Zero-Phase Phasor Fields for Non-Line-of-Sight Imaging

Pablo Luesia-Lahoz, Talha Sultan, Forrest B. Peterson, Andreas Velten, Diego Gutierrez, and Adolfo Muñoz

Abstract—Non-line-of-sight imaging employs ultra-fast illumination and sensing devices to reconstruct scenes outside their line of sight by analyzing the temporal profile of indirect scattered illumination on a secondary relay surface. Commonly, the NLOS methods transform the temporal domain into the frequency domain and operate on it, and then identify surface locations by locating the maxima in amplitude along the reconstruction volume. Phase information, which is virtual as it results from a Fourier transform, is very often discarded or ignored. We incorporate phase information into our novel Zero-Phase Phasor Fields imaging technique, which we derive for a confocal capture configuration. We show how, at positions that belong to the hidden geometry, we can ensure the phase is zero, so we can locate the hidden geometry with great precision by locating the zero crossings in the phase. This allows us to reconstruct at widely spaced locations and still achieve up to 125 micrometer depth precision, as our experimental validation shows with both synthetic and captured data, the latter publicly available. Moreover, the phase is robust to noise, as we demonstrate with decreasing signal-to-noise ratio using publicly available dataset captures of the same scene.

Index Terms—Computational Photography, Non-line-of-sight imaging, looking around corners, virtual wave optics.

1 Introduction

Recent advances in ultra-fast illumination and sensing devices make it possible to capture the propagation of light as a function of time, obtaining its time of flight (ToF) at an effective frame rate of up to trillions per second [1], [2], [3]. One of their multiple applications is non-line-of-sight (NLOS) imaging [4], [5], [6], [7], [8], [9], [10], [11], which aims to recover the geometry from partially or fully occluded scenes based on the indirect light that reaches a secondary surface or relay wall.

NLOS imaging algorithms usually reconstruct a predefined three-dimensional volume to search for hidden geometry. This volume is often expressed as a 3D point grid, which is equivalent to a regular set of 2D planes parallel to the relay wall. The reconstructed points with the highest values mark the approximate location of a scattering surface (hidden scene geometry). Higher precision is achieved by increasing resolution, either within all planes (horizontal and vertical resolution) or with the number of planes (depth resolution), at the cost of prohibitively increasing both memory usage and execution time.

A milestone in NLOS imaging methods is Phasor Fields (PF) [10], which poses NLOS reconstruction as a virtual wavefront propagation problem by virtually illuminating the hidden scene. This is done by transforming the time of flight to the frequency domain, which is more resistant to noise and increases the reconstruction speed due to the use of the Fast Fourier Transform. The reconstruction value at each point comes from the amplitude of the virtually-propagated waves, while phase information is usually discarded.

We introduce a novel method that leverages the virtual phase information of PF to interpolate between different reconstruction points. Thanks to the virtual wavefront propagation analogy (PF), we can utilize insights from imaging interferometry [12], [13], [14]. The key insight is that the phase of a propagated wavefront gives precise cues about the hidden surface position, even if computed at some distance from the true surface. This allows us to do coarser reconstructions, with a lower resolution in depth by using fewer planes, and then leverage phase information to correct the hidden surface positions according to it.

We make the key observation that we can control the virtual phase of the reconstructions at the hidden scene positions, which we set to zero, enabling us to perform the correction of the hidden surface locations by finding the nearest zero crossing in the phase information. For this reason, we name our method Zero-Phase Phasor Fields (ZPPF). In a first step, we apply standard PF with a low depth resolution. Then we develop a phase correction for confocal captures (i.e., capture and illumination located at the same position) [7] to correct the depth with the phase, obtaining a precision comparable to PF with a very dense depth resolution, which is orders of magnitude slower. In particular, our approach is sensitive to variations in depth as small as 150 micrometers, with the same execution times and complexity that previous methods employ for centimeter-scale results. Moreover, it outperforms previous methods in execution time and memory usage, while being extremely robust to noise in poor signal-to-noise (SNR) scenarios. We show results with publicly available real-world captures [9] and synthetic data, both captured in confocal configurations [7].

We believe that our approach represents a significant step forward in NLOS imaging techniques. Real-world applications where precise depth matters, such as non-invasive

- *P. Luesia-Lahoz, D. Gutierrez, and A. Muñoz are with the Graphics & Imaging Lab at the Universidad de Zaragoza - I3A.*
  E-mail: see http://p-luesia.github.io
- *T. Sultan, F. Peterson, and A.Velten are with the Computational Optics Group at the University of Wisconsin-Madison.*

medical imaging or exploration, may benefit from it in the future. Furthermore, its efficiency might make it suitable for larger-scale applications, such as car safety, by making the problem more manageable for embedded computing systems.

2 Related work

2.1 Non-Line-of-Sight Imaging

Transient Non-Line-of-Sight imaging, or just NLOS in this manuscript, employs ultrafast capture and illumination devices to leverage the time-of-flight of the indirect light to image hidden scenes [15]. The achievable depth resolution is fundamentally limited by the system's overall timing resolution [7]. In their work, Velten et al. demonstrated millimeter-scale depth of tracking of hidden objects over time, using picosecond-resolution systems [4]. In contrast, high-resolution depth imaging demands a dense depth map of an entire hidden scene, which presents greater challenges. By exploiting the first-returning photon geometry, previous approaches have achieved approximately 10 micrometer precision, at the cost of using femtosecond timing resolution and a computationally expensive optimization algorithm [16]. By limiting the capture to a confocal capture, i.e., the laser and capture are co-located, the Light Cone Transform [7] casts NLOS imaging as a deconvolution problem, getting approximately 2.5 millimeters depth precision via direct inversion. Another direct inversion is the fk-migration [9], which solves the problem efficiently in the frequency domain, using Stolt's interpolation, reaching 1.2 millimeter precision. Yet both approaches require a dense grid of reconstruction points and manipulating large Fourier-domain representations, making inversion impractical due to prohibitive memory demands. In this work, we present a linear inverse method with complexity comparable to LCT and fk-migration, achieving up to 125 micrometer scale precision, without incurring significant memory overhead. For that purpose, we build upon the Phasor fields framework [10], which poses the NLOS imaging problem as a wavefront propagation problem.

2.2 Phasor Field Framework

The Phasor Field (PF) framework [10] is a promising technique in NLOS imaging, effectively posing the relay wall as a virtual camera. The key observation is that by posing virtual wavefronts to represent the light propagation within the NLOS imaging scope, they can use well-known diffraction integrals from wave optics [17], [18], [19], [20], [21], [22], [23]. Thus, PF shares similarities with technologies applied in the line-of-sight scope, as the synthetic aperture radar community [24], specifically the THz community [25], [26]. Hidden scene geometry is reconstructed via computational back-propagation of these virtual waves [10], [11]. This formalism allows line-of-sight imaging concepts (e.g., from Fourier optics, cameras, phased arrays) to be employed for analyzing NLOS resolution [20], sampling criteria [27], extracting complex light transport phenomena [28] and finding mirror images of objects hidden around two corners [29]. Implementations include frequency-domain methods to improve computational speed [11], [30], [31] and adaptations for non-planar surfaces [32], [33] or memory-constrained scenarios [34], [35]. Notably, these PF reconstruction algorithms have primarily concentrated on estimating intensity distributions while ignoring phase information. We show that this information is pertinent for achieving high-precision depth recovery with fewer reconstructed samples.

