

GENERALIZED AND EFFICIENT NON-LINE-OF-SIGHT IMAGING

PABLO LUESIA-LAHOZ

SUPERVISED BY: ADOLFO MUÑOZ AND DIEGO GUTIERREZ

TESIS DOCTORAL - INGENIERÍA INFORMÁTICA

GRAPHICS AND IMAGING LAB
DEPARTAMENTO DE INFORMÁTICA E INGENIERÍA DE SISTEMAS
UNIVERSIDAD DE ZARAGOZA

JULY 2026

To all Luesia people,

to all Lahoz people,

to Alejandra,

to my friends,

and to the rest.

ABSTRACT

Non-Line-of-Sight (NLOS) imaging aims to see what is hidden around a corner, with a variety of compelling applications such as medical imaging, search and rescue operations, or car safety. The field emerged with the advent of transient imaging, which, utilizing an extremely high temporal resolution, offers a means to analyze light transport over time. By capturing indirect photons reflected from a visible *relay wall* and exploiting their time of flight, NLOS methods can reconstruct the geometry of hidden scenes.

Despite recent promising results, NLOS imaging is confined to laboratory-controlled situations, limiting its practical deployment. This thesis aims to progress NLOS imaging towards its practical implementation into everyday technology. To that end, we contribute in three different aspects that shape the structure of this thesis: generalization of NLOS for varying and complex setups, efficiency of reconstruction methods, and correct modeling of its transient capture.

Existing NLOS reconstruction methods rely on restrictive physical assumptions to make the problem tractable. This limits the robustness and applicability of NLOS in real-world scenarios. In the first part of this thesis, we tackle the mentioned limitations in reconstruction by relaxing common assumptions. Specifically, we show that NLOS reconstruction is feasible under dense scattering media, removing the common assumption of light traveling through a vacuum. On top of that, we propose a novel setup consisting of multiple relay walls that we can treat as a single optical system, as we demonstrate. This enhances the hidden scene visibility, breaking the planar-relay wall assumption in the process.

Additionally, NLOS reconstructions often demand considerable computational resources. To address this limitation, in the second part of the thesis, we propose two novel reconstruction algorithms focused on efficiency. The first simplifies the light propagation, offering a trade-off between quality and memory consumption of the reconstruction. Our second approach achieves micrometer depth precision in the same computational complexity as previous methods employ to achieve centimeter-scale precision, employing the virtual phase that was discarded by previous methods.

Besides reconstruction, NLOS imaging struggles in handling captures with suboptimal instrument response, such as noise and low temporal resolution. Existing methods ignore underlying physical events, assuming ideal captures (noiseless, delta-illuminated, and infinite temporal resolution). Therefore, experimental validation for reconstruction methods is limited to sophisticated hardware and controlled environments. For the third part of this thesis, we explicitly model the physical processes occurring during an NLOS capture, computationally enhancing the signal from lower-end devices and noisy environments. For that purpose, we build upon a formulation based on the Dirac notation, originally developed in a collaborative work, proposing a cohesive and compact representation of the light transport and instrument response of an NLOS scene.

In summary, this thesis generalizes NLOS imaging models, algorithms, and physical setups to facilitate practical deployment in real-world environments. We expect these contributions to support future advances such as adaptive NLOS systems and novel capture strategies.

RESUMEN

La formación de imagen sin línea de visión (NLOS, por sus siglas en inglés) establece su objetivo en ver lo que se oculta a la vuelta de la esquina, con variedad de aplicaciones tales como la imagen médica, operaciones de búsqueda y rescate, o seguridad automovilística. El campo surgió con el advenimiento de la imagen transitoria, la cual, utilizando una resolución temporal extremadamente alta, ofrece los medios para analizar el transporte de luz a lo largo del tiempo. Al capturar fotones indirectos reflejados desde una *superficie auxiliar* visible y aprovechar su tiempo de vuelo, los métodos NLOS pueden reconstruir la geometría de escenas ocultas.

Pese a resultados prometedores recientes, la aplicación de imagen NLOS se limita a escenarios controlados, restringiendo su despliegue práctico. Esta tesis tiene como objetivo avanzar en la imagen NLOS hacia su implementación práctica en la tecnología cotidiana. Para ello, contribuimos en tres aspectos diferentes que estructuran esta tesis: la generalización de NLOS para diversas configuraciones, la eficiencia de los métodos de reconstrucción y el modelado correcto de la captura transitoria.

Los métodos actuales de reconstrucción NLOS recurren a presunciones físicas restrictivas para hacer el problema manejable. Esto limita la robustez y aplicabilidad de NLOS en escenarios del mundo real. En la primera parte de esta tesis, abordamos las limitaciones mencionadas en la reconstrucción relajando presunciones comunes. Concretamente, demostramos que la reconstrucción NLOS es factible en medios participativos densos, desafiando la presunción común de que la luz viaja en el vacío. Además, proponemos una configuración novedosa consistente en múltiples *superficies auxiliares* que podemos tratar como un único sistema óptico, tal y como demostramos. Esto mejora la visibilidad de la escena oculta, rompiendo la suposición de *superficie auxiliar* planar en el proceso.

Además, los métodos de reconstrucción en NLOS demandan considerables recursos computacionales. Abordando esta limitación, en la segunda parte de la tesis, proponemos dos algoritmos de reconstrucción centrados en la eficiencia. El primero simplifica la propagación de la luz, ofreciendo un compromiso entre la calidad de la reconstrucción y el consumo de memoria. Nuestro segundo enfoque aprovecha la fase virtual, que es descartada por los métodos anteriores. Esto permite identificar variaciones de profundidad en la escena oculta a escala micrométrica, mientras que los enfoques anteriores solo alcanzan una precisión a escala centimétrica con una complejidad computacional comparable.

Aparte de la reconstrucción, la imagen NLOS tiene dificultades para manejar capturas con una respuesta instrumental subóptima, como ruido o baja resolución temporal. Métodos existentes ignoran los eventos físicos subyacentes, asumiendo capturas ideales. La validación en el campo se ve limitada a hardware sofisticado y entornos controlados. Para la tercera parte de la tesis, modelamos explícitamente los procesos físicos que ocurren durante una captura NLOS, refinando computacionalmente la señal de dispositivos de gamma baja o entornos ruidosos. Para ello, nos basamos en una formulación basada en la notación de Dirac, desarrollada originalmente en un trabajo colaborativo, proporcionando una representación cohesionada y compacta del transporte en la escena oculta.

En resumen, esta tesis generaliza los modelos de imagen NLOS, los algoritmos y las configuraciones físicas para facilitar su implementación práctica en entornos del mundo real. Esperamos que estas contribuciones sirvan de apoyo a futuros avances, como los sistemas NLOS adaptativos y las nuevas estrategias de captura.

ACKNOWLEDGEMENTS

The reader who is taking the time to read this probably already knows, but I feel compelled to say it anyway. This is a document that aims to collect the work I have performed over the last four and a half years. However, this has never been only work, but more like a journey. And like in every journey, I haven't accomplished it alone. Therefore, I am taking these pages to thank some of the people who helped, taught, supported, or just accompanied me in the materialization of this. This is a long list, but not an exhaustive one, so I am only highlighting some of these people.

Diego and Adolfo: You have been such a guide during these years, and I doubt any of these words would be possible without your assistance. You advised me any time I needed it (even if I did not ask for it), and taught me how to be a better researcher, teacher, and speaker. *Diego*, you always helped me to see the whole point of the problem, to step up. Where I am a depth searcher, you lifted me to see the overall point of view. Your ability to motivate has pushed me to cross some boundaries I was not aware I had. *Adolfo*, your humor, energy, and passion are contagious. You love what you do, and I will always remember those moments of lucubration of new ideas, although we never have the time to implement them. Thank you both for giving me the opportunity.

Andreas, I am just thankful for having been able to work with you. Your head is full of interesting, promising, and maybe crazy ideas that the research community loves. Thank you for accepting me under your wing during those three months I was abroad. Also, to the rest of the Computational Optics Group from the University of Wisconsin: *Talha, Forrest, David, Alex, and Jidong*. It was an amazing time for us to discuss anything anytime we saw each other, not necessarily about the topic that any of us were working on. I hope we manage to see each other more often in the near future.

To the people of the Graphics and Imaging lab. The number of members is so high that I am probably missing somebody. The environment, the confidence, the coffee breaks, everything is a perfect breeding ground for new researchers. The context allowed me to reach further than I could even think when I first arrived to the group. *Alfonso, Ana, Albert, Belen, Julio, and Sandra*, you all shine in your peculiar way, and your advice, discussion, and teachings have been the perfect company to my advisors. To the younger members of the group (and I am probably missing somebody here): *Óscar, Jorge García and Jorge Pina, María Peña and María Plaza, Sergio, Santiago, Carmen, Pilar, and Diego Bielsa*. Some of you will soon be at this point, and some of you have just started. I will always appreciate your fresh energy. Thank you all for being a great support for everybody, even if you did not do it on purpose.

Now, there is a small subgroup where it is difficult to draw a line between coworkers, colleagues, and friends. *Edurne* and our conversations during breaks. *dr. Oyo* (or *Diego*, as other people like to call you), and your infinite chillness. *Dario*, and the magic behind your photogenetics. *Juan Raúl*, and the difficult/uncomfortable questions you always managed to bring. *Néstor* and your not-so-good but constant humor (and your stubbornness to learn that tangerines are better than oranges). *Julia*, and your kindness behind your enormous responsibility, which bothers me. *David* and your incredible, optimistic, and intense board-gaming energy. *Víctor* and your acquired wealth. *Martincho* (el del pelo pincho), and, of course, your balls. *Dani* and the fact that you are like five years older than me, although we both were born in the same year. Thank you all for your amazing support over these years. I am sure the result would not have been the same without you.

To the members of the *Campa* or, as by the proper name, the *secta*. There are more than 30 people in this list, so I won't list you one by one. You are all well aware of how special the bond that we share is. I know that I have not been with you a lot lately, but, well, in your hands is the reason for that. Inside that group, I should mention the *Royal Elite* (+ *Julis*). Thank you, friends, for always being able to meet up, yet I could not go.

To *Carlos, Juuaan, Fernando C., Angel, David O., Lorena, Sandra, Natalia, Mario Rico, Mario Punk, and Marco*. You are probably at the top of this list in hearing my laments. Thank you for hearing them, for not supporting them, and for asking me, "When will you start working?" Joke apart, your presence has been invaluable to me. You were always capable of distracting me in my darkest times. I love you.

To my *parents, Laura, Dr. Juan, brother in law, and nephews*. Sorry for the non-Spanish speaker, but most of these people are unable to understand English: *Gracias por todo, por vuestra ayuda, por vuestro apoyo. Sabéis que os quiero, y que agradezco todo vuestro apoyo a lo largo de estos años. Aunque no entendáis todavía que hago. Y aunque no agradezco tener que haber cogido tantas olivas mientras escribia esta tesis.*

Finally, but not least, *Alejandra*. I am sure you are already enjoying reading all these fancy English words coming out of me. Although you haven't been here from the beginning, you have climbed your position in this list to become one of the most important people on the list. Your support has been indescribable. This thesis would have come up even if you weren't here, but my mental state would probably be done by now. Just, thank you for being here, and somehow, to become my partner (for now¹).

Thank you.

In terms of funding, this work has been supported 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 Innovación 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.

¹ This is an internal joke.

CONTENTS

Preliminaries.

1	Introduction.	3
1.1	Goal and overview.	6
1.2	Contributions and Measurable Results.	7
1.2.1	Publications.	7
1.2.2	Research Internship.	7
1.2.3	Published Code.	8
1.2.4	Supervised Students.	8
1.2.5	Miscellaneous: dissemination and reviewing.	8
2	Related Work.	9
2.1	Transient Non-line-of-sight Imaging.	9
2.1.1	Transient captures for Non-Line-of-Sight imaging.	9
2.1.2	Reconstructions methods for Non-Line-of-Sight imaging.	10
2.2	Phasor Fields.	11
2.2.1	Virtual wavefront abstraction.	11
2.2.2	Phasor Fields efficiency.	12
3	Background.	13
3.1	Non-Line-of-Sight imaging nomenclature.	13
3.2	The Phasor Fields formulation.	14

I Towards Generalized Reconstruction of Hidden Scenes.

4	Phasor Fields in the Presence of Scattering Media.	19
4.1	Introduction.	19
4.2	Experimental design.	20
4.3	Reconstruction.	21
4.4	Analysis.	21
5	Stereo Non-Line-of-Sight Imaging.	25
5.1	Introduction.	25
5.2	From NLOS to LOS.	26
5.3	Visibility limitation of NLOS imaging.	27
5.4	Our method.	27
5.4.1	Combination of different Phasor-Fields views.	28
5.5	Results.	29
5.5.1	Stereo combination experiments.	30
5.5.2	Combination of captured results.	31
5.5.3	Analysis on the relay walls contributions.	32
5.6	Discussion.	33

II Efficient Algorithms to Reconstruct the Hidden Scene.

6	Fresnel Zone Plates for Memory Constrained NLOS Imaging.	37
6.1	Introduction.	37
6.2	Zone plates.	38
6.3	Our method.	39
6.4	Results.	39
6.5	Discussion.	41
7	Zero-Phase Phasor Fields for NLOS.	43
7.1	Introduction.	43

7.2	Depth precision in NLOS.	44
7.3	Our method.	44
7.3.1	Analyzing phase.	45
7.3.2	Zero-Phase Phasor Fields.	47
7.4	Results.	47
7.4.1	Depth precision analysis.	48
7.4.2	Robustness to noise.	49
7.5	Discussion.	50
 III Enhancing the NLOS Capture with Cohesive and Correct Formulations.		
8	A Cohesive Framework for NLOS using Dirac Notation	55
8.1	Introduction.	55
8.2	Background.	56
8.2.1	Dirac notation and quantum mechanics.	57
8.3	Fundamental equations for NLOS imaging.	59
8.4	Derivation of backprojection formulation.	60
8.5	Derivation of Phasor fields formulation.	61
8.5.1	Derivation.	62
8.5.2	Rayleigh-Sommerfeld diffraction operator is the propagation operator in wave-based imaging methods.	62
8.6	Derivation of f-k migration formulation.	63
8.7	backprojection and wave-based imaging methods are equivalent point-to-point.	64
8.8	Application: NLOS with non-delta illumination.	67
8.9	Discussion and Conclusion.	68
9	Non-Line-of-Sight Imaging from Suboptimal Instrument Response.	71
9.1	Introduction.	71
9.2	Temporal PSF in NLOS imaging.	72
9.3	Our method.	72
9.3.1	Representing an NLOS scene with Dirac notation.	73
9.3.2	Point spread function in the scene operator.	74
9.3.3	Noise in the scene operator.	74
9.3.4	Estimating ideal capture from suboptimal capture.	75
9.3.5	Ideal estimation in the frequency domain.	76
9.4	Results	76
9.4.1	Removing the point-spread function.	77
9.4.2	Including photon-loss noise.	78
9.4.3	Arbitrary illumination reconstruction: train of delta pulses.	81
9.5	Discussion	81
 Conclusion.		
10	Discussion and Future Work.	85
10.1	English version.	85
10.2	Versión en español.	87
 Appendix		
A	Propagation operators are unitary.	91
B	Ideal signal derivation in the frequency domain.	93
 Bibliography		
		97

LIST OF FIGURES

Figure 1.1	Applications for Non-Line-of-Sight imaging.	3
Figure 1.2	Example of a NLOS setup.	4
Figure 2.1	Transient imaging capture.	10
Figure 2.2	Reconstructions using filtered backprojection, fk-migration and Phasor Fields	11
Figure 3.1	Non-Line-of-Sight setup and variable names.	13
Figure 3.2	NLOS capture configurations	14
Figure 3.3	Overview of the Phasor Fields formulation.	15
Figure 3.4	Phasor Fields reconstruction.	16
Figure 4.1	Scattering media scene setup.	20
Figure 4.2	Simulated scenes for imaging through scattering media.	20
Figure 4.3	Reconstructions of the Z-LETTER scene submerged in scattering media of varying properties.	23
Figure 4.4	Reconstruction of the Z-LETTER submerged in scattering media using different Phasor Fields parameters.	24
Figure 4.5	Reconstruction of the SHELF scene submerged in scattering media of varying properties.	24
Figure 4.6	Reconstruction of real scenes behind a thick diffuser using Phasor Fields.	24
Figure 5.1	Our stereo NLOS overview.	26
Figure 5.2	Overview of our stereo NLOS setup.	28
Figure 5.3	Stereo reconstruction with independent illumination.	30
Figure 5.4	Multipath interference in our stereo setup.	31
Figure 5.5	Stereo setup with real captures.	32
Figure 5.6	Stanford bunny reconstruction using dependant illumination of our approach.	33
Figure 5.7	Analysis of the dependant illumination under varying relay wall angles.	34
Figure 6.1	Fresnel Zone Plates for NLOS overview.	38
Figure 6.2	Results of synthesized scenes using different kernels.	42
Figure 7.1	Summary of methods: Zero-Phase Phasor Fields, Phasor Fields and dense Phsor Fields	44
Figure 7.2	Phase of Phasor Fields.	46
Figure 7.3	Results of our Zero-Phase Phasor Fields on real captures.	48
Figure 7.4	Results of our Zero-Phase Phasor Fields on micrometer displacements.	49
Figure 7.5	Results of our Zero-Phase Phasor Fields on local displacement at micrometer scale.	50
Figure 7.6	Results of our Zero-Phase Phasor Fields on real data with decreasing SNR.	52
Figure 8.1	Schematic of the backprojection formulation	60
Figure 8.2	Diagram showing the Dirac operators and their propagation directions.	65
Figure 8.3	Phasor Fields reconstruction of the dragon with increasing width of the pulse.	68
Figure 9.1	Dirac's notation scheme	73
Figure 9.2	Histograms of low temporal capture and our capture enhancement.	77
Figure 9.3	Backprojection reconstructions of low temporal capture and our capture enhancement.	78
Figure 9.4	Phasor Fields and fk-migration reconstructions of low temporal capture and our capture enhancement.	79
Figure 9.5	Histograms of low temporal and noisy captures and our capture enhancement.	79
Figure 9.6	Backprojection reconstructions of low temporal and noisy captures and our increased temporal resolution.	80
Figure 9.7	Phasor Fields and fk-migration reconstructions of low temporal and noisy captures, and our capture enhancement.	81

LIST OF FIGURES

- Figure 9.8 Fk-migration reconstructions of low temporal and noisy captures, and our capture enhancement, with real world data. [81](#)
- Figure 9.9 Backprojection reconstructions of arbitrary illumination and our capture enhancement. [82](#)

LIST OF TABLES

Table 6.1	RMSE of the reconstructions using different kernels.	40
Table 7.1	Metrics of our Zero-Phase Phasor Fields on micrometers local variations.	49
Table 8.1	Symbols used for the Dirac notation	57
Table 9.1	Scale invariant signal to distortion ratio of the original signal versus ours, compared with the ground truth (the higher the better).	77
Table 9.2	Scale invariant signal to distortion ratio of the original data versus ours, compared with the ground truth (the higher the better), on a system with a PSF of 314-ps width, and simulating the noise of different photon losses (columns).	79

PRELIMINARIES.

INTRODUCTION.

Everybody has fantasized about using superpowers at some point in their life. One of the most iconic is the x-ray vision, which fiction and comic stories allude to the ability to see through objects. Although this bears little relation to the physical properties of X-ray radiation, it points to more grounded and compelling problems where access to the direct line of sight is challenging or involves risk. Some examples are the identification of diseases using non-intrusive methods for medical diagnosis, the warning of occluded incoming dangers to drivers in car safety, exploration of areas where access is difficult or dangerous, or the localization of people in rescue operations after disasters (see Figure 1.1). Nowadays, technology allows for achievements that were once considered fantasy, and now it is possible to see what is hidden from our direct line of sight. This ability is known as *Non-line-of-sight* (NLOS) imaging.

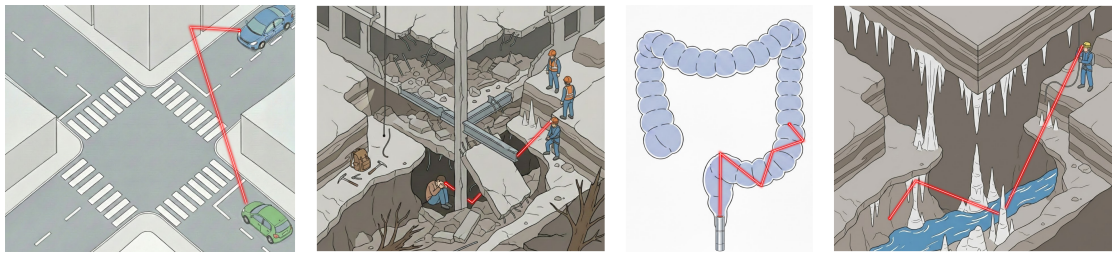


Figure 1.1: Applications in which the imaging targets lie outside the line of sight. From left to right: car safety, rescue operations, medical imaging, and underground exploration. The red light indicates a NLOS system, which captures light indirectly scattered from the hidden scene. *This figure was created with assistance from Google Gemini.*

NLOS imaging is a novel field that aims to reconstruct scenes that are hidden around a corner. Its emergence is related to the arrival of transient imaging, a technology that records light at an effective frame rate of trillions per second. At this frame rate, it is possible to record the light propagation over time in macroscopic scenes [51]. That allows researchers to employ the time of flight (ToF) of the light for NLOS imaging [21]. The key idea behind how it works is to employ the indirect illumination that scatters onto a visible Lambertian surface, or a *relay wall*. Thus, the relay wall acts as an auxiliary observation point of the hidden scene. Figure 1.2 depicts a typical NLOS imaging setup, where a laser indirectly illuminates the hidden scene, and a transient sensor records the light returning to the relay wall over time.

One of the first captures of transient imaging applied to an NLOS setting occurred thirteen years before the completion of this thesis [115]. Since then, the NLOS imaging landscape has seen rapid progression in both hardware and computational tools. The sensing devices are faster and more reliable, while methods for reconstructing NLOS scenes have emerged with varying efficiency and accuracy. Yet, NLOS technology is still in the early stages of development. The NLOS measurements are limited to controlled environments, such as laboratories, due to the requirements of sophisticated hardware and challenging calibration. Moreover, computational demands prevent the implementation in resource-constrained devices, such as onboard car computers, drone processing units, or smartphones. All these reasons hamper the transformation of NLOS imaging into a daily and accessible tool.

The ultimate goal of this thesis is to enable future NLOS imaging beyond controlled laboratory environments and towards real-world applications. We argue that current limitations in NLOS arise from three sources. First, reconstruction methods rely on restrictive physical assumptions, limiting robustness across diverse NLOS scenes. Second, even considering the aforementioned assumptions, calculations of NLOS reconstructions require expensive computational resources. This hinders the NLOS implementation into smaller and portable hardware, or limited-time applications. Third, existing NLOS approaches



Figure 1.2: Example of a common setup in NLOS imaging. Using a transient camera (dark camera) and controlled illumination (a red laser), the hidden scene (the dragon) can be imaged.

impose method-specific and independent formulations with ad-hoc assumptions. This often implies that methods ignore the modeling of the underlying physical events present at the capture of the NLOS signal. Consequently, current imaging solutions are highly sensitive to suboptimal capture conditions. These three problems motivate the structure of this thesis, which is divided into three parts: generalization of the NLOS reconstruction, efficient algorithms for NLOS reconstructions, and capture formulation and modeling. We introduce the contributions of this thesis to each of these challenges in the following.

TOWARDS GENERALIZED RECONSTRUCTION OF HIDDEN SCENES. Reconstructing in the NLOS context involves inverting the propagation of the light scattered from a hidden scene to a relay wall, leveraging its ToF. Addressing all possible physical events of the light makes the problem intractable. Researchers rely on physical assumptions of the NLOS scene to simplify the reconstruction complexity, such as light traveling in the vacuum or planar relay walls, making reconstructions tractable. Although these simplifications enable previous works to achieve promising results, they present drawbacks. Reality must adapt to the assumed scenario. Otherwise, reconstructions fail due to unforeseen physical events. This lack of robustness constrains practical NLOS deployment to in-the-wild scenarios, confining NLOS imaging to laboratory-controlled scenes.

In this first part of the thesis, we focus on the robustness of the NLOS reconstructions to expand NLOS applicability from the laboratory to the real world. Rather than specializing reconstruction algorithms to increasingly restrictive scenarios, we challenge common assumptions on light propagation and relay wall geometry. To that end, we leverage the Phasor Fields formulation [69], which poses the relay wall as a virtual wavefront propagation. This abstraction enables the use of well-known tools from wave and Fourier optics, posing the NLOS reconstruction as a virtual line-of-sight problem.

Based on the Phasor Field framework, we demonstrate that it is feasible to reconstruct scenarios that were previously considered impractical. Specifically, we show that it is feasible to reconstruct a hidden scene submerged in scattering media, challenging the common assumption that light travels in the vacuum. We also prove that it is possible to compose a generalized optical system from multiple relay walls, relaxing the planar relay wall requirement.

EFFICIENT ALGORITHMS TO RECONSTRUCT THE HIDDEN SCENE. In terms of reconstruction quality, current NLOS approaches achieve promising results for recovering hidden scenes. Existing methods employ a variety of strategies to estimate the position and the albedo of a hidden surface from the captured transient signal. Yet, this task may require a considerable amount of computational resources, even when incorporating the assumptions mentioned previously, with memory demands that exceed the capabilities of portable or embedded devices. Furthermore, in many NLOS approaches, reconstruction quality is tightly tied to computational resources. As a result, NLOS usage is restricted for memory-constrained devices, high-precision reconstructions become impractical, and scalability materializes as a significant challenge.

For the second part of this thesis, we focus on the efficiency of the NLOS reconstruction in both memory and time. To that end, we again leverage the virtual wavefront abstraction provided by the Phasor Fields formulation [69]. This abstraction serves us as a foundation upon which we build two reconstruction algorithms.

Our first algorithmic solution simplifies the wavefront propagation of Phasor Fields, offering a trade-off between quality and memory consumption of the reconstruction. This approach benefits memory-constrained devices by reducing computational costs of reconstructing irrelevant areas of the hidden scene. Our second approach achieves substantially higher depth precision than prior approaches while maintaining the same temporal complexity. This method is capable of isolating reconstruction quality from computational demands, paving the path for novel approaches tailored to larger scales or more complex scenarios.

ENHANCING THE NLOS CAPTURE WITH COHESIVE AND CORRECT FORMULATIONS. In recent years, the NLOS panorama has conceived an explosion of proposals whose development has occurred independently to each other. The approaches often rely on ad-hoc assumptions and method-specific formulations. Consequently, there is a fragmentation between methods, hindering integration and interoperability between them. Besides, previous methods' formulations often ignore underlying physical events present at the moment of the capture. For example, noise or the temporal effects derived from the hardware at the capture degrade the signal use for NLOS methods, and they are rarely considered. This conceals the usage of NLOS to sophisticated and specific hardware, and long acquisition times, in order to minimize those undesired effects.

The third part of this thesis addresses capture-related limitations in NLOS imaging. We seek to model the physical effects occurring at the capture, enabling lower-end hardware and relaxing the calibration challenges during acquisition. To that end, we draw inspiration from quantum mechanics. Similarities exist between the uncertainty principle and NLOS imaging, as reconstructions represent probability distributions due to inherent ambiguities and incomplete information present in the light transport of the hidden scene. In a collaborative work, we adopt the Dirac notation, commonly employed in quantum mechanics, and adapt it for NLOS imaging.

Building upon this representation, the Dirac notation offers a cohesive and compact mathematical formulation that encompasses existing NLOS methods. We build on top of this mathematical representation to offer a more complete model of the physics present at the moment of capture in a compact syntax. Hence, we are capable of deriving a mathematical formalism that relates the ideal signal desired for any NLOS approach to what is actually captured. As we show, we can enhance the capture from low temporal resolution or with high noise derived from low photon counts, enabling the usage of lower-end devices and lower acquisition times for NLOS imaging.

In conclusion, we present several contributions towards the applicability of NLOS. Our advances in reconstructions allow for the generalized application of NLOS into new configurations and conditions, while lowering the computational demands. Furthermore, by leveraging the NLOS formulation based on Dirac, we can represent the light propagation within the hidden scene in a compact format without losing generalization. This relaxes the strict capture conditions and enables imaging with lower-end devices in more noisy scenarios. Together, we bring better and more efficient strategies to the NLOS community, reducing the gap between current implementations and in-the-wild applications. We believe this work provides foundations for future researchers' endeavors in the area. Novel NLOS strategies may achieve, e.g.,

adaptive approaches that relate capture and reconstruction, or compressed sensing to enhance measurements in less acquisition time, building upon this thesis’s contributions.

In the following, we provide more insights about the structure and the detailed contributions of this thesis.

1.1 GOAL AND OVERVIEW.

As previously stated, this thesis pursues the improvement of NLOS imaging for its practical implementation on three fronts. Therefore, the structure of this document is divided into three parts accordingly:

- Part I sets its focal point into generalizing the reconstruction of NLOS into broader and more varying setups. For that purpose, we leverage the Phasor Fields formulation, which we describe in detail in Section 3.2 of Chapter 3. This part is composed of the following chapters:
 - Chapter 4. Previous methods fail in handling hidden scenes with the presence of dense fog, smoke, or dust due to the assumption of *light travelling through the vacuum*. In this chapter, we demonstrate that it is feasible to reconstruct a hidden scene submerged in scattering media using the virtual wavefront abstraction of Phasor Fields.
 - In Chapter 5, we devise a novel physical setup consisting on multiple relay walls. Under Phasor Fields, these relay walls are treated as a single optical system, as we prove. The outcome is an increase in the hidden scene visibility, diminishing the effect of the missing cone [67].
- In Part II, our goal is to propose novel reconstruction algorithms, tailored for efficiency. We use Phasor Fields again, but in this case we use it as a base for our novel approaches. Our proposals are described in these chapters:
 - Chapter 6 describes our novel algorithm, which offers a trade-off between memory consumption and reconstruction quality. For that purpose, we simplify the light transport of efficient Phasor Fields algorithms for parallel planes, allowing us to reduce up to 16 times the memory consumption.
 - In Chapter 7, we leverage the virtual phase of the Phasor Fields, which previous methods discard, to devise an efficient algorithm for NLOS reconstruction. We report that our method achieves up to micrometer-scale depth variations with the same time complexity as previous works require for achieving centimeter-scale depth variations.
- Part III shifts to a general view of NLOS imaging, and focuses on providing a generalized formulation and its capture part of the pipeline. This part consists of one chapter:
 - Chapter 8 recompiles a collaborative work of this thesis, where we propose a novel cohesive representation for NLOS. To that end, we leverage Dirac notation, setting some similarities between NLOS and quantum mechanics.
 - Chapter 9. Here, we factor in the notation from the previous chapter to define a precise model of capture and computationally refine the NLOS signal. Consequently, we enable lower-end devices, noisy scenarios, and desynchronized captures for NLOS imaging.

