128 sensor $\mathcal{S}_1$ points. The relay wall $\mathfrak{R}_1$ receives indirect illumination from the object (purple arrow), and from the top $\mathfrak{R}_2$ and right $\mathfrak{R}_3$ walls (cyan arrows).

Figure 8b shows our results. The first column shows the image obtained from classic third-bounce methods $f_{1,1}(\mathbf{v}) = \Phi_{\mathrm{cc}}(\mathbf{v}; H)$, where many features are missing from the bunny. The second and

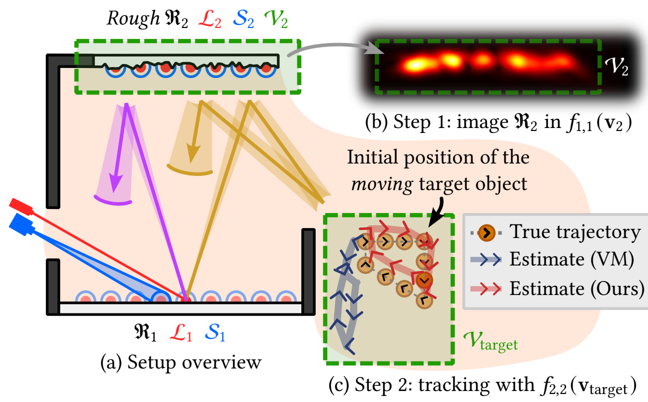


Fig. 9. We simulate tracking an object in a region $\mathcal{V}_{\mathrm{target}}$ hidden around two corners, not visible from the relay wall $\mathfrak{R}_1$. (a) Light from the target object is reflected on the rough hidden wall $\mathfrak{R}_2$ before reaching $\mathfrak{R}_1$, resulting in fifth-bounce illumination (yellow arrow). (b) First, we image the secondary wall $\mathfrak{R}_2$ from $\mathfrak{R}_1$ using third-bounce illumination (purple arrow). (c) The virtual mirrors (VM) approach [Royo et al. 2023] fails due to the roughness of $\mathfrak{R}_2$. Our cascaded NLOS imaging yields a good approximation of the trajectory of the hidden object by implementing a secondary imaging system at $\mathfrak{R}_2$.

third columns show images obtained using fourth-bounce illumination (Section 5.1), from the top $\mathfrak{R}_2$ and right $\mathfrak{R}_3$ walls respectively, applying Φ_{tc} from Equation 13 to compute $H'_{1,2}$ and $H'_{1,3}$ from H, and computing $f_{1,2}(\mathbf{v})$ and $f_{1,3}(\mathbf{v})$ as per Equation 14. We can see how, adding the top and right virtual imaging systems, we recover many of the previously missing features, such as the ears and the tail. The combined result in Figure 8c shows a greater coverage of the hidden geometry with respect to third-bounce methods, reducing depth RMSE from 3.5 cm to 2.9 cm. We use $\lambda_c = \sigma = 8$ cm for all cases.

7.3 Imaging and tracking around two corners

Figure 9a shows an NLOS setup where we detect and track an object hidden around two corners in the presence of a rough secondary relay wall $\mathfrak{R}_2$, with the centimeter-scale features one may find in regular brick or stone walls. Light from the object needs to bounce on $\mathfrak{R}_2$ before reaching the relay wall $\mathfrak{R}_1$, resulting in fifth-bounce illumination (yellow arrow) captured on H. We observe that the interaction of light with rough walls produces distinctive transport that affects NLOS imaging, as explained in Section 6. Concretely, rough surfaces decrease spatial imaging resolution but allow us to image parts of the hidden scene that would otherwise remain invisible to other NLOS methods, similar to how diffuse surfaces scatter the reflected light across a broader angular range.

We simulate the capture of H for 30 points in $\mathcal{L}_1$ and $\mathcal{S}_1$ on the 2 m relay wall $\mathfrak{R}_1$. Our method works in two steps. First, we image $\mathfrak{R}_2$ using conventional third-bounce illumination (purple arrow). In particular, we compute $f_{1,1}(\mathbf{v}) = \Phi_{\mathrm{cc}}(\mathbf{v}; H)$ (Equation 6) on points $\mathbf{v}$ in the bounding volume $\mathcal{V}_2$ using Φ_{cc}. We assign apertures $\mathcal{L}_2$ and $\mathcal{S}_2$ that follow our estimation of $\mathfrak{R}_2$. Note that $\mathcal{L}_2$ and $\mathcal{S}_2$ can be planar, i.e., they do not need to follow the irregularities of $\mathfrak{R}_2$.

Our second step uses the fifth-bounce method from Section 5.2 to image the target object. We compute Equations 11 and 14 with $a = b = 2$ to create a second virtual imaging system at $\mathfrak{R}_2$. However,

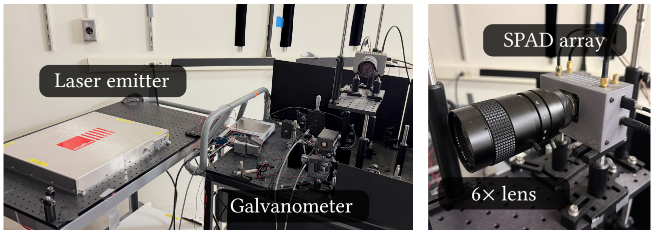


Fig. 10. Our hardware prototype. We use a 16×16 SPAD array, focused on the relay wall using a 6× Edmund Optics lens, and a two-mirror galvanometer (Thorlabs GVS012) to guide an ultra-fast laser emitter towards a 190 × 190 grid of points on the relay wall.

the roughness of $\mathfrak{R}_2$ also introduces virtual diffraction, which would create strong speckle patterns in the resulting image $f_{2,2}(\mathbf{v})$. To mitigate this, we instead take the *absolute value* of the virtual impulse response $H'_{2,2}$, which effectively removes the phase of the reconstructed wave, retaining only its magnitude. We use $\lambda_c = \sigma = 10$ cm for H and $\lambda_c = \sigma = 25$ cm for $H'_{2,2}$. We repeat this for different positions of the target object, tracking its location by finding a nearby peak in the image $f_{2,2}(\mathbf{v})$ at points $\mathbf{v} \in \mathcal{V}_{\text{target}}$. Figure 9c shows that we accurately follow the object's trajectory. The figure also shows how the virtual mirrors approach from Royo et al. [2023], which requires a planar secondary relay wall and images the mirrored region behind $\mathfrak{R}_2$, fails to accurately track the same object.

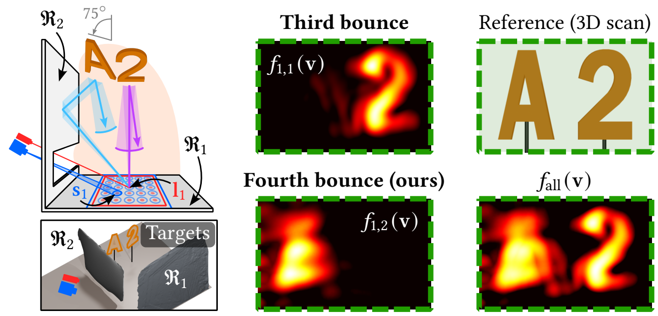
8 Results with our real prototype

We now present real-world results captured with our hardware prototype, imaging scenes using higher-order illumination scattered by secondary hidden surfaces. We evaluate our fourth-bounce imaging method in Section 8.1, and our fifth-bounce imaging method in Section 8.2, including imaging through rough walls, and without any prior knowledge of the hidden scene.

Hardware setup. Our NLOS imaging system (Figure 10) consists of a 16×16 SPAD array sensor [Riccardo et al. 2022], a laser emitter (Onefive KATANA HP) and a two-mirror galvanometer (Thorlabs GVS012). The galvanometer guides the laser towards a grid of 190×190 points on a 1.9 m × 1.9 m region of the relay wall, resulting in 1 cm spacing between adjacent laser points. We use an Edmund Optics 6× lens (12.5 mm, f/1.2) to focus the sensor array on a 50 cm × 32.5 cm region on the relay wall. The laser emits 532 nm pulses with a maximum pulse width of 40 ps, at an average repetition rate of 10 MHz. The sensor has a temporal resolution of approximately 60 ps Full-Width at Half Maximum (FWHM). Pictures of all our scenes are included in Appendix F of our Supplemental Material.