3 Background: The Phasor Fields framework

To understand our work, first, it is key to understand the Phasor Fields (PF) framework [10]. The original authors present dual domains in the PF framework: the real and the virtual.

The real domain is equivalent for all transient NLOS imaging methods [4], [7], [9], [36], [37], Figures 1a and b show an overview. An ultra-fast illumination device, i.e., a laser, emits a short light pulse to a set of illumination positions $\mathbf{x}_l \in \mathcal{L}$ on the visible relay wall (Figure 1a). The light scatters from those $\mathbf{x}_l$ positions to the hidden scene, which scatters part of the illumination back onto the relay wall. An ultra-fast capture device focused on a set of positions $\mathbf{x}_s \in \mathcal{S}$ captures the incoming light, decoupled by its time-of-flight (ToF) (Figure 1b). Note that both $\mathcal{L}$ and $\mathcal{S}$ can represent the same positions in the relay wall.

The distribution of $\mathbf{x}_l \in \mathcal{L}$ and $\mathbf{x}_s \in \mathcal{S}$ defines the capture configuration. In this work, we employ a regular grid of sensing points $\mathcal{S}$ and $\mathcal{L}$ on the relay wall in all cases, limiting the capture to those points where $\mathbf{x}_l = \mathbf{x}_s$.

The illumination device is assumed to emit a delta illumination pulse in time $\delta(t)$, thus, it is possible to capture the impulse response of the scene as $H(\mathbf{x}_l, \mathbf{x}_s, t)$. For simplicity, we omit the light ToF from the devices to the relay wall, which is known since the relay wall is in the line of sight of the physical devices. It is possible to shift $H(\mathbf{x}_l, \mathbf{x}_s, t)$ in time to account for these ToFs.

The original authors define a virtual illumination pulse $\mathcal{P}(t)$ as a phasor such that

$$\mathcal{P}(\mathbf{x}_l, \mathbf{x}_s, t) = \mathcal{P}(t) * H(\mathbf{x}_l, \mathbf{x}_s, t), \tag{1}$$

where $*$ is a convolution in time (Figure 1c). This computation shifts the system from the real domain to the virtual domain, since the phasor $\mathcal{P}(\mathbf{x}_l, \mathbf{x}_s, t)$ is equivalent to virtually illuminating the scene with the phasor $\mathcal{P}(t)$.

In this virtual domain, the authors define the virtual illumination pulse $\mathcal{P}(t)$ as a Gaussian wave packet with a central wavelength of $\lambda_g = 2\pi c/\omega_g$, being ω_g its angular frequency, c the speed of light, and ρ being a scale factor, as

$$\mathcal{P}(t) = \rho e^{-\frac{t^2}{2\sigma^2}} e^{i\omega_g t + \phi_0}, \tag{2}$$

where σ controls the width of the Gaussian.

With this framework, the NLOS problem can be posed as a virtual wavefront propagating at the relay wall, which enables imaging $I(\mathbf{x}_v)$ on a point $\mathbf{x}_v$ using a lens focusing operator Φ:

$$I(\mathbf{x}_v) = \Phi[\mathcal{P}(\mathbf{x}_l, \mathbf{x}_s, t), \mathbf{x}_v]. \tag{3}$$

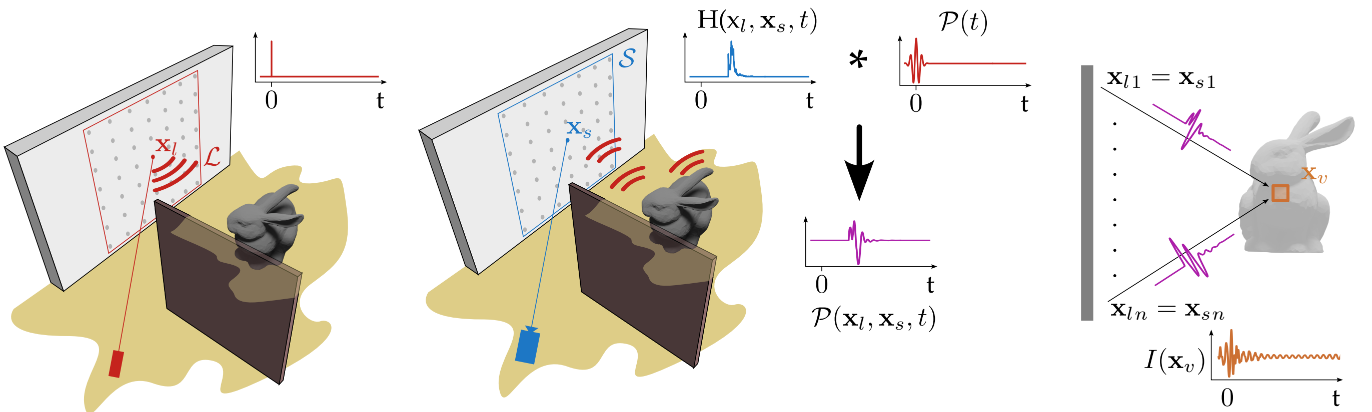


Fig. 1: Phasor Fields image formation overview. A delta pulse of light is emitted to a point $\mathbf{x}_l$ in the relay wall, where it scatters to the hidden scene (a). The hidden scene, scattered back to the relay wall part of the illumination, where a sensor (blue) captures its response over time (b). Computationally, the hidden scene is virtually illuminated by an input phasor, which transforms the problem into a virtual wavefront propagation (c). Phasor Fields reconstruct the hidden scene by propagating from the relay wall to an imaging point $\mathbf{x}_v$, which resembles the Rayleigh-Sommerfeld Diffraction integral [10] (d).

The authors propose a focusing for monochromatic sources based on the Rayleigh Sommerfeld Diffraction (RSD) integral as a forward operator as

$$\Phi\left[\hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega), \mathbf{x}_v\right] = \int_{\mathcal{L}} \frac{e^{i\omega\frac{r_l}{c}}}{r_l} \int_{\mathcal{S}} \frac{e^{i\omega\frac{r_s}{c}}}{r_s} \hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega) d\mathbf{x}_s d\mathbf{x}_l,$$
$$r_s = |\mathbf{x}_v - \mathbf{x}_s|$$
$$r_l = |\mathbf{x}_v - \mathbf{x}_l| \quad (4)$$

for a reconstruction point $\mathbf{x}_v$, where $\hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega)$ is a single ω frequency of the phasor $\mathcal{P}(\mathbf{x}_l, \mathbf{x}_s, t)$. The authors named Equation (4) as a confocal camera model (do not confuse this with confocal capture [7], [9]), and it brings the hidden geometries into focus.