Besides this introduction, we also incorporate a review of previous works related to the thesis in Chapter 2, which each chapter may expand for the specific topic treated there, and a short insight into the mathematical concepts and notation treated in this thesis in Chapter 3. Finally, we also present our conclusions and future work in Chapter 10.

While I am the leading author of most of the works presented in this thesis, they have been done in collaboration with different colleagues. For the sake of readability and completeness, the full description of the work is included in this document. At the beginning of each chapter, the works it refers to are contextualized and, when needed, my contribution is explicitly described.

1.2 CONTRIBUTIONS AND MEASURABLE RESULTS.

1.2.1 *Publications.*

This section details the publications that support this thesis's contributions. In particular, this thesis has led to three JCR publications, and two conference communications, that we order chronologically by publication date:

- (Chapter 4, Part I) Non-Line-Of-Sight Imaging in the Presence of Scattering Media using Phasor Fields.
 - *Pablo Luesia-Lahoz**, *Miguel Crespo**, *Adrian Jarabo*, and *Albert Redo-Sanchez*.
 - Published in Optics Letters in 2022.
 - Impact factor of 3.6, SJR index: 0.971 (Q1). Data from 2024.
- (Chapter 6, Part II) Zone Plate Virtual Lenses for Memory-Constrained NLOS Imaging.
 - *Pablo Luesia-Lahoz*, *Diego Gutierrez*, and *Adolfo Muñoz*.
 - Published in IEEE International Conference on Acoustics, Speech and Signal Processing (Proceedings of IEEE ICASSP 2023).
 - iCORE B. Data from 2025.
- (Chapter 8, Part III) A Cohesive Framework for Non-Line-of-Sight Imaging based on Dirac Notation.
 - *Albert Redo-Sanchez*, *Pablo Luesia-Lahoz*, *Diego Gutierrez*, and *Adolfo Muñoz*.
 - Published in Optics Express in 2024.
 - Impact factor of 3.3, SJR index: 0.930 (Q1). Data from 2024.
- (Chapter 7, Part II) 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*.
 - Published in IEEE International Conference on Computational Photography (Proceedings of IEEE ICCP 2025).
 - iCORE B. Data from 2025.
- (Chapter 5, Part I) Stereo Non-Line-of-Sight Imaging.
 - *Pablo Luesia-Lahoz*, *Sergio Cartiel*, and *Adolfo Muñoz*.
 - Published in the Visual Computer in 2026.
 - Impact factor of 2.9, SJR index: 0.637 (Q2). Data from 2024.

Besides, one work is currently under submission progress:

- (Chapter 9, Part III) Non-Line-of-Sight Imaging from Suboptimal Instrument Response.
 - *Pablo Luesia-Lahoz*, *Albert Redo-Sanchez*, *Diego Gutierrez*, and *Adolfo Muñoz*.
 - Under submission for revision to Optics Express.

1.2.2 *Research Internship.*

A research internship, for 3 months, has been carried out during this thesis:

- May 2024 - July 2024: Research stay at *University of Wisconsin-Madison*, Madison (Wisconsin, USA). Supervisor: Prof. Andreas Velten. Project: Zero Phase Phasor Fields. As a result of this stay, a

publication was presented at the International Conference of Computational Photography (ICCP) 2025.

1.2.3 *Published Code.*

During the development of this thesis, I have contributed to publishing the following code repository, which is made publicly available for future research:

- *y-tal: (Yours)-Transient Auxiliary Library*. Toolkit for simulation and analysis of time-resolved light transport captures. Contributors: *Diego Royo*, and *Pablo Luesia-Lahoz*.
 - Available at <https://github.com/diegoroyo/tal>. Starred 28 times.

Besides, I have contributed to a lesser extent to the development of the following code repository:

- *mitransient*. Code for *Non-line-of-sight transient rendering* in Mitsuba 3 - Full implementation of transient path tracing [95]. Authors: *Diego Royo*, *Miguel Crespo*, and *Jorge Garcia-Pueyo*.
 - Additional contributors: *Diego Bielsa*, *Pablo Luesia-Lahoz*, *Óscar Pueyo-Ciudad*, *Guillermo Enguita-Lahoz*, and "lyd405121" (<https://github.com/lyd405121>)
 - Available at <https://github.com/diegoroyo/mitransient>. Starred 58 times.

1.2.4 *Supervised Students.*

During the development of this thesis, I have co-supervised the following students:

- Jul 2023. Ayush Gupta. Internship. *Frequency Analysis of Information given Transient Signals*.
- Jul 2023. Isaac Velasco. Internship. *3D optimizations for Non-Line-of-Sight Imaging with Multiple Points of View*.
- Sept. 2024. Jorge Solán Morote. Bachelor thesis (Trabajo Fin de Grado). *Simulation of Ultra-Fast Photon Capture Hardware in Light Transportation*. Grade 9/10.

1.2.5 *Miscellaneous: dissemination and reviewing.*

During my PhD, I have been a reviewer for the following international journals:

- Optics Express
- Optics Letters

Also, I have collaborated in the following dissemination activities:

- Young Researchers Day edition XII (organized by I3A).
- European Researchers' Night 2022, 2023, and 2025.
- Engineering and Architecture Week edition XV (organized by *Universidad de Zaragoza*).
- Presentation in situ of our work "Zero-Phase Phasor Fields for Non-Line-of-Sight imaging" in the ICCP conference held in Vancouver in 2024.

RELATED WORK.

In this chapter, we review the most relevant literature relevant to this thesis. First, Section 2.1 provides a general introduction to previous work on transient imaging and its application in Non-Line-of-Sight imaging, focusing on capture and reconstruction methods. Second, we focus on the state-of-the-art of the Phasor Fields framework in Section 2.2, from which we build upon in Parts I and II. Additionally, each chapter may incorporate a brief literature review specific to the work described there.

2.1 TRANSIENT NON-LINE-OF-SIGHT IMAGING.

Non-Line-of-Sight imaging aims to reconstruct scenes that are partially or fully occluded from the direct line of sight, employing indirect illumination that reaches visible surfaces from the hidden scene. Within NLOS imaging, a common division exists between passive and active methods. Passive methods employ either existing illumination of the hidden scene in the form of ambient light, or light emitted by the hidden objects themselves [5, 101, 58, 77]. These approaches are appealing because they do not require specialized hardware for the illumination, yet they are generally limited by low signal-to-noise ratio. Moreover, the lack of light modulation restricts the quality of their results. On the other hand, active NLOS methods employ controlled illumination and then analyze the indirect reflectance of the hidden scene. Transient NLOS imaging is a subcategory of active NLOS, which uses high temporal resolution hardware to image the hidden scene using the time-of-flight of the light. We refer the interested reader to recent surveys [76, 21, 30]. The content of this thesis lies in transient NLOS imaging category. We briefly discuss literature related to transient captures and reconstruction for NLOS in the following.

2.1.1 *Transient captures for Non-Line-of-Sight imaging.*

Transient imaging aims to capture the light propagation in motion, capturing trillions of frames per second using a combination of ultra-fast sensors, illumination devices, and computational tools [51]. Although the term was coined with the work of Velten et al. in 2013 [116], one of the first methods to report capturing the light propagation over time was achieved with interferometry in 1978 by Abramson [1]. This early approach employed two short laser pulses, one travelling through the scene and one as a reference. When recorded in a hologram, the pulses produce constructive interference, allowing to spatially differentiate the light propagation over time in the scene.

Hardware has advanced since then, until Velten et al. directly record the temporal time stamp using ultra-fast lasers and *streak cameras* [116]. One result of this capture can be seen in Figure 2.1. These devices sacrifice one spatial dimension, which transforms into a temporal map of the incoming light, achieving effective resolution of picoseconds. Soon this technology was used in NLOS imaging [36, 115]. Other alternatives employed *photonic mixer devices*, that, modulating continuous waves, offered a lower budget solution at the cost of temporal resolution [40, 23]. When applied in NLOS, this temporal resolution was sufficient for location, but not to achieve the geometry of the hidden scene [42].

The proposal to use single-photon avalanche diodes (SPAD) to capture the transient signal appeared in 2015 [27]. These devices trigger an electron avalanche at a photon arrival, offering an alternative with lower power of the illumination source while reducing the capture time. Shortly, these devices were applied in NLOS scenes [28, 7]. SPADs application reports multiple achievements in the NLOS imaging field, such as capturing NLOS scenes at kilometer distances [121], capturing complete scenes in real-time [81, 64], and capture using of polarization for NLOS [89].



Figure 2.1: A transient imaging capture. A pulse of light enters from the bottom of the bottle (left), and we can see its propagation to the cup of the bottle (right). The effective frame rate of this capture is trillions of frames per second. *Results adapted from Velten et al. work [116].*

Recently, other strategies have emerged, which are named transient renderers [95, 50]. These works simulate the light transport over time computationally. Consequently, these methods offer an agreement between realism and controllability of a scene [78], without the requirement of precise calibration or expensive hardware for its capture. Royo et al. proposed specific optimizations for rendering NLOS scenes in a transient context [96]. This provides a powerful tool for prototyping and validating novel NLOS approaches.

In this thesis, we rely on rendered NLOS scenes to validate our contributions, which appear across all chapters. To ensure robustness, in Chapters 4, 5, 7 and 9, we also evaluate our methods on publicly available datasets captured with SPAD sensors, which present more challenging conditions due to noise, calibration errors, and electronic artifacts.

2.1.2 Reconstructions methods for Non-Line-of-Sight imaging.

The idea of using photons’ time of flight to image a hidden scene using a relay wall was originally conceived by Kirmani et al. [55, 56]. This approach draws parallels between NLOS imaging and computed tomography problems in medical imaging. As a result, one of the first high-quality 3D reconstructions of the hidden scene employed filtered ellipsoidal back projection (FBP) [115]. This motivated later approaches, which iteratively estimate and refine the reconstruction using error FBP [59], or Laplacian and Gaussian FBP [61]. These methods are conceptually straightforward. Yet, they struggle with noise, limited sampling, and complex light interactions [36].

Later, O’Toole et al. restricted the NLOS acquisition to *confocal configurations* [86], in which the laser illumination and capture points coincide on the relay wall. This allows for simplifying the light transport in the hidden scene, reformulating the NLOS problem as a linear inverse problem solvable with deconvolution-based techniques. Yet, these techniques consume considerable temporal resources. Lindell et al. addressed this limitation with their *fk-migration* approach [66]. They solved the problem in the frequency domain, drawing inspiration from seismology, and using Stolt’s interpolation, achieving much higher temporal efficiency at the cost of storing the whole capture in memory. These two approaches yield high-quality reconstructions, although they degrade their performance on non-confocal configurations.

Other methods employ the Fermat path theory to estimate the geometry of an NLOS scene [124], although they neglect the albedo information. Alternatively, some others employ a global inversion of the transient light transport [41, 125] at the cost of higher computational cost. Inverse rendering techniques have also been explored [48], with recent advances in differentiable rendering [127, 88]. The span of data-driven methods has led numerous works to adopt neural network-based approaches that combine simulation and real captures to reinterpret transient responses from hidden scenes [9, 10, 11, 49, 39, 126].

The Phasor Fields framework for NLOS was introduced in 2019 [69], which effectively poses the relay wall as a virtual camera. This formulation has led to robust and efficient reconstruction methods. We review the literature of Phasor Fields in greater detail in Section 2.2.

Throughout this thesis, we employ three reconstruction approaches from previous works: filtered backprojection [115], *fk-migration* [66] and Phasor Fields [69]. In Figure 2.2 we show the results of reconstructing with these three methods (figure adapted from Lindell et al. work [66]). Many of the

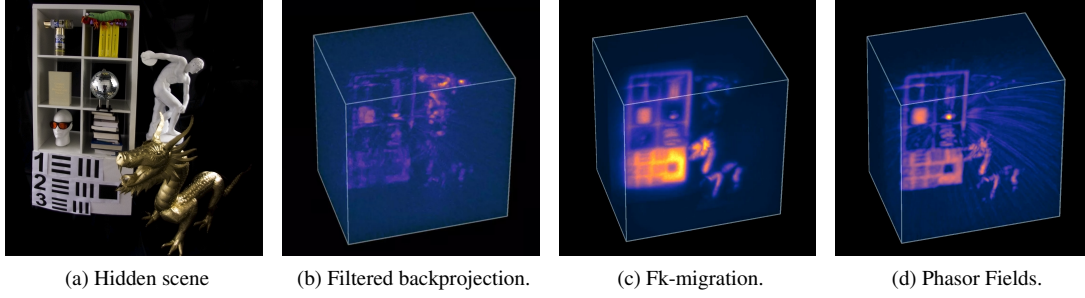


Figure 2.2: Results of the three reconstruction methods employed in this thesis. From left to right: conventional picture of the hidden scene, reconstruction of filtered backprojection, reconstruction of fk-migration, and reconstruction of Phasor Fields. *Results adapted from Lindell et al. work [66].*

contributions of this thesis build upon Phasor Fields, as we detail in Section 2.2, particularly the Chapters 4 to 7. Additionally, we use filtered backprojection for comparison analysis in Chapters 5 and 9. Finally, in Chapter 9 we employ all three approaches to illustrate the impact of our proposed formulation to enhance capture, especially in the case of filtered backprojection, which is more sensitive to noise and capture artifacts.

2.2 PHASOR FIELDS.

The Phasor Fields framework [69] formulates the NLOS imaging problem as a *virtual* wavefront propagation problem, interpreting the transient signal arriving at the relay wall as waves reaching on a wall-sized aperture. In contrast to previous approaches that rely on ad hoc strategies to invert the light transport of the hidden scene, Phasor Fields leverages well-known concepts from Fourier optics in the line-of-sight domain to image the hidden scene [94, 92, 93, 110, 17, 19, 18, 107]. For this reason, Phasor Fields represents a milestone in NLOS imaging. Across this thesis, we build upon the Phasor Fields formulation for experimental validation, and in Parts I and II we further explore its insights to establish our contributions.

2.2.1 Virtual wavefront abstraction.

Dove and Shapiro showed that, within the Phasor Fields framework, imaging the hidden scene is equivalent to focusing using a Fresnel diffraction integral [17, 19] or a Rayleigh–Sommerfeld diffraction (RSD) integral [18]. Hence, this formulation establishes the relay wall as a virtual camera, enabling the migration of concepts from line-of-sight imaging to NLOS imaging, such as resolution analysis [70]. Subsequent works extended this formulation to account for more complex light interactions within the hidden scene. Marco et al. demonstrated the differentiation of direct and indirect illumination in synthesized scenes using light transport matrices [79], a result later corroborated in real-world scenarios by Sultan et al. [106]. Royo et al. further observed that the virtual wavefronts in Phasor Fields exhibit mirror-like behavior, allowing imaging beyond the third bounce by exploiting fifth-bounce photons from the hidden scene [98]. Additionally, Phasor Fields has demonstrated notable flexibility and robustness with respect to relay-wall geometry, mitigating the effects of dynamic surfaces that undergo motion during capture [60].

In this thesis, we draw motivation by the Phasor Fields robustness, and we adopt state-of-the-art implementations in Part I. By modeling the hidden scene as a virtual wavefront propagation, we exploit well-known wave optics phenomena to enable imaging through scattering media in Chapter 4, whereas prior Phasor Fields and NLOS approaches assumed light propagation in vacuum. Furthermore, in Chapter 5 we build upon the concept of the relay wall as a virtual line-of-sight system to introduce stereo NLOS imaging. In doing so, we demonstrate that the Phasor Fields formulation naturally supports the combination

of multiple independent relay walls into a single optical system, whereas previous methods used planar relay-walls, or required constraining assumptions to handle three-dimensional relay-wall geometries [34].

2.2.2 *Phasor Fields efficiency.*

In the original Phasor Fields work, the authors employed a temporal-domain integral equivalent to filtered backprojection [115], requiring high $\mathcal{O}(n^5)$ complexity to image a hidden scene. Shortly after, more efficient implementations emerged by exploiting the diffraction-like behavior of the RSD [68]. These works proposed kernels for imaging planes parallel to the relay wall, enabling efficient convolution-based implementations via spatial Fourier transforms, and reducing its complexity to $\mathcal{O}(n^3 \log n)$. This line of research enabled NLOS imaging at up to five frames per second [81], combining specialized hardware configurations with optimized computational pipelines to achieve low latency in meter-scale scenes. Jiang et al. later showed that the computational cost of generating RSD kernels could be reduced by exploiting their radial structure [52]. When combined with dedicated FPGA implementations, this approach enabled reconstruction rates of up to 25 frames per second [64]. Extending these plane-based kernels, Gu et al. proposed a formulation for non-planar relay walls using spatial 3D Fourier transforms [34].

In Part II, we further explore the virtual wavefront propagation paradigm of Phasor Fields and propose two novel algorithmic solutions. In Chapter 6, we introduce a propagation method based on Fresnel zone plates that enables a memory–quality trade-off for reconstruction, addressing a limitation of prior methods that primarily focused on temporal efficiency. Additionally, drawing inspiration from interferometry, we exploit the virtual phase of the propagated wavefront, discarded by previous approaches, to refine the spatial positioning of the hidden scene without increasing reconstruction complexity in Chapter 7.

BACKGROUND.

3.1 NON-LINE-OF-SIGHT IMAGING NOMENCLATURE.

NLOS imaging aims to reconstruct a hidden scene that is occluded from the direct line of sight. Researchers exploit the indirect illumination to gain access to the hidden scene information, using an auxiliary surface where the light reflects into and from, commonly named relay wall. As mentioned in Section 2.1 from Chapter 2, this thesis falls into the category of active NLOS imaging, specifically transient NLOS imaging. Therefore, as common to works in the field, we employ a pair of ultrafast devices: a short pulse emitter and a transient sensor. Combined, both devices can record the light propagation multiplexed over time. Pointing these devices to the relay wall, it is possible to obtain the transient indirect illumination of the hidden scene. In this manner, the pulse emitter illuminates a set of points $\mathbf{x}_l \in \mathbf{L}$ in the relay wall, from which light is scattered into the hidden scene. This light interacts with the hidden scene, which scatter back to the relay wall. The sensor, focused on a set of points $\mathbf{x}_s \in \mathbf{S}$, receives part of this light, capturing its time of flight, which is a time-resolved signal.

NLOS imaging approaches assume that the temporal profile of the emitted light is a differential pulse, a delta function in time $\delta(\mathbf{x}_l, t)$. Thus, under this assumption, the recorded time-resolved signal represents the impulse response of the hidden scene as a function of time as $H(\mathbf{x}_l, \mathbf{x}_s, t)$. However, delta functions are practically impossible to achieve in the real world in practice, primarily due excitation and decay time of materials that emit photons. To minimize this effect, researchers employ sophisticated and expensive lasers whose pulse widths are in the range of picoseconds.

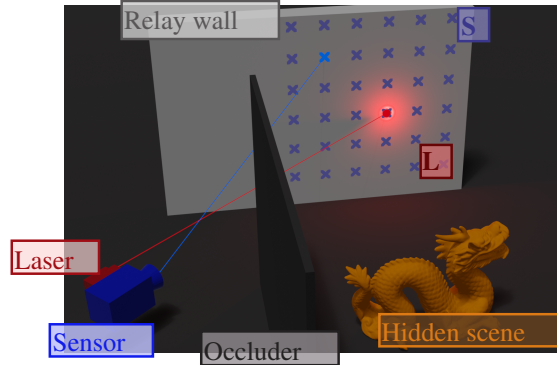


Figure 3.1: **Non-Line-of-Sight setup.** A laser illuminates a set of points $\mathbf{L}$ in the relay wall with a short light pulse. The relay wall scatters part of the light into the hidden scene, which consequently scatters part of this illumination back to the relay wall. A sensor focused on a set of points $\mathbf{S}$ in the relay wall acquires the temporal profile of light.

Figure 3.1 shows a common NLOS imaging setup. Usually, the acquired signal is shifted in time to discard the time of flight of the light from the devices to the relay wall. Therefore, the set of points $\mathbf{L}$ defines the illumination points, and $\mathbf{S}$ the capture points. Note this is a simplification.

The distribution of the illumination points $\mathbf{x}_l \in \mathbf{L}$ and capture points $\mathbf{x}_s \in \mathbf{S}$ in the relay wall defines the capture configuration. In this thesis, we distribute the capture points $\mathbf{S}$ on a regular grid. Yet, depending on the positioning of the illumination points with respect to this grid, we define different configurations (also see Figure 3.2):

- *Confocal* capture configuration refers to collocating sensor and illumination points ($\mathbf{x}_l = \mathbf{x}_s$). The illumination point also acts as a capture point, allowing some reconstruction algorithms to exploit simplification assumptions [86, 66].
- *Non-confocal* captures distribute the points $\mathbf{x}_l \in \mathbf{L}$ and $\mathbf{x}_s \in \mathbf{S}$ independently. In other words, the captured signal of the scene may be the result of the illumination of any other point in the relay wall. Inside the non-confocal configuration, we can also identify two sub-configurations:
 - *Exhaustive* configuration, where $\mathbf{x}_l \in \mathbf{L}$ are also distributed in a regular grid, independent of the capture grid $\mathbf{S}$.
 - *Single* configuration consisting of a single illumination point $\{\mathbf{x}_l\} = \mathbf{L}$, usually placed at the center of the relay wall.

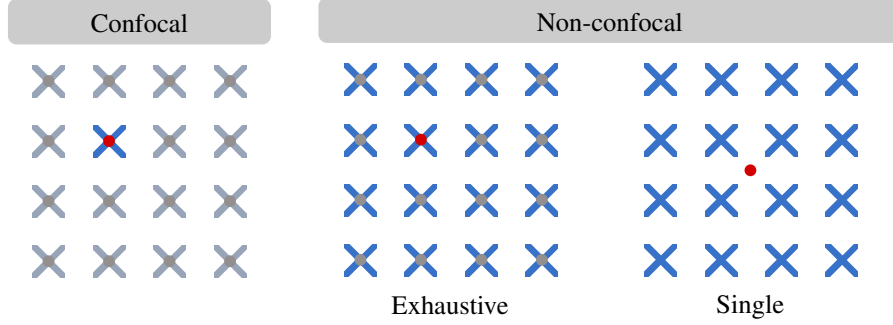


Figure 3.2: **NLOS capture configurations.** Distribution of capture points (blue X) and illumination points (red dots). Capture and illumination points that are not active at the capture are represented in gray. In the confocal configuration, capture and illumination points are located in the same position. In non-confocal, capture and illumination points are placed independently, allowing for exhaustive configurations (all capture points can acquire hidden scene data illuminated from any of the illumination points) or single configurations (there is only a single illumination point at the center of the relay wall).

In this thesis, we primarily employ the single configuration, that is non-confocal, in Chapters 4 to 6 and 9. Chapters 4, 7 and 9 also employ confocal configurations. Along with the results presented in each chapter, we specify the configuration of the scenes.

3.2 THE PHASOR FIELDS FORMULATION.

A considerable contribution of this thesis builds upon the Phasor Fields formulation [69]. In this section, we introduce Phasor Fields and its insights, and we define the notation used in subsequent chapters.

Like many previous approaches, Phasor Fields assumes that the emitted light from $\mathbf{x}_l$ is a delta function $\delta(\mathbf{x}_l, t)$. Note that, while we follow this assumption for Parts I and II, this introduces some limitations, that we later addressed in Part III. In a convolution, a delta function acts as an identity operation. Therefore, any function convolved with $H(\cdot, t)$ under this delta assumption results in a capture as if the scene were virtually illuminated with such function. Thus, Liu et al. define an arbitrary input signal $\mathcal{P}_g(t)$ as a virtual illumination temporal profile to yield

$$\mathcal{P}(\mathbf{x}_l, \mathbf{x}_s, t) = \mathcal{P}_g(t) *_t H(\mathbf{x}_l, \mathbf{x}_s, t), \quad (3.1)$$

being $*_t$ a convolution in time, and $\mathcal{P}(\mathbf{x}_l, \mathbf{x}_s, t)$ the captured signal as the hidden scene was virtually illuminated with $\mathcal{P}_g(t)$. Following the original work, we model the input signal as a Gaussian-envelope wave pulse, with central wavelength λ_g , and a standard deviation σ . Thus,

$$\mathcal{P}_g(t) = \rho e^{-\frac{t^2}{2\sigma^2}} e^{i\omega_g t + \psi_0}, \quad (3.2)$$

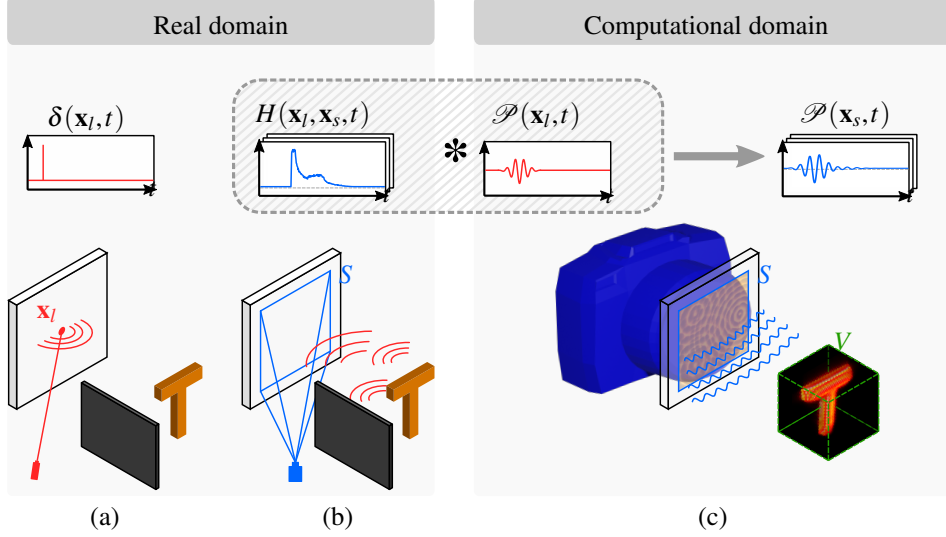


Figure 3.3: **Phasor Fields formulation overview.** In an NLOS setup in the real world (a), a laser emits a delta pulse to a point $\mathbf{x}_l$ in the relay wall that illuminates the hidden scene. Then, an ultra-fast capture device (b) focused at an area S in the relay wall acquires the impulse response $H(\mathbf{x}_l, \mathbf{x}_s, t)$ from the hidden scene. In the computational domain (c), an input signal $\mathcal{P}(\mathbf{x}_l, t)$ serves as a virtual illumination of the hidden scene when convolving with the impulse response $H(\cdot)$. The outcome is a field of phasors $\mathcal{P}(\mathbf{x}_s, t)$ placed at the area S of the relay wall. The field of phasors is equivalent to a planar wavefront, and thus it is possible to use well-known Fourier optics to reconstruct a volume $\mathbf{V}$, posing the relay wall as a virtual camera. *Figure adapted from Royo et al. 2023 [98].*

being ρ a scale factor, ψ_0 the defined phase at $t = 0$, c the speed of light, and ω_g the central angular frequency which is related to the wavelength such as $\lambda_g = 2\pi c / \omega_g$.

For the notation of this thesis, we define $\hat{O}(\cdot, \omega)$ as a single component ω in the frequency-domain, i.e. after a Fourier transform, from a time-domain signal $O(\cdot, t)$. Note then that $\hat{\mathcal{P}}(\mathbf{x}_s, \mathbf{x}_l, \omega)$ is a phasor: a single frequency component from the time-domain signal $\mathcal{P}(\mathbf{x}_s, \mathbf{x}_l, t)$. Therefore, considering the whole grid of $\mathbf{x}_s \in \mathbf{S}$ in the relay wall behaves as a phasor field (i.e., a wavefront with frequency ω), which gives the formulation its name. Equation (3.1) can be expressed for a specific frequency component ω as:

$$\hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega) = \hat{\mathcal{P}}_g(\omega) \hat{H}(\mathbf{x}_l, \mathbf{x}_s, \omega). \quad (3.3)$$

In this computational domain, the authors interpret the phasor field formed by $\hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega)$ as out-of-focus illumination of the hidden scene. Hence, the original light distribution can be observed by bringing that out-of-focus illumination into focus at each specific point. Under this premise, the authors can use well-known operators from classic Fourier optics to model the propagation of the virtual wavefronts. **Therefore, the Phasor Fields framework poses the NLOS imaging problem as a virtual LOS problem, virtually setting the relay wall as a virtual camera.** We define our imaging operator Φ to compute the in-focus value of a point $\mathbf{x}_v \in \mathbf{V}$ of a reconstruction volume $\mathbf{V}$ as

$$\hat{I}(\mathbf{x}_v, \omega) = \Phi_\omega [\mathbf{x}_v, \hat{\mathcal{P}}(\mathbf{x}_s, \mathbf{x}_l, \omega)]. \quad (3.4)$$

assuming $\hat{I}$ as the image formation function (in the Fourier domain), where Φ is a computational focusing operator equivalent to the lens of a traditional camera. This permitted the authors to make the analogy of the relay wall as a virtual camera, thus posing the NLOS problem as a virtual LOS one. Here, we employ $I(\mathbf{x}_v, t)$ to denote the temporal domain image formation, while $\hat{I}(\mathbf{x}_v, \omega)$ is its Fourier-domain counterpart.