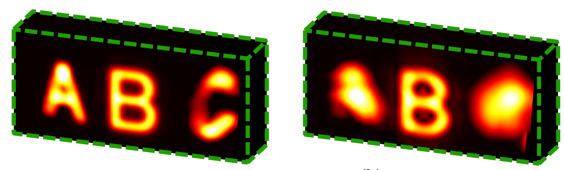
8.1 Fourth-bounce imaging results

In Figure 11a, we show a scene similar to Figure 5. We focus on two target hidden objects: a *2*-shaped object which can be imaged from the relay wall $\mathfrak{R}_1$ using third-bounce methods, and an *A*-shaped object which has limited visibility to third-bounce methods due to its orientation. Again, we capture the impulse response H at the relay wall $\mathfrak{R}_1$ (see hardware setup). To image the hidden A, we leverage fourth-bounce illumination from H coming from the left wall $\mathfrak{R}_2$.



(a) Scene setup (b) Imaging results (c) Combined image

Fig. 11. Real-world demonstration of NLOS imaging with fourth-bounce light. (a) We use a setup similar to Figure 5. (b) While the 2 can be imaged using third-bounce methods ($f_{1,1}(\mathbf{v})$), the orientation of the A requires fourth-bounce illumination ($f_{1,2}(\mathbf{v})$), enabled by our cascaded approach (analogous to our simulations in Figure 5b and Figure 5c). (c) Combined image $f_{\text{all}}(\mathbf{v}) = f_{1,1}(\mathbf{v}) + f_{1,2}(\mathbf{v})$.



(a) 32 × 32 SPAD, 2 m × 2 m aperture (b) 16 × 16 SPAD, 50 cm × 32.5 cm aperture

Fig. 12. We repeat in simulation the same imaging procedure from Figure 5e, changing the size and resolution of the illumination aperture $\mathcal{L}_1$ to match the SPAD array from our prototype. (a) Original result. (b) Reducing both the resolution to 16 × 16 points and covered area to 50 cm × 32.5 cm. The fourth-bounce imaging results of the A letter match the resolution of our real capture in Figure 11. Moreover, as explained in Section 8.2, the fifth-bounce results of the C letter also match our real-world, fifth-bounce experiments.

We use our fourth-bounce imaging method: first, we compute our virtual impulse response $H'_{1,2}$ from H by applying Equation 13. We sample a grid of 40×40 points $\mathbf{s}_2 \in \mathcal{S}_2$ on $\mathfrak{R}_2$. Figure 11b shows our imaging results: while the third-bounce-method image $f_{1,1}(\mathbf{v})$ computed from H only shows the 2 at points $\mathbf{v}$ in the bounding volume $\mathcal{V}$, our fourth-bounce-method $f_{1,2}(\mathbf{v})$ computed from $H'_{1,2}$ successfully images the A. We use $\lambda_c = \sigma = 6$ cm for computing $f_{1,1}(\mathbf{v})$, and $\lambda_c = \sigma = 8$ cm for computing $H'_{1,2}$ and $f_{1,2}$.

Resolution analysis. Imaging quality depends heavily on the resolution and size of the captured laser $\mathcal{L}_1$ and sensor $\mathcal{S}_1$ grids, which we discuss further in Appendix B of our Supplemental Material. Our simulations in Section 7 mimic a 32×32 SPAD array covering a 2 m × 2 m area of the relay wall. While higher-resolution SPAD arrays are progressively becoming available, our prototype uses a lower-resolution, state-of-the-art SPAD array with 16×16 points, covering a smaller area on the relay wall of 50 cm × 32.5 cm. These limitations affect the fidelity of the real-world results, introducing factors such as higher noise and reduced imaging resolution that account for the observed simulation-to-real gap.

In Figure 12, we validate our method by modifying our simulation parameters in the scene of Figure 5 to match our real prototype, both in aperture size and SPAD resolution. Figure 12a shows the

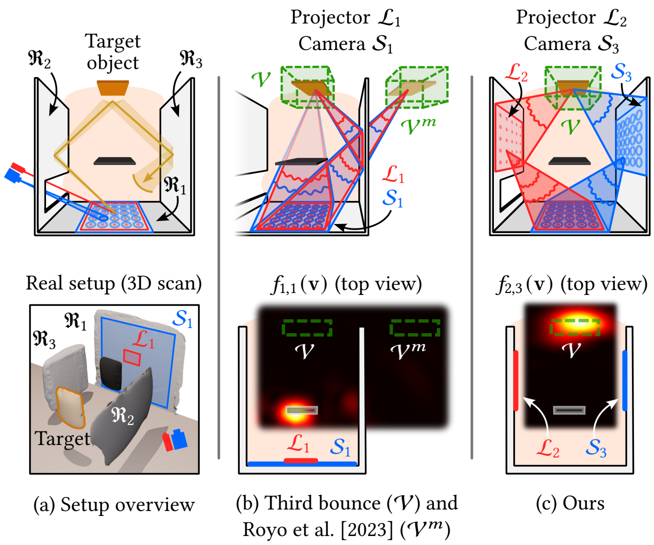


Fig. 13. Overcoming two-corner visibility limits using our hardware prototype. (a) The left $\mathfrak{R}_2$ and right $\mathfrak{R}_3$ walls reflect higher-order illumination to and from the target object. (b) Due to the walls' position and orientation, neither third-bounce methods nor the virtual mirrors approach [Royo et al. 2023] can image the target object (marked by $\mathcal{V}$ and $\mathcal{V}^m$). (c) Our cascaded NLOS fifth-bounce method images the object at $\mathcal{V}$ using $\mathcal{L}_2$ and $\mathcal{S}_3$.

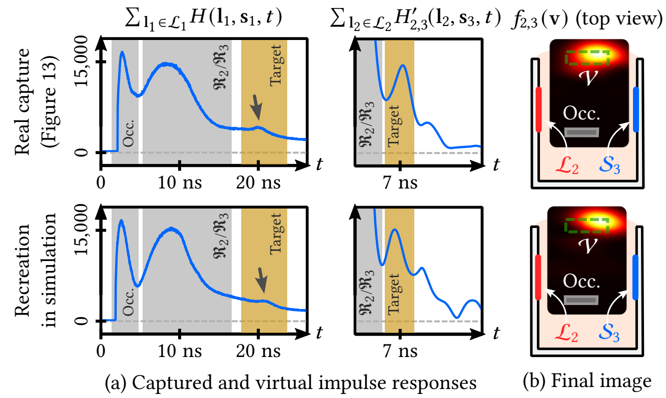


Fig. 14. We compare our real-world experiment (top) from Figure 13 with its simulated counterpart (bottom), mimicking our 16×16 SPAD. Both the captured H and virtual $H'_{2,3}$ impulse responses (a), as well as the resulting image $f_{2,3}(\mathbf{v})$ (b) closely match, with small differences due to geometric discrepancies between the real and the simulated scenes.

combined result of our original simulation with a 32×32 SPAD array capturing a $2\,\text{m} \times 2\,\text{m}$ area on the relay wall. In Figure 12b, we change our simulated SPAD array to match the 16×16 resolution and $50\,\text{cm} \times 32.5\,\text{cm}$ area covered by our real prototype. The simulated result degrades consistently. The B and A letters closely match the captured third- and fourth-bounce results in Figure 11. Moreover, as we will see in Section 8.2, the fifth-bounce image of the C letter also matches the resolution of our real fifth-bounce results.

8.2 Fifth-bounce imaging results

Two-corner imaging in unfavorable orientations. Figure 13a illustrates an object hidden around the corner (highlighted in orange) with an additional occluder (black) in front, preventing third- and even fourth-bounce illumination from reaching $\mathfrak{R}_1$ and the sensor. We capture the impulse response H on $\mathfrak{R}_1$, which includes fifth-bounce illumination from walls $\mathfrak{R}_2$ and $\mathfrak{R}_3$.

As Figure 13b shows, conventional third-bounce methods fail to image the object, which should appear in the volume $\mathcal{V}$. Moreover, the recent *virtual mirrors* approach [Royo et al. 2023], which images the mirrored volume $\mathcal{V}^m$ and relies on fifth-bounce illumination, also fails: their method imposes strict conditions on the position and orientation of both the hidden objects and the secondary walls, which need to align with favorable mirror directions. Note that Royo et al. originally demonstrated their approach using only a single illumination point $\mathbf{l}_1$. To provide a fair comparison, we have enhanced their method to use our 16×16 illumination grid $\mathcal{L}_1$.