PF takes all frequencies focused on $\mathbf{x}_v$, and, with a Fast Fourier Transform, yields the temporal image (Figure 1d). When evaluated at $t = 0$, this results in the hidden geometry acting as an emitter of a scaled phasor. Finally, to segment the empty space, PF relies on the amplitude value of those $\mathbf{x}_v$, assuming higher values correspond to a hidden geometry.

4 Our Method

As we describe in Section 3, the Phasor Fields framework virtually illuminates the scene with a Gaussian wave packet, which allows the original authors to pose a NLOS problem as a virtual wavefront propagation. Therefore, they image a single reconstruction point with a wavefront focus operator, as described by Equation (4). PF employs a grid of reconstruction points to image the complete hidden scene, assuming a high value in the amplitude denotes a nearby geometry, discarding its phase. Consequently, increasing the precision is costly since it involves incrementing the resolution (the number of computed reconstruction points).

Inspired by interferometry and especially the usage of phase for achieving nanometer-scale [12], [13], [14], we reconsider using the virtual phase obtained by PF for a novel reconstruction technique which refines the position of the reconstruction point to nearby geometry. Our proposal requires prior knowledge of the phase at the hidden surface's position, for estimating the distance correction of reconstruction points. For that purpose, we mathematically probe a key observation in Section 4.1: by ensuring certain properties of the virtual Gaussian wave packet, we can assert zero phase at the hidden geometry positions. Searching for those zero-phase positions increases our accuracy, which gives our novel approach the name Zero-Phase Phasor Fields (ZPPF).

Our ZPPF is a two-step algorithm that uses PF maxima to select candidate points near the hidden geometry, and then refines their position using phase information to match those zero phase values (see Section 4.2). We leverage PF efficient implementations for reconstruction points grouped by parallel planes to the relay wall [11], which we refer to as reconstruction planes. Figure 2 shows a summary of PF (left) and our ZPPF (right). In both cases, the distance between the reconstruction planes is in the range of centimeters, which we name sparse planes. We also term *dense* PF as a case of PF that achieves high precision in placing reconstruction planes on a micrometer scale. We show that our ZPPF computational cost is the same as PF for sparse reconstruction time, achieving an accuracy comparable to dense PF (micrometer-scale depth precision), as proven by our experiments in Section 5.1. On top of that, phase propagation is very robust to noise, as we later show in our experimental validation (see Section 5.2).

4.1 Analyzing phase

In this section, we analyze the phase behavior on a reconstruction point when it matches a hidden scene geometry, imaged with the virtual wavefront of the PF framework. Later, we prove we can ensure zero phase at those points, controlling the virtual Gaussian wave packet.

We start by the frequency domain version of Equation (1), for a particular frequency ω:

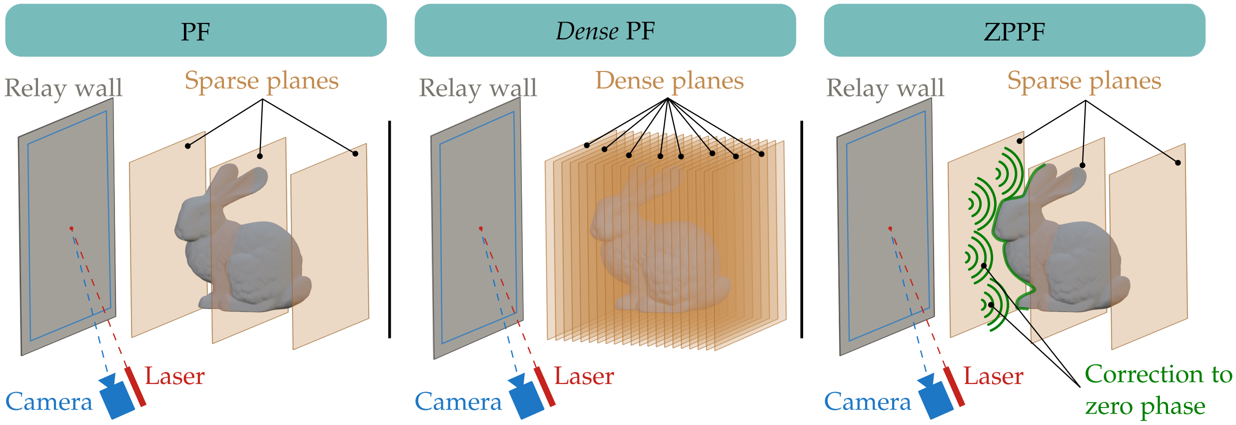
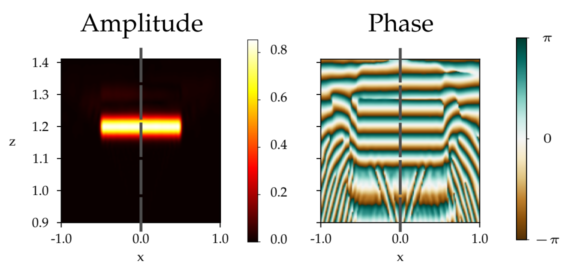
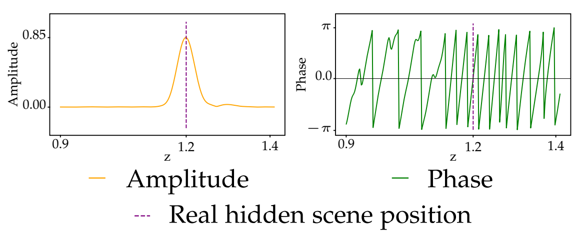


Fig. 2: Summary of our novel ZPPF (right) compared to previous PF (left). Efficient PF implementations optimized for parallel planes achieve centimeter scale using sparse reconstruction planes, i.e., spacing the planes by centimeters. To increase the PF depth precision, we establish the reconstruction planes in a dense configuration so the distance between them is in terms of micrometers. We name that configuration *dense* PF (center), and it is prohibitively costly compared with standard PF. In contrast, we present our ZPPF, which, in the same complexity as PF, achieves the same precision as in dense PF. To do so, ZPPF employs the phase information to refine the hidden geometry position given the reconstruction on the sparse planes.