In their work, Liu et al. employ a function to propagate their virtual phasors as

$$\hat{R}(\mathbf{x}_a, \mathbf{x}_b, \omega) = \frac{e^{i2\pi\omega|\mathbf{x}_a - \mathbf{x}_b|/c}}{|\mathbf{x}_a - \mathbf{x}_b|}, \quad (3.5)$$

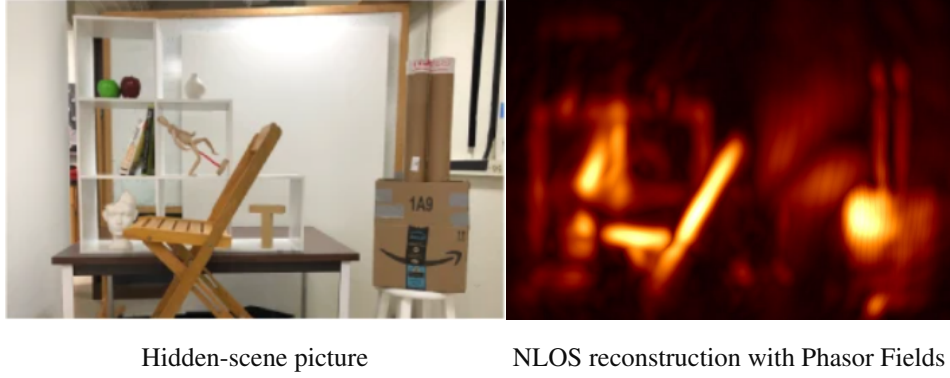


Figure 3.4: **Phasor Fields reconstruction.** Right, a picture of the hidden scene from the relay wall as seen with a traditional camera. Left, Phasor Fields reconstruction of the NLOS capture. Phasor Fields transforms the NLOS capture into a virtual wavefront propagation problem, enabling Fourier-optics reconstructions. Consequently, Phasor Fields can successfully reconstruct complex hidden scenes.

hence, they can model Φ_ω with a function that resembles the Rayleigh-Sommerfeld Diffraction (RSD) integral:

$$\Phi_\omega^{(c)}[\mathbf{x}_v, \hat{\mathcal{P}}(\mathbf{x}_s, \mathbf{x}_l, \omega)] = \int_{\mathbf{L}} \hat{R}(\mathbf{x}_v, \mathbf{x}_l, \omega) \int_{\mathbf{S}} \hat{R}(\mathbf{x}_v, \mathbf{x}_s, \omega) \hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega) d\mathbf{x}_s d\mathbf{x}_l, \quad (3.6)$$

where c is the speed of light and $|\cdot|$ is the Euclidean distance. The authors named the image formation described by Equation (3.6) as *confocal camera* [69] (do not mistake with confocal configuration for capture [86]), which we represent as $\Phi_\omega^{(c)}[\cdot]$, and it is characterized for bringing into focus both illumination and capture to a point $\mathbf{x}_v$.

Employing Equation (3.6) for imaging formation, and applying an inverse Fourier transforms, allows obtaining at time $t = 0$ the geometry at the exact moment it scatters the light. Alternatively, the authors also present an alternative named *transient camera*. This other model omits the propagation from the laser position, resulting in

$$\Phi_\omega^{(t)}[\mathbf{x}_v, \hat{\mathcal{P}}(\mathbf{x}_s, \mathbf{x}_l, \omega)] = \int_{\mathbf{S}} \hat{R}(\mathbf{x}_v, \mathbf{x}_s, \omega) \hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega) d\mathbf{x}_s d\mathbf{x}_l. \quad (3.7)$$

This second model, when transformed into time domain, represents at $t = 0$ the moment the light scatters from the relay wall, and $I(\mathbf{x}_v, t)$ yields the light propagating at the point $\mathbf{x}_v$.

In practice, the acquired signal at $\mathbf{S}$ is discrete over time, as points in $\mathbf{S}$ and $\mathbf{L}$ are discrete over space. Hence, the computation of Equation (3.6) and Equation (3.7) results in summations instead of integrals over the voxel grid $\mathbf{V}$, yielding a complete 3D reconstruction.

This thesis makes use of Phasor Fields in a greater or lesser extent throughout all its chapters. Understanding Parts I and II requires thoughtful knowledge of Phasor Fields: Chapters 4 and 5 use the confocal camera of Phasor Fields for imaging. Chapter 4 focuses on turbid media, while Chapter 5 addresses imaging using multiple relay walls. Chapter 6 adapt the transient camera to reduce the memory usage based on Fresnel Zone Plates, and Chapter 7 builds upon the confocal camera and uses the virtual phase of Phasor Fields to achieve micrometer depth variations. Additionally, in Part III we make use of Phasor Fields to illustrate the Dirac notation cohesion to any imaging method, among others, in Chapters 8 and 9.

Part I

TOWARDS GENERALIZED RECONSTRUCTION OF HIDDEN
SCENES.

ABOUT THIS CHAPTER

In this work, we challenge a common assumption in NLOS imaging: light travels in the vacuum. We submerge the hidden scene into dense scattering media, and demonstrate that it is feasible to reconstruct it using Phasor Fields under its assumption of virtual wavefront propagation.

This chapter corresponds to a work published at the *Optics Letters* [71]. As joined-first author, I contributed to the experimental validation, discussion, writing, and revision of the manuscript.

P. Luesia, M. Crespo, A. Jarabo, & A. Redo-Sanchez
Non-line-of-sight imaging in the presence of scattering media using phasor fields
 Optics Letters

4.1 INTRODUCTION.

Most of NLOS techniques assume that light travels in the vacuum with no scattering media and it is only scattered by the relay wall. This assumption hinders the applicability of NLOS imaging to scenes in which the presence of scattering media (e.g., smoke or fog) might be significant. Imaging through scattering media is challenging even in LOS setups [90], where many techniques have been proposed such as optical coherence tomography (OCT) [47, 85, 37, 31], diffuse optical tomography (DOT) [128, 83, 65, 75], time-of-flight (ToF) measurements [43, 112], coherent light modulation using spatial light modulators [114], or polarization [14, 22, 123].

Given the amount of incoherent light due to scattering, we cannot rely on traditional inversion methods for NLOS imaging. Instead, we leverage recent advances on wave-based NLOS imaging, and in particular on the Phasor Fields framework proposed by Liu and colleagues [69]. Phasor Fields enables creating virtual light sources and sensors on the visible surfaces, resulting in robust reconstruction of multiple incoherent scattered light sources [93, 17, 110] (more details in Section 2.2). In this chapter, we empirically analyze the performance of Phasor Fields in NLOS scenes submerged in scattering media (Figure 4.1). We show reconstructions of NLOS scenes submerged in scattering media of increasing density, showing that Phasor Fields is capable of reconstructing scenes in very challenging visibility conditions using a single-frequency laser grid for illumination projected on a single plane. We hope this work will expand the capability of NLOS imaging to see through scattering media and foster new avenues of work in the NLOS imaging field.

While Phasor Fields have demonstrated very effective at reconstructing NLOS scenes in standard ambient conditions, little has been investigated in relation to its performance in reconstructing scenes submerged in scattering media, which results in more challenging visibility conditions for the experimental equipment operating in the visible range. In the following, we analyze the performance of Phasor Fields in scenes hidden behind a visible diffuser and submerged in scattering media of increasing density (see Figure 4.1). In this setup, the diffuser, which transmits the light in and out of the scene, acts as the relay wall, as opposed to a traditional NLOS setup in which the relay wall reflects the light into the scene, generating a geometry for looking around corners. The setup can be considered as a LOS setup through heterogeneous media in which the diffuser acts as an occluder of the scene.

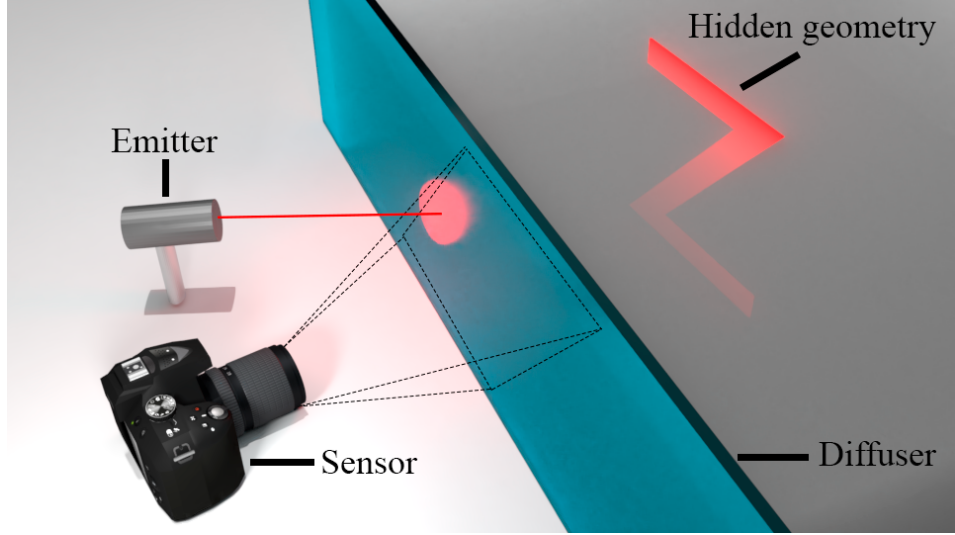


Figure 4.1: **Scene setup.** The scene is hidden behind a diffuser, with the scene being submerged in a scattering medium.

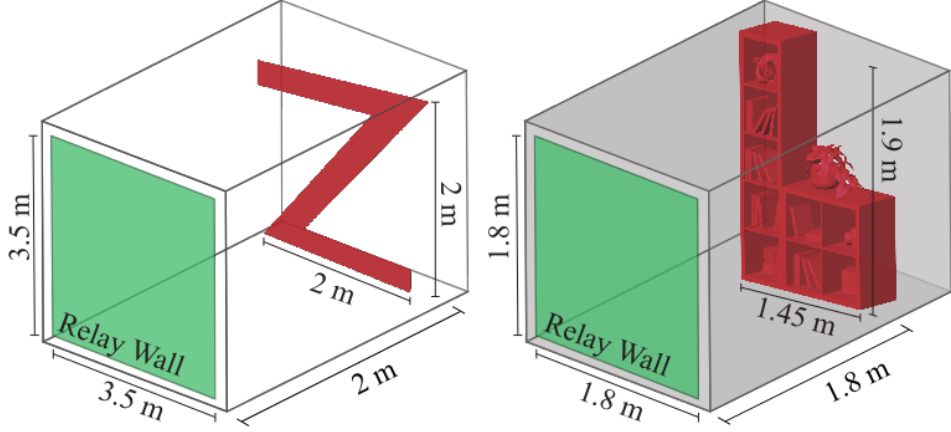


Figure 4.2: **Simulated scenes.** We use two simulated scenes: (Left) a single planar letter behind the diffuser (green), and (Right) a closed room with a shelf (red) at the back.

4.2 EXPERIMENTAL DESIGN.

We perform our experiments using both simulated data and experimental data. The simulated scenes allow controlling the properties of the scattering media, while the experimental data allows assessing the behavior in real-world scenarios.

For **simulation**, we use two different scenes (Figure 4.2): A simple scene with a planar Z-letter located at 2 meters away from the diffuser, and a complex scene containing a shelf located 1.8 meters away from the diffuser and completely enclosed in a room (the diffuser is one of the walls of the room). Phasor Fields requires a relay wall surface. The diffuser is a planar surface in both scenes that acts as the relay wall, which also serves as boundary between the scattering media and the outside world. We avoid undesired reflections, being the only mismatch between the Fresnel index at the relay wall. The Z-LETTER scene allows studying how the medium affects the reconstructions. The SHELF scene is more challenging due to feature occlusions and multi-path effects and it is useful to validate the results derived from the Z-LETTER scene in more complex scenarios.

We characterize the scattering media by using their bulk optical parameters. In particular, we use the extinction coefficient $\mu_t = \mu_a + \mu_s$ [in m^{-1}], and the scattering albedo $\alpha = \mu_s / \mu_t$ [unitless], with μ_a and μ_s the absorption and scattering coefficients [in m^{-1}], respectively. Intuitively, the extinction is

related to the medium density and the albedo is related to the strength of the scattering. We model the directionality of the scattering by using the Henyey-Greenstein [44] phase function, which models the directionality using an anisotropy factor $g \in (-1, 1)$, where $g > 0$, $g = 0$, and $g < 0$ for forward, isotropic, and backward scattering, respectively. Unless stated otherwise, we assume an isotropic phase function $g = 0$ and homogeneous media.

We use a publicly available transient renderer software [50] to compute the impulse response $H(\mathbf{x}_p \rightarrow \mathbf{x}_c, t)$ of the hidden scenes. We compute each impulse response for 128^2 laser positions $\mathbf{x}_p$ equally spaced in a grid along the visible diffuser, and captured in a single SPAD point $\mathbf{x}_c$. The temporal resolution is 4096 bins, each with a temporal resolution of about 3 ps. We use the **experimental** data captured by Lindell and Wetzstein [65], which corresponds to the impulse response of a scene behind a polyurethane foam slab using lasers and SPAD sensors. They captured 32^2 equally-spaced positions $\mathbf{x}_c$ using a confocal configuration ($\mathbf{x}_c = \mathbf{x}_p$). The scenes are located behind a 2.54 cm thick slab that acts as both the diffuser and a dense scattering media. The scene itself is not submerged in any scattering medium. The estimated scattering properties values for the slab are $\mu_t \approx 262.52 \text{ [m}^{-1}\text{]}$ and $\alpha \approx 0.99$ [unitless].

4.3 RECONSTRUCTION.

We reconstruct a volumetric representation of the scenes by sequentially focusing the virtual imaging system on each voxel. The voxelization resolution is $73 \times 59 \times 73$ for the Z-LETTER scene, 142^3 for the SHELF scene, and 32^3 for all the CAPTURED scenes from Lindell and Wetzstein. We use the Matlab solver provided by the authors [69]. On an Intel Xeon E5 with 256 GB RAM, reconstructions took between 5 to 254 seconds.

We employ the confocal camera model (Equation (3.6)) for our results. Following Liu et al. [69], we set the Gaussian pulse (Equation (3.2)) with a central wavelength $\lambda_g = 4\Delta_c$ with Δ_c , being the distance between sampled points on the relay wall. The pulse has a complete width of 4λ in the simulated data, and 8λ in the experimental data ($\sigma = 4\lambda_g/6$ and $\lambda_g = 8\lambda/6$ respectively). Similar to LOS Fourier optics, the wavelength defines the maximum resolution of the reconstruction. Later, we analyze the effect of λ_g in depth of the reconstructions. For visualization, we remove the contribution of the extinction of the medium by filtering the final reconstruction in post-process. We exploit the spatial information of the voxels, as well as the known properties of the medium, and apply the following scale factor $K(\mathbf{x}_v)$ for each voxel

$$K(\mathbf{x}_v) = \frac{(1 - \alpha e^{-|\mathbf{x}_v - \mathbf{x}_c| \alpha})}{(e^{-|\mathbf{x}_v - \mathbf{x}_c| \mu_t})}, \quad (4.1)$$

which depends on the bulk optical properties of the medium. We define this scaling factor as motivated by the attenuation term of the radiative transfer equation. On a real-world scenario, the media properties are usually unknown, but could be obtained from the measurements using inverse scattering [32].

4.4 ANALYSIS.

Figure 4.3b shows the results of Phasor Fields when reconstructing the Z-LETTER scene as a function of different medium parameters (the extinction coefficient μ_t , and scattering albedo α). For low scattering albedo, as the medium thickness increases, the amplitude of the signal is attenuated, although the coherence remains and, hence, it is still possible to reconstruct a sharp image. In contrast, a higher scattering albedo results into less coherence and, therefore, into the loss of high-frequency information.

Figure 4.3c shows the effect on the reconstructions of the phase function's anisotropy. For forward scattering, the phasor penetrates deeper into the medium, whereas backward scattering results into more scattering near the virtual camera and, therefore, poorer visibility conditions.

Now we analyze the effect of the phasor wavelength λ_g on the reconstruction depth. Figure 4.4 compares the reconstruction in a highly scattering medium ($\mu_t = 1 \text{ m}^{-1}$, $\alpha = 0.83$) for increasing λ_g . In particular, we double and triple the optimal central wavelength defined by $\text{centrallambda} = \lambda_0 = 4\Delta_c$. This results into a resolution loss, visible as smoother reconstructions. However, we still can recover the geometry hidden by the scattering medium.

We also analyze how our experiments extrapolate to scenes with significantly more complexity. Figure 4.5 shows the NLOS reconstruction of the SHELF scene. The scene contains surface-to-surface multi-path, and multiple scattering. This results in larger indirect contribution to the medium, thus resulting in more challenging visibility conditions. As our results show, even for a relatively thick medium, we can reconstruct the scene with good quality.

Finally, we analyze the performance of Phasor Fields on the confocal data captured by Lindell and Wetzstein [65]. Similar to ours, their method is based on an inversion technique used for NLOS imaging (fk migration [66]). The qualitative results shown in Figure 4.6 are on-par with Lindell and Wetzstein’s method, without any modification of the Phasor Fields framework, demonstrating its flexibility even in challenging conditions.

In conclusion, we have empirically analyzed the performance of Phasor Fields in the presence of scattering media. We have leveraged the capability of Phasor Fields in transforming a NLOS problem into a virtual LOS one, in which the scattering media acts as an additional occluder. Our empirical analysis shows that it is feasible to reconstruct NLOS scenes in very challenging visibility conditions under a wide range of media properties. Moreover, we have also shown that, by increasing the reconstruction wavelength, we can partially remove the contribution of the medium at the cost of spatial resolution. However, our work is limited by the same challenges present in LOS systems and, therefore, we can expect that off-the-shelf Phasor Fields for very dense scattering media will not be optimal at reconstructing a hidden scene. However, given the flexibility of Phasor Fields for developing virtual complex optical systems, a possible avenue of work is to translate current LOS methods for vision through scattering media into Phasor Fields.

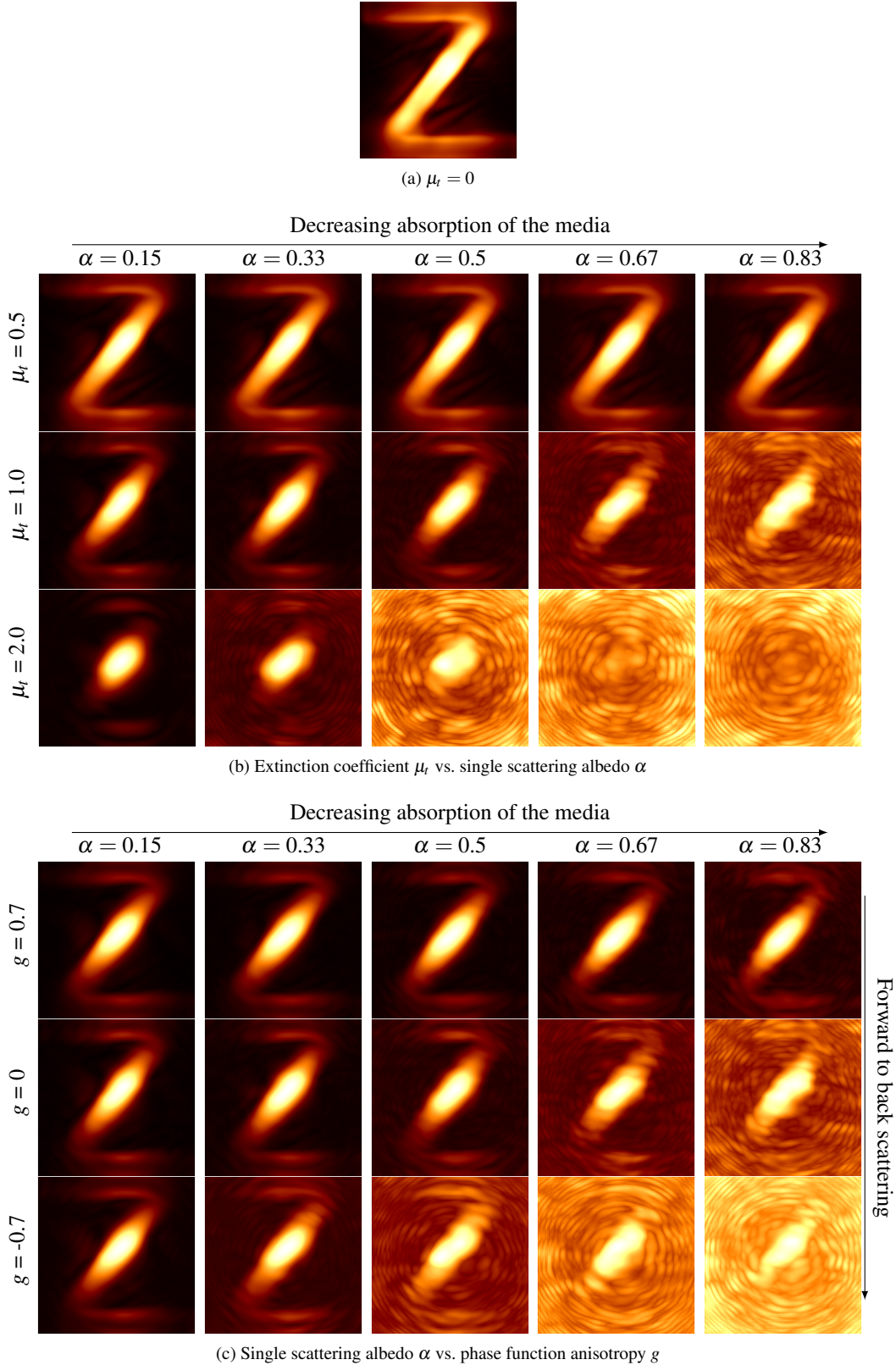


Figure 4.3: Reconstructions of the Z-LETTER scene. 4.3a) scene with no scattering media. 4.3b) reconstructions with scattering media of varying extinction μ_t (in m^{-1}) and single scattering albedo α . 4.3c) reconstructions with scattering media of fixed extinction $\mu_t = 1 m^{-1}$, varying scattering albedo α , and phase function's anisotropy g . Phasor Fields is able to reconstruct the scene even in the presence of highly scattering media.

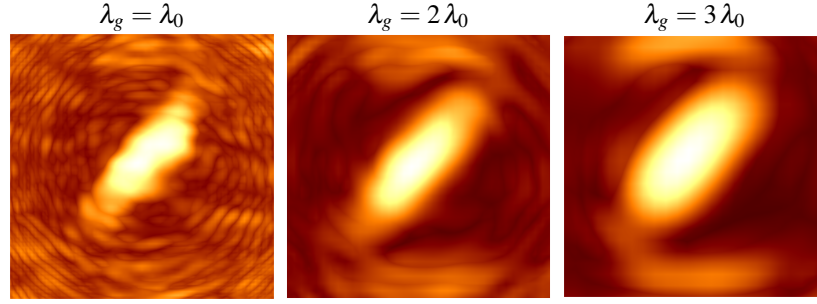


Figure 4.4: Reconstructions of the Z-LETTER scene in the presence of a scattering medium ($\mu_t = 1m^{-1}$ and $\alpha = 0.83$) for increasing wavelength λ_g , with baseline $\lambda_0 = 4\Delta_c$ and $\Delta_c = 0.11m$. Higher values of λ_g result into deeper penetration through the medium, at the cost of lower spatial.

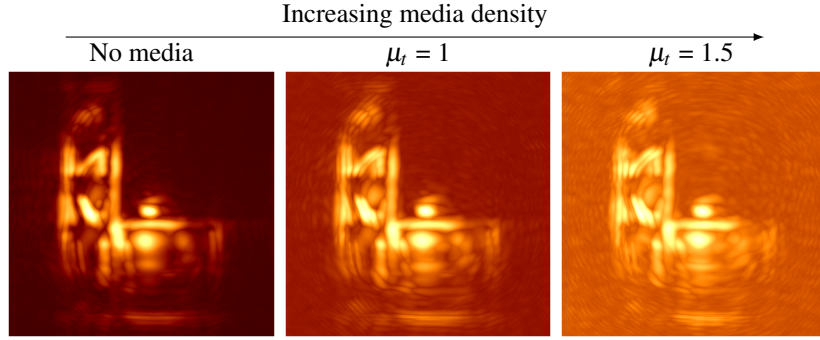


Figure 4.5: Reconstruction of the SHELF scene submerged in a medium of increasing density: $\mu_t = 0$ (no media), $\mu_t = 1$, and $\mu_t = 1.5$. In all cases, the scattering albedo is $\alpha = 0.5$.



Figure 4.6: Reconstructions of experimental data. Each column is a different scene. Top: images of the scenes behind the diffuser. Middle: Lindell and Wetzstein reconstructions [65] (CDT). Bottom: our reconstructions using Phasor Fields. The higher-frequency of CDT is related to the deconvolution to compensate scattering at the diffuser.

ABOUT THIS CHAPTER

In this chapter, we propose a physical setup for NLOS imaging, that serves us to enhance the visible features of the reconstructed scene. We show that we can combine different relay walls at different positions and orientations into a single generalized optical system, challenging the common assumption of planar relay wall in NLOS.

The work presented in this chapter has been published at the *The Visual Computer* [72]. As first author, I contributed to the theoretical formulation, method development, experimental design, experimental validation, literature revision, writing, manuscript revision and discussion.

P. Luesia-Lahoz, S. Cartiel, & A. Muñoz
Stereo non-line-of-sight imaging
 The Visual Computer

5.1 INTRODUCTION.

Besides its promising results, Phasor Fields and other Non-Line-of-Sight imaging methods have limitations regarding their ability to reconstruct surfaces at specific locations and orientations. The problem is ill-posed and presents uncertainties and occlusions making state-of-the-art approaches unable to reconstruct specific geometries in the scene, even if some signal is captured. This limitation is often called the missing-cone problem and it is ubiquitous in NLOS measurements (we refer the interested reader to previous works [67]).

Following the analogy between the relay wall and a virtual camera, in this chapter, we explore the possibility of using classical LOS techniques from computer vision to reconstruct the geometry of the hidden scene. In particular, we analyze stereo vision. For that purpose, we devise scenes with not one, but two relay walls, in which each of those walls acts as a virtual camera in the computational domain, and apply a stereo-vision-based reconstruction technique (see Figure 5.1). We demonstrate that multiple captures from different relay walls can be combined under Phasor Fields into a single and generalized optical system. This approach leads to a better-posed NLOS problem, where we capture the light scattered from one and reaching the other, besides their independent capture, which would be lost otherwise. As we show, our approach not only retrieves geometry but is also able to enhance reconstruction positions and orientations when compared to any of the individual relay walls, even in real-world captures.

Classical stereo-vision algorithms often require the identification of common feature points on both views. However, the missing cone limits the visibility of the NLOS reconstructions, as it occurs for specific positions and orientations of surfaces with respect to the relay wall [67]. Two different relay walls will more likely have non-overlapping missing-cone regions because the same geometrical features of the hidden object have different orientations with respect to each of the individual relay walls. We therefore cannot leverage information about common features from both perspectives (as given by the two relay walls) because the missing cone leads to finding very few of them in both virtual cameras.

In contrast, thanks to the knowledge about the missing cone, the fact that a particular surface is visible from one of the relay walls is actually giving some cues about the orientation of such surface (otherwise it would not be visible). We leverage these missing-cone orientation cues as a feature, instead of a limitation.

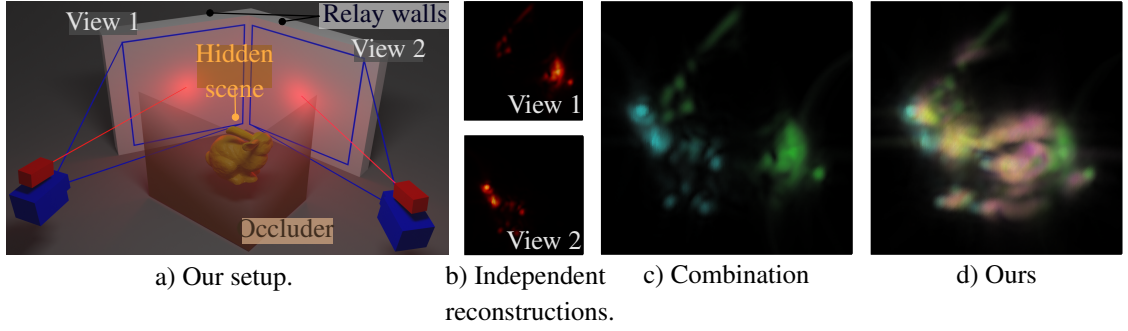


Figure 5.1: We present a novel stereo configuration for non-line-of-sight imaging using the time of flight of the light. Two relay walls illuminate and capture a hidden scene (a). The missing cone limits the visibility in the independent reconstructions, regardless of the imaging method employed. Thus, we observe limited visibility when applying the Phasor Fields framework in independent reconstructions (b). Our proposal poses a combination of Phasor Fields virtual cameras of two views as a generalized single camera (c), and we can employ the limited visibility to extract orientation cues of the hidden surface by color coding the contribution from each wall. Moreover, our setup allows us to gain additional illumination and capture signals between relay walls, increasing the reconstruction’s overall visibility (d).

We label the reconstructed surfaces according to their orientations, following such orientation cues, giving a more detailed image of the reconstruction with additional orientation information.

Our approach combines different views into a single reconstruction, enhancing the visibility of the reconstructed hidden scene and gaining orientation information of the surfaces, as our results show. We employ the Phasor Fields framework to image the hidden scene under a novel setup consisting of different apertures. We devise a combination method for multiple relay walls in contrast to the single aperture of previous Phasor Fields applications. Hence, this work offers a step towards a generalized model of extended non-planar apertures in the NLOS imaging scenario.

5.2 FROM NLOS TO LOS.

One of the most classical technologies in computer vision [108] is stereo vision, whose target consists of mimicking human depth perception using dual vision. The first application of this technology in constructing topographic elevation maps from aerial images [38] has evolved and is still in use to estimate depth or 3D structure from a scene. Traditional techniques involve block matching and correlation-based to generate disparity maps [103], that are more sensible to occlusions, regions without texture and efficiency [46], and, nowadays approaches, exploit machine and deep learning [54].

We have already introduced the analogy of the relay wall from the NLOS domain to a virtual LOS camera using Phasor Fields (see Section 3.2 in Chapter 3). Our work gets influenced by stereo vision from computer vision to present a stereo setup of NLOS, saving some differences between the two domains later explained in Section 5.4.

Similar to stereo vision, our setup allows us to obtain more information than two independent captures on their own. Specifically, we are capable to obtain orientation information of the surfaces from the hidden scene. Some researchers have indirectly tackled this problem in the NLOS domain, generating geometry through inverse rendering [88, 48, 127] or retrieving an estimation of the normals in reconstruction time [129]. In contrast, we directly target the orientations in NLOS imaging, making use of the missing cone (see Section 5.3).

5.3 VISIBILITY LIMITATION OF NLOS IMAGING.

In this section, we provide insights into limitations of NLOS methods, and, specifically, the missing cone.

The limited size of the relay wall leans to posing the NLOS imaging as an ill-posed problem. Only a part of the scattered light from the hidden scene is captured, which causes certain features to be impossible to image in the reconstructions, regardless of the imaging method employed [67]. This problem is present in other imaging fields such as medical computed tomography which originally served as a baseline for looking around corners with transient data [115], where it receives the name of the missing cone. In NLOS imaging, the lack of modulation of the signal in the Fourier domain makes it impossible to reconstruct those surfaces in the missing cone, even if some radiance reaches the relay wall. Under the Phasor Fields scope, the missing cone has also been explored by analyzing near-specular paths of light on the hidden scene [98].