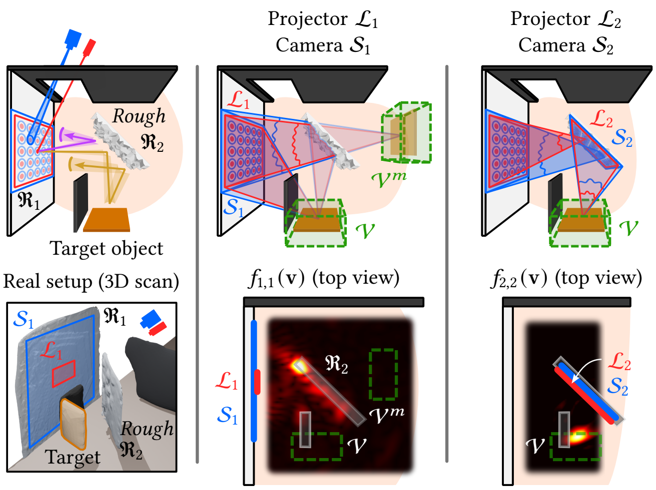
In contrast, our fifth-bounce method (Section 5.2) leverages a broader range of specular paths, and thus is able to image the hidden object at the correct location marked by $\mathcal{V}$, by creating both a virtual illumination aperture $\mathcal{L}_2$ on $\mathfrak{R}_2$, and a virtual camera aperture $\mathcal{S}_3$ on a different wall $\mathfrak{R}_3$ (Figure 13c). Specifically, we compute our virtual impulse response $H'_{2,3}(\mathbf{l}_2, \mathbf{s}_3, t)$ (Equation 11 with $a = 2$ and $b = 3$), effectively illuminating points $\mathbf{l}_2 \in \mathcal{L}_2$ and computing the time-resolved response at points $\mathbf{s}_3 \in \mathcal{S}_3$. For both *secondary* apertures, we sample grids of 8×6 points with 12.5 cm spacing. We then use $H'_{2,3}$ as input for the confocal camera operator Φ_{cc} (Equation 14), obtaining an image $f_{2,3}(\mathbf{v})$, which reveals the target object at the correct location. We use $\lambda_c = \sigma = 30$ cm for H and $H'_{2,3}$.

Figure 14 further validates our method, recreating the same scene in simulation, mimicking the 16×16 SPAD from our prototype. The captured impulse response H and virtual impulse response $H'_{2,3}$, as well as the resulting image $f_{2,3}(\mathbf{v})$, closely match our physical capture. Note that we show the sum over all illumination positions for H and $H'_{2,3}$, which smooths SPAD shot noise. Early-arriving photons preceding the third-bounce signal are gated out following standard practice. Other small differences are mainly due to geometric discrepancies between the real and the simulated scenes.

Two-corner imaging through rough walls. One key advantage of our method is that it works in the presence of rough relay walls. We illustrate this in our next example, imaging a scene around two corners without previous knowledge of the secondary relay wall.

Figure 15a shows our setup. As in the previous example, both third-bounce methods and our enhanced version of virtual mirrors [Royo et al. 2023] (which leverages our 16×16 points for $\mathcal{L}_1$) fail to image the target at $\mathcal{V}$ and $\mathcal{V}^m$, respectively (Figure 15b).

We first image the secondary wall $\mathfrak{R}_2$ by computing a conventional third-bounce image $f_{1,1}(\mathbf{v})$ on points $\mathbf{v} \in \mathcal{V}$ (as shown in Figure 15b). Leveraging our cascaded approach, we then compute a secondary impulse response $H'_{2,2}$ (Equation 11) by creating virtual illumination and camera apertures $\mathcal{L}_2$ and $\mathcal{S}_2$ (Figure 15c, top) at the revealed location of $\mathfrak{R}_2$ in $f_{1,1}(\mathbf{v})$ (see Appendix C of our Supplemental Material for further details). We sample grids of 6×6 points for both apertures with 12.5 cm spacing. Finally, we use $H'_{2,2}$



(a) Setup overview (b) Third bounce ($\mathcal{V}$) and Royo et al. [2023] ($\mathcal{V}^m$) (c) Ours

Fig. 15. Imaging around two corners using our hardware prototype. (a) The rough hidden wall $\mathfrak{R}_2$ reflects light to and from the target object. (b) Both conventional methods and virtual mirrors [Royo et al. 2023] fail to image the target object (marked by $\mathcal{V}$ and $\mathcal{V}^m$, respectively). (c) In contrast, our method identifies the target at the correct location $\mathcal{V}$ by creating a secondary virtual imaging system at $\mathfrak{R}_2$.

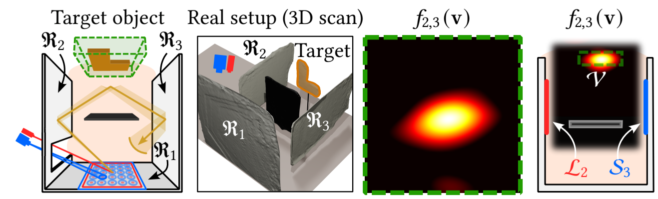
to image the target object by computing $f_{2,2}(\mathbf{v})$ (Equation 14). Figure 15c, bottom, shows that $f_{2,2}(\mathbf{v})$ reveals the target object hidden around two corners on its location marked by $\mathcal{V}$. Our algorithm leverages roughness in the secondary relay wall by modulating λ_c accordingly; in particular, we use $\lambda_c = \sigma = 15$ cm for H, and a coarser $\lambda_c = \sigma = 25$ cm for $H'_{2,2}$.

9 Conclusions and discussion

Our novel cascaded NLOS imaging methods allow us to create and concatenate secondary virtual imaging systems. This enables new NLOS imaging beyond the three-bounce assumption used by conventional methods, leveraging fourth- and fifth-bounce illumination. We have provided results in simulation and with a hardware prototype, showcasing the potential of our approach in several scenarios.

We have shown results cascading two imaging systems although, in theory, multiple cascading NLOS imaging systems could be concatenated. We could compute another virtual impulse response $H''_{a,b}(\mathbf{l}_a, \mathbf{s}_b, t)$ on points $\mathbf{l}_a, \mathbf{s}_b$ on a *tertiary* relay wall, using *fifth-* to *seventh-bounce* illumination, expanding even further the possibilities of future NLOS imaging systems. However, each concatenation imposes a resolution loss due to the reduced signal, limited aperture sizes, longer focus distances, and increased multi-path interference.

In particular, our current prototype uses a 16 × 16 SPAD array [Riccardo et al. 2022] covering a 50 cm × 32.5 cm area. This results in a spatial resolution Δx for the virtual impulse response of about 20–40 cm depending on the setup (see Appendix B for details on the derivation of both spatial and temporal resolution), which in turn limits the features that can be effectively resolved. This is shown in Figure 16, where we have repeated the same setup from Figure 13, this time on an asymmetric, L-shaped target (Figure 16a).



(a) Setup overview (b) Front view (c) Top view

Fig. 16. (a) We repeat the setup from Figure 13 using an L-shaped target. (b) Front view of our imaging result using our real prototype (16 × 16 SPAD covering a 50 cm × 32.5 cm aperture). (c) Top view of the same result.

We use the imaging procedure explained in Section 8.2, with filtering parameters $\lambda_c = \sigma = 20$ cm for H and $H'_{2,3}$. Figure 16b shows a front view of $f_{2,3}(\mathbf{v})$, clearly imaging the target object in the right location, but failing to resolve its asymmetric features. Figure 16c displays a top view of $f_{2,3}(\mathbf{v})$, showing how the target is also imaged at the correct depth. Despite the current resolution limitations, note that existing NLOS methods fail to image the target, as we have shown in this paper.

As Figure 12 illustrates, as larger SPADs become available, the resolution of our cascaded systems will increase. From our analysis in Appendix B, a future 128 × 128 SPAD array will theoretically yield a much higher resolution of around 1–2 cm, without changes in our algorithm. Further research on wave scattering properties may also help; for instance, recent research has found that a photon's Fisher information does not decrease with increasing scattering [Hüpfl et al. 2024]. Another interesting path for further analysis is the use of wavefront shaping [Cao et al. 2022] to potentially leverage diffuse paths in smoother secondary surfaces without centimeter-scale features. Last, as more modern laser and sensor technology becomes available, we foresee a progressive increase in imaging capabilities enabled by our cascaded framework.