(a) Top view



(b) 1D view

Fig. 3: Dense PF reconstruction of a plane at 1.2 meters from the relay wall. In (a), we illustrate the amplitude as a heatmap and the phase as a blue-white-brown. (b) plots the same values from the dashed line of (a), with the real ground truth position of the geometry (purple dashed). We used a Gaussian wave packet for the virtual illumination pulse of PF that has zero phase at $t = 0$. Hence, the reconstruction also has zero phase at the ground truth hidden geometry location, as Section 4.1 proves.

$$\hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega) = \hat{\mathcal{P}}(\omega)\hat{H}(\mathbf{x}_l, \mathbf{x}_s, \omega), \tag{5}$$

The virtual illumination pulse $\hat{\mathcal{P}}(\omega)$ now has the form

$$\hat{\mathcal{P}}(\omega) = A(\omega)e^{i\phi(\omega)} \tag{6}$$

where $A(\omega)$ represents amplitude and $\phi(\omega)$ phase, obtained from the Gaussian wave packet in the Fourier domain, described in Equation (2). In a confocal camera model (Equation (4)), the propagation of the waves is given by

$$\begin{aligned} &\Phi\left[\hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega), \mathbf{x}_v\right] \\ &= \int_{\mathcal{L}} \frac{e^{\frac{i\omega r_l}{c}}}{r_l} \int_{\mathcal{S}} \frac{e^{\frac{i\omega r_s}{c}}}{r_s} A(\omega)e^{i\phi(\omega)}\hat{H}(\mathbf{x}_l, \mathbf{x}_s, \omega) d\mathbf{x}_s d\mathbf{x}_l \\ &\qquad = A(\omega)e^{i\phi(\omega)}\Phi\left[\hat{H}(\mathbf{x}_l, \mathbf{x}_s, \omega), \mathbf{x}_v\right], \\ &\qquad\qquad r_s = |\mathbf{x}_v - \mathbf{x}_s|, \\ &\qquad\qquad r_l = |\mathbf{x}_v - \mathbf{x}_l|. \end{aligned} \tag{7}$$

which means that the RSD-based propagation of PF is independent of the virtual illumination pulse. Usually, a straightforward optimization discards the propagation of frequencies whose amplitude $A(\omega)$, due to the Gaussian wave packet, is close to zero.

Let's now analyze what happens to the phase at points $\mathbf{x}_t$ on the surface of the hidden geometry. Without loss of generality, let's assume a hidden scene consisting of a single point $\mathbf{x}_t$. Note that its impulse response would be a δ function in time, and therefore its frequency-domain impulse response is

$$\hat{H}(\mathbf{x}_l, \mathbf{x}_s, \omega) = \epsilon_h \, e^{-i\omega\frac{|\mathbf{x}_t - \mathbf{x}_l| + |\mathbf{x}_t - \mathbf{x}_s|}{c}}, \tag{8}$$

where ϵ_h is a scale factor for each $\mathbf{x}_l$ and $\mathbf{x}_s$ of $\hat{H}(\cdot)$ that accounts for the attenuation of light, and c is the speed of light.

Plugging Equation (8) into Equation (7), we get

$$\begin{aligned} &\Phi\left[\hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega), \mathbf{x}_v\right] = \\ &\int_{\mathcal{L}} \frac{e^{\frac{i\omega r_l}{c}}}{r_l} \int_{\mathcal{S}} \frac{e^{\frac{i\omega r_s}{c}}}{r_s} A(\omega)e^{i\phi(\omega)}\epsilon_h \, e^{-i\omega\frac{|\mathbf{x}_t - \mathbf{x}_l| + |\mathbf{x}_t - \mathbf{x}_s|}{c}} d\mathbf{x}_s d\mathbf{x}_l. \end{aligned} \tag{9}$$

When imaging point $\mathbf{x}_t = \mathbf{x}_v$, the phase shifts in Equation (9) (the exponents of e) from the impulse response cancel the phase shifts due to the RSD propagation, yielding

$$\Phi\left[\hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega), \mathbf{x}_v\right] = A(\omega)e^{i\phi(\omega)} \int_{\mathcal{L}} \frac{1}{r_l} \int_{\mathcal{S}} \frac{1}{r_s} \epsilon_h \, d\mathbf{x}_s d\mathbf{x}_l. \tag{10}$$

Therefore, the only phase shift corresponds to the virtual illumination pulse $\phi(\omega)$, the Gaussian wave packet.

As we control the Gaussian wave packet, we can define such a pulse that ensures phase zero at the reconstruction point $\mathbf{x}_v = \mathbf{x}_t$. Considering that PF ultimately applies an inverse Fourier transform for a reconstruction point $\mathbf{x}_v$, Equation (10) results in a scaled version of the original Gaussian wave packet in the time domain. Intuitively, this is equivalent to having the geometry of the hidden scene as an emitter of our virtual illumination pulse. Therefore, by ensuring that the virtual Gaussian wave packet has zero phase ($\phi(\omega) = 0$) at $t = 0$, we can assert that the imaged phase at $\mathbf{x}_t$ is also zero. We show an example of this with a dense PF reconstruction in Figure 3, establishing a zero phase for the reconstruction surface.

4.2 Zero-Phase Phasor Fields

Conventionally, amplitude-based PF needs to rely on very dense reconstruction planes to yield precise reconstructions; this in turn imposes very large computational costs. Our Zero-Phase Phasor Fields (ZPPF) method works with a regular, small number of reconstruction planes. Once we have obtained an approximate solution from amplitude values (as in conventional PF methods) we refine it by looking for nearby zeroes on the phase of the virtual wavefront. Figure 2 shows a summary comparing our ZPPF with PF.

Virtual Gaussian wave packet definition. To ensure the geometry in the hidden scene is imaged with zero values, we set our virtual Gaussian wave packet with zero phase at $t = 0$ (see Section 4.1). To represent the 99.73% of the Gaussian wave packet, we use the 3σ rule from Gaussian distributions, where σ is the standard deviation of the distribution. We therefore set n, a real positive number of cycles of the wavelength, such $n\lambda_g = 6\sigma$, where λ_g is the central wavelength of the packet, to represent a pulse of width 6σ.

Sparse reconstruction planes. To make sure that for every reconstructed plane there is only one nearby zero crossing in phase, we need to make sure that there is less than one wave period between planes. We thus set a maximum distance between sampled planes of $\lambda_g/2$. This is still a sparse reconstruction, as λ_g is usually defined in a range of centimeters.

Phase correction. We then identify the maxima in amplitude in the sparse reconstruction. Note that, as we leverage efficient PF by planes, our maxima $\mathbf{x}_v$ correspond to the maximum amplitude along the depth. For an imaged reconstruction point $\mathbf{x}_v$ one of those maxima in amplitude, the phase $\phi(\mathbf{x}_v) \in (-\pi, \pi]$ is the phase of the reconstruction $\Phi\left[\hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega), \mathbf{x}_v\right]$. In order to find the closest zero crossing in phase ($\phi(\mathbf{x}_t) = 0$), we apply Netwon's method for a scalar function with a vector argument, starting from $\phi(\mathbf{x}_v)$:

$$\begin{aligned} \mathbf{x}_0 &= \mathbf{x}_v \\ \mathbf{x}_{k+1} &= \mathbf{x}_k - \frac{\phi(\mathbf{x}_k)}{||\nabla\phi(\mathbf{x}_k)||^2}\nabla\phi(\mathbf{x}_k), \end{aligned} \tag{11}$$

where $\mathbf{x}_t \approx \mathbf{x}_n$ is obtained after n iterations. In our experiments we have only needed a single iteration ($n = 1$).