In practice, due to the missing-cone problem, surfaces of the hidden scene can only be reconstructed for a subset of positions and orientations with respect to the relay wall, which is a general limiting factor for NLOS reconstructions. However, in this work, we factor in that subset of positions and orientations and use them as additional orientation cues. We use such orientation cues to label the reconstructions of the hidden scene, providing additional orientation information that resembles a normal map (see Section 5.5.1).

5.4 OUR METHOD.

Stereo in computer vision algorithms can recreate depth, transforming two different 2D views of the same scene into a 3D representation using line-of-sight (LOS) captures. Inspired by the same idea, we aim to use two different relay walls. Under Phasor Fields, these two walls can be seen as two cameras that provide two views in the NLOS domain of the same scene. Figure 5.2 shows the outline of our setup. However, there are two key differences between computer vision and NLOS imaging:

1. Our views consist of Phasor-Fields reconstructions, which already contain 3D information about the scene. In contrast with classical LOS imaging, they may contain some errors such as out-of-focus artifacts and limited visibility.
2. The missing-cone problem limits our feature visibility from the relay walls in the NLOS domain. That prevents us from finding common feature points, which are crucial for most stereo-vision techniques.

As classical stereo algorithms are not applicable in this NLOS virtual domain, we propose an alternative approach to combine the contribution from different relay walls in the same reconstruction space. This combination diminishes the lack-of-visibility effect from the missing cone on individual views if compared to only using independent relay walls on their own. However, we also enable the acquisition of the hidden scene’s impulse response by illuminating and capturing in different relay walls, thereby increasing the recovered information.

The combined aperture resolution contrasts with the LOS multi-view approaches, from which we draw inspiration, where detached aperture information is crucial for gaining geometrical reconstructions. In this section, we prove that different reconstructions can be combined into a generalized single one, where each relay wall exhibits different visibility constraints. That leads us to the following statement: the strict constraints of visibility from the missing cone allow us to label surface orientations of the reconstructed objects depending on the contribution from each relay wall. The missing cone transforms from a constraint to a feature.

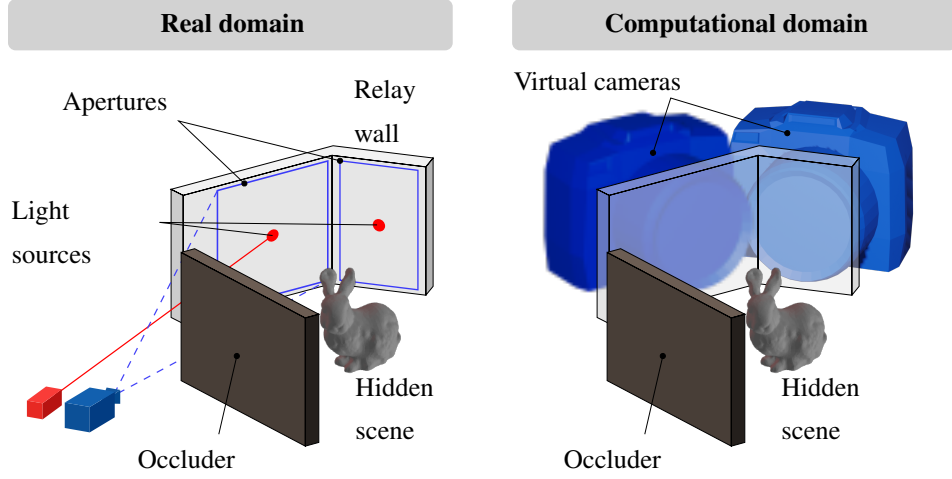


Figure 5.2: **Our Stereo setup.** Our setup consists of two relay walls for illuminating and capturing in the real domain. In the computational domain, our setup is equivalent to placing a camera on each relay wall. Our approach allows us to combine both cameras as a single generalized one.

5.4.1 Combination of different Phasor-Fields views.

Under the scenario of K different relay walls ($K = 2$ in our experiments), we can define a general impulse response $H(\mathbf{x}'_l, \mathbf{x}'_s, t)$, where $\mathbf{x}'_l \in \mathbf{L}'$ and $\mathbf{x}'_s \in \mathbf{S}'$. $\mathbf{L}'$ and $\mathbf{S}'$ represent all illumination and capture points of all relay walls. This way we represent the entire impulse response illuminated and captured by our setup, even from the illumination point of a relay wall m to a capture point in another one n .

By applying the virtual illumination of Phasor Fields (Equation (3.1)), and transforming it into the frequency domain, we can define a general phasor $\hat{\mathcal{P}}(\mathbf{x}'_l, \mathbf{x}'_s, \omega)$. As a result, we can directly apply the image formation for a single frequency, using Equation (3.6). However, since the acquired data is not continuous and both the illumination and capture points are discrete due to the current capture technologies, we represent Equation (3.6) integral as a summation as

$$\hat{I}(\mathbf{x}_v) = \sum_{\mathbf{x}'_l \in \mathbf{L}'} \sum_{\mathbf{x}'_s \in \mathbf{S}'} R(\mathbf{x}_v, \mathbf{x}'_s, \omega) R(\mathbf{x}_v, \mathbf{x}'_l, \omega) \hat{\mathcal{P}}(\mathbf{x}'_l, \mathbf{x}'_s, \omega), \quad (5.1)$$

to define the combined image formation $\hat{I}(\mathbf{x}_v, \omega)$ into the point $\mathbf{x}_v$ of all the relay wall.

Recalling Equation (3.3), we express Equation (5.1) as

$$\hat{I}(\mathbf{x}_v) = \sum_{\mathbf{x}'_l \in \mathbf{L}'} \sum_{\mathbf{x}'_s \in \mathbf{S}'} R(\mathbf{x}_v, \mathbf{x}'_s, \omega) R(\mathbf{x}_v, \mathbf{x}'_l, \omega) \hat{\mathcal{P}}_g(\omega) \hat{H}(\mathbf{x}'_l, \mathbf{x}'_s, \omega). \quad (5.2)$$

In this chapter, we assume the capture points $\mathbf{x}'_s \in \mathbf{S}'$ are grouped in grids $\mathbf{x}_s^{(k)} \in \mathbf{S}_k$ for each relay wall k , thus $\mathbf{S}_k \subseteq \mathbf{S}'$. In contrast, there is only a single illumination point $\mathbf{x}_l^{(k)} \in \mathbf{L}'$ at the center of each grid $\mathbf{S}_k$. We define

$$H^{(m,n)}(\mathbf{x}_l^{(i)}, \mathbf{x}_s^{(k)}, t) = H(\mathbf{x}_l^{(j)}, \mathbf{x}_s^{(k)}, t) \delta(m-j) \delta(n-k) \quad (5.3)$$

as the contribution of our impulse response into pairs of grids $\mathbf{S}_n$ and illumination points $\mathbf{x}_l^{(m)}$ of two relay walls n and m respectively. Equation (5.3) allows us to identify two different cases: i) the illumination and capture are both in the same relay wall ($n = m$), that we will refer to as *independent illumination*, and ii) the illumination and capture are in different relay walls ($n \neq m$), which we name *dependent illumination*.

The summation of all pairs of illumination-grid $H^{(i,j)}(\cdot)$ will result in the original generalized capture. Therefore, factoring in Equation (5.3) in to Equation (5.2) yields

$$\begin{aligned}\hat{I}(\mathbf{x}_v) &= \sum_{m,n \in K} \hat{f}^{(m,n)}(\mathbf{x}_v), \\ \hat{f}^{(m,n)}(\mathbf{x}_v) &= \sum_{\mathbf{x}'_l \in \mathbf{L}'} \sum_{\mathbf{x}'_s \in \mathbf{S}'} R(\mathbf{x}_v, \mathbf{x}'_s) R(\mathbf{x}_v, \mathbf{x}'_l) \hat{\mathcal{P}}(\mathbf{x}'_l) \hat{H}_\Omega^{(m,n)}(\mathbf{x}'_l, \mathbf{x}'_s).\end{aligned}\tag{5.4}$$

Equation (5.4) poses how to combine different and independent relay walls into a generalized view under Phasor Fields. We leverage the separated contribution by pairs of illumination-capture points in our stereo context to define the image formation from two views as a summation of their reconstructions. Note that we also gain the information of the dependent illumination captures, previously lost.

The direct calculation of Equation (5.4) may be expensive in time. Therefore, for efficiency's sake, we employ a Phasor Fields implementation optimized for reconstruction volumes formed by parallel planes to the relay wall [68]. By leveraging the regular grid of the captures $\mathbf{S}_k$ for each independent capture, we define a reconstruction volume for each $\hat{f}^{(m,n)}(\mathbf{x}_v, \omega)$. Later, we project all the reconstruction planes into a common area of matching points $\mathbf{x}_v$ of the reconstruction volume V . As mismatches may occur in the projection, we use a 3D cubic interpolation, whose values are later used for the final summation of Equation (5.4).

The missing cone establishes a strict visibility condition, which implies that high values in $\hat{f}^{(m,n)}(\mathbf{x}_v, \omega)$ of the relay wall k correspond to the view of a feature from a surface of the hidden-scene oriented towards that specific relay wall. Factoring the known information of orientation of the relay walls, and the missing cone, the labeling allows us to extract orientation cues of the hidden scene as additional information. We color-code the contribution for each relay wall as

$$\hat{I}(\mathbf{x}_v, \omega) = \sum_{m,n \in K} c^{(m,n)} \hat{f}^{(m,n)}(\mathbf{x}_v, \omega),\tag{5.5}$$

where $c^{(m,n)}$ corresponds to the chosen color for the illumination point $\mathbf{x}_l^{(m)}$ and capture grid S_n . As Section 5.5.1 shows, this coding provides clear hints about hidden surfaces' orientation.

5.5 RESULTS.

Our experimental validation demands precise control of the position and rotation of the hidden objects, sensors, illumination, and relay walls. Therefore, we opt to synthesize the transient data required to validate our proposal, as it provides sufficient flexibility in managing and controlling the scene, while maintaining the quality and realism to validate our proposal. Previous work [78] has already shown that there is a good agreement between real captures and time-resolved light transport simulations. Specifically, we use a transient renderer optimized for the NLOS configuration [96] to generate the time-resolved responses. The only exception is the experiment represented in Figure 5.5, which has been generated from measured transient data.

The use of synthetic data enables us to evaluate performance using quantitative metrics derived from the ground truth. Specifically, we report the Peak Signal-to-Noise Ratio (PSNR) and the Structural Similarity Index Measure (SSIM) for our synthesized results. PSNR is computed on the reconstructed volume after removing values close to zero in both the ground truth and the reconstruction. SSIM is calculated on the projected 2D images presented in this paper.

To simulate our setup, we place two relay walls, one adjacent to the other, at a particular angle (90° in our experiments unless specified otherwise). Each relay wall has a size of 2×2 meters, with 256×256 capture points distributed in a regular grid, and a single illumination point at the center.

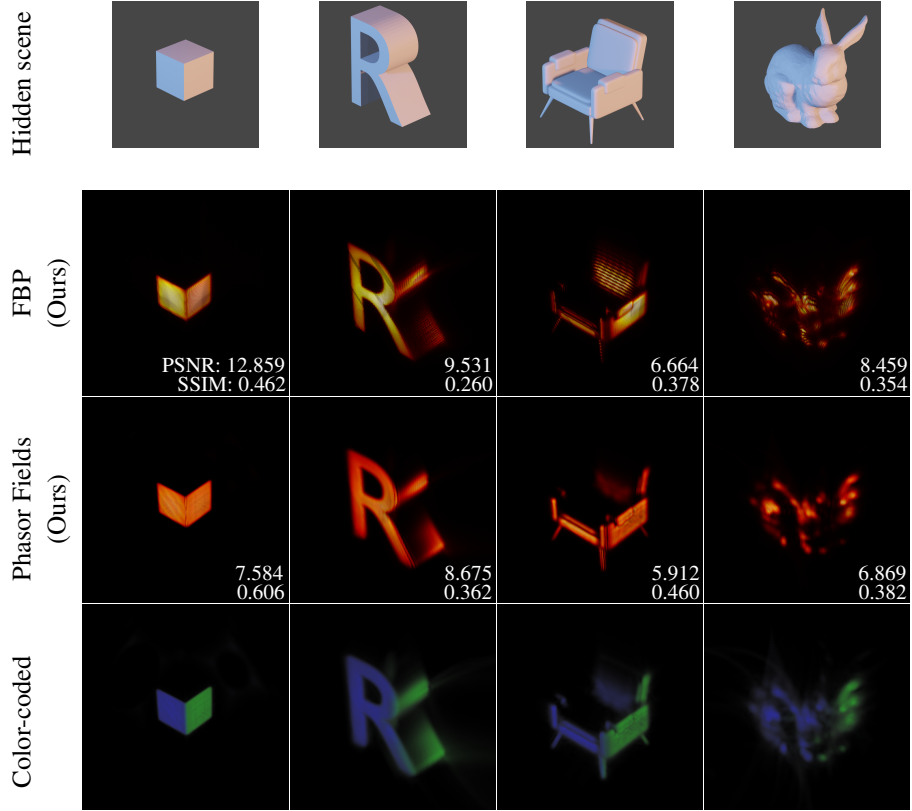


Figure 5.3: **Reconstructions of two views with independent illumination.** From top to bottom: the hidden geometry, the reconstruction with filtered backprojection (FBP), the reconstruction from two views according to Equation (5.4), and the reconstruction color-coded from each relay wall (Equation (5.5)) in blue and green. Our combination improves the visual recognition of the hidden geometries when using two independent illuminations, i.e., illuminating and capturing on each relay wall independently. Furthermore, by simply color-coding by contribution, we can identify the surfaces by their orientation.

5.5.1 Stereo combination experiments.

Section 5.4.1 demonstrates that, under Phasor Fields, we can sum independent reconstructions into a common volume, and this combination behaves as a single optical system. This mathematical analysis stands for backprojection [115], which is a linear aggregation model. In their original work, Velten et al. filtered the reconstruction by applying a Laplacian filter to increase the visibility of the hidden-scene features, resulting in filtered backprojection (FBP). In this section, we present the FBP approach for experimental validation. We apply the Laplacian filter independently, following the original work, and combine the views as discussed in Section 5.4.1. We do not provide a mathematical model that proves that this approach can generalize to multiple views, nor devise a specific filter that would work after combining multiple views, both of which would be interesting avenues for future work.

We reconstruct four hidden scenes with our combination method: a cube, a chair, an R letter with volume, and the Stanford bunny. We show the results of using only the *dependent illumination* in Figure 5.3, and they support our claims for Phasor Fields and FBP. Part of the surfaces that are not visible due to the missing cone from one aperture are visible from the other. The outcome is that, at the combination (second and third rows of Figure 5.3), the reconstruction increases the overall visibility if compared with any of the individual apertures. Consequently, it is more intuitive to understand the geometry of the hidden object. The bunny is a particularly challenging case due to the number of surfaces in the missing cone of both apertures. Yet, it is possible to identify the general shape of the bunny, with the face and body, even with missing information (e.g., the bunny ears).

The fourth row of Figure 5.3 shows the result colored depending on the contribution from each relay wall. We can observe that, by simply color-coding the image, we obtain similarities with a normal map. Furthermore, we observe that this matches the orientations of the visible surfaces from the meshes. This is consistent with our key observation of using the missing cone as a feature, which allows us to label the orientation of a surface knowing the contribution and orientation of the relay wall that reconstructs it.

The aggregation of the independent reconstructions into the common projection volume is a linear function that incorporates a temporal overhead of $\mathcal{O}(n^3)$, where n is the lateral size of the reconstruction volume. Thus, the complete temporal analysis for reconstructing yields $\mathcal{O}(k(h + n^3))$, being k the number of relay walls, and h the reconstruction time for each independent reconstruction. Existing efficient implementations of Phasor Fields for parallel planes report $h = n^3 \log n$, whereas for FBP this is $h = n^5$. For Figure 5.3, the Phasor Fields approach takes on average 104 s per reconstruction, while FBP takes 156788 s.

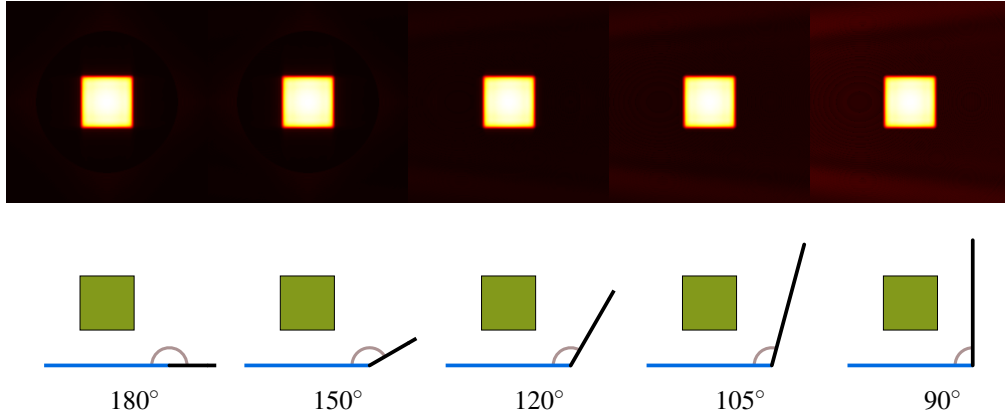


Figure 5.4: **Visibility under multi-path interference:** From left to right: reducing the rotation of a secondary wall to the relay wall. Top: reconstruction front view. Bottom: schematic top view. The secondary wall adds some background radiance due to multi-path interference when reducing its rotation. Even so, the reconstruction is sharp and clear in all cases, and the noise is negligible compared to the signal.

The incorporation of distinct multiple walls also produces multi-path interference in the areas near the corner. Light scatters between the walls as with the hidden scene, and this radiance is all captured at the relay walls. Intuitively, this leads to a degradation in the visibility features of the reconstruction as new uncertainties in the signal appear. Yet, our Phasor Fields reconstructions only bring into focus the third bounce of the light (see Section 2.2). The influence from the secondary wall belongs to the relay wall’s missing cone. Thus, the reconstruction does not show any feature or artifact corresponding to this effect, but a low-frequency noise in the background, corresponding to the energy leakage at shorter angles. We show this effect in Figure 5.4, where we simulated a single reconstruction scenario (with one relay wall) with a cube as a hidden scene, and a secondary wall at different rotations to the relay wall. In all the cases, the cube features are clear and sharp, and the background noise is negligible compared to them.

5.5.2 Combination of captured results.

Our generalized camera model from Equation (5.4) supports any relay wall configuration, regardless of rotation or position. So far, we only considered scenarios where both views are simultaneously present at the moment of capture. However, as stated before, the contribution to the reconstruction of multi-path interference is a minor background noise. Consequently, we can establish an equivalent setup using a single physical relay wall, which, under rotation and displacement between captures, acts as two views of the same hidden scene with independent illumination and without multi-path interference.

In Figure 5.5, we present both the independent and combined reconstruction results of a patch rotated and displaced between the two real captures. By estimating the relative transformations of the relay wall

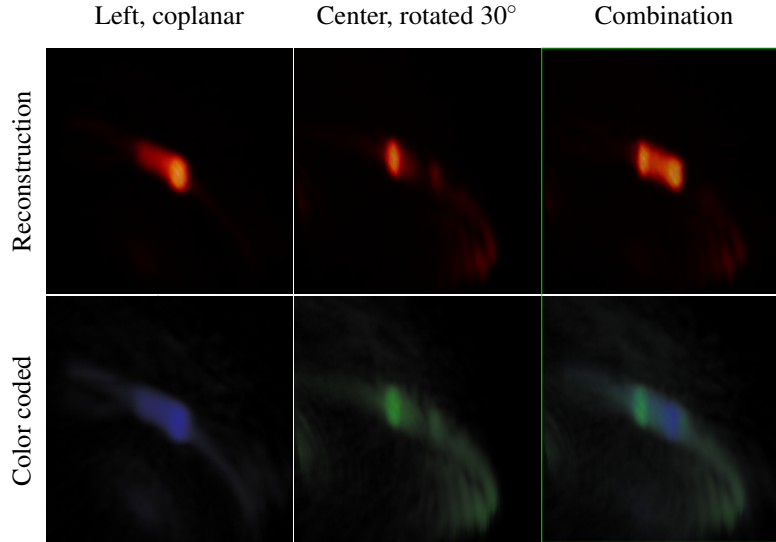


Figure 5.5: **Real capture combinations.** We frame a rotated and translated relay wall with respect to the scene as an equivalence to our approach. We use different independent illumination captures to increase the visible surfaces of a real captured patch, increasing the identifiable boundaries at our combination.

with respect to the hidden scene, we set this as a well-suited problem for our approach. As a result, our stereo approach gains complementary visible features from both relay walls, compensating from one view the missing cone of the other. Hence, the patch boundaries become more identifiable in the combination, whereas they are not in their individual reconstructions.

5.5.3 Analysis on the relay walls contributions.

Although our approach with independent illumination enhances the visibility of the hidden scene, certain features are not visible in the reconstruction. Yet, our approach enables us to capture the dependent illumination between relay walls, otherwise lost, that we leverage as additional captures from the hidden scene.

In Figure 5.6, we evaluate the impact of dependent illumination in our combination method with the bunny, where features are hidden in our previous experiments, e.g., the ear, the face, and the paws. Using two relay walls with a rotation of 120° , surfaces that remain hidden from both views under independent illumination (first column) become visible with the dependent illumination (second and third columns). Our approach using Equation (5.4) supports the combination of all views (fourth column) to significantly enhance overall visibility when compared to using only independent illumination. The SSIM and PSNR metrics support these claims, achieving maximum value at all views combination. Additionally, these new dependent illumination views can also be color-coded, which provides additional orientation hints.

In Figure 5.7, we vary the rotation angle between the two relay walls to analyze the performance of our combined reconstruction method. The rotation angle influences the contribution of each illumination-capture pair. The overall reconstruction (second row) remains qualitatively similar for all the experiments. Yet, we observe a shift in the separated contributions in the color-coded (third row). Note that the variation from yellow dominant in the experiment of 165° , to pink dominant in the 90° experiment, indicates a variation in the dependent illumination. The causes are the visibility change of the features in the hidden scene, and distance attenuation from the relay wall to the geometry. This variation directly impacts the SSIM and PSF of the final reconstruction. The 120° experiment achieves the best metrics because the bunny is aligned with the bisector of this angle. For the analysis on the Phasor Fields behavior under different aperture sizes, we refer the interested reader to previous work [35].

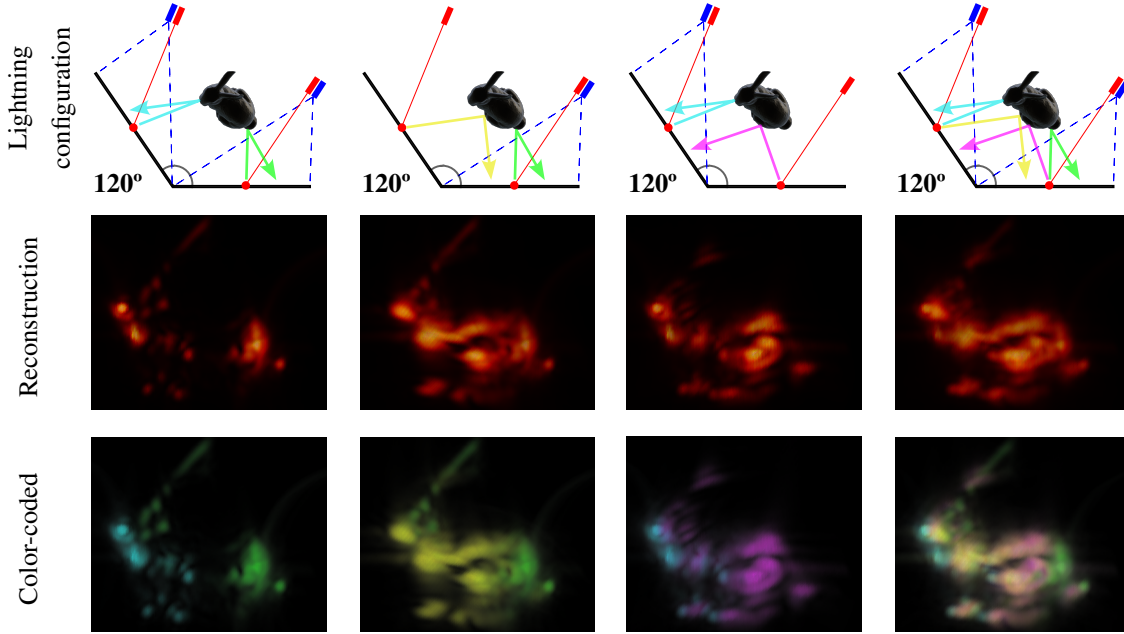


Figure 5.6: **Bunny scene reconstructions under dependent and independent illumination.** Simulated bunny scene with two relay walls forming 120° between them. From left to right: reconstruction using only independent illumination, reconstruction capturing at the first relay wall using both illuminations, reconstruction capturing at the second relay wall using both illuminations, and reconstruction using all captures and illuminations. In the color-coded (second row), we show the contribution to the visibility from each pair of illumination-capture. Our approach enables us to capture the dependent illumination of one relay wall reaching the other, otherwise lost, and to increase the overall visibility of the hidden scene.

5.6 DISCUSSION.

Drawing inspiration from the stereo vision technology from computer vision, we introduce a novel configuration consisting of two walls acting as apertures in the NLOS computational domain. This dual-aperture setup increases the amount of captured light reflected on the hidden scene if compared with any of the individual virtual apertures.

Our proposal does come with some limitations, as capturing and processing the additional independent and dependent views requires more acquisition and computational time respectively. Nonetheless, our results demonstrate that our approach can be applied in similar real-world scenarios. Moreover, exploring the interactions between non-independent views increases the overall visibility, effectively quadrupling the available views with two relay walls and making the NLOS imaging problem less ill-posed.

As the missing-cone problem generally limits the reconstruction to planar surfaces parallel to the relay wall, current real captures consist of hidden scenes whose elements are mainly planar and lack volume. Our method, being based on multiple relay walls, enables the possibility of obtaining some volume in the hidden scene, with the additional benefit of providing color-coded surface orientations.

Our experiments support our claims. The increased number of apertures helps to retrieve more information from the hidden scene. The direct effect is the decrease in the missing-cone effect and the emergence of more visible features in the reconstruction. We manage to recover further information by exploiting the missing cone to recover surface orientations from the hidden scene. In the future, researchers may delve into sophisticated geometry reconstruction algorithms that benefit from these orientation cues and overcome some current parameterizations of the NLOS imaging methods, i.e., the volume voxelization.

The previous statements come along with several applications for our setup. In the short term, the independence between relay walls enables the use of different devices for capturing. This implies that our approach is highly parallelizable by relay walls, even in concurrent and distributed systems. Further

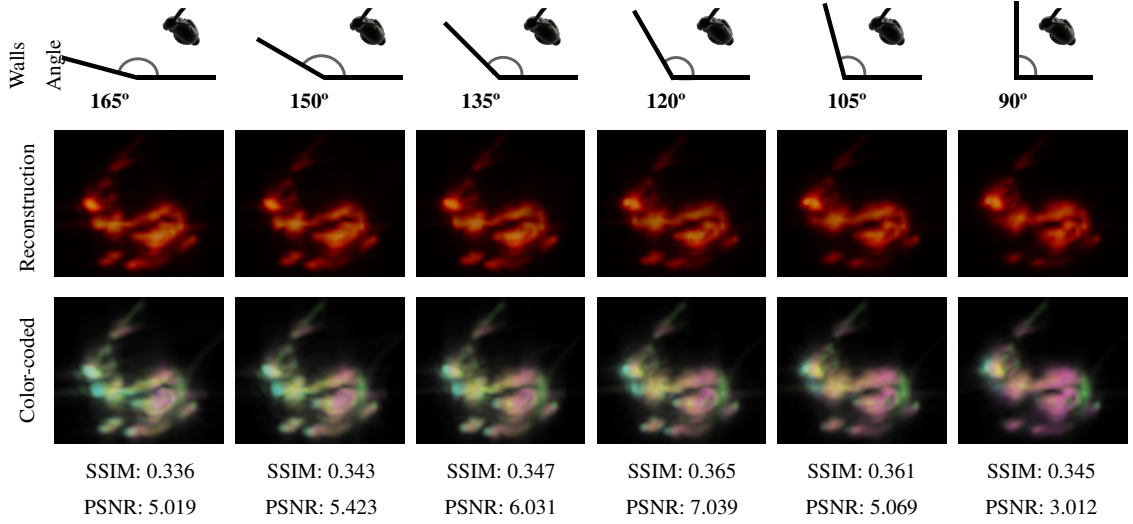


Figure 5.7: **Bunny scene reconstruction under varying relay wall angles.** Simulated bunny scene with two relay walls forming different angles between them. (Middle row) The reconstructed scene geometry using the combined illumination and capture data from both RWs. (Bottom row) A color-coded visualization showing the contribution to the final reconstruction from each illumination-capture pair (colors correspond to the scheme in Figure 5.6).

optimizations for Phasor Fields could be applied, e.g., parallelizing by planes, frequencies, or voxels. Other researchers' endeavors, such as the utilization of Phasor Fields with non-planar surfaces [34], complement our work, and together they serve as a step towards completely generalized apertures with arbitrary geometries, which are closer with respect to real-world scenarios, e.g., medical imaging, underground exploration, and car safety. On top of that, there is potential for foster avenues to expand upon our efforts to integrate multiple different apertures as a single composited one to form an extended aperture paradigm. If we take a look at recent advances in the NLOS domain to image beyond the third bounce [98], recursive imaging to see around many corners with the extended aperture paradigm may obtain astonishing results in the NLOS imaging scope.

In conclusion, we have applied a well-known technology, such as stereo vision, in the NLOS imaging scope. The Phasor Fields framework pursues an assumption of a classic optics system, and now our work helps to reduce the gap between more traditional optics and NLOS imaging. In the following chapter we delve further into this idea, which allows us to employ LOS lenses into the NLOS scope for imaging with memory constrained devices.

Part II

EFFICIENT ALGORITHMS TO RECONSTRUCT THE HIDDEN SCENE.

FRESNEL ZONE PLATES FOR MEMORY CONSTRAINED NLOS IMAGING.