Acknowledgments

This work has been supported by grants PID2022-141539NB-I00 and PID2025-169453NB-I00, funded by MICIU/AEI/10.13039/501100011033 and by ERDF, EU; and by grant FA9550-26-1-B169, funded by the Air Force Office of Scientific Research. María Peña was supported by the FPU23/03646 predoctoral grant and Diego Royo was supported by the Government of Aragon CUS/803/2021 predoctoral grant. This material is partially based upon work supported by DE-NA0004196 and through the Enabling Technologies & Innovation Graduate Fellowship, funded by the Department of Energy/National Nuclear Security Administration. The authors would like to thank Talha Sultan and the members of the Graphics and Imaging Lab for useful discussions.

References

Victor Arellano, Diego Gutierrez, and Adrian Jarabo. 2017. Fast Back-Projection for Non-Line of Sight Reconstruction. *Optics Express* 25, 10 (2017).

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

Ruizhi Cao, Frederic de Goumoens, Baptiste Blochet, Jian Xu, and Changhuei Yang. 2022. High-resolution non-line-of-sight imaging employing active focusing. *Nature Photonics* 16, 6 (2022), 462–468.

Kiseok Choi, Inchul Kim, Dongyoung Choi, Julio Marco, Diego Gutierrez, and Min H. Kim. 2023. Self-Calibrating, Fully Differentiable NLOS Inverse Rendering. In *Proceedings of ACM SIGGRAPH Asia 2023*.

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

Justin Dove and Jeffrey H Shapiro. 2020. Nonparaxial phasor-field propagation. *Optics Express* 28, 20 (2020), 29212–29229.

Jorge Garcia-Pueyo and Adolfo Muñoz. 2025. Forward and inverse diffraction in phasor fields. *Optics Express* 33, 5 (2025), 11420–11441.

Javier Grau Chopite, Matthias B. Hullin, Michael Wand, and Julian Iseringhausen. 2020. Deep Non-Line-of-Sight Reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Chaoying Gu, Talha Sultan, Khadijeh Masumnia-Bisheh, Laura Waller, and Andreas Velten. 2023. Fast non-line-of-sight imaging with non-planar relay surfaces. In *2023 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 1–12.

Jakob Hüpfl, Felix Russo, Lukas M Rachbauer, Dorian Bouchet, Junjie Lu, Ulrich Kuhl, and Stefan Rotter. 2024. Continuity equation for the flow of Fisher information in wave scattering. *Nature Physics* 20, 8 (2024), 1294–1299.

Adrian Jarabo, Belen Masia, Julio Marco, and Diego Gutierrez. 2017. Recent advances in transient imaging: A computer graphics and vision perspective. *Visual Informatics* 1, 1 (2017), 65–79.

Ahmed Kirmani, Tyler Hutchison, James Davis, and Ramesh Raskar. 2009. Looking around the corner using transient imaging. In *2009 IEEE 12th International Conference on Computer Vision*. IEEE, 159–166.

Yue Li, Jiayong Peng, Juntian Ye, Yueyi Zhang, Feihu Xu, and Zhiwei Xiong. 2023. NLOST: Non-Line-of-Sight Imaging With Transformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 13313–13322.

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

Xiaochun Liu, Sebastian Bauer, and Andreas Velten. 2019a. Analysis of feature visibility in non-line-of-sight measurements. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 10140–10148.

Xiaochun Liu, Sebastian Bauer, and Andreas Velten. 2020. Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems. *Nature Communications* 11, 1 (2020), 1645.

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

Pablo Luesia-Lahoz, Diego Gutierrez, and Adolfo Muñoz. 2023. Zone Plate Virtual Lenses for Memory-Constrained NLOS Imaging. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE.

Pablo Luesia-Lahoz, Talha Sultan, Forrest B. Peterson, Andreas Velten, Diego Gutierrez, and Adolfo Muñoz. 2025. Zero-Phase Phasor Fields for Non-Line-of-Sight Imaging. In *2025 IEEE International Conference on Computational Photography (ICCP)*. 1–9.

Tomohiro Maeda, Guy Satat, Tristan Swedish, Lagnojita Sinha, and Ramesh Raskar. 2019. Recent Advances in Imaging Around Corners. *arXiv preprint arXiv:1910.05613* (2019).

Julio Marco, Adrian Jarabo, Ji Hyun Nam, Xiaochun Liu, Miguel Ángel Cosculluela, Andreas Velten, and Diego Gutierrez. 2021. Virtual light transport matrices for non-line-of-sight imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2440–2449.

Ji Hyun Nam, Eric Brandt, Sebastian Bauer, Xiaochun Liu, Eftychios Sifakis, and Andreas Velten. 2020. Real-time Non-line-of-Sight imaging of dynamic scenes. *arXiv preprint arXiv:2010.12737* (2020).

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

María Peña, Diego Gutierrez, and Julio Marco. 2025. Looking Around Flatland: End-to-End 2D Real-Time NLOS Imaging. *IEEE Transactions on Computational Imaging* 11 (2025), 189–200.

Adithya Pediredla, Akshat Dave, and Ashok Veeraraghavan. 2019. Snlos: Non-line-of-sight scanning through temporal focusing. In *2019 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 1–13.

Oscar Pueyo-Ciutad, Julio Marco, Stephane Schertzer, Frank Christnacher, Martin Laurenzis, Diego Gutierrez, and Albert Redo-Sanchez. 2024. Time-Gated Polarization for Active Non-Line-of-Sight Imaging. In *Proceedings of ACM SIGGRAPH Asia 2024*.

Syed Azer Reza, Marco La Manna, Sebastian Bauer, and Andreas Velten. 2019. Wave-like properties of phasor fields: experimental demonstrations. *arXiv preprint arXiv:1904.01565* (2019).

Simone Riccardo, Enrico Conca, Vincenzo Sesta, Andreas Velten, and Alberto Tosi. 2022. 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* 22, 17 (2022), 16874–16885.

Diego Royo, Jorge Garcia-Pueyo, Miguel Crespo, Guillermo Enguita, Óscar Pueyo-Ciutad, and Diego Bielsa. 2025. mitransient: Transient light transport in Mitsuba 3. *arXiv preprint arXiv:2510.25660* (2025).

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

Siyuan Shen, Zi Wang, Ping Liu, Zhengqing Pan, Ruiqian Li, Tian Gao, Shiying Li, and Jingyi Yu. 2021. Non-line-of-Sight Imaging via Neural Transient Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43, 7 (July 2021), 2257–2268.

Talha Sultan, Alex Bocchieri, Chaoying Gu, Xiaochun Liu, Pavel Polynkin, and Andreas Velten. 2025. Optimized Sampling for Non-Line-of-Sight Imaging Using Modified Fast Fourier Transforms. *arXiv preprint arXiv:2501.05244* (2025).

Talha Sultan, Eric Brandt, Khadijeh Masumnia-Bisheh, Simone Riccardo, Pavel Polynkin, Alberto Tosi, and Andreas Velten. 2026. Iterating the transient light transport matrix for non-line-of-sight imaging. *Nature Communications* (2026).

Jeremy A Teichman. 2019. Phasor field waves: a mathematical treatment. *Optics Express* 27, 20 (2019), 27500–27506.

Chia-Yin Tsai, Kiriakos N Kutulakos, Srinivasa G Narasimhan, and Aswin C Sankaranarayanan. 2017. The geometry of first-returning photons for non-line-of-sight imaging. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 7216–7224.

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

Zewei Wang, Xiaoyin Li, Mingbo Pu, Lianwei Chen, Fei Zhang, Qi Zhang, Zhibin Zhao, Longfei Yang, Yinghui Guo, and Xiangang Luo. 2024. Vectorial-Optics-Enabled Multi-View Non-Line-Of-Sight Imaging with High Signal-To-Noise Ratio. *Laser & Photonics Reviews* (2024), 2300909.

Shumian Xin, Sotiris Nousias, Kiriakos N Kutulakos, Aswin C Sankaranarayanan, Srinivasa G Narasimhan, and Ioannis Gkioulekas. 2019. A theory of Fermat paths for non-line-of-sight shape reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 6800–6809.

Juntian Ye, Yu Hong, Xiongfei Su, Xin Yuan, and Feihu Xu. 2024. Plug-and-Play Algorithms for Dynamic Non-line-of-sight Imaging. *ACM Transactions on Graphics* 43, 5 (June 2024), 155:1–155:12.