To solve Equation (11) we need to calculate the gradient of the phase $\nabla\phi(\mathbf{x}_v)$, which is related to the wave vector. In the general case, calculating $\nabla\phi(\mathbf{x}_v)$ involves differentiating the propagation of Equation (4). In our captures, we use a confocal setup, in which illumination and capture occurs on the same positions ($\mathbf{x}_l = \mathbf{x}_s$). For confocal capture setups, the propagation is perpendicular to the relay wall ($\mathbf{z}$ direction), and so the whole integral is simplified to the product of two planar waves that follow the same direction, so that $\nabla\phi(\mathbf{x}_v) = -2\omega_g\mathbf{z}/c$. We can therefore correct the phase from a single iteration of Equation (11) as

$$\begin{aligned} \mathbf{x}_t = \mathbf{x}_v + \frac{\phi(\mathbf{x}_v)c^2}{4\omega_g^2}\frac{2\omega_g\mathbf{z}}{c} &= \mathbf{x}_v + \phi(\mathbf{x}_v)\mathbf{z}\frac{c}{2\omega_g} \\ &= \mathbf{x}_v + \phi(\mathbf{x}_v)\mathbf{z}\frac{\lambda_g}{4\pi}, \end{aligned} \tag{12}$$

which we use to propagate per-plane maxima at $\mathbf{x}_v$ towards the closest zero-crossing in phase. This results in a corrected position of the reconstruction point, without additional imaging computation.

5 Results and Validation

We next show results illustrating two key characteristics of our approach: increased depth precision and increased robustness in low signal-to-noise (SNR) conditions. The virtual illumination employed in all reconstructions consists of a Gaussian wave packet with a central wavelength of $\lambda_g = 0.08$ m, and a width of $n = 5$ cycles (see Section 4.2). For this parametrization, we set a distance between reconstruction planes as 0.03 m (smaller than $\lambda_g/2$) for the sparse reconstruction planes of ZPPF. We use the same sparse planes to reconstruct with PF for same-time comparison purposes. We also define a *dense* PF, with a reconstruction plane every 100 μm, that we use as a baseline and as a same-precision comparison to our ZPPF.

5.1 Depth precision analyisis

To analyze depth precision, we employ publicly available captured data [9] of two different scenes. These scenes are the *statue*, and the *dragon*, which both consist of a 2x2 m relay wall with a regular grid of 512x512 capture and illumination points in a confocal capture configuration.

We compare our ZPPF depth estimation with the PF on this dataset in Figure 4, using the dense PF as a baseline. ZPPF outperforms the previous PF, with a noticeably smaller error in the depth estimation (columns three and four), and with the same execution time.

In addition, our ZPPF method yields comparable results to the dense PF we used as baseline (depth estimation in the second column), using 300 times fewer reconstruction

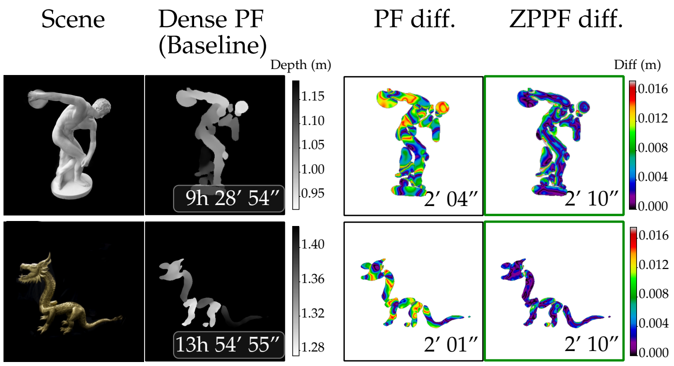


Fig. 4: From left to right: picture of the hidden scene, dense PF depth used as baseline, PF depth difference to the baseline, and ZPPF depth difference to the baseline. From top to bottom, the *statue* and the *dragon* scenes. Insets indicate the computation time required for each reconstruction shown. ZPPF reports subtle depth variations in sparse reconstruction planes, achieving similar results to the dense PF in a much faster execution time (almost 15 hours vs 2 minutes). In contrast, the previous PF employs the same time for much coarser depth estimations.

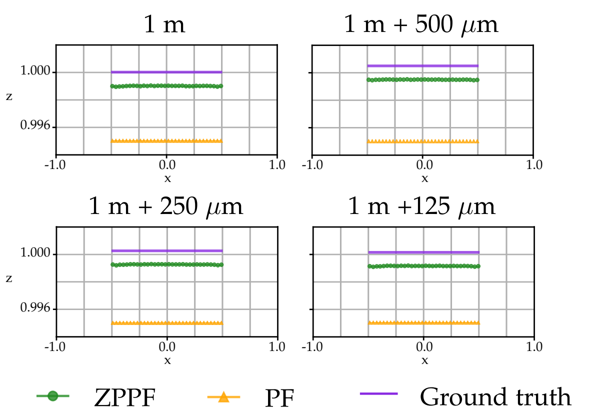


Fig. 5: Width (x) and depth (z) coordinates in meters of the PF and our ZPPF depth estimation of a 0.5 m-sized squared plane at varying distances in the micrometer scale compared to the ground truth. Ours achieves better depth approximations, using the sparse planes, with a small constant error, maintaining a coherent estimation when compared between experiments. Horizontal and vertical axes are not on the same scale for visualization purposes.

points, and being almost 400 times faster (a bit over two minutes vs almost fourteen hours).

We further analyze the sensitivity of our method to small variations in depth, on the order of micrometers. We use a baseline synthetic scene [38] consisting of a 0.5 x 0.5 m plane, placed parallel to the relay wall at 1 m. We progressively offset variations of 500, 250, and 125 μm in depth between experiments. Figure 5 shows the resulting estimated depths for ZPPF (green), PF (orange), and the ground truth (purple). PF does not capture such fine-grained variations in depth due to sticking to the sparse reconstruction planes. Leveraging phase information allows ZPPF to capture such subtle variations, with the exact same number of reconstruction planes as PF, while requiring the same 2.7-second reconstruction time for the same 2d slice.

Depth offset (μm)	Estimation diff. (μm)
500	492.60
250	246.21
125	123.23

TABLE 1: Mean estimated difference of our ZPPF between a 0.5 m plane located at 1 m of the relay wall, and the same plane offset 500, 250, and 125 μm in depth.

As Table 1 shows, our ZPPF estimates the relative depth variations between experiments with precision, even if a small error appears between our ZPPF estimation and the ground truth due to near-field diffraction affecting small phase shifts near the hidden geometry.