ABOUT THIS CHAPTER

In this chapter, we propose an algorithmic proposal that offers a trade-off between memory consumption and reconstruction quality in NLOS imaging. For that purpose, we simplify the light propagation at reconstructions, drawing inspiration from Fresnel-Zone Plates and the line-of-sight analogies from Phasor Fields.

This chapter corresponds to a work presented in the *2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, and it can be found in its proceedings [73]. As first author, I contributed to the theoretical formulation, method development, experimental design, experimental validation of the method, literature revision, writing and discussion.

P. Luesia-Lahoz, D. Gutierrez, & A. Muñoz
Zone Plate Virtual Lenses for Memory-Constrained NLOS Imaging
 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

6.1 INTRODUCTION.

As we mention in Section 3.2 of Chapter 3, Phasor Fields is a wave-based method that generalized NLOS imaging by treating the relay surface as a virtual lens, effectively transforming an NLOS problem into a computational representation of a traditional line-of-sight (LOS) system [69]. Current efficient implementations of this technique rely on RSD-based kernels to simulate this propagation with 2D convolutions [68] (see Section 2.2 of Chapter 2). These kernels depend on the parameters of the imaging setup, but not on the transient data (the hidden scene) itself. In other words, it makes sense to pre-compute the kernels for a given imaging setup, even if the objects to be imaged around the corner change. Unfortunately, the required memory to store the kernels scales with the number of Fourier components used for propagation, the resolution of the data, and the reconstruction size. For instance, a 256^2 sampling grid at the relay wall using a single illumination point, with a 256^3 voxel grid for reconstruction and 80 Fourier components in the phasor field requires more than 80 GB of memory. And even in the case of less complex reconstructions with a single Fourier component, the kernel still needs around 1 GB of storage. This impractical memory consumption is due to the complex values that the RSD kernel requires, which increase storage compared to floating precision or integers values.

To overcome this problem, we investigate the use of alternative lenses in the virtual camera of Phasor Fields approach, to perform the propagation of the light wave. We specifically propose the use of zone plates (ZP) as virtual lenses, which use diffraction instead of refraction to bring a wavefront into focus [119]. The simplicity of manufacturing these lenses has made them very useful in coherent light scenarios, such as microscopic analysis with x-rays [3]. In our virtual light transport domain, this translates into simpler implementations with binary values, which require orders of magnitude less memory to store.

Our binary zone plate requires up to 16 times less memory than the RSD-based kernels while being also faster to compute. Our results show that the reconstructions using our ZP kernels provide an attractive trade-off between efficiency and accuracy, being a valid approximation for the exact wave diffraction of

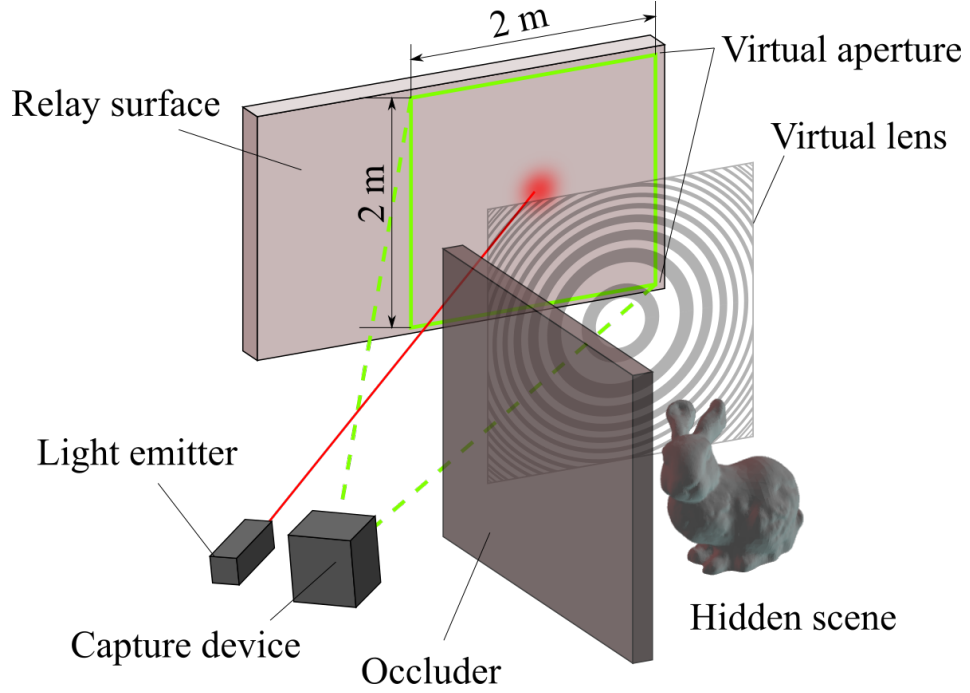


Figure 6.1: The NLOS setup. A paired system of pulsed light and SPAD (single-photon avalanche diode) illuminate and capture a relay surface, to obtain the time-of-flight of the light scattered from an occluded hidden scene. In this chapter, we explore the idea of using virtual lenses based on zone plates to reconstruct the hidden geometry.

the RSD at a much smaller cost. Our approach represents a step forward toward the integration of NLOS imaging systems in embedded and small devices, e.g. in car driving systems or medical imaging.

6.2 ZONE PLATES.

A zone plate is a planar lens for spherical wavefronts. It is made up of concentric rings that focus light onto a focal point situated in a perpendicular direction from the center of the lens at a distance f , depending on the wavelength. Instead of refraction, the lens structure forces the incoming wavefront to diffract and generate the Arago spot on the focal point. Figure 6.1 shows an illustrative example, where the ZP acts as a virtual lens on the relay surface.

Given a specific focus distance f and wavelength λ , we determine a value $\mathcal{M}(\cdot)$ for a point $\mathbf{x}_z \in Z$, where Z represents the virtual lens, as

$$\mathcal{M}(\mathbf{x}_z, \lambda, f) = \frac{2}{\lambda} \left(\sqrt{r_{\mathbf{x}_z}^2 + f^2} - f \right), \quad (6.1)$$

where $r_{\mathbf{x}_z}$ is the distance of $\mathbf{x}_z$ to the center of the lens Z . The traditional Fresnel zone plate (FZP) is made up of alternating opaque and transparent rings, as

$$FZP(\mathbf{x}_z, \lambda, f) = \mathcal{M}(\mathbf{x}_z, \lambda, f) \mod 2. \quad (6.2)$$

The opaque rings of the FZP absorb part of the contribution from the incoming wavefront. A simple solution for increasing the energy that arrives at the focal point is to transform the opaque rings from the FZP into phase-shifting rings. The so-called phase Fresnel zone plate, or binary zone plate (bZP) shares the ring structure with the FZP. For the transparent rings, the value is 1 in both lenses, while the opaque rings with a value of 0 are set to -1 in the phase-shifting rings. So, a bZP is given by

$$bZP(\mathbf{x}_z, \lambda, f) = 2(\mathcal{M}(\mathbf{x}_z, \lambda, f) \bmod 2) - 1. \quad (6.3)$$

6.3 OUR METHOD.

As Phasor Fields set the reconstruction as a lens focusing equivalence (Equation (3.4) from Chapter 3), we modify its imaging pipeline substituting the virtual lens with a bZP to reduce the memory consumption. For that purpose, we leverage implementations of Phasor Fields highly optimized for parallel reconstruction planes $\mathbf{p}_v \in V$ and equally spaced sensor grids $\mathbf{x}_s \in \mathbf{S}$ [68]. The propagation of the transient camera (Equation (3.7) of Chapter 3) from $\mathbf{S}$ only differs in the z axis for each of the reconstruction planes $\mathbf{p}_v$. By assuming Cartesian coordinates of $\mathbf{x}_v = (x_v, y_v, z_v)$ and $\mathbf{x}_s = (x_s, y_s, 0)$, where $\mathbf{x}_v \in \mathbf{p}_v$, we define a RSD kernel as

$$\mathcal{R}(x_s, y_s, z_v, \omega) = \frac{e^{-i\omega\sqrt{(x_s^2+y_s^2+z_v^2)}/c}}{\sqrt{(x_s^2+y_s^2+z_v^2)}}, \quad (6.4)$$

and thus, the propagation of a wave to a plane $\mathbf{p}_v$ at distance z_v from $\mathbf{S}$ is

$$\Phi_{\omega}^{(t)}(\mathbf{p}_v, \omega) = \hat{\mathcal{P}}(x_s, y_s, 0, \omega) * \mathcal{R}(x_s, y_s, z_v, \omega), \quad (6.5)$$

where $*$ represents a 2D spatial convolution, which can be efficiently computed with a 2D Fourier transform.

The image formation for the reconstruction as given in Equation (3.4) poses the propagation from the plane $\mathbf{S}$ to the volume V as a LOS problem, transforming the relay surface into a virtual camera that focuses using the exact diffraction propagation of Equation (3.6). Simulating the effect of changing the lens of a physical camera device, our proposal involves replacing the RSD kernel from Equation (6.5) with a bZP-based kernel in order to estimate the propagation of the wavefront into the scene through software. Assuming Cartesian coordinates $\mathbf{x}_z = (x_z, y_z, 0)$ and $\mathbf{x}_z = \mathbf{x}_s$, we can use the Equation (6.3) definition to generate our bZP kernel as

$$\Phi_{\omega}^{(f)}[\mathbf{p}_v, \hat{\mathcal{P}}(\mathbf{x}_s, \mathbf{x}_l, \omega)] = \mathcal{P}(x_s, y_s, 0, \omega) * bZP(x_s, y_s, 0, \omega, z_v), \quad (6.6)$$

where $\Phi(\mathbf{p}_v, \omega)^{(f)}$ is our novel image formation model for the plane $\mathbf{p}_v$.

While the RSD kernel values are complex numbers, our bZP kernel values cover only two integers (-1 and 1), which can be represented as a single signed byte without any precision loss. At the complete kernel size scale, the proportion of memory requirements is kept for the same number of values between the kernels, representing 16 times less memory for our complete bZP kernel storage.

6.4 RESULTS.

We have synthesized data to analyze the reconstructions using our bZP virtual lens, and to compare with the RSD propagation as ground truth. The synthesized data provides us flexibility while presenting a good agreement with the real captured data [78]. By using a transient renderer optimized for NLOS [96], we have generated four different scenes, with a common setup of an equally spaced sensor grid S of 2×2 m, with 256^2 points $\mathbf{x}_s \in S$ in the relay surface, and a single illumination point $L = \mathbf{x}_l$ at the center of the grid. The hidden scenes consist of four different geometries: a letter R, the four letters of NLOS at different distances, focusing on each one in sequence, a cube rotated 15° , and an office chair. We have used a single frequency propagation of wavelength $\lambda = 4\Delta\mathbf{x}_s$, where $\Delta\mathbf{x}_s$ is the minimum distance between samples in S , to reconstruct a 256^3 volume V . We have also trimmed an RSD kernel to provide fair comparisons, matching its memory storage to our bZP kernel.

Figure 6.2 (first three lines) shows the results for the plane of focus of the hidden geometries comparing the exact RSD solution (1GB of memory), the equal-memory trimmed RSD version, and our proposed

	Trimmed RSD (64 MB)	Our bZP (64 MB)	sZP (512 MB)
R	1817.58	1248.21	1045.09
NLOS letters focus N	2851.79	679.97	755.59
NLOS letters focus L	3111.89	484.46	554.07
NLOS letters focus O	2907.99	448.92	477.37
NLOS letters focus S	3184.99	469.38	525.62
15° rotated cube	2491.67	1242.42	1174.90
Office chair	1584.89	327.88	271.48

Table 6.1: RMSE of the reconstructions using different kernels relative to the RSD kernel of 1 GB reconstruction as ground truth. Our bZP kernel obtains better metrics than the trimmed RSD kernel of same memory in all the experiments. Note that, while obtaining similar values of the metric between our bZP and the sZP kernels, ours require 8 times less memory.

bZP virtual lens (the last two needing only 64MB of memory). Compared to the exact diffractive wave propagation of the RSD kernels, focusing with our bZP kernel allows imaging the hidden scene at the cost of introducing noise in the empty areas. Despite this noise, the geometry of the hidden objects is clearly identifiable, with sharp edges. Note that we have not filtered the noise to better illustrate the outcome of our method; this kind of filtering is common even in well-established techniques like backprojection, and could be applied to our results as well. In contrast, the equal-memory trimmed RSD kernel reconstruction fails to properly reconstruct the hidden geometry, generating unfocused surface artifacts due to the smaller aperture used for the wave propagation. The assumption is supported by the numerical comparison of the Root Mean Square Error (RMSE) values presented in the two first rows of Table 6.1: the reconstructions of our bZP kernels obtain lower values in all the scenes compared with the trimmed RSD.

Additionally, since the bZP introduces abrupt discontinuities whereas a wavefront presents a smooth phase evolution, we also show the results of a continuous *sinusoidal* zone plate (sZP) for completeness purposes, defined as

$$sZP(\mathbf{x}_z, \lambda, f) = \cos(\pi \mathcal{M}(\mathbf{x}_z, \lambda, f)), \quad (6.7)$$

following the notation for Equation (6.7). Equivalent to our Equation (6.6) bZP kernel, the new sZP kernel would result into the image formation model $\Phi_{\omega}^{(b)}[\mathbf{p}_v, \hat{\mathcal{P}}(\mathbf{x}_s, \mathbf{x}_l, \omega)]$ as

$$\Phi_{\omega}^{(b)}[\mathbf{p}_v, \hat{\mathcal{P}}(\mathbf{x}_s, \mathbf{x}_l, \omega)] = \mathcal{P}(\mathbf{x}_s, \omega) * sZP(x_s, y_s, 0, \omega, z_v). \quad (6.8)$$

As the fourth line in Figure 6.2 shows, the results do not improve our bZP lens. Furthermore, Table 6.1 present a comparison for the two lenses metrics (second and third rows) with polarized results. Moreover, the sZP continuous kernel values now span the real range $[-1..1]$, which requires float-precision values. Thus, its memory requirements discourage the use of the sZP as an attractive alternative to our proposed bZP.

The propagation of all the kernels requires similar computational time, since its implementation is the same convolution in the 2D Fourier space. In terms of kernel generation, our bZP is somewhat faster; the exact RSD kernel takes 3.5 s, while our bZP kernel only needs 1 s. For the sZP kernel, the time is increased to 1.4 s. The trimmed RSD kernel only takes 0.2 s, but at the cost of not being able to image the hidden scene. All these times have been obtained in a single thread of a i7-11700KF (3.6 GHz) processor, with 32 GB of RAM.

6.5 DISCUSSION.

The Phasor Fields framework allows to theoretically transform the relay wall into a virtual lens, thus making a NLOS problem into a LOS one. This in turn allows researchers to exploit the wealth of knowledge of Fourier optics to come up with variations of the basic framework. In this work, we have tackled the problem of memory consumption when using RSD-based kernels to propagate light with convolutions.

Our proposal overcomes these memory consumption problems by introducing zone plates as virtual lenses. As we have shown, this offers an attractive trade-off between quality and needed resources, making it a viable option for small and portable devices with strong memory constraints, such as embedded and small devices integrated in car navigation systems, or medical exploration devices.

There are still many interesting avenues of research. Some other diffraction operators, lenses or kernels could be adapted to the virtual camera system to improve the reconstructions in terms of time, memory or quality. Future work could try these new operators, for example for non-parallel volume reconstructions.

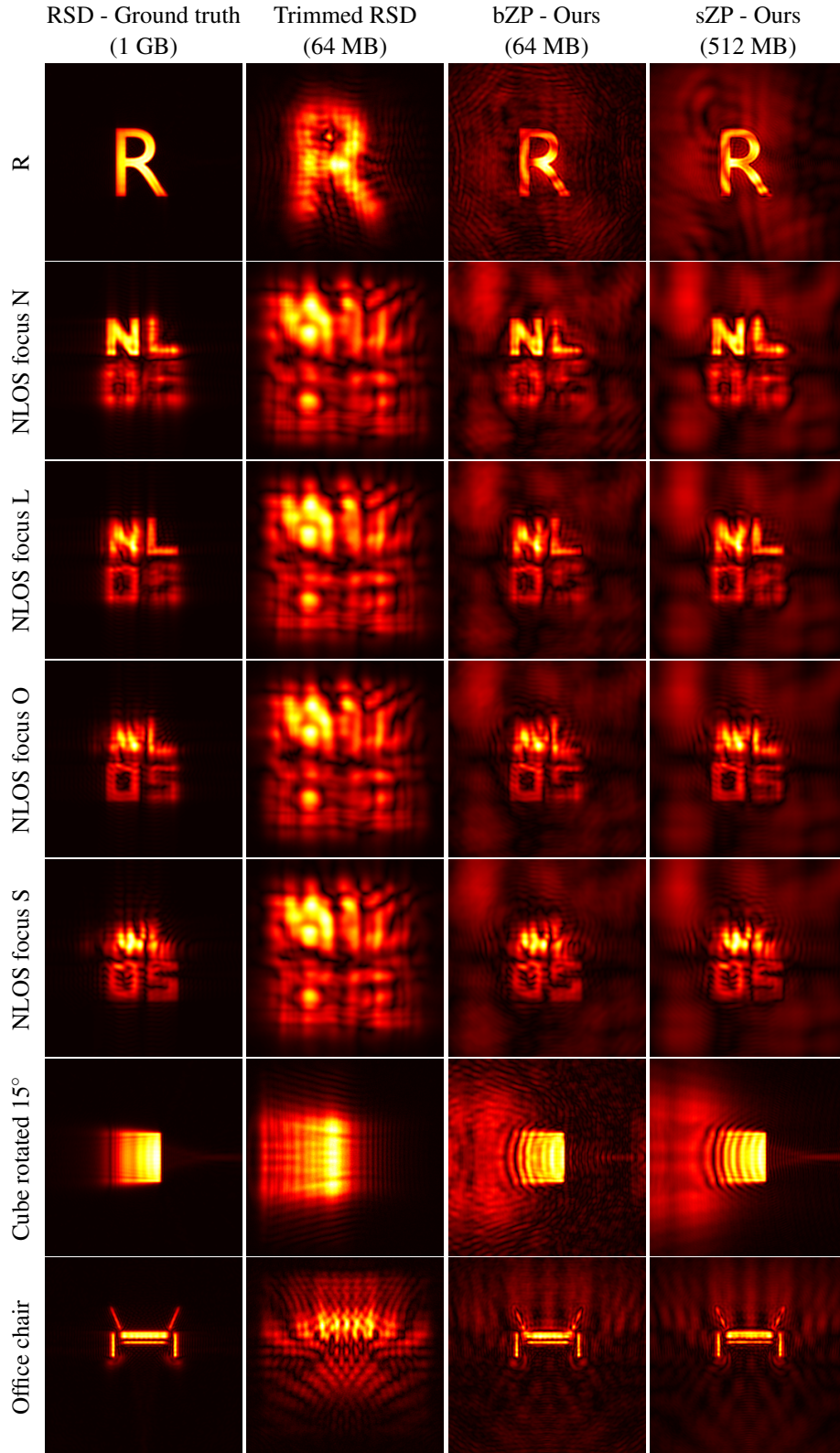


Figure 6.2: Reconstructions of the synthesized hidden scenes using different kernels. From top to bottom: the exact RSD kernel (1GB), a trimmed RSD kernel (64MB), our bZP kernel (64 MB), and a sZP kernel (512 MB). The trimmed RSD kernel fails to properly reconstruct the hidden geometry, and the sZP kernels do not improve the reconstructions of the bZP, at the cost of expensive memory consumption. In contrast, our bZP kernel allows to image the hidden scene, with clearly identifiable geometry and sharp edges, despite the residual noise in empty areas, presenting an attractive memory trade-off compared with the RSD.

ZERO-PHASE PHASOR FIELDS FOR NLOS.

ABOUT THIS CHAPTER

In this chapter, we present a novel approach for NLOS imaging, which can achieve micrometer-scale depth variations, in the time complexity that previous methods achieve centimeter-scale depth variations. For that purpose, we build on top of Phasor Fields, employing the virtual phase that previous methods discard, achieving higher precision than the voxelization employed.

The work described in this chapter has been presented orally and in poster in the 2025 *IEEE International Conference on Computational Photography*, and it is published in the 2025 ICCP proceedings [74]. This work has been conducted in collaboration with the Computational Optics Group from the University of Wisconsin-Madison, where I did a three-month internship. As first author, I contributed to the development, experimental design, experimental validation, theoretical formulation, literature revision, writing, manuscript revision, and discussion.

P. Luesia-Lahoz, T. Sultan, F. B. Peterson, A. Velten, D. Gutierrez, & A. Muñoz
Zero-Phase Phasor Fields for Non-Line-of-Sight Imaging
 2025 IEEE International Conference on Computational Photography (ICCP)

7.1 INTRODUCTION.

NLOS imaging algorithms usually reconstruct a pre-defined 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.

Phasor Fields (PF) only employs the amplitude of the virtually-propagated wave for the reconstruction value, 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 [13, 29, 6]. 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) [86] 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 [66] and synthetic data, both captured in confocal configurations [86].

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 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.

7.2 DEPTH PRECISION IN NLOS.

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 [113]. By limiting the capture to a confocal capture, i.e., the laser and capture are co-located, the Light Cone Transform [86] casts NLOS imaging as a deconvolution problem, getting approximately 2.5 millimeters depth precision via direct inversion. Another direct inversion is the fk-migration [66], 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 chapter, 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 [69], which poses the NLOS imaging problem as a wavefront propagation problem.

7.3 OUR METHOD.

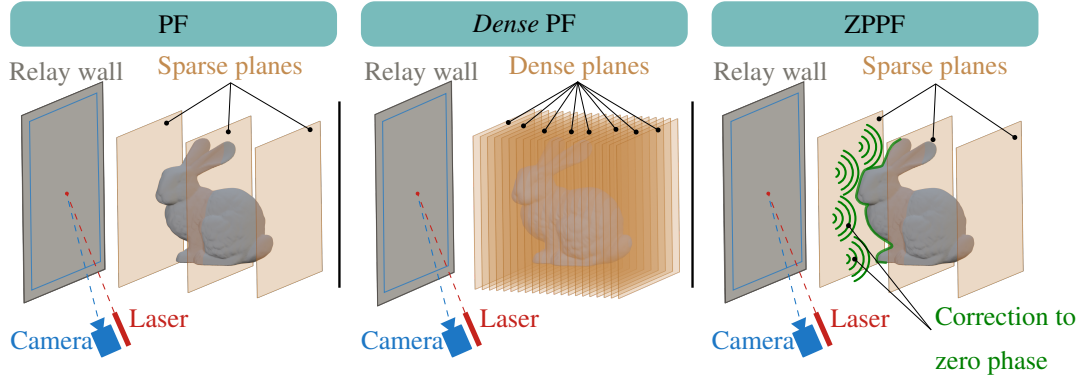


Figure 7.1: Summary of our novel ZPPF (right) compared to previous PF (left). Efficient PF implementations optimized for parallel planes achieve centimeter scale using sparse reconstruction planes, i.e., spacing the planes by centimeters. To increase the PF depth precision, we establish the reconstruction 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.

As we describe in Section 2.2, 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 (3.6). 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 [13, 29, 6], 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 prove a key observation in Section 7.3.1: 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 7.3.2). We leverage PF efficient implementations for reconstruction points grouped by parallel planes to the relay wall [68], which we refer to as reconstruction planes. Figure 7.1 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 shown by our experiments in Section 7.4.1. On top of that, phase propagation is very robust to noise, as we later show in our experimental validation (see Section 7.4.2).

7.3.1 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 a phasor (Equation (3.3)), for a particular frequency ω . We represent virtual illumination pulse $\hat{\mathcal{P}}(\omega)$ now has the form

$$\hat{\mathcal{P}}(\omega) = A(\omega)e^{i\psi(\omega)} \quad (7.1)$$

where $A(\omega)$ represents amplitude and $\psi(\omega)$ phase in the Fourier domain, obtained from the Gaussian wave packet in the Fourier domain, described in Equation (3.2). Adapting the confocal camera model (Equation (3.6)), with Equation (7.1) the propagation of the waves is given by

$$\begin{aligned} \Phi_{\omega}^{(c)}[\mathbf{x}_v, \hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega)] &= \\ &= \int_{\mathbf{L}} \hat{R}(\mathbf{x}_v, \mathbf{x}_l, \omega) \int_{\mathbf{S}} \hat{R}(\mathbf{x}_v, \mathbf{x}_s, \omega) A(\omega) e^{i\psi(\omega)} \hat{H}(\mathbf{x}_l, \mathbf{x}_s, \omega) d\mathbf{x}_s d\mathbf{x}_l \\ &= A(\omega) e^{i\psi(\omega)} \Phi_{\omega}^{(c)}[\hat{H}(\mathbf{x}_l, \mathbf{x}_s, \omega), \mathbf{x}_v]. \end{aligned} \quad (7.2)$$

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 $\mathbf{x}_l$ on the surface of the hidden geometry. Without loss of generality, let's assume a hidden scene consisting of a single point $\mathbf{x}_l$. Note that its impulse response would be a δ function in time, and therefore its frequency-domain impulse response is

$$\hat{H}(\mathbf{x}_l, \mathbf{x}_s, \omega) = \varepsilon_h e^{-i\omega \frac{|\mathbf{x}_l - \mathbf{x}_l| + |\mathbf{x}_l - \mathbf{x}_s|}{c}}, \quad (7.3)$$

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.

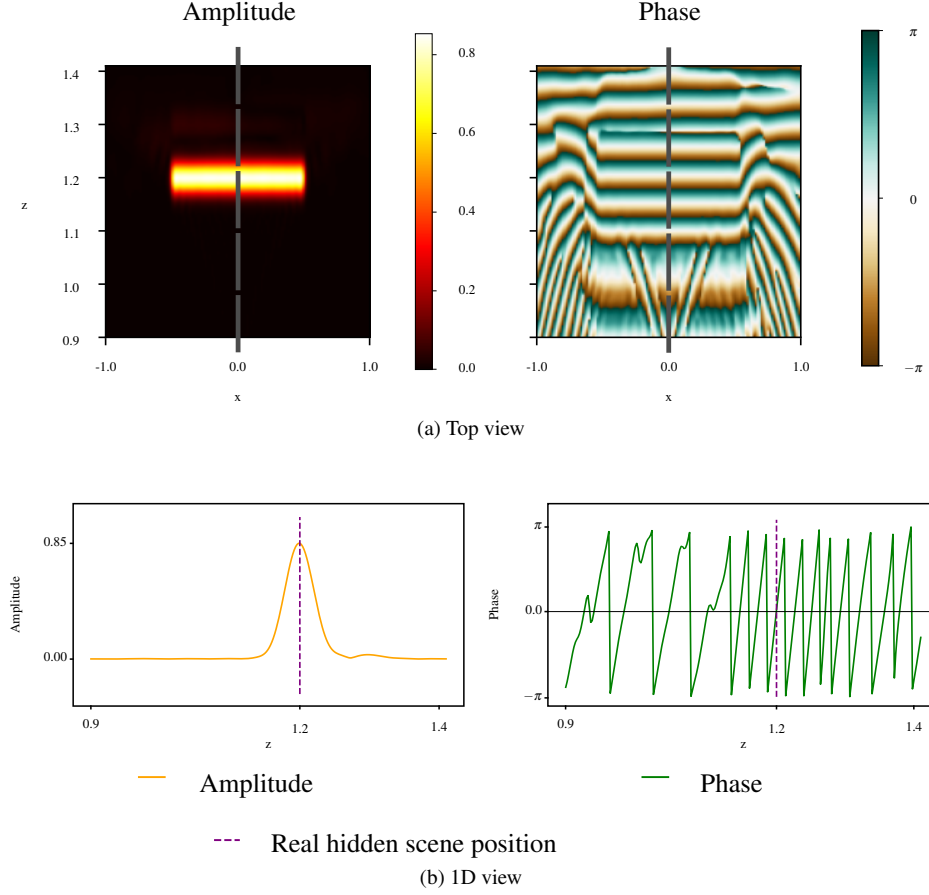


Figure 7.2: Dense PF reconstruction of a plane at 1.2 meters from the relay wall. In (a), we illustrate the amplitude as a heatmap and the phase as a blue-white-brown. (b) plots the same values from the dashed line of (a), with the real ground truth position of the geometry (purple dashed). We used a 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 7.3.1 proves.

Plugging Equation (7.3) into Equation (7.2), we get

$$\begin{aligned} \Phi^{(c)}[\mathbf{x}_v, \hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega)] &= \\ &= \int_{\mathbf{L}} \hat{R}(\mathbf{x}_v, \mathbf{x}_l, \omega) \int_{\mathbf{S}} \hat{R}(\mathbf{x}_v, \mathbf{x}_s, \omega) e^{i\psi(\omega)} \varepsilon_h e^{-i\omega \frac{|\mathbf{x}_l - \mathbf{x}_l| + |\mathbf{x}_l - \mathbf{x}_s|}{c}} d\mathbf{x}_s d\mathbf{x}_l. \end{aligned} \quad (7.4)$$

If we now expand Equation (7.4) with Equation (3.5), when imaging point $\mathbf{x}_l = \mathbf{x}_v$, the phase shifts (the exponents of e) from the impulse response cancel the phase shifts due to the RSD propagation, yielding

$$\Phi^{(c)}[\mathbf{x}_v, \hat{\mathcal{P}}(\mathbf{x}_l, \mathbf{x}_s, \omega)] = A(\omega) e^{i\psi(\omega)} \int_{\mathbf{L}} \frac{1}{|\mathbf{x}_v - \mathbf{x}_l|} \int_{\mathbf{S}} \frac{1}{|\mathbf{x}_v - \mathbf{x}_s|} \varepsilon_h d\mathbf{x}_s d\mathbf{x}_l. \quad (7.5)$$

Therefore, the only phase shift corresponds to the virtual illumination pulse $\psi(\omega)$, 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 $\mathbf{x}_v = \mathbf{x}_l$. Considering that PF ultimately applies an inverse Fourier transform for a reconstruction point $\mathbf{x}_v$, Equation (7.5) 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

($\psi(\omega) = 0$) 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 7.2, establishing a zero phase for the reconstruction surface.