Shinyoung Yi, Donggun Kim, Kiseok Choi, Adrian Jarabo, Diego Gutierrez, and Min H Kim. 2021. Differentiable Transient Rendering. *ACM Transactions on Graphics* 40, 6 (2021), 1–11.

SUPPLEMENTAL MATERIAL
Cascaded Non-Line-of-Sight Imaging

DIEGO ROYO*✉, Universidad de Zaragoza–I3A, Spain
MARÍA PEÑA*, Universidad de Zaragoza–I3A, Spain
FORREST B. PETERSON, University of Wisconsin–Madison, United States of America
ANDREAS VELTEN, University of Wisconsin–Madison, United States of America
JULIO MARCO, Universidad de Zaragoza–I3A, Spain
DIEGO GUTIÉRREZ, Universidad de Zaragoza–I3A, Spain

A Fast computation of our virtual H' using convolutions

The computation of $H'_{a,b}(\mathbf{l}_a, \mathbf{s}_b, t)$ (Equation 11) could require heavy computation if implemented naively, due to the large dimensionality of the impulse response $H(\mathbf{l}_1, \mathbf{s}_1, t)$. Instead, we formulate both $\mathcal{L}_1$ and $\mathcal{S}_1$ propagations simultaneously with convolutions [Liu et al. 2020]. This allows us to efficiently compute $H'_{a,b}(\mathbf{l}_a, \mathbf{s}_b, t)$ for all pairs of points $\mathbf{l}_a, \mathbf{s}_b$ in parallel.

We consider the XYZ coordinates of points on the relay wall $\mathfrak{R}_1$, noted as $\mathbf{l}_1 = [l_x, l_y, 0]$ and $\mathbf{s}_1 = [s_x, s_y, 0]$, on a plane at depth $z = 0$. The corresponding points on the secondary relay walls, $\mathbf{l}_a = [l'_x, l'_y, l'_z]$ on $\mathfrak{R}_a$ and $\mathbf{s}_b = [s'_x, s'_y, s'_z]$ on $\mathfrak{R}_b$, are both on planes at depths l'_z and s'_z. With this, we define a 4D convolution kernel $\widehat{C}$ using RSD operators $\widehat{R}$ from Equation 5:

$$\begin{aligned} &\widehat{C}(d_{lx}, d_{ly}, d_{sx}, d_{sy}; l'_z, s'_z, \Omega) \\ &\quad \equiv \widehat{R}\left(\sqrt{d_{lx}^2 + d_{ly}^2 + (l'_z)^2}/c, \Omega\right) \widehat{R}\left(\sqrt{d_{sx}^2 + d_{sy}^2 + (s'_z)^2}/c, \Omega\right), \end{aligned} \tag{S.1}$$

with $d_{lx} = l'_x - l_x$, $d_{ly} = l'_y - l_y$, $d_{sx} = s'_x - s_x$, and $d_{sy} = s'_y - s_y$ representing distances in the XY dimensions, and parameterized by the plane depths l'_z, s'_z and imaging frequency Ω. This allows us to express the frequency-domain counterpart of Equation 11 combining both 2D propagations into a 4D spatial convolution as:

$$\begin{aligned} &\widehat{H'}_{a,b}(l'_x, l'_y, l'_z, s'_x, s'_y, s'_z, \Omega) \\ &\quad = \widehat{\Phi}_{\mathrm{tc}}\left(s'_x, s'_y, s'_z, \Omega; \widehat{\Phi}_{\mathrm{tp}}\left(l'_x, l'_y, l'_z, \Omega; \widehat{H}\right)\right) \\ &\quad = \iiiint_{-\infty}^{+\infty} \widehat{C}\left(l'_x - l_x, l'_y - l_y, s'_x - s_x, s'_y - s_y; l'_z, s'_z, \Omega\right) \\ &\qquad\qquad \widehat{H}(l_x, l_y, 0, s_x, s_y, 0, \Omega)\, \mathrm{d}l_x\, \mathrm{d}l_y\, \mathrm{d}s_x\, \mathrm{d}s_y \\ &\quad = \widehat{C}(l_x, l_y, s_x, s_y; l'_z, s'_z, \Omega) *_s \widehat{H}(l_x, l_y, 0, s_x, s_y, 0, \Omega), \end{aligned} \tag{S.2}$$

where $*_s$ denotes a 4D spatial convolution over the XY dimensions of both the camera $\mathcal{S}_1$ and illumination $\mathcal{L}_1$ apertures l_x, l_y, s_x, s_y, which becomes a multiplication in the Fourier domain.

Note that this propagation approach also works if the two walls are not coplanar. Assuming for instance that two walls $\mathfrak{R}_1$ and $\mathfrak{R}_2$ are perpendicular (as in Figure 4), we can create a series of virtual walls $\mathfrak{R}_{2i}$ parallel to $\mathfrak{R}_1$, where each one includes a column of points of $\mathfrak{R}_2$. We then apply our propagation over multiple $\mathfrak{R}_{2i}$, discarding points that do not belong to $\mathfrak{R}_2$.

B Resolution limits in cascaded NLOS imaging

We analyze the resolution of our cascaded imaging systems, extending prior work on third-bounce NLOS imaging systems [Buttafava et al. 2015; Liu et al. 2019b; O'Toole et al. 2018]. We split our analysis in two parts. First, in Section B.1, we reason about the spatial and temporal resolution of our virtual impulse response H'. Second, in Section B.2, we analyze how the resolution of our virtual impulse response affects the imaging resolution of secondary virtual imaging systems and thus of the entire cascaded imaging system.

B.1 Resolution of our virtual impulse response

First, the primary relay wall acts as a virtual imaging system that reconstructs the virtual impulse response H' at a secondary relay wall (Equation 11). We model the resulting imaging resolution of the primary relay wall using the Rayleigh criterion, akin to conventional third-bounce methods.

Spatial and temporal resolution. The Rayleigh criterion relates the angular resolution θ of an imaging system to its aperture size D and the imaging wavelength λ:

$$\theta = k\frac{\lambda}{D}, \tag{S.3}$$

where k is a constant that depends on the shape of the aperture (e.g., for a circular aperture $k \approx 1.22$). We can use a small-angle approximation ($\sin\theta \approx \theta$) to establish a bound for the lateral resolution Δx of our imaging system at a given depth z from the relay wall:

$$\Delta x \approx k\frac{\lambda z}{D}. \tag{S.4}$$

For our virtual imaging systems, apertures are typically square-shaped with a side length of D, where $k = 1/\sqrt{2}$ has empirically provided a good estimate of the resolution.

We use Δx from Equation S.4 to model the spatial resolution of our virtual impulse response $H'(\mathbf{l}_a, \mathbf{s}_b, t)$. Thus, when reconstructing H' at several points $\mathbf{l}_a$ or $\mathbf{s}_b$ these can be separated by a distance of Δx. Empirically, we found that the temporal resolution of H' can be

*Both authors contributed equally to this research.

Authors' Contact Information: Diego Royo, droyo@unizar.es, Universidad de Zaragoza–I3A, Spain; María Peña, mpena@unizar.es, Universidad de Zaragoza–I3A, Spain; Forrest B. Peterson, fpeterson2@wisc.edu, University of Wisconsin–Madison, United States of America; Andreas Velten, velten@wisc.edu, University of Wisconsin–Madison, United States of America; Julio Marco, juliom@unizar.es, Universidad de Zaragoza–I3A, Spain; Diego Gutiérrez, diegog@unizar.es, Universidad de Zaragoza–I3A, Spain.

modeled as $\Delta t = \Delta x / c$, with c the speed of light. Later, we show an experiment which validates this last claim.

Scanning spacing constraints. The spacing of points $\mathbf{s}_1$ at the relay walls sets a theoretical minimum for imaging wavelengths λ. By the Nyquist-Shannon sampling theorem, to reproduce a phasor field $\widehat{H}(\mathbf{l}_1, \mathbf{s}_1, \Omega)$ with wavelength $\lambda = c/\Omega$ over a sparsely sampled aperture (here $\mathbf{s}_1 \in \mathcal{S}_1$ on the relay wall), points $\mathbf{s}_1$ must be spaced by a distance $\Delta \mathbf{s}_1 \leq \lambda/2$. For the phasor-field framework, the scanning resolution effectively provides a lower bound for λ_c in the filter $K(t; \lambda_c, \sigma)$ (Equation 3). In practice, λ typically exceeds this limit to avoid spatial aliasing of the virtual phasor waves. For the cases where we also focus the illumination aperture $\mathcal{L}_1$ (e.g., Figures 5d, 13 and 15), the same analysis applies to points $\mathbf{l}_1 \in \mathcal{L}_1$ and their spacing $\Delta \mathbf{l}_1$.