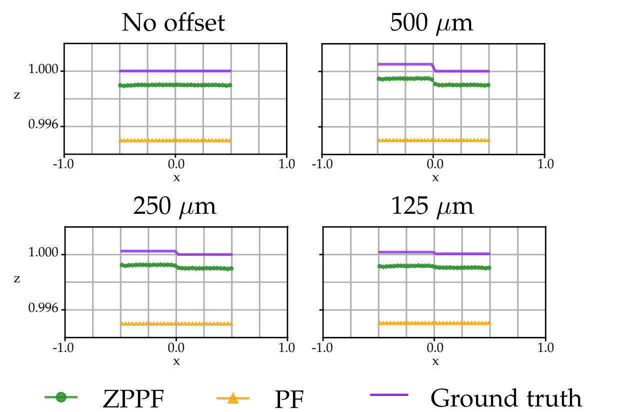
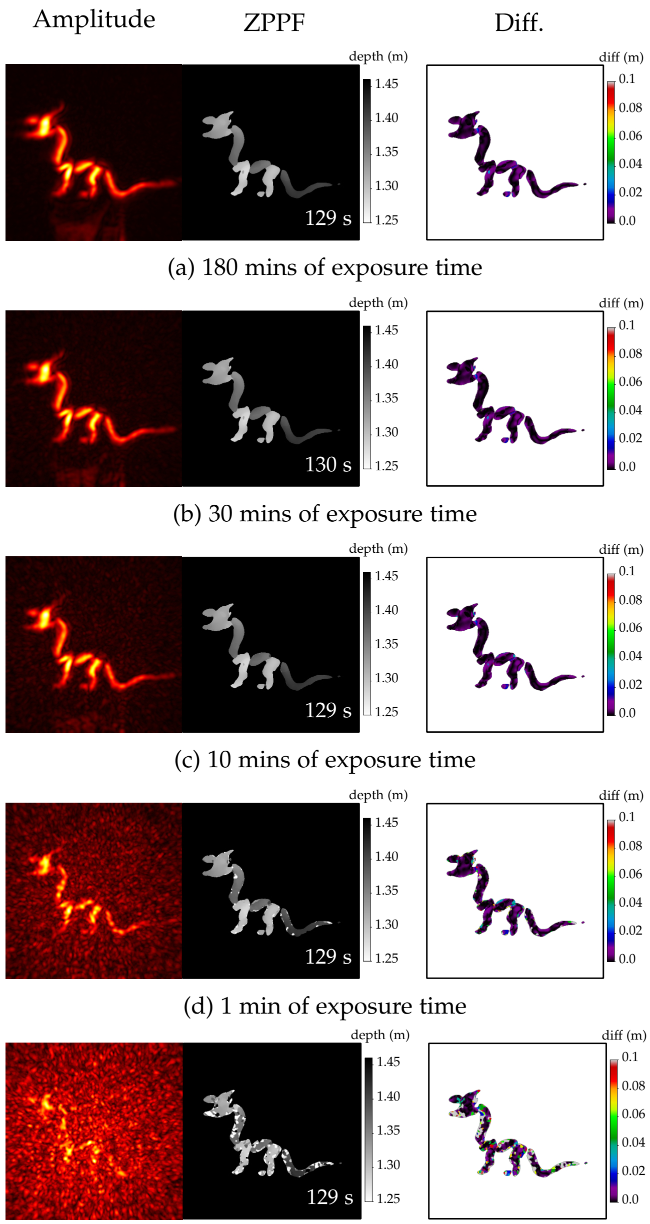


Fig. 6: Width (x) and depth (z) coordinates in meters of the PF and our ZPPF depth estimation of two 0.25 m squared planes, maintaining one fixed at 1 m, and applying a distance offset to the second one. Our ZPPF can detect the slope between the planes using sparse reconstruction planes, where previous PF cannot, even in the extreme case of 125 μm. Horizontal and vertical axes are not on the same scale for visualization purposes.

We next analyze *local* depth variations in the plane. We employ the same scene as the previous experiment, but splitting the plane in two equal parts and displacing one halfway along the z-axis 500, 250, and 125 μm, respectively. As Figure 6 shows, our method is robust at detecting the displacement that appears between both planes, even in the extreme case of a 125 μm offset, while PF estimations do not detect such small changes.

5.2 Robustness to noise

We next evaluate the robustness of our ZPPF estimations in challenging scenarios with low SNR. We use the *dragon* scene, which has been captured under decreasing exposure times of 180, 30, 10, one minute, and 15 seconds. For baseline comparison, we employed the dense PF on a high SNR capture (180 minutes exposure), providing a reference for evaluating the differences. As Figure 7 shows, our method achieves consistent and reliable depth estimations across all cases. At 15-second exposure, the low SNR constrains the



(a) 180 mins of exposure time

(b) 30 mins of exposure time

(c) 10 mins of exposure time

(d) 1 min of exposure time

(e) 15 sec of exposure time

Fig. 7: Depth estimation of our ZPPF for different captures of the *dragon* scene of decreasing exposure times. Our ZPPF depth estimation is coherent and reliable in all cases, showing some errors in the geometry borders of the 1-minute and 15-second exposure. We used the dense PF reconstruction for the difference baseline.

depth estimation from the amplitude, limiting the phase refinement of our ZPPF. Nevertheless, even for the challenging case of 1-minute exposure time, where reduced SNR leads to localized error at the geometry borders due to a decreased lateral resolution, our method maintains robust performance.

6 Discussion

We have proposed a novel NLOS imaging technique, named Zero-Phase Phasor Fields, which can leverage the phase information usually discarded in previous NLOS imaging methods. Our approach consists of a first coarse reconstruction (similar to conventional Phasor Fields), which we then refine by finding zero-crossings in phase. In contrast to previous NLOS imaging methods, including standard Phasor Fields, our method yields accurate depth reconstructions without increasing the voxel resolution of the hidden scene region, and therefore without imposing prohibitive computational constraints. As our experimental validation shows, ZPPF is sensitive to depth variations as small as 100 μm. While standard PF requires dense reconstruction planes and hours of computation to achieve the same precision, ours runs in a little over two minutes. Furthermore, ZPPF is robust to noise, as we showed in our experiments with low SNR scenarios. We believe that in the future, this will enhance the applicability of NLOS imaging into real-world scenarios, where noise, time, and memory constraints are challenges to overcome. We focus on the micrometer scale, which might be practical for medical imaging. Yet, future endeavors on longer virtual wavelengths would make the problem suitable for larger-scale applications, such as car safety or topography.

We have only focused on confocal capture configuration for our phase correction, and non-confocal captures require more complex estimation to refine their position correctly. Yet, when imaging far enough from the relay wall, the virtual wavefronts behave as planar even in a non-confocal configuration. Therefore, we presume that our estimation would perform similarly as we present here in that scenario.

Since our method involves a first standard PF reconstruction, our technique naturally benefits from all optimizations already implemented for PF, which have reached reconstructions as fast as five frames per second [30], [31]. Additionally, our technique could be adapted to other, more sophisticated PF-based techniques, including imaging beyond the third bounce [29].

Our method does not require a dense sampling of the reconstruction volume, which makes it suitable for large scenes, thus overcoming a common limitation in the size of the hidden scene for most NLOS imaging methods. The only limit to resolution comes from the wavelength of the virtual illumination; an iterative version of our approach could be used, starting from very large wavelengths, finding the candidate locations of the hidden geometry and progressively increasing resolution near such candidate locations until a certain precision is achieved. We have developed a prototype of this iterative technique, and our preliminary results show that hidden scenes with sparse geometry can be reconstructed up to 2.5 times faster than our standard ZPPF implementation.

Our results show very small variations with respect to the ground truth in the simulated data, which we believe are due to near-field virtual diffraction effects. Although such variations are negligible, it would be interesting to explore them further in future work. Other frequency-domain techniques, such as fk-migration [9], might also benefit from explicitly incorporating phase information. In this sense, we hope that our work inspires future research on more sophisticated NLOS imaging methods.