7.3.2 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 7.1 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 7.3.1). 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 (Equation (3.2)), 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 $\psi(\mathbf{x}_v) \in (-\pi, \pi]$ is the phase of the reconstruction $\Phi_\omega^{(c)}[\mathcal{P}(\mathbf{x}_l, \mathbf{x}_s, \omega), \mathbf{x}_v]$. In order to find the closest zero crossing in phase ($\psi(\mathbf{x}_t) = 0$), we apply Newton's method for a scalar function with a vector argument, starting from $\psi(\mathbf{x}_v)$:

$$\begin{aligned} \mathbf{x}_0 &= \mathbf{x}_v \\ \mathbf{x}_{k+1} &= \mathbf{x}_k - \frac{\psi(\mathbf{x}_k)}{\|\nabla\psi(\mathbf{x}_k)\|^2} \nabla\psi(\mathbf{x}_k), \end{aligned} \quad (7.6)$$

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 (7.6) we need to calculate the gradient of the phase $\nabla\psi(\mathbf{x}_v)$, which is related to the wave vector. In the general case, calculating $\nabla\psi(\mathbf{x}_v)$ involves differentiating the propagation of Equation (3.6). 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 $\nabla\psi(\mathbf{x}_v) = -2\omega_g\mathbf{z}/c$. We can therefore correct the phase from a single iteration of Equation (7.6) as

$$\begin{aligned} \mathbf{x}_t &= \mathbf{x}_v + \frac{\psi(\mathbf{x}_v)c^2}{4\omega_g^2} \frac{2\omega_g\mathbf{z}}{c} = \mathbf{x}_v + \psi(\mathbf{x}_v)\mathbf{z} \frac{c}{2\omega_g} \\ &= \mathbf{x}_v + \psi(\mathbf{x}_v)\mathbf{z} \frac{\lambda_g}{4\pi}, \end{aligned} \quad (7.7)$$

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.

7.4 RESULTS.

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 7.3.2). 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.

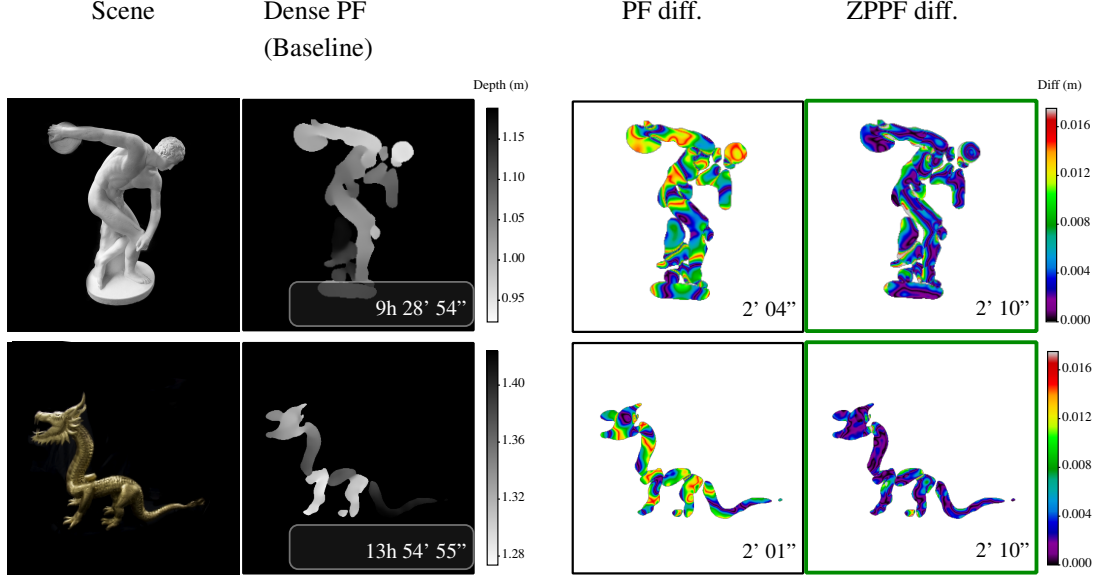


Figure 7.3: From left to right: picture of the hidden scene, dense PF depth used as baseline, PF depth difference to the baseline, and ZPPF depth difference to the baseline. From top to bottom, the *statue* and the *dragon* scenes. Insets indicate the computation time required for each reconstruction 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.

7.4.1 Depth precision analysis.

To analyze depth precision, we employ publicly available captured data [66] 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 7.3, 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 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 [96] 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 7.4 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.

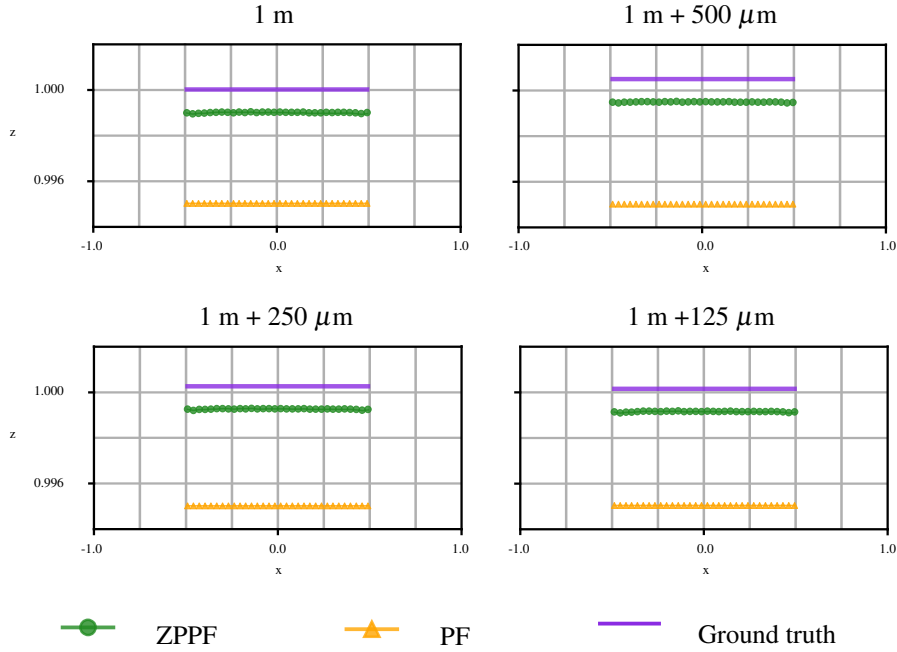


Figure 7.4: Width (x) and depth (z) coordinates in meters of the PF and our ZPPF depth estimation of a 0.5 m-sized squared plane at varying distances in the micrometer scale compared to the ground truth. Ours achieves better depth approximations, using the sparse planes, with a small constant error, maintaining a coherent estimation when compared between experiments. Horizontal and vertical axes are not on the same scale for visualization purposes.

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

Table 7.1: 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 7.1 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.

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 7.5 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.

7.4.2 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

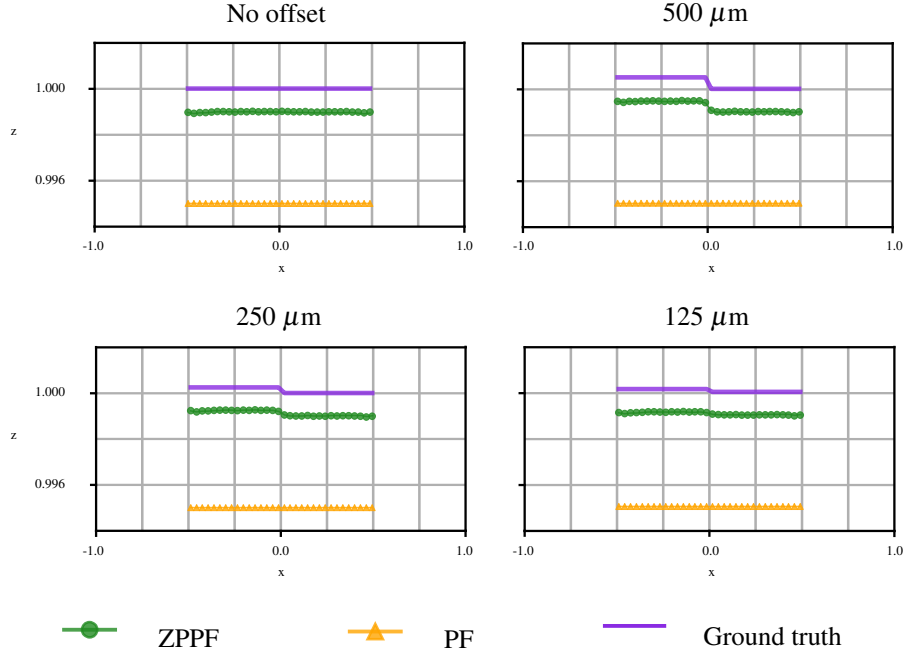


Figure 7.5: Width (x) and depth (z) coordinates in meters of the PF and our ZPPF depth estimation of two 0.25 m squared planes, maintaining one fixed at 1 m, and applying a distance offset to the second one. Our ZPPF can detect the slope between the planes using sparse reconstruction planes, where previous PF cannot, even in the extreme case of 125 μm . Horizontal and vertical axes are not on the same scale for visualization purposes.

exposure), providing a reference for evaluating the differences. As Figure 7.6 shows, our method achieves consistent and reliable depth estimations across all cases. At 15-second exposure, the low SNR constrains the 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.

7.5 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 [81, 64]. Additionally, our technique could be adapted to other, more sophisticated PF-based techniques, including imaging beyond the third bounce [98].

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 [66], 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.

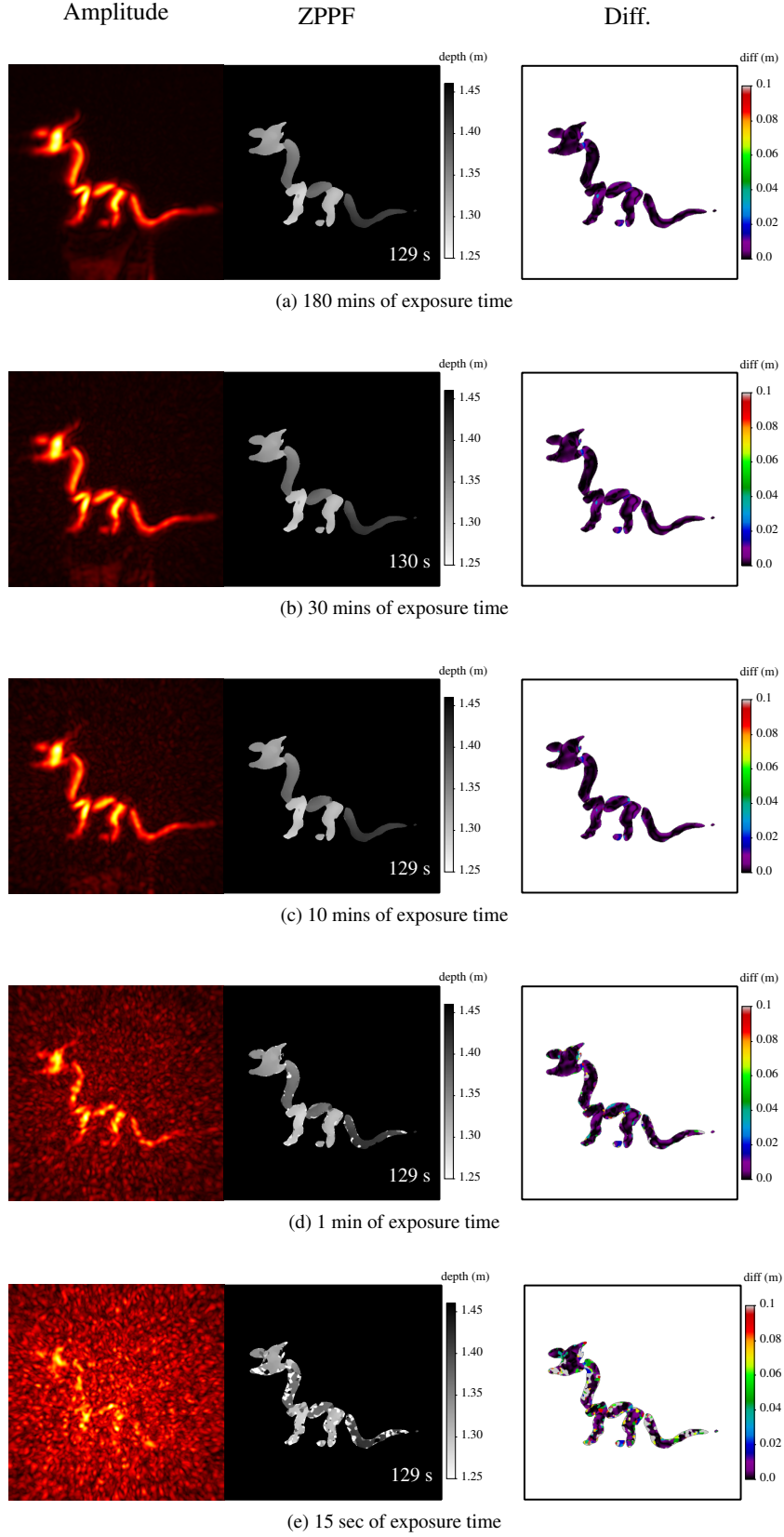


Figure 7.6: Depth estimation of our ZPPF for different captures of the *dragon* scene of decreasing exposure times. Our ZPPF depth estimation is coherent and reliable in all cases, showing some errors in the geometry borders of the 1-minute and 15-second exposure. We used the dense PF reconstruction for the difference baseline.

Part III

ENHANCING THE NLOS CAPTURE WITH COHESIVE AND
CORRECT FORMULATIONS.

ABOUT THIS CHAPTER

In this chapter, we propose a novel mathematical formulation for Non-Line-of-Sight Imaging that lays the foundations for a unified representation of existing NLOS imaging methods. To this end, we draw inspiration from quantum mechanics and employ the Dirac notation. We provide formulation examples for three of the most representative NLOS methods, demonstrate that the Rayleigh-Sommerfeld diffraction operator is the appropriate propagation for wave-based imaging methods, and that back-propagation and wave-based imaging formulations are equivalent.

This work is published in the *Optics Express* [91]. While most of the mathematical theory was developed by Albert Redo-Sanchez, as second author, I specifically contributed to the experimental design, experimental validation, writing, revision, and discussion. For completeness purposes, and to establish the mathematical insights of the notation, this chapter closely follows the published work. In Chapter 9, we further analyze the experimental implications of this chapter thoroughly.

A. Redo-Sanchez, P. Luesia-Lahoz, A. Muñoz & D. Gutierrez
Cohesive framework for non-line-of-sight imaging based on Dirac notation
 Optics Express

8.1 INTRODUCTION.

The NLOS imaging landscape has seen rapid progression in recent years, as experimental setups have become more reliable, faster, and more accurate, which has generated an effervescence in reconstruction methods with varying accuracy and reconstruction time. However, this body of reconstruction methods has been formulated independently, with ad-hoc assumptions and usually simplified light transport models adapted to particular cases. Two of the major formulations are backprojection and wave-based propagation [69, 66], the latter including Phasor Fields (see Section 2.2 in Chapter 2). Backprojection formulations model the transport of light within the hidden scene and aim at recovering the geometry inverting the transport model [115, 62, 2, 59, 124]. Wave-based propagation formulations aim at modeling the time-of-flight of light within the phase of a complex function, applying well-known Fourier optics to image the hidden scene [86, 66, 69, 92, 79].

A key part of this chapter's contribution is the formulation of a cohesive theoretical framework that not only encloses several reconstruction methods but can also explain their limitations and virtues in a cohesive way. One of the key insights that has inspired our work is the realization that the magnitudes propagated in both major formulations (time-resolved real numbers in backprojection and phasors in wave-based propagation) are, indeed, elements in a Hilbert space $\mathcal{H}$ connected through a linear operator.

We build on top of the idea that reconstruction methods are unable to accurately predict the hidden geometry due to insufficient information and the inherent ambiguities of light transport within the hidden scene. Therefore, the output of such methods can be interpreted as a probability distribution, for each point in the volume, of being part of the hidden geometry. This leads to another key insight inspiring our

work: thinking about the measurement of a hidden scene as the measurement of an operator in quantum mechanics. This in turn links with quantum mechanics theory and the uncertainty principle, in which it is impossible to predict particular outcomes of events, but it is possible to estimate probabilities of such events.

A fundamental piece in the formulation of quantum mechanics is the use of Dirac notation to describe quantum systems [15, 16]. As we show, the use of Dirac notation in an NLOS context provides a cohesive theoretical framework that will explain why the methods proposed so far work the way they do, as well as to connect the reconstruction methods with the fundamental physics of light propagation.

In this chapter, we will demonstrate the following theoretical contributions i) the derivation of back-projection, Phasor Fields, and f-k migration formulations for NLOS imaging from a single equation (Section 8.4, Section 8.5, and Section 8.6), ii) the demonstration that the Rayleigh-Sommerfeld diffraction operator is the propagation operator for wave-based imaging methods (Section 8.5), and iii) the demonstration that backprojection and wave-based imaging methods are equivalent point-to-point (Section 8.7), which is supported by a prior demonstration that propagation operators are unitary (Chapter A). This chapter also brings other contributions, such as modeling general temporal profiles for light source emission (most methods assume delta pulses), as well as novel insights on the general properties of propagation operators, including the Rayleigh-Sommerfeld diffraction operator commonly used in Phasor Fields. Other NLOS imaging formulations [101, 87, 8, 109, 104, 120, 57] may also be derived from our proposed framework, as we discuss at the end of this manuscript.

Based on this work, we expect to enable future research including, but not limited to i) bringing the mathematical toolset of quantum mechanics to NLOS imaging, which can result in more optimal and higher performing computational methods; ii) implement other imaging and reconstruction methods that may model the interaction of light with the hidden scene more accurately by finding an appropriate Hilbert space representation and specific propagation operator; iii) define optimal scene sampling and reconstruction strategies based on the expected properties of the operators, for example, their rank, dimensionality, and symmetry; and iv) use the generality of our proposed framework to devise algorithms that will work with all NLOS imaging methods.

8.2 BACKGROUND.

Two of the major formulations to realize NLOS imaging are backprojection and wave-based propagation. The backprojection formulation creates a voxelization of the NLOS scene and estimates the presence of geometry on a particular voxel based on the cumulative contributions of light intensities captured by the time-resolve sensor at specific timestamps [115, 2, 59]. The timestamps are computed based on the estimated time-of-flight of the light according to the relative position of the voxel with respect to the illumination and sensor points on the relay wall. Backprojection uses the computed timestamps to look-up the intensity of light captured at the corresponding time index on the time-resolved signal. A more mathematical background on backprojection is provided at the beginning of Section 8.4.

Wave-based propagation formulations aim at modeling the time-of-flight of light within the phase of a complex function so that virtual light sources and cameras can be modeled and well-known Fourier wave optics methods can be used to image the hidden scene. Two of the most popular wave-based formulations are Phasor Fields [69] (see Section 2.2 in Chapter 2), and f-k migration [66].

In the rest of this section, our goal is to familiarize the reader with the terminology and some fundamental concepts of the Dirac notation and quantum mechanics. More mathematical backgrounds on Phasor Fields and f-k migration are provided at the beginning of Section 8.5 and Section 8.6 respectively.

Table 8.1: Table describing the symbols used throughout this chapter. ⁽¹⁾Thorough this chapter we will use $\mathbf{x}_l$ and n indistinctly to refer to an illumination point in $\mathbf{L}$. ⁽²⁾Thorough this chapter we will use $\mathbf{x}_s$ and m indistinctly to refer to a sensing point in $\mathbf{S}$.

Symbol	Description
$\mathbf{L}$	Surface containing the illumination points.
n, n'	⁽¹⁾ Index identifying an illumination point at position $\mathbf{x}_l$. The apostrophe indicates that it is used as an index for a summation.
$\langle l_n l_n \rangle$	The <i>bra</i> and <i>ket</i> representing the state of light at illumination point n .
$\mathbf{S}$	Surface containing the sensing points.
m, m'	⁽²⁾ Index identifying a sensing point at position $\mathbf{x}_s$. The apostrophe indicates that it is used as an index for a summation.
$\langle s_m s_m \rangle$	The <i>bra</i> and <i>ket</i> representing the state of light at sensing point m .
$\mathbf{x}$	Vector describing the position of an arbitrary point.
t	Time dimension.
ω	Frequency of the light.
$\mathbf{k}$	Wavevector of the light.
$[\hat{T}_{m,n}]$	Scene operator.
$[T_{m,n}]$	Observable corresponding to scene operator $[\hat{T}_{m,n}]$.
$\hat{U}_{m,n}$ $\hat{U}_{m,n'}$	Propagation operator from illumination point n to sensor point m . The apostrophe indicates that it is used as an index for a summation.

8.2.1 Dirac notation and quantum mechanics.

Quantum mechanics is a fundamental theory in physics developed in the first half of the 20th century that aims at describing the physical properties of matter at the scale of atoms and subatomic particles. Although being controversial and often conflicting with our common-sense intuition derived from our experience in the world, quantum mechanics has been successful in providing correct predictions and results in every situation to which it has been applied.

A key component in quantum mechanics is the use of Dirac notation to represent the different states of a quantum system and its dynamic evolution. Dirac created this notation to simplify the formulation of quantum mechanics in mathematical terms [15, 16] and quickly became as *de-facto* standard notation in the field. The scope of this section is to provide enough background in terms of terminology an intuition of the key mathematical elements in Dirac notation. In particular, we aim at describing the state of a pulse of light and how this pulse propagates and interacts with a hidden scene. We refer the interested reader to other sources for more in-depth explanations [26, 20, 24, 33, 100].

8.2.1.1 Kets and Bras

Dirac notation is also known as *bra-ket* notation. A *ket* represents the state of a quantum system and it is expressed as $|u\rangle$ (any letter other than u may be used to identify the *ket*). This *ket* is a vector, or element, in a Hilbert space $\mathcal{H}$. In general, the term Hilbert space refers to an infinite-dimensional space where the inner product is complete. This definition also includes finite-dimensional spaces, which automatically satisfy the condition of completeness. Without loss of generality, we can think of a *ket* as a column vector with a dimension equal to K in which the components are complex:

$$|u\rangle = (u_1, \dots, u_K)^T \quad (8.1)$$

A *bra* is the adjoint of a *ket*. Therefore, the *bra* is the transposed and conjugated version of the *ket*. Without loss of generality, the *bra* is expressed as a row vector with dimension K in which the components are the conjugated components of the *ket*:

$$|u\rangle^\dagger = (u_1^*, \dots, u_K^*) \equiv \langle u| \quad (8.2)$$

Notice that the transposed conjugated $|\cdot\rangle^\dagger$ is denoted as $\langle\cdot|$ in Dirac notation. Note that the dimension K of these vectors can be infinite, provided, as stated before, that their inner product is complete.

8.2.1.2 Inner product.

The inner product defines the contraction of a *bra* with a *ket*, which is expressed as:

$$\langle v|u\rangle = |v\rangle^\dagger |u\rangle = (v_1^*, \dots, v_K^*) (u_1, \dots, u_K)^T = v_1^* u_1 + \dots + v_K^* u_K = \sum_{k=1}^K v_k^* u_k \quad (8.3)$$

In practical terms, the inner product is the scalar product of two vectors in which the components are complex. The inner product of a *ket* with its *bra* is the square of the norm of the *ket*:

$$\langle u|u\rangle = (u_1^* u_1 + \dots + u_K^* u_K) = \|u\|^2 \quad (8.4)$$

Note that the inner product is expressed as a bracket of $\langle v|$ and $|u\rangle$ and that both $\langle v|$ and $|u\rangle$ have the same dimension. Hence, the name of *braket* notation for this formalism [15]. The generalization of Equation (8.3) to an infinite dimensional space is the integral:

$$\langle v|u\rangle = \int_{-\infty}^{\infty} dx v^*(x) u(x) \quad (8.5)$$

Kets play a prominent role in representing the states of quantum systems. *Kets* are the states upon which operators act; for instance, $A|u\rangle$ means that operator A is acting on $|u\rangle$, because in standard algebra operators are placed on the left side of the element they act on. *Bras* are considered the dual representation of the *Kets* and, thus, adjoint operators act on the element placed on their left side. For instance, $\langle u|A^\dagger$ means that the adjoint operator of A is acting on $\langle u|$. Considering matrices and vectors as particular cases, we can consider $A|u\rangle$ as the product of matrix A with column vector $\mathbf{u}$, whereas we can consider $\langle u|A^\dagger$ as the product of row vector $[\mathbf{u}^*]^T$ with matrix $A^\dagger = [A^*]^T$.

8.2.1.3 Outer product.

The outer product is defined as the cartesian product of a *ket* with a *bra*:

$$|u\rangle \langle v| = \begin{bmatrix} u_1 v_1^* & \dots & u_1 v_L^* \\ \vdots & \ddots & \vdots \\ u_K v_1^* & \dots & u_K v_L^* \end{bmatrix} \quad (8.6)$$

where K is the dimension of $|u\rangle$ and L is the dimension of $\langle v|$. Notice that the dimensions can be different and, therefore, the matrix defined in Equation (8.6) is not necessarily square. The outer product behaves like an operator. The generalization of Equation (8.6) to infinite dimensional spaces is:

$$(|u\rangle \langle v|) = (u \otimes v), \quad (8.7)$$

which could be represented as an operator kernel, such as

$$(|u\rangle \langle v|)(a, b) = u(a) v^*(b). \quad (8.8)$$

8.3 FUNDAMENTAL EQUATIONS FOR NLOS IMAGING.

Our target in this chapter is to present a cohesive framework that encompass and derives for all existing NLOS imaging methods. In this section we propose two simple equations based on quantum mechanics and Dirac notation that will allow us deriving backprojection and wave-based imaging methods such as Phasor Fields and f-k migration.

Let us start by defining the *kets* that represent the state of light at both the illumination and sensing points. Let us define $|l_n\rangle$ as the state of light at an arbitrary illumination point $\mathbf{x}_l$ with index n located on the virtual light source surface $\mathbf{L}$, and $|s_m\rangle$ as the state of light at an arbitrary sensing point $\mathbf{x}_s$ with index m located in the virtual camera surface $\mathbf{S}$. We can then define an operator $\hat{T}_{m,n}$ as the outer product between the *kets* representing light measured at the sensing points on $\mathbf{S}$ and the *bras* representing light at the illumination points on $\mathbf{L}$ as:

$$\hat{T}_{m,n} = |s_m\rangle \langle l_n| \quad (8.9)$$

which allows us to define a matrix operator $\hat{T}$ that contains all the illumination-sensing pairs such that:

$$\hat{T} = [\hat{T}_{m,n}] = [|s_m\rangle \langle l_n|] \quad (8.10)$$

Equation (8.10) defines a simple, comprehensive operator to model the interaction of light with a hidden scene in a NLOS context; we call $\hat{T}$ the **scene operator**. In the following, we will use individual elements of the scene operator $\hat{T}_{m,n}$ instead of its matrix form $\hat{T} = [\hat{T}_{m,n}]$.

Equation (8.10) can be expressed more explicitly by introducing **propagation operators**, which actually model the propagation of light from point-to-point, whereas $\hat{T}_{m,n}$ models a convenient mathematical relationship between the state of light at the illumination point and the state of light at the sensing point.

Let's define a propagation operator $\hat{U}_{m,n}$ that models the interaction of light with the hidden scene by propagating light from an illumination point n on $\mathbf{L}$ to a sensing point m on $\mathbf{S}$. The contributions from all illumination points to an arbitrary sensing point m are given by:

$$|s_m\rangle = \sum_{n'=1}^N \hat{U}_{m,n'} |l_{n'}\rangle \quad (8.11)$$

where N is the total number of illumination points and n' is the summation index over all n illumination points.

Plugging Equation (8.11) into Equation (8.10) yields:

$$\hat{T}_{m,n} = \sum_{n'=1}^N \hat{U}_{m,n'} |l_{n'}\rangle \langle l_n| \quad (8.12)$$

which **defines the fundamental equation that comprehends the general formulation in solving any NLOS scene using propagation operators**. We propose the construction of $\hat{T}_{m,n}$ as an operator that will allow us to develop and state the claims of the chapter. $\hat{T}_{m,n}$ contains information about the measurement of the scene through the information of the captured light contained in $|s_m\rangle$ and the set of propagation operators $\hat{U}_{m,n}$. However, $\hat{T}_{m,n}$ is not a measurement operator.

Equation (8.10) will allow us to demonstrate a series of general properties of the different NLOS reconstruction methods and it is the starting point to i) derive the backprojection, Phasor Fields, and f-k migration formulations from the same Equation (8.10) in Section 8.4, Section 8.5, and Section 8.6, respectively, thus setting evidence that our proposed theoretical framework represents a cohesive formulation for the NLOS problem; ii) demonstrate that the Rayleigh-Sommerfeld diffraction operator is the propagation operator for wave-based imaging methods, such as Phasor Fields and f-k migration in Section 8.5; and iii) demonstrate that backprojection and wave-based imaging methods are equivalent in Section 8.7.

Both Equation (8.10) and Equation (8.12) are general, and be seen as sheet equations under which several NLOS reconstruction methods can be formally described. To realize this, we need to i) identify

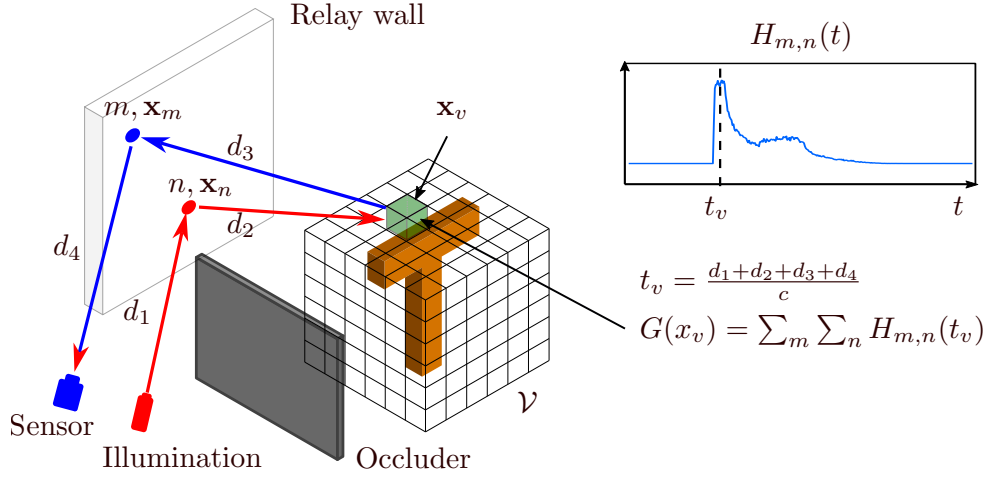


Figure 8.1: Schematic of the backprojection formulation. A mesh of voxels partitions the imaging space containing the hidden geometry. The signal for each voxel $G(\mathbf{x}_v)$ is computed as the contributions from all time-resolved signals $H_{m,n}(t)$ corresponding to time t_v , which corresponds to the total time-of-flight d_1 of light from the illumination source to illumination point n on the relay wall, time-of-flight d_2 from illumination point n to a voxel $\mathbf{x}_v \in \mathbf{V}$, time-of-flight d_3 from voxel $\mathbf{x}_v$ to sensing point m on the relay wall, and time-of-flight d_4 back to the sensor from sensing point m . $\mathbf{x}_l$ and $\mathbf{x}_s$ describe the position coordinates of illumination point n and sensing point m respectively.

the form that *kets* take in the particular Hilbert space that describes light under a given method. For example, wave-based imaging methods are conveniently described in frequency space (phasors), whereas time-domain methods (backprojection) are described in a real time-resolved space; and ii) find the adequate propagation operator $\hat{U}_{m,n}$ that acts on the *kets* in a particular Hilbert space based on the specific restrictions of the reconstruction method. Other NLOS reconstruction methods may be derived from our proposed framework as well, following this methodology. For the following three sections, we derive backprojection, Phasor Fields and fk-migration from our notation.