Resolution constraints using our real prototype. Our current prototype relies on a 16×16 SPAD array for its illumination aperture $\mathcal{L}_1$. Due to physical constraints in our laboratory, our SPAD array only covers a 50 cm $\times$ 32.5 cm area of the relay wall (i.e., $D = 32.5$ cm). Given that imaging with higher-order illumination requires longer imaging wavelengths ($\lambda \geq 10$ cm for our scenes and capture prototype), plugging both D and λ into Equation S.4 gives us a resolution of $\Delta x \approx 22$ cm at $z = 1$ m. In practice, that resolution allows us to image simple objects in our harder higher-order imaging tasks (see Section 8 of the main paper). Even if our 16×16 SPAD array covered the full 1.9 m $\times$ 1.9 m relay wall, the loss in detail from longer wavelengths would offset a large part of the resolution gains the larger aperture would have provided. For comparison, our simulations use a 32×32 SPAD array spanning the entire 2 m $\times$ 2 m relay wall, resulting in a sharper resolution of $\Delta x \approx 3.5$ cm. As we discuss in Section 9 of the main paper, our prototype demonstrates a promising proof-of-concept that will scale significantly with future SPAD technology.

Validation. Here, we validate our virtual impulse response functions computed at secondary relay walls under experimental data captured with our real hardware prototype, and support it with simulations and our theoretical resolution analysis. Figure S.1a (top) shows an overview of the real NLOS scene we use for this validation, with the relay wall $\mathfrak{R}_1$, a hidden secondary wall $\mathfrak{R}_2$, and a hidden T-shaped object. We capture the impulse response H at $\mathfrak{R}_1$. Following Section 4.2 of the main paper, we reconstruct the virtual impulse response $H'_{2,2}$ on the secondary wall $\mathfrak{R}_2$.

To validate our reconstructed $H'_{2,2}$, we simulate an equivalent captured impulse response function $H_{2,2}$ as if $\mathfrak{R}_2$ itself were the *first* relay wall, by pointing the laser and sensor directly to $\mathfrak{R}_2$, as shown in Figure S.1a (bottom). As stated in the main paper and detailed in Section E, all our simulations include the characteristic temporal uncertainty and noise of real capture systems.

Figure S.1b (top) shows a sample of our virtual impulse response function $H'_{2,2}$ reconstructed at $\mathfrak{R}_2$. Figure S.1b (bottom) shows the corresponding sample of the equivalent captured impulse response function $H_{2,2}$ at $\mathfrak{R}_2$. We verify that both $H'_{2,2}$ and its equivalent $H_{2,2}$ have peaks aligned in time, that correspond to fifth-bounce illumination at $H'_{2,2}$ (top, yellow) and third-bounce illumination at $H_{2,2}$

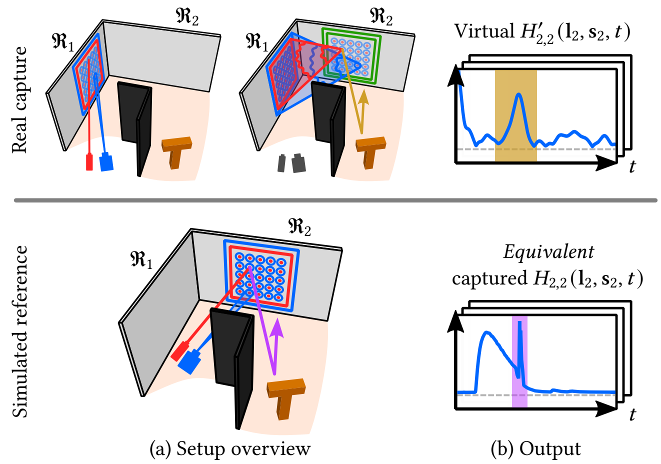


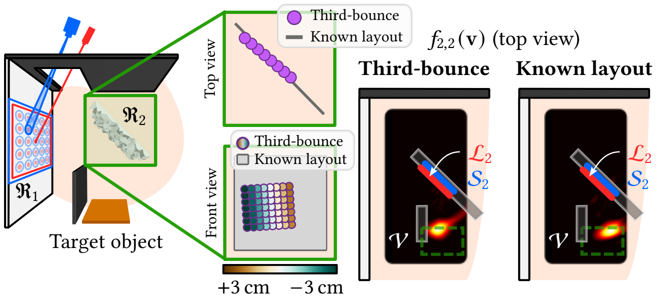
Fig. S.1. Comparison of our reconstructed virtual impulse response $H'_{2,2}$ at the wall $\mathfrak{R}_2$ with the equivalent captured impulse response $H_{2,2}$ for the setup in Figure 4 of the main paper. (a) Imaging procedures to obtain each impulse response. The top row virtually illuminates and captures points $\mathbf{l}_2$ and $\mathbf{s}_2$ of $\mathfrak{R}_2$ using our transient projector Φ_{tp} and camera Φ_{tc} operators, respectively. The virtual impulse response uses fifth-bounce illumination from the T letter, highlighted in yellow. In the bottom row, we simulate an equivalent scene where we move the laser and SPAD to point directly towards $\mathfrak{R}_2$. This results in third-bounce illumination from the T letter, highlighted in purple. (b) Resulting impulse responses for each case.

(bottom, purple) from the T-shaped object. The peak in $H_{2,2}$ (bottom, purple) has an 11 cm Full-Width at Half Maximum (FWHM), which depends on the system temporal uncertainty of our simulated NLOS imaging system pointed at $\mathfrak{R}_2$. The peak in our virtual impulse response function $H'_{2,2}$ (top, yellow), computed from the real captured H using $\lambda_c = 15$ cm, has a 40 cm FWHM. This wider peak of $H'_{2,2}$ follows Equation S.4: plugging $k = 1/\sqrt{2}$ (square aperture), $\lambda = \lambda_c = 15$ cm, $z = 1.2$ m (the distance between $\mathfrak{R}_1$ and $\mathfrak{R}_2$) and $D = 32.5$ cm (shortest side of the area covered by the SPAD array), yields $\Delta x \approx 39.1$ cm, which closely matches the measured 40 cm.

B.2 Resolution of our cascaded imaging systems

Our cascaded imaging systems use our virtual impulse response H' at secondary relay walls to form secondary imaging systems (Equation 14). We measure the resolution of the entire cascaded system using again the Rayleigh criterion (Equation S.4) for the secondary imaging system instead, while also taking into account aperture size constraints and resolution constraints from the resolution of our virtual impulse response H'.

Scanning spacing constraints. The Nyquist-Shannon sampling theorem also applies to the scanning spacing at secondary relay walls, when focusing secondary apertures $\mathcal{S}_b$ or $\mathcal{L}_a$. In our cascaded imaging systems, experiments showed that using a scanning resolution $\Delta \mathbf{s}_b = \Delta \mathbf{s}_1$ and $\Delta \mathbf{l}_a = \Delta \mathbf{l}_1$ for secondary apertures $\mathcal{S}_b$ and $\mathcal{L}_a$ works well in most cases. Finer sampling with shorter $\Delta \mathbf{s}_b$ or $\Delta \mathbf{l}_a$ did not seem to provide any improvements. For experiments with large values of λ_c, we use coarser resolution with larger $\Delta \mathbf{s}_b$ and $\Delta \mathbf{l}_a$ to reduce execution times.



(a) Setup overview (b) Estimation error (c) Final images

Fig. S.2. (a) We replicate the scene from Figure 15 in the main paper. (b) We compare our estimation of the secondary relay wall $\mathfrak{R}_2$ using a third-bounce method against a known layout obtained via 3D scans. (c) We analyze the impact of estimation error in the final cascaded imaging results $f_{2,2}(\mathbf{v})$, and observe both methods show very similar results.