ACKNOWLEDGMENTS

We would like to thank the people from the Graphics & Imaging Lab for their assistance with some figures, especially Santiago Jiménez and María Plaza. This work has received funding from the Government of Aragon's Departamento de Ciencia, Universidad y Sociedad del Conocimiento through the Reference Research Group "Graphics and Imaging Lab" (ref T34_23R), and by Ministerio de Ciencia Innovacion y Universidades/Agencia Estatal de Investigación/10.13039/501100011033 through the project PID2019-105004GB-I00, which also provide Pablo Luesia-Lahoz with the contract referenced PRE2020-096096. This work was also supported by the Air Force Office for Scientific Research (FA9550-21-1-0341)

REFERENCES

[1] F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich, "Low-budget transient imaging using photonic mixer devices," *ACM Transactions on Graphics (ToG)*, vol. 32, no. 4, pp. 1–10, 2013.

[2] A. Velten, D. Wu, A. Jarabo, B. Masia, C. Barsi, C. Joshi, E. Lawson, M. Bawendi, D. Gutierrez, and R. Raskar, "Femto-photography: capturing and visualizing the propagation of light," *ACM Transactions on Graphics (ToG)*, vol. 32, no. 4, pp. 1–8, 2013.

[3] M. Buttafava, J. Zeman, A. Tosi, K. Eliceiri, and A. Velten, "Non-line-of-sight imaging using a time-gated single photon avalanche diode," *Optics express*, vol. 23, no. 16, pp. 20 997–21 011, 2015.

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

[5] V. Arellano, D. Gutierrez, and A. Jarabo, "Fast back-projection for non-line of sight reconstruction," in *ACM SIGGRAPH 2017 Posters*, 2017, pp. 1–2.

[6] M. La Manna, F. Kine, E. Breitbach, J. Jackson, T. Sultan, and A. Velten, "Error backprojection algorithms for non-line-of-sight imaging," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 7, pp. 1615–1626, 2018.

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

[8] S. Xin, S. Nousias, K. N. Kutulakos, A. C. Sankaranarayanan, S. G. Narasimhan, and I. Gkioulekas, "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*, 2019, pp. 6800–6809.

[9] D. B. Lindell, G. Wetzstein, and M. O'Toole, "Wave-based non-line-of-sight imaging using fast fk migration," *ACM Transactions on Graphics (ToG)*, vol. 38, no. 4, pp. 1–13, 2019.

[10] X. Liu, I. Guillén, M. La Manna, J. H. Nam, S. A. Reza, T. Huu Le, A. Jarabo, D. Gutierrez, and A. Velten, "Non-line-of-sight imaging using phasor-field virtual wave optics," *Nature*, vol. 572, no. 7771, pp. 620–623, 2019.

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

[12] A. D. Cronin, J. Schmiedmayer, and D. E. Pritchard, "Optics and interferometry with atoms and molecules," *Reviews of Modern Physics*, vol. 81, no. 3, pp. 1051–1129, 2009.

[13] G. Gasbarri, A. Belenchia, M. Carlesso, S. Donadi, A. Bassi, R. Kaltenbaek, M. Paternostro, and H. Ulbricht, "Testing the foundation of quantum physics in space via interferometric and non-interferometric experiments with mesoscopic nanoparticles," *Communications Physics*, vol. 4, no. 1, p. 155, 2021.

[14] S. R. Brueck, "Optical and interferometric lithography-nanotechnology enablers," *Proceedings of the IEEE*, vol. 93, no. 10, pp. 1704–1721, 2005.

[15] D. Faccio, A. Velten, and G. Wetzstein, "Non-line-of-sight imaging," *Nature Reviews Physics*, vol. 2, no. 6, pp. 318–327, Jun. 2020, number: 6 Publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/s42254-020-0174-8

[16] C.-Y. Tsai, A. C. Sankaranarayanan, and I. Gkioulekas, "Beyond Volumetric Albedo — A Surface Optimization Framework for Non-Line-Of-Sight Imaging," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA: IEEE, Jun. 2019, pp. 1545–1555. [Online]. Available: https://ieeexplore.ieee.org/document/8953504/

[17] S. A. Reza, M. L. Manna, S. Bauer, and A. Velten, "Phasor field waves: A huygens-like light transport model for non-line-of-sight imaging applications," *Opt. Express*, vol. 27, no. 20, pp. 29 380–29 400, Sep 2019. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?URI=oe-27-20-29380

[18] S. A. Reza, M. La Manna, S. Bauer, and A. Velten, "Phasor field waves: experimental demonstrations of wave-like properties," *Optics Express*, vol. 27, no. 22, p. 32587, Oct. 2019. [Online]. Available: https://www.osapublishing.org/abstract.cfm?URI=oe-27-22-32587

[19] J. A. Teichman, "Phasor field waves: a mathematical treatment," *Optics Express*, vol. 27, no. 20, pp. 27 500–27 506, Sep. 2019, publisher: Optical Society of America. [Online]. Available: https://www.osapublishing.org/oe/abstract.cfm?uri=oe-27-20-27500

[20] J. Dove and J. H. Shapiro, "Paraxial theory of phasor-field imaging," *Optics Express*, vol. 27, no. 13, pp. 18 016–18 037, Jun. 2019. [Online]. Available: https://www.osapublishing.org/oe/abstract.cfm?uri=oe-27-13-18016

[21] ——, "Paraxial phasor-field physical optics," *Optics Express*, vol. 28, no. 14, pp. 21 095–21 109, Jul. 2020, publisher: Optical Society of America. [Online]. Available: https://www.osapublishing.org/oe/abstract.cfm?uri=oe-28-14-21095

[22] ——, "Nonparaxial phasor-field propagation," *Optics Express*, vol. 28, no. 20, pp. 29 212–29 229, Sep. 2020, publisher: Optical Society of America. [Online]. Available: https://www.osapublishing.org/oe/abstract.cfm?uri=oe-28-20-29212

[23] T. Sultan, S. A. Reza, and A. Velten, "Towards a more accurate light transport model for non-line-of-sight imaging," *Optics Express*, vol. 32, no. 5, pp. 7731–7761, Feb. 2024, publisher: Optica Publishing Group. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?uri=oe-32-5-7731

[24] C. Rambour, A. Budillon, A. C. Johnsy, L. Denis, F. Tupin, and G. Schirinzi, "From interferometric to tomographic sar: A review of synthetic aperture radar tomography-processing techniques for scatterer unmixing in urban areas," *IEEE Geoscience and Remote Sensing Magazine*, vol. 8, no. 2, pp. 6–29, 2020.

[25] J. Ding, M. Kahl, O. Loffeld, and P. H. Bolívar, "Thz 3-d image formation using sar techniques: Simulation, processing and experimental results," *IEEE Transactions on Terahertz Science and Technology*, vol. 3, no. 5, pp. 606–616, 2013.