8.4 DERIVATION OF BACKPROJECTION FORMULATION.

The backprojection formulation and reconstruction was one of the first methods proposed to reconstruct a hidden scene from an arbitrary number of illumination and sensing points [115]. In practical terms, illumination is realized using laser pulses whereas sensing is realized using time-resolved detectors. In the backprojection formulation, for every voxel $\mathbf{x}_v$ of a hidden scene inside the imaging volume $\mathbf{V}$, the image signal corresponding to that voxel $G(\mathbf{x}_v)$ from a set of time-resolved measurements $H_{m,n}(t)$ on the relay wall is defined as:

$$G(\mathbf{x}_v) = \sum_m \sum_n H_{m,n}(t_v) \quad (8.13)$$

where t_v corresponds to the total time-of-flight of light from the illumination source to the illumination point n (d_1), from illumination point n to voxel $\mathbf{x}_v$ (d_2), from voxel $\mathbf{x}_v$ to sensing point m (d_3), and back from sensing point m to the detector (d_4). Figure 8.1 illustrates the backprojection formulation in which the hidden geometry is contained within a mesh of voxels in $\mathbf{V}$, and the different light paths from the illumination source to the sensor. $\mathbf{x}_l$ describes the position coordinates of illumination point n , while $\mathbf{x}_s$ describes the position coordinates of sensing point m .

In quantum mechanics, an *observable* is an experimental measurement of an operator. We will show that the backprojection formulation described in Velten et al. [115] can be derived by inverting the observable corresponding to the operator defined in Equation (8.12). One of the fundamental postulates of quantum mechanics is that it is not possible to measure all possible outcomes of an operator simultaneously; instead, we can only measure the probability of all the possible outcomes of such operator. The Born rule [100]

defines how to measure the observable O corresponding to an operator $\hat{O}$, which defines the observable O as the expected value of operator $\hat{O}$ using suitable basis functions $\{|b_a\rangle\}$ representing the system's state:

$$O = \sum_a \langle b_a | \hat{O} | b_a \rangle \quad (8.14)$$

The way that the expected value of an operator is defined by the Born rule determines that the result is a Real number or function. Although our scene operator is not based on an actual physical quantum mechanics system, it is formally defined as one and it is a Hermit operator and, thus, it benefits from the mathematical body applicable to quantum mechanics. The reason to use the observable of operator Equation (8.12) in deriving the backprojection formulation is that the observable is a Real magnitude and backprojection uses the intensity of the captured light, which is also a Real magnitude.

The extension of Equation (8.13) to the continuum domain leads to Equation S2 in the supplemental material of Velten et al. [115], which states that:

$$I(\mathbf{x}_s, t) = \int_{\mathbb{R}} dt' \int_{\mathbb{R}^2} \frac{d\mathbf{S}_v}{\pi \|\mathbf{r}_v\|^2} \delta(\|\mathbf{r}_v\| - ct + ct') I(\mathbf{x}_l, t') \quad (8.15)$$

where $\mathbf{r}_v = \mathbf{x}_v - \mathbf{x}_s$, $\int_{\mathbb{R}^2} d\mathbf{S}_v$ is the surface integral of the hidden volume in $\mathbf{V}$, the term $\frac{1}{\pi \|\mathbf{r}_v\|^2} \delta(\|\mathbf{r}_v\| - ct + ct')$ is the diffuse Lambertian reflection of a point in $\mathbf{V}$ illuminated by a point source, and c is the speed of light in a vacuum. As stated in Velten et al. [115], the dependency of the reflection on the surface normal's in $\mathbf{V}$ is neglected.

We can evaluate the observable corresponding to operator Equation (8.12) for a single illumination-sensor pair using the Born rule as:

$$T_{m,n} = \langle b | \hat{T}_{m,n} | b \rangle = \langle b | \hat{U}_{m,n} | l_n \rangle \langle l_n | b \rangle \quad (8.16)$$

where $|b\rangle$ is a suitable basis, for example, Gaussian packets. We can then establish the following identification:

$$I(\mathbf{x}_s, t) \equiv T_{m,n} \quad , \quad I(\mathbf{x}_l, t') \equiv \langle l_n | b \rangle \quad (8.17)$$

where $I(\mathbf{x}_s, t)$ is the intensity captured at sensor position m as a function of time t , and $I(\mathbf{x}_l, t')$ is the intensity of illumination position n as function of time t' .

The propagation term $\langle b | \hat{U}_{m,n} | l_n \rangle$ is:

$$\int_{\mathbb{R}^2} \frac{d\mathbf{S}_v}{\pi \|\mathbf{r}_v\|^2} \delta(\|\mathbf{r}_v\| - ct + ct') \equiv \langle b | \hat{U}_{m,n} | l_n \rangle \quad (8.18)$$

Plugging Equation (8.17) into Equation (8.18) yields:

$$I(\mathbf{x}_s, t) = \int_{\mathbb{R}} dt' \int_{\mathbb{R}^2} \frac{d\mathbf{S}_v}{\pi \|\mathbf{r}_v\|^2} \delta(\|\mathbf{r}_v\| - ct + ct') I(\mathbf{x}_l, t') \quad (8.19)$$

where $\int_{\mathbb{R}} dt'$ implements the outer product defined in Equation (8.16). The result is an equation that is exactly Equation (8.15), which is the starting equation for the backprojection formulation described in Velten et al. [115]. Therefore, the identification expressed by Equation (8.17) and Equation (8.18), coming originally from Equation (8.12), demonstrate that the backprojection formulation can be formally derived from our proposed framework.

8.5 DERIVATION OF PHASOR FIELDS FORMULATION.

In this section we show how the Phasor Fields formulation can be derived from Equation (8.12). Prior to demonstrate the derivation, we refer to Section 3.2 for background, although in this chapter we focus on the notation as it appears in the original work [69].

8.5.1 *Derivation.*

ToF data is captured in the time domain. However, any signal captured in the time domain can be expressed as the superposition of its frequency components. To derive Phasor Fields from Equation (8.10) we thus need to express the signals captured in the time-domain as the inverse Fourier transform of the corresponding signals in the frequency-domain. Moreover, a convenient Hilbert space to describe wave-based propagation methods is the frequency space, or Fourier space. Other Hilbert spaces may be used as well and, indeed, our proposed framework enables that possibility.

The *ket* representing the captured signal in the time-domain can be expressed as the inverse Fourier transform of the signal in the frequency-domain as:

$$|s_m\rangle = \frac{1}{2\pi} \int_{\mathbb{R}} d\omega S(\mathbf{x}_s, \omega) e^{i\omega t} \quad (8.20)$$

where $S(\mathbf{x}_s, \omega)$ is the captured signal expressed in the frequency domain.

The time component of the *bra* representing the illumination has to be expressed as the conjugate of the inverse Fourier transform as:

$$\langle l_n| = \frac{1}{2\pi} \int_{\mathbb{R}} d\omega' L^*(\mathbf{x}_l, \omega') e^{-i\omega' t} \quad (8.21)$$

Using these expressions of the *ket* and *bra* in Equation (8.10) results in:

$$\hat{T}_{m,n} = \int_{\mathbb{R}^2} d\mathbf{x}_l \frac{1}{2\pi} \int_{\mathbb{R}} d\omega S(\mathbf{x}_s - \mathbf{x}_l, \omega) e^{i\omega t} \frac{1}{2\pi} \int_{\mathbb{R}} d\omega' L^*(\mathbf{x}_l, \omega') e^{-i\omega' t} \quad (8.22)$$

Now, assuming that the illumination is a delta in time with a Gaussian distribution of energy in space results in:

$$L(\mathbf{x}_l, \omega') = L_n^{(0)} e^{-\left(\frac{\|\mathbf{x}_l\|}{\sigma_n}\right)^2} e^{-i\omega' t_n^{(0)}} \quad (8.23)$$

and plugging Equation (8.23) into Equation (8.22) yields:

$$\hat{T}_{m,n} = \int_{\mathbb{R}^2} d\mathbf{x}_l \frac{1}{2\pi} \int_{\mathbb{R}} d\omega' L_n^{(0)} e^{-\left(\frac{\|\mathbf{x}_l\|}{\sigma_n}\right)^2} e^{-i\omega'(t-t_n^{(0)})} \frac{1}{2\pi} \int_{\mathbb{R}} d\omega S(\mathbf{x}_s - \mathbf{x}_l, \omega) e^{i\omega t} \quad (8.24)$$

where $\int_{\mathbb{R}^2} d\mathbf{x}_l$ implements the outer product between Equation (8.20) and Equation (8.21) as required in defining $\hat{T}_{m,n}$ in Equation (8.10). For delta-type illumination pulses, $S(\mathbf{x}_s - \mathbf{x}_l, \omega) = \tilde{H}_{m,n}(\omega) = \mathcal{F}[H_{m,n}(t)]$, where $\tilde{H}_{m,n}(\omega)$ is the Fourier transform of the impulse response of the scene $H_{m,n}(t)$.

Note that Equation (8.24) is equivalent to the original Phasor Fields formulation in Equation (3.6), as follows:

$$\begin{aligned} \mathcal{P}(\mathbf{x}_s, t) &\equiv \hat{T}_{m,n} \\ \mathcal{P}(\mathbf{x}_l, t) &\equiv \frac{1}{2\pi} \int_{\mathbb{R}} d\omega' L_n^{(0)} e^{-\left(\frac{\|\mathbf{x}_l\|}{\sigma_n}\right)^2} e^{-i\omega'(t-t_n^{(0)})} \\ H_{m,n}(t) &\equiv \frac{1}{2\pi} \int_{\mathbb{R}} d\omega S(\mathbf{x}_s - \mathbf{x}_l, \omega) e^{i\omega t} \end{aligned} \quad (8.25)$$

where $\mathcal{P}(\mathbf{x}_s, t)$ is the phasor at the virtual camera $\mathbf{S}$, $\mathcal{P}(\mathbf{x}_l, t)$ is the phasor at the virtual illumination source surface $\mathbf{L}$, and $H_{m,n}(t)$ is the scene impulse response of the illumination-sensor pair $\{m, n\}$. This shows how our quantum-mechanics framework includes the Phasor Fields formulation.

8.5.2 *Rayleigh-Sommerfeld diffraction operator is the propagation operator in wave-based imaging methods.*

The Schrödinger equation describes the temporal evolution of a quantum system, which is expressed as:

$$i \frac{\hbar}{2\pi} \frac{\partial}{\partial t} |\psi(\mathbf{x}, t)\rangle = \hat{H}(\mathbf{x}, t) |\psi(\mathbf{x}, t)\rangle \quad (8.26)$$

where h is Planck's constant, $|\psi(\mathbf{x}, t)\rangle$ is the wavefunction describing the state of the quantum system, and $\hat{H}(\mathbf{x}, t)$ is the Hamiltonian operator, which is a unitary operator that describes the physics of the quantum system. Note that the dependencies in position $\mathbf{x}$ and time t are expressed explicitly in the *kets* and operators.

The solution of the Schrödinger equation at an arbitrary position $\mathbf{x}$ and time t for a free particle, in our case a photon, is described by a wavefunction $|\psi(\mathbf{x}, t)\rangle$ such that [20, 100]:

$$|\psi(\mathbf{x}, t)\rangle = \int_{\mathbb{R}^3} d\mathbf{k} \frac{e^{i[\mathbf{k}(\mathbf{x}-\mathbf{x}_0)-\omega(t-t_0)]}}{\|\mathbf{x}-\mathbf{x}_0\|^2} |\psi(\mathbf{x}_0, t_0)\rangle \quad (8.27)$$

where $\mathbf{k}$ is the wavenumber vector in the direction of propagation of the photon, ω its frequency, and $|\psi(\mathbf{x}_0, t_0)\rangle$ the wavefunction describing its state at initial position $\mathbf{x}_0$ and time t_0 . We can set $t_0 = 0$ as our temporal reference and, thus, redefine $t \equiv t - t_0$.

Defining $|\psi(\mathbf{x}_l)\rangle \equiv |\psi(\mathbf{x}_0, t_0)\rangle$ as the wavefunction describing the state of light at the virtual light source surface $\mathbf{L}$ at $t_0 = 0$ and $|\psi(\mathbf{x}_s, t)\rangle$ as the wavefunction describing the state of the captured light at the virtual camera $\mathbf{S}$ at a later time t , we obtain:

$$|\psi(\mathbf{x}_s, t)\rangle = \int_{\mathbb{R}^3} d\mathbf{k} \frac{e^{i\mathbf{k}(\mathbf{x}_s-\mathbf{x}_l)-i\omega t}}{\|\mathbf{x}_s-\mathbf{x}_l\|^2} |\psi(\mathbf{x}_l)\rangle \quad (8.28)$$

and defining the following identification yields:

$$\begin{aligned} \mathcal{P}(\mathbf{x}_l) &\equiv |\psi(\mathbf{x}_l)\rangle \\ \mathcal{P}(\mathbf{x}_s, t) &\equiv |\psi(\mathbf{x}_s, t)\rangle \end{aligned} \quad (8.29)$$

which transforms Equation (8.28) into:

$$\mathcal{P}(\mathbf{x}_s, t) = \int_{\mathbb{R}^3} d\mathbf{k} \frac{e^{i\mathbf{k}(\mathbf{x}_s-\mathbf{x}_l)-i\omega t}}{\|\mathbf{x}_s-\mathbf{x}_l\|^2} \mathcal{P}(\mathbf{x}_l) \quad (8.30)$$

The contribution from all the point sources lying on the virtual light surface $\mathbf{L}$ results in:

$$\boxed{\mathcal{P}(\mathbf{x}_s, t) = \int_{\mathbf{L}} d\mathbf{x}_l \int_{\mathbb{R}^3} d\mathbf{k} \frac{e^{i\mathbf{k}(\mathbf{x}_s-\mathbf{x}_l)-i\omega t}}{\|\mathbf{x}_s-\mathbf{x}_l\|^2} \mathcal{P}(\mathbf{x}_l)} \quad (8.31)$$

which is equivalent to Equation S.8 and S.9 in Liu et al. [69], representing the polychromatic wave as a superposition of monochromatic waves. Different from the original formulation, which was limited to delta-like illumination functions, our framework allows the modeling of any *arbitrary* light signals.

8.6 DERIVATION OF F-K MIGRATION FORMULATION.

The f-k migration formulation is applied to confocal measurements in which the measurement of each illumination point happens at the same position of the sensing point, i.e. $\mathbf{x}_s = \mathbf{x}_l = \mathbf{x}$. Considering a single illumination-sensing pair located at the same position and writing the dependency on position and time explicitly, turns Equation (8.12) into:

$$\hat{T}_{m,n} = |s_m(\mathbf{x}, t)\rangle \langle l_n(\mathbf{x}, t')| \quad (8.32)$$

where *kets* $|l_n(\mathbf{x}, t)\rangle$ and $|s_m(\mathbf{x}, t')\rangle$ contain the explicit dependency on position $\mathbf{x}$ and time t and describe the state of light at the confocal position $\mathbf{x}$ at times t and t' for the illumination and sensing instants, respectively.

The propagation of light from t to t' can be expressed as:

$$|s_m(\mathbf{x}, t)\rangle = \hat{U}_{m,n}(\mathbf{x}, t, t') |l_n(\mathbf{x}, t')\rangle \quad (8.33)$$

where $\hat{U}_{m,n}(\mathbf{x}, t, t')$ is the propagation operator containing the explicit dependency on confocal position $\mathbf{x}$ and time propagating the light from the illumination point at time t' to the hidden scene and back to the sensing point at t . Plugging Equation (8.33) into Equation (8.32) yields:

$$\hat{T}_{m,n} = \hat{U}_{m,n}(\mathbf{x}, t, t') |l_n(\mathbf{x}, t')\rangle \langle l_n(\mathbf{x}, t')| \quad (8.34)$$

Multiplying Equation (8.34) by $|l_n(\mathbf{x}, t)\rangle$ on both sides we obtain:

$$\hat{T}_{m,n} |l_n(\mathbf{x}, t)\rangle = \hat{U}_{m,n}(\mathbf{x}, t, t') |l_n(\mathbf{x}, t')\rangle \langle l_n(\mathbf{x}, t') | l_n(\mathbf{x}, t)\rangle \quad (8.35)$$

and considering that the illumination states are orthogonal in the time axis:

$$\langle l_n(\mathbf{x}, t') | l_n(\mathbf{x}, t)\rangle = \delta(t' - t) \quad (8.36)$$

which leads to $t = t'$ and, thus:

$$\hat{U}_{m,n}(\mathbf{x}, t, t') = \hat{U}_{m,n}(\mathbf{x}, t, t) \quad (8.37)$$

which allows us to redefine $\hat{U}_{m,n}(\mathbf{x}, t, t')$ as $\hat{U}_{m,n}(\mathbf{x}, t, t') \equiv \hat{U}_{m,n}(\mathbf{x}, t)$ and, therefore:

$$(\hat{T}_{m,n} - \hat{U}_{m,n}(\mathbf{x}, t)) |l_n(\mathbf{x}, t)\rangle = 0 \quad (8.38)$$

The propagation equation that is used as the starting point for the f-k migration formulation is Equation 1 in Lindell et al. [66], which states that:

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \Psi(\mathbf{x}, t) = 0 \quad (8.39)$$

where Ψ is the complex-valued scalar wave field, the expression in front of Ψ is the wave equation operator, and c is the speed of light.

Now, using the following identification:

$$\Psi(\mathbf{x}, t) \equiv |l_n(\mathbf{x}, t)\rangle \quad (8.40)$$

$$(\hat{T}_{m,n} - \hat{U}_{m,n}(\mathbf{x}, t)) \equiv \left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \quad (8.41)$$

we can exactly match Equation (8.39) with Equation (8.38).

The solution for Equation (8.39) relies on the boundary condition:

$$\Psi(\mathbf{x}, t)|_{z=0} = \Psi(\mathbf{x}, t)|_{t=0} \quad (8.42)$$

This results shows how the f-k migration formulation can also be derived from our framework within the context of confocal measurements and the boundary conditions described by Lindell et al [69]. The physical interpretation of Equation (8.40) is the virtual wavefront that complies with the boundary condition set by the f-k model.

8.7 BACKPROJECTION AND WAVE-BASED IMAGING METHODS ARE EQUIVALENT POINT-TO-POINT.

As we explained in Section 8.2, backprojection and wave-based propagation are the two most used formulations to image hidden scene in NLOS scenarios. While backprojection voxelizes the imaging space and computes the signal contributions corresponding to the time of flight between the illumination and sensing points and a particular voxel, wave-based propagation models the propagation defining virtual light

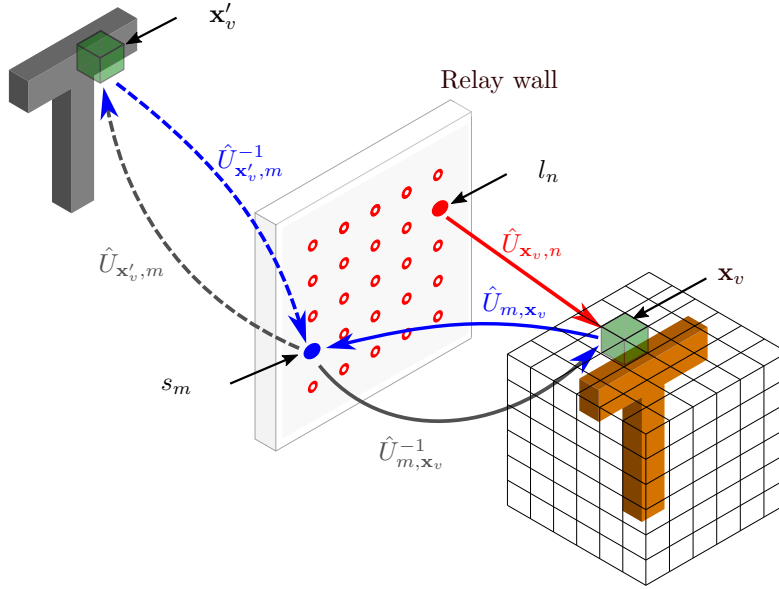


Figure 8.2: Diagram showing the operators and their propagation directions for both backprojection and forward propagation to help visualizing the algebra developed in Section 8.7. $|l_n\rangle$ represents the illumination at point n , $|s_m\rangle$ represents the signal at sensing point m , $|\mathbf{x}_v\rangle$ represents the signal at the voxel $\mathbf{x}_v$ in the backprojection model, and $|\mathbf{x}'_v\rangle$ represents the signal at the virtual voxel $\mathbf{x}'_v$ in the forward propagation model. The result expressed by Equation (8.49) demonstrates that backprojection and forward propagation are equivalent point-to-point.

sources and virtual cameras converting the NLOS geometry in a line-of-sight geometry (hence, wave-based propagation formulations are also referred as forward-propagation formulations).

At first sight, it may look as if backprojection and wave-based methods are quite different; however, our proposed formulation allows us to demonstrate formally that both are actually equivalent. The demonstration of this equivalence is based on the fact that propagation operators are unitary, which is demonstrated in Chapter A. A unitary operator is an operator such that $\hat{U}\hat{U}^\dagger = \hat{U}\hat{U}^{-1} = \mathbb{I}$.

Figure 8.2 depicts a diagram showing the operators and their propagation directions for both backprojection and forward propagation, which will help visualizing the algebra developed in this section. The state of light at point m is represented by $|s_m\rangle$, the state of light at point $\mathbf{x}_v$ is represented by $|\mathbf{x}_v\rangle$, the state of light at point n is represented by $|l_n\rangle$, and the state of light at the virtual point $\mathbf{x}'_v$ is represented by $|\mathbf{x}'_v\rangle$.

The propagation from n to m goes through $\mathbf{x}_v$, therefore:

$$|s_m\rangle = \hat{U}_{m,\mathbf{x}_v} |\mathbf{x}_v\rangle \quad (8.43)$$

where $\hat{U}_{m,\mathbf{x}_v}$ propagates the light from voxel $\mathbf{x}_v$ to point m . The inverse of Equation (8.43) is:

$$|\mathbf{x}_v\rangle = \hat{U}_{m, \mathbf{x}_v}^{-1} |s_m\rangle \quad (8.44)$$

In general, we will consider that $\hat{U}_{b,a}$ is the propagation operator of light from a generic point a to a generic point b .

Now, the propagation from point m to point $\mathbf{x}'_v$ is given by:

$$|\mathbf{x}'_v\rangle = \hat{U}_{\mathbf{x}'_v, m} |s_m\rangle \quad (8.45)$$

and the inverse is:

$$|s_m\rangle = \hat{U}_{\mathbf{x}'_v, m}^{-1} |\mathbf{x}'_v\rangle \quad (8.46)$$

Plugging Equation (8.43) into Equation (8.45) yields:

$$|\mathbf{x}'_v\rangle = \hat{U}_{\mathbf{x}',m} \hat{U}_{m,\mathbf{x}_v} |\mathbf{x}_v\rangle \quad (8.47)$$

We can define the propagation operator from $\mathbf{x}_v$ to $\mathbf{x}'_v$ as:

$$\hat{U}_{\mathbf{x}'_v, m} \hat{U}_{m, \mathbf{x}_v} \equiv \hat{U}_{\mathbf{x}'_v, \mathbf{x}_v} \quad (8.48)$$

Now, using the property that $\hat{U}_{\mathbf{x}'_v, \mathbf{x}_v}$ must be unitary:

$$\boxed{\hat{U}_{\mathbf{x}'_v, \mathbf{x}_v} \hat{U}_{\mathbf{x}'_v, m} \hat{U}_{m, \mathbf{x}_v} = \mathbb{I} \Rightarrow \hat{U}_{\mathbf{x}'_v, m} = \hat{U}_{m, \mathbf{x}_v}^{-1}} \quad (8.49)$$

and the only way this is possible is if $|\mathbf{x}'_v\rangle = |\mathbf{x}_v\rangle$, which means that the state of light at both $\mathbf{x}_v$ and $\mathbf{x}'_v$ voxels is identical.

This result demonstrates that propagating the state of light at point m using the inverse of the forward operator from point m to point $\mathbf{x}'_v$ in Equation (8.45) is the same as inverting the propagation operator from point $\mathbf{x}_v$ to point m defined in Equation (8.43). In other words, we have demonstrated quite naturally that both backprojection and wave-based imaging methods are equivalent point-to-point. This equivalence may open the possibility to develop hybrid formulations to optimize imaging scenes based on certain features of the scene and the experimental setup.

This result imposes certain conditions on the properties of the transport matrix $[\hat{U}_{m, n}]$, which considers the propagation among all illumination and sensing points. For instance, Equation (8.9) describes the propagation of light from an illumination point $\mathbf{x}_l$ to a sensor point $\mathbf{x}_s$. However, the signal collected at the sensor point will be the contribution of all the illumination points and, hence:

$$|s_m\rangle = \sum_{n'} \hat{U}_{m, n'} |l_{n'}\rangle \quad (8.50)$$

Similarly, the inverse of Equation (8.50) is:

$$|l_n\rangle = \sum_{m'} \hat{U}_{m', n}^{-1} |s_{m'}\rangle \quad (8.51)$$

which can be written in terms of Equation (8.51) as:

$$|s_m\rangle = \sum_{n'} \sum_{m'} \hat{U}_{m, n'} \hat{U}_{m', n}^{-1} |s_{m'}\rangle \quad (8.52)$$

Equation (8.51) is not the direct inverse of Equation (8.50). Equation (8.50) expresses the contributions of all illumination points on a sensing point m , whereas Equation (8.51) represents the contribution of all sensing points to an illumination point n if we were to invert the propagation from all sensing points to illumination point n .

Using the fact that each component (m, n) of transport matrix $\hat{U}_{m, n}$ is unitary, i.e. $\hat{U}_{m, n}^{-1} = \hat{U}_{m, n}^\dagger = \hat{U}_{n, m}^*$, Equation (8.52) can be further expressed as:

$$|s_m\rangle = \sum_{n'} \sum_{m'} \hat{U}_{m, n'} \hat{U}_{n, m'}^* |s_{m'}\rangle \quad (8.53)$$

which leads to:

$$\sum_{n'} \sum_{m'} \hat{U}_{m, n'} \hat{U}_{n, m'}^* = \delta_{m, m'} \delta_{n', n} \quad (8.54)$$

where $\delta_{m, m'}$ and $\delta_{n', n}$ are Kronecker deltas for sensing point index m and illumination point index n . The Kronecker deltas are introduced because the states representing the illumination points and sensing points are orthogonal by definition.

Equation (8.54) implies that the trace of the matrix resulting from the product in Equation (8.54) is equal to one (or a Real number if $[\hat{U}_{m, n}]$ is not normalized):

$$\boxed{\text{tr}([\hat{U}_{m, n}][\hat{U}_{n, m}^*]) = 1} \quad (8.55)$$

Although the transport matrix $[\hat{U}_{m, n}]$ as a whole might not be invertible, it must comply with Equation (8.55) and, therefore, it can be used to invert it by using diagonalization or matrix decomposition methods.

8.8 APPLICATION: NLOS WITH NON-DELTA ILLUMINATION.

Equation (8.10) and Equation (8.12) provide the fundamental description of light interaction and propagation in a NLOS context using Dirac notation. This description is general enough to represent a varied range of NLOS reconstruction methods, such as backprojection, Phasor Fields and f-k migration. The use of this notation not only serves as a general theoretical framework, but it enables derivations that are independent of the reconstruction algorithm as well. In this section we illustrate this with one of such derivations, in which we lift one of the main assumptions of current NLOS methods: the use of delta pulses of light as illumination source. This serves as foundations for later Chapter 9 in which we delve further.

As stated in Equation (8.11), the propagation operator $U_{m,n}$ relates the state of light on capture $|s_m\rangle$ and the state of the illumination $|l_n\rangle$. The information about the hidden scene lies within this propagation operator $U_{m,n}$ and NLOS reconstruction algorithms assume the illumination $|l_n\rangle$ to be a delta in time. However, real-world illumination hardware does not emit delta but near Gaussian pulses, accounting for excitation and decay. The width (standard deviation) of such Gaussian pulses is extremely small for sophisticated, expensive lasers, being larger (and therefore impractical for NLOS reconstructions) for lower-cost emitters. Using our framework, we can disentangle the effect of the illumination from the capture, enabling the use of Gaussian distributed temporal models for emitters (cheaper lasers), instead of delta models, which in turn enables the use of cheaper time-resolved sensors as well.

Similar to other techniques [69, 79], we assume a single illumination point in $|l_n\rangle$ with known temporal emission profile, which transforms Equation (8.12) into:

$$\hat{T}_{m,1} = \hat{U}_{m,1} |l_1\rangle \langle l_1| \quad (8.56)$$

and multiplying either side with $|l_1\rangle \equiv \hat{U}_{m,1} \|l_1\|^2$ results in:

$$\hat{T}_{m,1} |l_1\rangle = \hat{U}_{m,1} |l_1\rangle \langle l_1 | l_1\rangle = \hat{U}_{m,1} \|l_1\|^2 |l_1\rangle \rightarrow (\hat{T}_{m,1} - \hat{U}_{m,1} \|l_1\|^2) |l_1\rangle = 0 \quad (8.57)$$

which implies that:

$$\hat{T}_{m,1} = \hat{U}_{m,1} \|l_1\|^2 \quad (8.58)$$

if Equation (8.57) has to hold true $\forall |l_1\rangle$.