Effective aperture constraints. Consider the case where we reconstruct $H'_{a,b}(\mathbf{l}_a, \mathbf{s}_b, t)$ at points $\mathbf{s}_b \in \mathcal{S}_b$ on a secondary relay wall with aperture $\mathcal{S}_b$ with side length D. Because surfaces can exhibit *virtual specular* behavior (see Appendix D), some points $\mathbf{s}_b$ might reflect light away from the hidden object. Thus, only part of $\mathcal{S}_b$ has usable signal, which we denote as the effective aperture size $\hat{D}$. To get the exact imaging resolution, one can plug $\hat{D}$ into Equation S.4 instead of D. However, note that exact computation of $\hat{D}$ depends on the scene and the choice of secondary aperture $\mathcal{S}_b$. For most scenes we tested in our work, our rule of thumb is that $\hat{D}$ is roughly half of D.

Resolution constraints. Similar to how the temporal resolution of H imposes a bound on imaging resolution [O'Toole et al. 2018], the spatial Δx and temporal Δt resolutions of our virtual impulse response H' affect our cascaded imaging resolution. Exact computation of imaging resolution as in Equation S.4 depends on the layout of the hidden scene, and thus we do not give a one-size-fits-all formula. To design our experiments, we estimate a lower bound of the minimum resolvable distance $\Delta x'$ of the entire system as $\Delta x' \geq c\Delta t/2 \approx \Delta x/2$.

C Impact of estimation error of secondary relay walls

In Section 4.3 of the main paper, we discuss our strategies to select secondary relay walls for our cascaded imaging systems without needing to know the layout of the hidden scene beforehand. These strategies use third-bounce methods to image potential secondary relay walls from the primary relay wall $\mathfrak{R}_1$. Here, we evaluate whether errors in these location estimates actually affect the final image quality, using the NLOS scenario from Figure 15 of the main paper as a practical example.

We show an overview of the setup in Figure S.2a. To obtain the position of the secondary wall $\mathfrak{R}_2$, we test two settings: (i) we image the hidden scene using third-bounce methods (Equation 6), and then post-process the resulting image to obtain a set of points $\mathbf{l}_2, \mathbf{s}_2$ corresponding to $\mathfrak{R}_2$ (represented using triangles in Figure S.2b), and (ii) we manually specify the exact positions of $\mathbf{l}_2, \mathbf{s}_2$, based on prior knowledge of $\mathfrak{R}_2$ (here obtained via a 3D scan) but not of the target object. As shown in Figure S.2b, our third-bounce estimations fall within just ±3 cm of error.

For both settings, we use the same captured impulse response H and follow the same imaging procedure (please see Section 8.2 for the full procedure), only changing the position of $\mathbf{l}_2 \in \mathcal{L}_2$ and $\mathbf{s}_2 \in \mathcal{S}_2$. The imaging results from both methods are quite similar, as seen in Figure S.2c. In both cases, an object with similar shape appears on the expected location marked with $\mathcal{V}$ corresponding to the target hidden object.

D Virtual surface reflectance and the missing cone

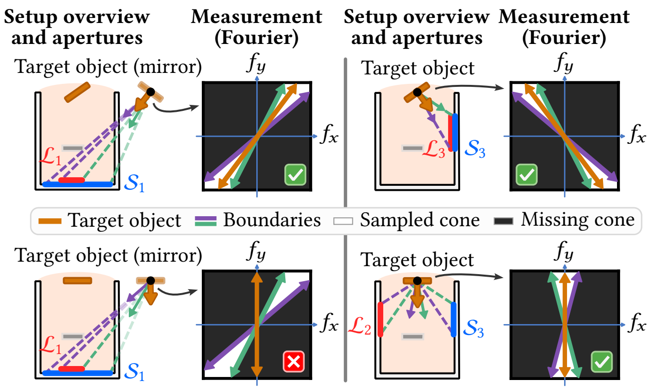
In this section, we expand on the analysis from Section 6 of the main paper. In our NLOS imaging setups, although all surfaces are diffuse under visible light, they may exhibit different reflectance properties when illuminated using phasor-field virtual waves. This effect, which we term *virtual surface reflectance*, depends on the subset of imaging frequencies Ω used in the NLOS imaging process. This virtual surface reflectance in general differs from the *real* reflectance properties of the same surfaces, which respectively depend on their micro- and nano-scale structure. Under the phasor-field formulation, the impulse response $H(\mathbf{l}_1, \mathbf{s}_1, t)$ (captured using nanometer-scale laser pulses) is typically convolved with a Morlet wavelet virtual illumination function $K(t; \lambda_c, \sigma)$ (Equation 3 of the main paper) of central wavelength λ_c and standard deviation σ. This function acts as a band-pass filter, with a Gaussian-shaped frequency spectrum centered at $\Omega_c = c/\lambda_c$, and width inversely proportional to σ. Due to limited sampling of the relay wall $\Delta\mathbf{s}_1$ (Section B) and temporal uncertainty of the imaging system (Section E), practical λ_c and σ values (expressed as spatial instead of temporal measures) are typically in the centimeter range [Liu et al. 2020, 2019b; Royo et al. 2023]. As a result of this convolution, the imaged objects exhibit virtual surface reflectance corresponding to a narrow set of centimeter-range wavelengths centered at λ_c, which is usually determined by the centimeter-scale structure of such surfaces.

In Section 6 of the main paper, we present two simulated experiments which illustrate the different reflectance behaviors between real and phasor illumination, and their relationship to the missing-cone problem and surface visibility. These insights are especially useful to understand our simulated and real NLOS imaging experiments in Sections 7 and 8 of our paper, respectively.

D.1 Missing cone under higher-order illumination

In this section, we extend the missing-cone analysis from Liu et al. [2019a] to our cascaded imaging systems, and show how to leverage our transient projector Φ_{tp} and camera Φ_{tc} operators.

Figure S.3 illustrates the scene originally presented in Figure 13 of the main paper. There, we show an overview comparing the NLOS imaging approach from Royo et al. [2023] with our own. We test for two orientations of the target object, only one of which is favorable to Royo et al.'s method. Specifically, we highlight the different illumination $\mathcal{L}$ and camera $\mathcal{S}$ apertures used for imaging. While the method proposed by Royo et al. can only use the apertures $\mathcal{L}_1$ and $\mathcal{S}_1$ in the primary relay wall $\mathfrak{R}$, we use our transient projector Φ_{tp} and camera Φ_{tc} operators to form new apertures $\mathcal{L}_2$, $\mathcal{L}_3$, and $\mathcal{S}_3$ on walls $\mathfrak{R}_2$ and $\mathfrak{R}_3$. We display the corresponding Fourier domain



(a) Virtual mirrors [Royo et al. 2023] (b) Our cascaded imaging systems

Fig. S.3. We extend the missing cone analysis from Liu et al. [2019a] to our scene in Figure 13 of the main paper, where we rotate the target object. The virtual mirrors method from Royo et al. [2023] and our method use different apertures $\mathcal{L}$ and $\mathcal{S}$, and as a result each samples different Fourier cones of the NLOS measurement function. (a) The virtual mirrors method is fixed to $\mathcal{L}_1$ and $\mathcal{S}_1$. In this case, the sampled Fourier cone only works for setups with specific orientations, and thus fails on the bottom row. (b) Our method can create apertures at different parts of the hidden scene ($\mathcal{L}_2$, $\mathcal{L}_3$ and $\mathcal{S}_3$), which in turn sample different parts of the NLOS measurement function. Note how our method can adapt to each configuration, lifting previous restrictions about specific orientations.

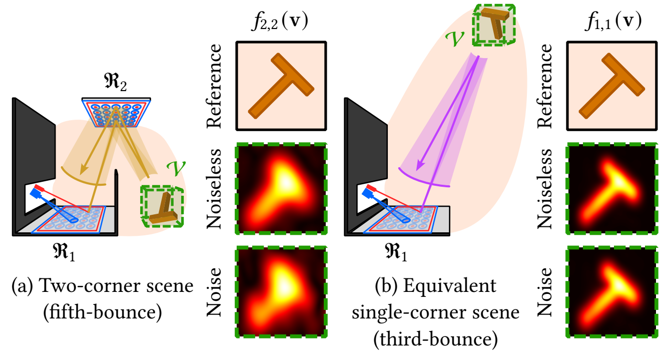


Fig. S.4. (a) We compare the resolution of our fifth-bounce imaging using $\mathfrak{R}_2$ and (b) the corresponding third-bounce imaging where the object at $\mathcal{V}$ is at a similar distance from the relay wall $\mathfrak{R}_1$. We simulate both captures with and without noise. The imaging resolution loss from using multiple apertures in $\mathfrak{R}_1$ and $\mathfrak{R}_2$ in the two-corner scene (compared to the single-corner scene) is much more notable than the effect of noise.

representation of the NLOS measurement function for the hidden target [Liu et al. 2019a]. While the exact measured time response is a projection along various ellipsoids with foci at points $\mathbf{l} \in \mathcal{L}$ and $\mathbf{s} \in \mathcal{S}$ (corresponding to purple/green dashed lines), we follow Liu et al. and approximate these projections using planes (corresponding to purple/green arrows). Due to finite aperture sizes, large portions of the Fourier spectrum (represented in black) are not contained in any measurements.