[26] B. Baccouche, P. Agostini, S. Mohammadzadeh, M. Kahl, C. Weisenstein, J. Jonuscheit, A. Keil, T. Löffler, W. Sauer-Greff, R. Urbansky *et al.*, "Three-dimensional terahertz imaging with sparse multistatic line arrays," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 23, no. 4, pp. 1–11, 2017.

[27] X. Liu and A. Velten, "The role of Wigner Distribution Function in Non-Line-of-Sight Imaging," in *2020 IEEE International Conference on Computational Photography (ICCP)*, Apr. 2020, pp. 1–12, iSSN: 2472-7636.

[28] J. Marco, A. Jarabo, J. H. Nam, X. Liu, M. Á. Cosculluela, A. Velten, and D. Gutierrez, "Virtual light transport matrices for non-line-of-sight imaging," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 2440–2449.

[29] D. Royo, T. Sultan, A. Muñoz, K. Masumnia-Bisheh, E. Brandt, D. Gutierrez, A. Velten, and J. Marco, "Virtual mirrors: Non-line-of-sight imaging beyond the third bounce," *ACM Transactions on Graphics (TOG)*, vol. 42, no. 4, pp. 1–15, 2023.

[30] Z. Liao, D. Jiang, X. Liu, A. Velten, Y. Ha, and X. Lou, "FPGA Accelerator for Real-Time Non-Line-of-Sight Imaging," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 69, no. 2, pp. 721–734, Feb. 2022, conference Name: IEEE Transactions on Circuits and Systems I: Regular Papers.

[31] J. H. Nam, E. Brandt, S. Bauer, X. Liu, M. Renna, A. Tosi, E. Sifakis, and A. Velten, "Low-latency time-of-flight non-line-of-sight imaging at 5 frames per second," *Nature communications*, vol. 12, no. 1, p. 6526, 2021.

[32] M. L. Manna, J.-H. Nam, S. A. Reza, A. Velten, and A. Velten, "Non-line-of-sight-imaging using dynamic relay surfaces," *Optics Express*, vol. 28, no. 4, pp. 5331–5339, Feb. 2020, publisher: Optica Publishing Group. [Online]. Available: https://opg.optica.org/oe/abstract.cfm?uri=oe-28-4-5331

[33] C. Gu, T. Sultan, K. Masumnia-Bisheh, L. Waller, and A. Velten, "Fast Non-line-of-sight Imaging with Non-planar Relay Surfaces," in *2023 IEEE International Conference on Computational Photography (ICCP)*, Jul. 2023, pp. 1–12, iSSN: 2472-7636.

[34] D. Jiang, X. Liu, J. Luo, Z. Liao, A. Velten, and X. Lou, "Ring and Radius Sampling Based Phasor Field Diffraction Algorithm for Non-Line-of-Sight Reconstruction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 11, pp. 7841–7853, Nov. 2022, conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence.

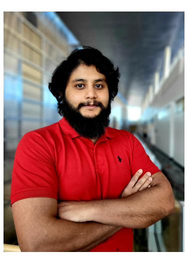
[35] P. Luesia-Lahoz, D. Gutierrez, and A. Muñoz, "Zone plate virtual lenses for memory-constrained nlos imaging," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–5.

[36] S. I. Young, D. B. Lindell, B. Girod, D. Taubman, and G. Wetzstein, "Non-line-of-sight surface reconstruction using the directional light-cone transform," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 1407–1416.

[37] F. Heide, M. O'Toole, K. Zang, D. B. Lindell, S. Diamond, and G. Wetzstein, "Non-line-of-sight imaging with partial occluders and surface normals," *ACM Transactions on Graphics (ToG)*, vol. 38, no. 3, pp. 1–10, 2019.

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

Pablo Luesia-Lahoz Pablo Luesia-Lahoz is a doctoral student at the Graphics & Imaging Lab at the Universidad de Zaragoza-I3A. He obtained his Computer Engineering and master's degrees from the Universidad de Zaragoza. Also, he was granted a competitive grant FPI 2020 from the *Agencia Estatal de Investigación-Ministerio de Ciencia, Innovación y Universidades* for his thesis, which focuses on computational imaging and graphics.

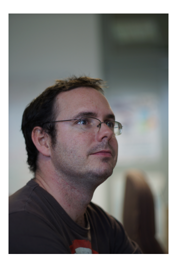


Talha Sultan Talha Sultan (Member, IEEE) is a Ph.D. candidate in the Electrical and Computer Engineering Department University of Wisconsin–Madison, Madison, WI. He received his BS degree in chemical engineering & applied mathematics from the University of Wisconsin–Madison, Madison, WI. His current research interests include computational optics and computational imaging. He specifically focuses on utilizing emerging sensing technologies to advance non-line-of-sight imaging systems.

Forrest Beck Peterson Forrest Beck Peterson is a PhD student in the Department of Electrical and Computer Engineering at the University of Wisconsin-Madison working in the Computational Optics Lab. He received his BS in Electrical Engineering from Calvin University in Grand Rapids, Michigan. His research focuses on computational optics and imaging, working with high speed, single photon detecting SPAD cameras for applications in none-line-of-sight imaging and low-light environments.

Andreas Velten Andreas Velten is Associate Professor at the Department of Biostatistics and Medical Informatics and the Department of Electrical and Computer Engineering at the University of Wisconsin-Madison and directs the Computational Optics Group. He obtained his PhD in Physics at the University of New Mexico in Albuquerque and was a postdoctoral associate of the Camera Culture Group at the MIT Media Lab. He has been included in the MIT TR35 list of the world's top innovators under the age of 35, is a senior member of NAI and SPIE, and is a fellow of Optica. He is the co-founder of Onlume, a company that develops surgical imaging systems, and Ubicept, a company developing single photon imaging solutions.

Diego Gutierrez Diego Gutiérrez is a Full professor at the Universidad de Zaragoza, where he leads the Graphics and Imaging Lab, a research group he founded in 2008, working on Computer Graphics and Computational Imaging. His areas of interest include computer graphics, computational imaging, and virtual reality. Included in the Top 100 most influential scholars in the decade in Computer Graphics 2020, he has published over 100 publications in ranked journals (including one in Nature), with an h-index of 40. He has been Editor in Chief of ACM Transactions on Applied Perception for three years, and is (or has been) an Associate Editor for all major journals in his fields of research. He has been Program Chair of 12 international conferences, Guest Editor of five special issues, and has served on more than 70 international committees for top conferences (including 12 times at SIGGRAPH or SIGGRAPH Asia).



Adolfo Muñoz Adolfo Muñoz is a tenured Associate Professor with the Department of Computer Science and Systems Engineering, Universidad de Zaragoza, Spain, where he belongs to the Graphics and Imaging Lab. His research interests include light transport simulation and capture and modeling of material appearance, as well as computational and transient imaging. He has co-authored publications in high-impact journals: 20 publications in indexed journals, 13 of them in the first quartile. He has also collaborated in 21 research projects, either from national and international calls or from collaborations with industry, being Principal Investigator in 2 international projects, 3 national projects, and 2 transference contracts.