The method proposed by Royo et al. and our own utilize different aperture configurations, resulting in the sampling of distinct regions within the NLOS measurement function. Notably, only our cascaded imaging system samples the specific Fourier components relevant to the hidden target for all cases. Consequently, while our method does not inherently solve the missing-cone problem, we can circumvent it thanks to our Φ_{tp} and Φ_{tc} operators, whereas the approach by Royo et al. remains susceptible to it.

E Realistic capture noise models for simulation

This section details how we model realistic noise in our simulations to mimic real capture conditions. In particular, we mimic photon counts, hardware response, and calibration errors from previous works [Hernandez et al. 2017; Liu et al. 2019b; Royo et al. 2023] and from our own hardware described in Section 8 of the main paper.

We observe that our cascaded NLOS imaging methods are robust to capture noise. In Figure S.4a, we simulate an object hidden around two corners and image it using a secondary virtual imaging system at $\mathfrak{R}_2$. In Figure S.4b we simulate the equivalent single-corner setup, placing the hidden object at the same distance so that values in Equation S.4 match as much as possible between both scenes. We simulate captures with and without noise. We show that both third-bounce imaging methods and our fifth-bounce imaging method remain robust to noise, while the main change in imaging resolution comes from the fact that fifth-bounce imaging is a much harder problem than the conventional third-bounce problem.

Photon counts and signal-to-noise ratio. Following prior NLOS imaging work [Liu et al. 2019b; Royo et al. 2023] and measurements from our own hardware prototype, we assume that for each illuminated position $\mathbf{l}_1$ the sensor captures approximately 10^9 photons; in those works the entire capture process requires about three minutes of exposure time. In our simulated experiments, we add Poisson noise to our simulated impulse response function $H(\mathbf{l}_1, \mathbf{s}_1, t)$ to mimic similar capture conditions. We experimentally observed that the resulting SNR under such conditions is high enough for our cascaded NLOS imaging method, as seen in Figures 13 and 15 of the main paper.

Temporal uncertainty. Photon arrival times estimated by NLOS imaging devices have inherent temporal uncertainty due to the limitations of laser and sensor hardware. We mimic temporal uncertainty on our simulated data by applying a convolution over the time domain between the impulse response function $H(\mathbf{l}_1, \mathbf{s}_1, t)$, the real measured response of a $20\mu m$ CMOS single photon avalanche diode (SPAD) with a temporal resolution of 25 ps of full width at half maximum (FWHM) [Hernandez et al. 2017], and a Gaussian-shaped laser pulse with 35 ps of FWHM [Royo et al. 2023]. The resulting virtual impulse response functions $H'_{a,b}(\mathbf{l}_a, \mathbf{s}_b, t)$ calculated by our cascaded imaging method are thus affected by this temporal uncertainty, along with any inaccuracies inherent to NLOS imaging resolution (Appendix B).

Calibration error. The real laser emission and sensor pixels of capture devices are not targeted at perfectly zero-dimensional points $\mathbf{l}_1, \mathbf{s}_1$ on the relay wall, but instead have a finite footprint that may introduce spatial inaccuracies during RSD propagation. We model these inaccuracies by randomly jittering the laser and sensor locations $\mathbf{l}_1, \mathbf{s}_1$ on the relay wall using a Gaussian distribution with standard deviation $\sigma_E = 1$ cm, following real capture conditions of previous works [Royo et al. 2023].

Multi-path interference. NLOS capture devices measure only photon arrival times, leading to multi-path interference (MPI) when two photons follow different paths but arrive simultaneously. Our simulations include MPI from all possible light paths in the hidden scene (see Figures 5 and 8 of the main paper), and our cascaded method remains robust to those, even in real-world cases (Figures 13 and 15 of the main paper).

F Setup photographs

Here we include photographs of each scene setup in Figure S.5. Our hardware prototype, visible in Figure S.5a, uses a 16×16 SPAD array sensor focused on a $50\,\mathrm{cm} \times 32.5\,\mathrm{cm}$ area on $\mathfrak{R}_1$ (red rectangle in the scans in Figures S.5b and c), and a laser that covers a $1.9\,\mathrm{m} \times 1.9\,\mathrm{m}$ area on $\mathfrak{R}_1$ (blue rectangle in the scans in Figures S.5b and c). We use colored tape on the floor to mark the position of hidden objects; this does not have any practical effect on the captured signal.

References

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

Quercus Hernandez, Diego Gutierrez, and Adrian Jarabo. 2017. A computational model of a single-photon avalanche diode sensor for transient imaging. *arXiv preprint arXiv:1703.02635* (2017).

Xiaochun Liu, Sebastian Bauer, and Andreas Velten. 2019a. Analysis of feature visibility in non-line-of-sight measurements. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* 10140–10148.

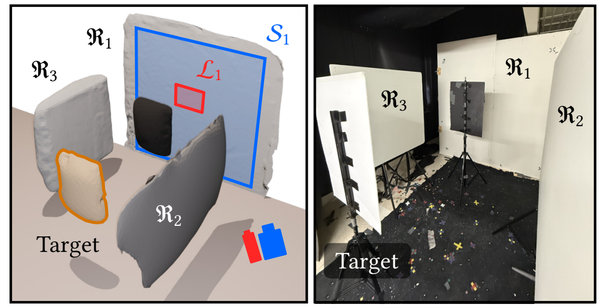
Xiaochun Liu, Sebastian Bauer, and Andreas Velten. 2020. Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems. *Nature Communications* 11, 1 (2020), 1645.

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

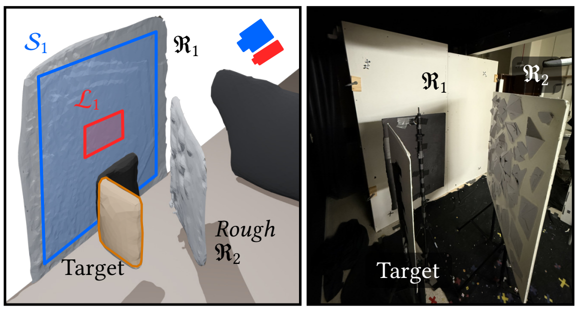
Matthew O'Toole, David B Lindell, and Gordon Wetzstein. 2018. Confocal non-line-of-sight imaging based on the light-cone transform. *Nature* 555, 7696 (2018), 338.

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

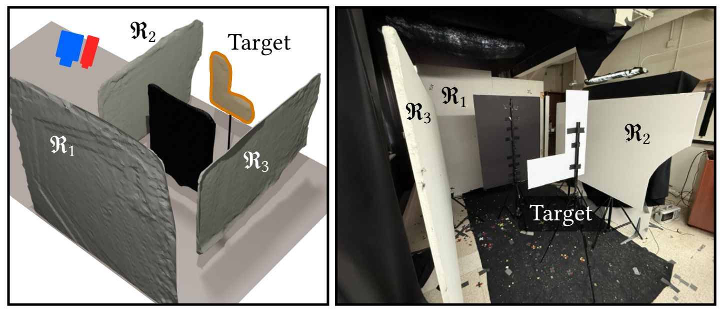
(a) Setup for Figure 11 (third- and fourth-bounce imaging)



(b) Setup for Figure 13 (fifth-bounce imaging)



(c) Setup for Figure 15 (fifth-bounce imaging)



(d) Setup for Figure 16 (fifth-bounce imaging)

Fig. S.5. 3D scans and photographs for the setups using our real prototype in Section 8. Our hardware prototype, visible in (a), is pointed towards the relay wall $\mathfrak{R}_1$. In the 3D scans, we highlight the hidden target objects in orange, and the areas where the SPAD array (red, $\mathcal{L}_1$) and laser (blue, $\mathcal{S}_1$) are focused (b, c).