Because Phasor Fields operates in the frequency space, we can express $\|l_1\|^2$ as the summation of the square amplitude of each frequency component of the signal as:

$$\|l_1\|^2 = \sum_v \|l_{1,v}\|^2 \quad (8.59)$$

where $\|l_{1,v}\|^2$ is the square amplitude of frequency component v . Notice that decomposition expressed in Equation (8.59) is not restricted to the frequency domain alone but other domains can benefit from such decomposition as well.

A delta pulse can be normalized such that:

$$\|l_1\|^2 = \sum_v \|l_{1,v}\|^2 = 1 \quad (8.60)$$

Because of the superposition principle, we can decompose Equation (8.58) for each frequency component v as well. Then, dividing Equation (8.58) on sides by $\|l_{1,v}\|^2$, we actually obtain a delta-coupled scene operator for each frequency component v :

$$\hat{T}'_{m,1,v} = \frac{\hat{T}_{m,1,v}}{\|l_{1,v}\|^2} \quad (8.61)$$

Now, we can add all frequency components of the scene operator as:

$$\hat{T}'_{m,1} = \sum_v \frac{\hat{T}_{m,1,v}}{\|l_{1,v}\|^2} \equiv \hat{U}_{m,1} \quad (8.62)$$

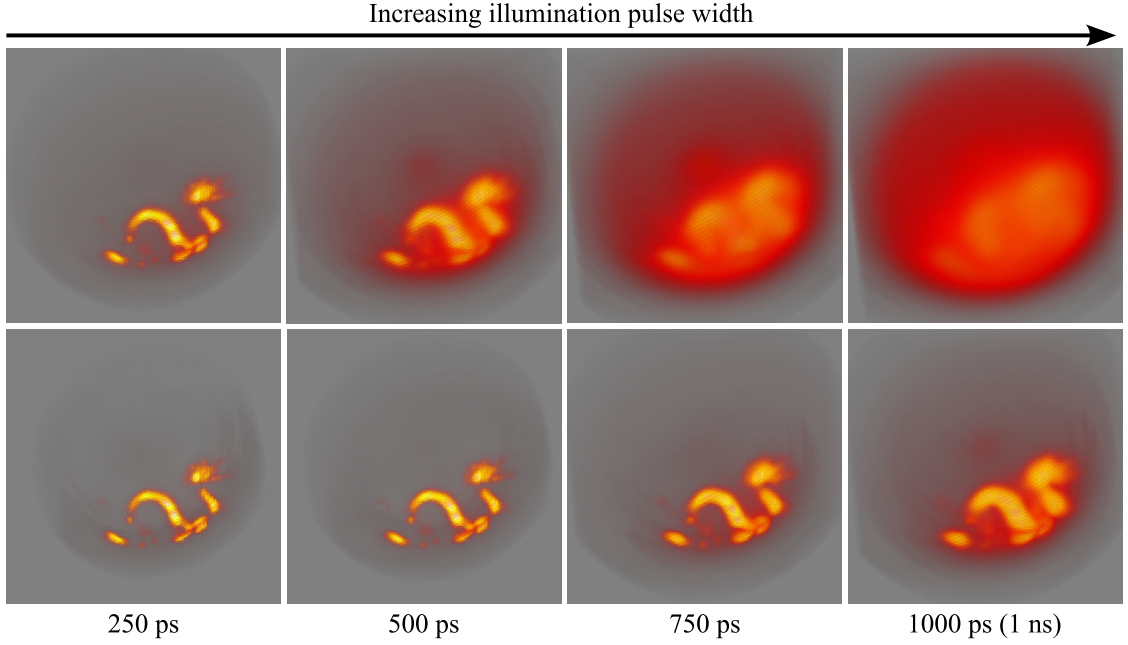


Figure 8.3: Results of the reconstruction of a dragon with Phasor Fields using a Gaussian light pulse of increasing width: 250, 500, 750, and 1000 picoseconds. Top: reconstruction of the raw signal with Gaussian illumination. As the width of the pulse increases the reconstructions become blurrier, to the point where the dragon cannot be recognized. Bottom: reconstruction using our framework. Even in the presence of very wide pulses, we can still obtain reasonable reconstructions.

We then use this delta-decoupled operator for the NLOS reconstructions instead of the original signal. In other words, we can achieve the same effect of a delta-like illumination from an arbitrary time-resolved illumination by applying Equation (8.62) and using $\hat{T}'_{m,1}$ as input for the NLOS reconstructions.

To validate our results, we have synthesized a hidden scene consisting of a Stanford dragon with a transient renderer optimized for NLOS [96]. The dragon is at 1 meter of distance to a relay wall of 2×2 meters, with a sensing grid of 128×128 points $\mathbf{S}$, and a single illumination point $\mathbf{L}$ at the center of the relay wall. In Figure 8.3 we show the results of reconstructing the hidden scene with the Phasor Fields framework [69], with the temporal model of a Gaussian pulse with different widths: 250, 500, 750, and 1000 picoseconds. As the top row shows, as the width of the pulse increases, the reconstruction becomes more and more blurry until no figure can be recognized. Using our framework, however, we obtain reasonable reconstructions (where the figure of the dragon is clearly recognizable) even in the presence of very wide Gaussian pulses.

Aspects of the signal such as the effect of noise and decreasing temporal resolution can be explored under our proposed framework following on the steps of the example described in this section. However, such exploration lies outside the scope of this chapter, which aims to establish a foundation for such exploratory works, as we perform in Chapter 9.

8.9 DISCUSSION AND CONCLUSION.

This work draws inspiration from Dirac himself, who said: *"if one is working from the point of view of getting beauty in one's equations, and if one has really a sound insight, one is on a sure line of progress"*, and: *"a theory with mathematical beauty is more likely to be correct than an ugly one that fits some experimental data"*. We have focused on laying down the theoretical groundwork to provide a *comprehensive* formulation of the different mechanisms and imaging methods in NLOS imaging. In particular, formulating the NLOS problem in terms of the Dirac notation and defining a novel scene operator (Equation (8.10)) has allowed us i) to derive the backprojection, the Phasor Fields, and the

f-k migration formulations from a unified framework; ii) to demonstrate that the Rayleigh-Sommerfeld diffraction operator is the propagation operator for wave-based propagation imaging formulations; and iii) to demonstrate that backprojection and wave-based imaging formulations are equivalent. We have demonstrated that the unitary property of propagation operators is necessary so that the experimental measurements are real-valued magnitudes, reversible in time, and consistent with Helmholtz reciprocity principle. The development of this work has been based on the key realization that phasors defined in the wave-based formulation can be considered as projections of *kets* on different representations and operators in a Hilbert space that model the interaction of light between a relay wall and a hidden scene. The different projections of the *kets* yields the different formulations, including backprojection, Phasor Fields, and f-k migration. The conclusions of this chapter are consistent with traditional wave optics framework. However, we provide a different perspective that is more formal and general in the context of formulating NLOS imaging. We actually express the concept of time-resolved NLOS imaging in a way that is agnostic to the reconstruction algorithm and, therefore, our conclusions are not tied to particular algorithms.

Other NLOS methods could also be derived from our framework. Although a full formal derivation falls out of the scope of this thesis, we provide here some meaningful insights towards this goal. For instance, the technique proposed in Saunders et al. [101] is based on measuring the shifts of the shadows caused by an occluder that is placed between the NLOS scene and the relay wall. Using our framework the shift of the measurement points on the relay wall could be modeled as an affine transformation of the propagation operator $\hat{U}$ connecting the illumination and the sensed points. Such affine transformation would be directly related to the relative position of the occluder and the hidden scene. The technique proposed in Pei et al. [87] is based on modeling arbitrary illumination patterns represented by different point-spread-function patterns, and performing an optimization process to minimize noise and select the most representative sampling points. One of the strengths of our framework is that it allows arbitrary illumination functions, therefore Pei et al. fits naturally in it. The technique proposed in Cao et al. [8] is based on actively changing the focus of the illumination using wavefront shaping. Being a wave-based method it could be derived similarly to the derivation of phasor-fields and f-k migration, with operator $\hat{U}$ being mapped to the operator in charge of wavefront shaping and focusing.

Finally, inspired by another Dirac's quote stating that "*I understand an equation when I can predict the properties of its solutions, without actually solving it*", we have laid down the foundation of a cohesive mathematical framework to formulate NLOS challenges at large and we expect it will provide a deeper understanding of the aspects that are most important in imaging NLOS scenes as well as to devise more optimal and fast scene sampling and computational methods. This effort may include the formulation of a generalized light transport operator to model properties such as polarization, scattering, dispersion, or birefringence. The key process for incorporating those properties would be to represent them in a suitable Hilbert space with the corresponding propagation operators. This work is the first stepping stone for future research directions involving both theoretical and experimental results.

NON-LINE-OF-SIGHT IMAGING FROM SUBOPTIMAL INSTRUMENT RESPONSE.

ABOUT THIS CHAPTER

In this chapter, we propose a generalized formulation to model suboptimal instrument responses in NLOS captures, based on Dirac notation. We explicitly account for noise and point-spread functions present during acquisition, and establish a mathematical relationship whose solutions allow estimating an ideal capture from a real measurement, as demonstrated in our results. This work can be understood as a continuation of the experiments presented in Chapter 8, expanding those results with a more thorough analysis.

At the time of this thesis submission, the work described in this chapter is in preparation for submission to Optics Express. As first author, I contributed to the theoretical foundations, method development, experimental design, experimental validation, literature revision, writing and revision of the manuscript.

9.1 INTRODUCTION.

Ideal NLOS reconstructions are obtained from a clean noiseless impulse response of the hidden scene, which demands an infinitely high temporal resolution, and assume a perfect delta as the instrument response of the sensor-laser pair. However, sensor and laser devices, nowadays mainly based on single-photon avalanche diodes (SPAD) [27, 7], present a suboptimal instrument response. The instrument response of the sensor presents a specific point-spread function (PSF) due to electronics [45], with temporal resolution ranging from tens to a few hundreds of picoseconds. Emitters also present a PSF as a result of the excitation and decay time of the materials to produce the light. These *sensor* PSF and *illumination* PSF combine in the capture, deteriorating the instrument response with a Gaussian-shaped PSF for the captures. Lower-end devices produce a wider PSF, with worse spatial resolution in NLOS reconstructions [86, 80]. Furthermore, another effect that deteriorates the instrument response comes from photon loss: a SPAD sensor requires absorbing a photon to trigger the detection, with a probability modeled by the photon-detection efficiency [45, 102]. At large, this affects capture with the appearance of a Poisson-like noise due to the loss of photons, which becomes more problematic for shorter capture times (less photons) or in photon-starved environments.

Therefore, the instrument response is suboptimal, compared to the ideal clean impulse response that is usually assumed by NLOS reconstruction algorithms, affected by the system PSF and the photon-loss noise. To enhance the capture quality, researchers employ devices with minimal PSF, restricting them to expensive and specialized hardware. Also, they increase capture time to maximize the number of photons captured by the SPAD sensors [7]. Overall, this limits NLOS applications to laboratory-controlled conditions, where even calibration already presents a challenge, hindering the practical usability and widespread adoption of in-the-wild NLOS techniques.

In this chapter, our goal is to overcome the limitations coming from a suboptimal instrument response by computationally transforming a noisy temporal signal, such as the ones obtained with fewer-photon scenarios (lower capture times), with a wide PSF of the system, associated with lower-end devices, into a cleaner higher-frequency output, as if it were captured with higher photon counts and higher-end devices. We propose a generalized solution for computationally reversing hardware-induced artifacts in the signal

captured from the hidden scene. As a result, we obtain less noisy and more precise reconstructions, independently of the specific NLOS reconstruction method. We achieve this by employing the Dirac notation, broadly applied in quantum mechanics, which we use to represent transient NLOS imaging in Chapter 8. This notation enables us to establish a mathematical base that models the ideal capture and relates it to the expected noise of the SPAD, along with the system PSF, from which we derive a way to computationally reverse both effects.

The work presented in Chapter 8 assumes that captures from the hidden scene are noiseless. In contrast, in this chapter we exploit Dirac notation to explicitly model both noise and the point-spread function (PSF) in NLOS captures. Note that in Section 8.8 we already introduced this idea. However, in this chapter we provide a more thorough analysis of the generalized notation, model the capture process more realistically by accounting for noise, and establish a mathematical relationship between the ideal and the real capture.

We first showcase our approach with NLOS simulated data, in which we can control the shape of the PSF as well as arbitrarily degrade the time-resolved signal with Poisson noise. In addition, simulated data provide a precise ground truth. We show how our method can reverse the instrument response from low temporal resolution, up to 1225 picoseconds (in contrast, the temporal resolution of common NLOS captures is in the range of tens of picoseconds), and noise associated with up to 90% photon loss. We also show how our method improves the results of real-world noisy captures obtained from publicly available datasets. Our work therefore relaxes the strict hardware conditions of current NLOS systems, making existing techniques more practical, by reversing the effect of the instrument response.

Our proposal is flexible enough to support any arbitrary PSF of the system, including any arbitrary illumination pulses. We demonstrate that previous methods correctly reconstruct a hidden scene illuminated with a consecutive train of deltas when using our method (instead of a single delta pulse assumed by previous methods), failing otherwise. This suggests that our method could inspire novel capture pipelines, free of laser-sensor synchronization, reducing current calibration procedures, or even setting up the base of novel capture methods based on coded illumination.

9.2 TEMPORAL PSF IN NLOS IMAGING.

Despite their promising performance, single-photon avalanche diodes (see Section 2.1 in Chapter 2) and similar devices present time-jittering due to electronics [86, 21, 122], which, together with the laser time of excitation and decay, results in temporal histogram errors that can be understood as the system PSF [4]. The size and shape of the PSF define the temporal resolution of the system, and low temporal resolution degrades the quality of the reconstruction [86]. Previous works tackle this by attempting to increase the resolution [122, 84]. For NLOS reconstructions, this is addressed using modified hardware to enhance the temporal signal [117], or software solutions after reconstructions [121]. Jin et al. estimate the PSF of the capture and applied a histogram deconvolution of the capture from this estimation [53]. Later, Wang et al. modeled the problem as a Poisson convolution and then deconvolved the temporal signal [118]. Recently, Tian et al. developed an iterative algorithm to achieve super resolution in NLOS captures [111].

In this chapter, we present a novel analytical solution that incorporates a model for this PSF, with the novelty of also accounting for photon-loss noise [45, 97]. For this goal, we apply previous Dirac notation for NLOS imaging, which offers a comprehensive and compact notation [91], in which we represent our system PSF and noise model. This model enables us to devise a technique to undo the effect of PSF and noise. Our solutions are agnostic to the reconstruction method later employed. Thus, we use three NLOS imaging methods (filtered backprojection [115], fk-migration [66], and Phasor Fields [69]) to analyze the effects of our method on different NLOS reconstructions.

9.3 OUR METHOD.

This section presents a unified and generalized mathematical framework for removing undesired artifacts derived from suboptimal NLOS captures. We specifically account for the temporal PSF from lower-end

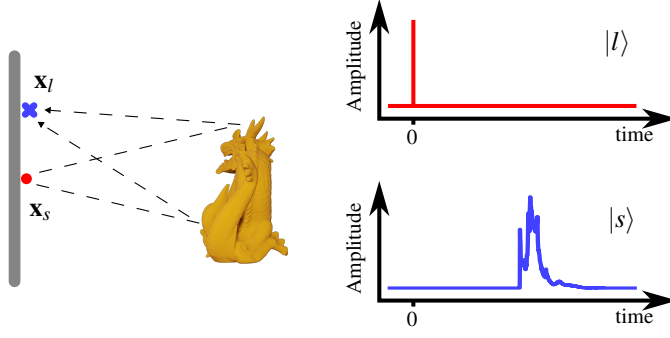


Figure 9.1: The time-resolved light signal reflected from the relay wall at a point $\mathbf{x}_l \in \mathbf{L}$ is represented as a state with the ket $|l\rangle$, while the time-resolved signal scattered from the hidden scene reaching a point $\mathbf{x}_s \in \mathbf{S}$ is represented as a state with the ket $|s\rangle$.

hardware and scene-induced noise such as photon loss. We employ the Dirac notation for NLOS following the ideas and formulation of the collaborative work we presented in Chapter 8, which we succinctly introduce and adapt for this chapter in Section 9.3.1.

Our mathematical derivation begins in Section 9.3.2, where we establish a mapping between an expected ideal capture and its measurement modulated by the system PSF. Later, in Section 9.3.3, we incorporate the noise into our derivation. From both, in Section 9.3.4, we obtain a relationship between the perfect ideal signal and the realistic signal from a suboptimal instrument response, which includes the effect of both PSF and noise. By solving this relationship, we can obtain this perfect ideal signal from the realistic one, and therefore apply NLOS reconstruction algorithms on the perfect ideal signal instead of the suboptimal one. There are different solutions for our proposal; thus, in Section 9.3.5, we propose an example solution in the frequency domain.

9.3.1 Representing an NLOS scene with Dirac notation.

As we previously pointed out in Chapter 8, similarities exist between the uncertainties of predicting a particular outcome from quantum mechanics and probabilities of the presence of geometry in the hidden scene in an NLOS context. This observation motivates them to employ the Dirac notation, broadly used in modern physics, in an NLOS context. As they claim, any NLOS imaging method can be represented under the Dirac notation, and they prove that to be the case for filtered backprojection [115], fk-migration [66], and Phasor Fields [69]. Thus, they provide a cohesive theoretical framework that encompasses all NLOS reconstruction methods. Through this subsection, we provide insights into their proposal adapted to this work. For the reader who is not familiar with Dirac notation, we refer to Section 8.2.1 and Chapter 8.

In the particular case of NLOS imaging, time-resolved light signals are represented by states. The state of light when reflected from a point $\mathbf{x}_l \in \mathbf{L}$ in the relay wall is defined as the ket $|l\rangle$, while the state of light as sensed from point $\mathbf{x}_s \in \mathbf{S}$ in the relay to the sensor wall is represented as $|s\rangle$. Figure 9.1 represents a schema of these kets.

Both kets are related by an operator $\hat{U}$ that represents the change of state due to the interaction of light with the hidden scene, so that

$$|s\rangle = \hat{U} |l\rangle. \quad (9.1)$$

$\hat{U}$ represents the propagation operator of the hidden scene, relating the sensed state to a combination of all illumination states from a single illumination point.

Additionally, the authors define the *scene operator* $\hat{T}$,

$$\hat{T} = |s\rangle \langle l| \quad (9.2)$$

which represents the relationship between the states of light at the illumination and capture points, indirectly embedding the effect that the hidden scene has for each pair of points. Note that $\hat{T}$ represents a mathematical abstraction, accounting for the complete light transport within the hidden scene. In their work, the authors illustrate the generality of Dirac notation by representing several NLOS imaging methods deriving from the scene operator, devising specific propagation and scene operators for each of them, from which the NLOS method emerges naturally.

In the following subsections of this section, we utilize both the propagation and scene operators to devise a general methodology for undoing the effect of non-delta light pulses in noisy capture pipelines. This approach, as in previous work, is general and agnostic to specific NLOS reconstruction methods.

9.3.2 Point spread function in the scene operator.

To fulfill the assumptions of most NLOS reconstruction algorithms, $\hat{T}$ from Equation (9.2), is assumed to be ideal. Therefore, $|l\rangle$ is a perfect delta pulse scattered to the relay wall, and $|s\rangle$ is the ideal noiseless impulse response of the hidden scene. We will refer to this operator as the *ideal delta scene operator*.

However, as we mentioned before, actual illumination and sensors are not delta in time, producing a temporal PSF in the captured signal. We define a new set of kets accounting for the PSF effects: $|l_p\rangle$ as the state of light reaching the relay wall from the emitter, and $|s_p\rangle$ as the state of the light captured by the sensor from the relay wall. Note that we do not incorporate noise into the derivation for the moment, which will be later included in Section 9.3.3. Thus, we define a *noiseless non-delta scene operator* $\hat{T}_p$ as

$$\hat{T}_p = |s_p\rangle\langle l_p|. \quad (9.3)$$

Without loss of generality, we attribute the overall system PSF, i.e., the combination of sensor PSF and emitter PSF, to $|l_p\rangle$. This equivalence is possible because the propagation operator $\hat{U}$ from Equation (9.1) is unitary (see Chapter A). Intuitively, this can be understood as attributing all electronic effects that affect the PSF to the emitter. This way, the two sensing kets $|s\rangle$ and $|s_p\rangle$ are related by the operator $\hat{P} = |l_p\rangle\langle l_p|$ (see Section 8.2.1.3 from Chapter 8) as

$$|s_p\rangle = \hat{P}|s\rangle = |l_p\rangle\langle l_p|s\rangle. \quad (9.4)$$

The operation of Equation (9.4) is an affine transformation, where we assume $\langle l_p|l_p\rangle = 1$ as the scale of the signal can be normalized. Therefore, the scene operators $\hat{T}_p$ can be thought of as different projections of the ideal scene operator $\hat{T}$ for the same hidden geometry. Recalling the unitary property of Equation (9.1), and the affine transformation of Equation (9.4), we can use the *spectral theorem* [12, 99] to establish the following

$$\langle l_p|\hat{T}_p|l_p\rangle = \langle l_p|\hat{T}|l_p\rangle, \quad (9.5)$$

where represents $|l_p\rangle$ the PSF of the system, but it could be any suitable illumination (or PSF) basis. In other words, Equation (9.5) tells us that the expected value of two unitary operators related through an affine transform is the same using suitable basis functions for both.

9.3.3 Noise in the scene operator.

In previous Section 9.3.2, we model the capture with PSF and its relation with the ideal capture, using their scene operators. In this section, we incorporate noise. This allows us to define a mathematical representation that relates the ideal noiseless capture and the effects of suboptimal instrument response derived from noise and PSF.

We define a ket $|s_{p\eta}\rangle$ which incorporates both the effect of the system PSF and the capture noise as:

$$|s_{p\eta}\rangle = |s_p\rangle + |r\rangle, \quad (9.6)$$

where $|r\rangle$ represents the noise reaching or induced by the sensor. The noise is defined with respect to the captured light $|s_p\rangle$ as

$$|r\rangle = \hat{N}|s_p\rangle, \quad (9.7)$$

where $\hat{N}$ is a noise operator that can be different for each capture.

Now, we define a *realistic scene operator*, incorporating the PSF and the noise, as

$$\begin{aligned} \hat{T}_{p\eta} &= |s_{p\eta}\rangle\langle l_p| = \\ &= (|s_p\rangle + |r\rangle)\langle l_p| = \\ &= |s_p\rangle\langle l_p| + |r\rangle\langle l_p| = \\ &= \hat{T}_p + \hat{\Delta}, \end{aligned} \quad (9.8)$$

where $\hat{\Delta} = |r\rangle\langle l_p|$ is the operator that represents the relation between the noise captured at the sensor and the illumination pulse. Note that our realistic scene operator $\hat{T}_{p\eta}$ is a result of an addition between the noise operator and the noiseless non-delta scene operator $\hat{T}_p$ from Equation (9.3). Our goal in following Section 9.3.4 is to relate our realistic capture to the ideal one, using their associated operators.

9.3.4 Estimating ideal capture from suboptimal capture.

In this section, we define the relationship between the ideal scene operator and the realistic scene operator that accounts for both the PSF (Section 9.3.2) and noise (Section 9.3.3). For that purpose, we focus on making the realistic capture dependent on the ideal capture, rather than the noiseless non-delta capture. We can apply the spectral theorem similarly to Equation (9.5) with our realistic operator $\hat{T}_{p\eta}$ from Equation (9.10). Thus, we employ $|l_p\rangle$ as a suitable basis, and establish that

$$\begin{aligned} \langle l_p | \hat{T}_{p\eta} | l_p \rangle &= \langle l_p | \hat{T}_p | l_p \rangle + \langle l_p | \hat{\Delta} | l_p \rangle = \\ &= \langle l_p | \hat{T} | l_p \rangle + \langle l_p | \hat{\Delta} | l_p \rangle. \end{aligned} \quad (9.9)$$

We now expand $\langle l_p | \hat{\Delta} | l_p \rangle$ using Equation (9.7) to substitute $|r\rangle$, which yields

$$\begin{aligned} \langle l_p | \hat{\Delta} | l_p \rangle &= \langle l_p | r \rangle \langle l_p | l_p \rangle = \\ &= \langle l_p | \hat{N} | s_p \rangle \langle l_p | l_p \rangle = \\ &= \langle l_p | \hat{N} | s_p \rangle, \end{aligned} \quad (9.10)$$

in which we consider that $\langle l_p | l_p \rangle = 1$ similarly as we do in Section 9.3.2. We factor in Equation (9.4) to substitute $|s_p\rangle$ in Equation (9.10), thus

$$\langle l_p | \hat{\Delta} | l_p \rangle = \langle l_p | \hat{N} | l_p \rangle \langle l_p | s \rangle. \quad (9.11)$$

We can now define a relationship between the ideal and realistic captures and operators, by plugging Equation (9.11) into Equation (9.5):

$$\langle l_p | \hat{T}_{p\eta} | l_p \rangle = \langle l_p | \hat{T} | l_p \rangle + \langle l_p | \hat{N} | l_p \rangle \langle l_p | s \rangle, \quad (9.12)$$

where we can expand $\hat{T}_{p\eta}$ and $\hat{T}$ (the ideal delta operator) using Equation (9.2) and Equation (9.8), respectively. Then, rearranging terms, Equation (9.12) becomes

$$\boxed{\langle l_p | l \rangle \langle s | l_p \rangle + \langle l_p | \hat{N} | l_p \rangle \langle l_p | s \rangle - \langle l_p | s_{p\eta} \rangle = 0.} \quad (9.13)$$

which relates the ideal PSF $|l\rangle$ and the ideal noiseless capture $|s\rangle$ with the realistic capture $|s_{p\eta}\rangle$, i.e., noisy and under the PSF effects. Our target is to estimate the ideal capture $|s\rangle$. For that purpose, we assume we have prior knowledge of the system PSF, formed by the illumination and sensor PSFs, as the devices are

physically accessible. Recall that, as we mentioned before in Section 9.3.2, we attribute the system PSF to the illumination ket $|l\rangle$, for mathematical tractability, and we assume the sensor is perfect. We also assume that we can estimate a mean noise power density from the sensors.

Once we know the system PSF as well as the mean noise power density, we can obtain the ideal, noiseless impulse response $|s\rangle$, free of both PSF and capture noise, by solving Equation (9.13). We then use $|s\rangle$ for applying the specific NLOS technique, replacing $|s_{p\eta}\rangle$ as input, implicitly removing both noise and the PSF from the NLOS solution.

This equation is general, and serves for any basis that represents the signals. For its application, we need to specify this basis. In the following Section 9.3.5, we propose a specific solution for the frequency domain, although other solutions are possible, also for other bases.

9.3.5 Ideal estimation in the frequency domain.

This subsection presents a solution example for our Equation (9.13) in the frequency domain, which we will later utilize in our results in Section 9.4. We define

$$|s_{p\eta}\rangle = \int d\omega S_{p\eta}(\omega) |\omega\rangle \quad |l_p\rangle = \int d\omega L_p(\omega) |\omega\rangle, \quad (9.14)$$

where $S_{p\eta}(\omega)$ and $L_p(\omega)$ denote the realistic sensing and arbitrary illumination values, respectively (in this case, the illumination is the system PSF), for a frequency ω in the frequency domain. Equivalently, the definition of the ideal capture kets in the frequency domain is

$$|s\rangle = \int d\omega S(\omega) |\omega\rangle \quad |l\rangle = L_0 \int d\omega |\omega\rangle, \quad (9.15)$$

where $|l\rangle$ is the frequency representation of a delta illumination pulse in the time domain at $t = 0$ of amplitude L_0 , and $S(\omega)$ is the ideal sensing value of the hidden scene at the frequency ω .

Substituting Equation (9.14) and Equation (9.15) in Equation (9.13), and developing as we describe in Chapter B, it yields

$$\int d\omega S(\omega) [L_p(\omega)^2 + N(\omega)L_p^*(\omega)] - L_p^*(\omega)S_{p\eta}(\omega) = 0, \quad (9.16)$$

where $N(\omega)$ is the mean density spectrum of the noise and $L_p^*(\omega)$ is the complex conjugate of $L_p(\omega)$.

Recall that, in this work, we aim to estimate the ideal impulse-response function $S(\omega)$ from the original captured function $S_{p\eta}(\omega)$, which is coupled with the system PSF $L_p(\omega)$ and the noise $N(\omega)$. **Many solutions are possible as long as the result of the integral in Equation (9.16) is zero.** Different solutions would yield new, different methods for processing the captured signal, suitable for a more general case.

Notice, however, that in a particular and ideal case, where the addition of the system PSF and noise is non-zero for all frequencies, we can solve Equation (9.16) by setting the integrand equal to zero for all frequencies, which becomes a Wiener deconvolution:

$$S(\omega) = \frac{L_p^*(\omega)}{|L_p(\omega)|^2 + L_p^*(\omega)N(\omega)} S_{p\eta}(\omega). \quad (9.17)$$

We employ this specific solution for the rest of the manuscript. In practice, we make use of a regularization term in the denominator of Equation (9.17) ($\epsilon \sim 10^{-8}$) to define a complete domain of the solution, i.e., to avoid divisions by zero.

9.4 RESULTS

In this section, we present our experimental validation. We show how our method estimates an ideal capture, removing the instrument response effects, comprehending both system PSF (Section 9.4.1) and

noise (Section 9.4.2), and its improvement in terms of quality for the reconstructions. Additionally, we also show the capabilities of our method with any arbitrary shape of the PSF, e.g., from an arbitrary illumination (Section 9.4.3). To demonstrate that our method is general enough to provide clear benefits for a variety of previous methods, we apply three of the most common NLOS algorithms: filtered backprojection [115], fk-migration [66], and Phasor Fields [69].

9.4.1 Removing the point-spread function.

Here, we demonstrate how our method yields a better estimation of the ideal capture by decoupling the effect of the PSF of the system in a noiseless scenario. We use a light transport simulator optimized for transient NLOS [96, 95] to synthesize three different scenes: *Two Planes*, *USAF*, and *Dragon*. These simulations consist of 128×128 capture points on a 2×2 meter relay wall, with a single illuminated point at the center (non-confocal configuration), placing the elements of the hidden scenes at one meter from the relay wall.

The PSF of the NLOS systems usually follows a Gaussian distribution in time, followed by an exponential tail, which is usually negligible compared to the Gaussian contribution. Thus, we define the system PSF as a Gaussian distribution whose full width at half maximum establishes its temporal width. We apply our technique, as described in Section 9.3.5, for simulated signals with three widths for the system PSF: 314 ps, 628 ps, and 1225 ps. For reference, common NLOS captures have a PSF width in the range of tens of picoseconds.

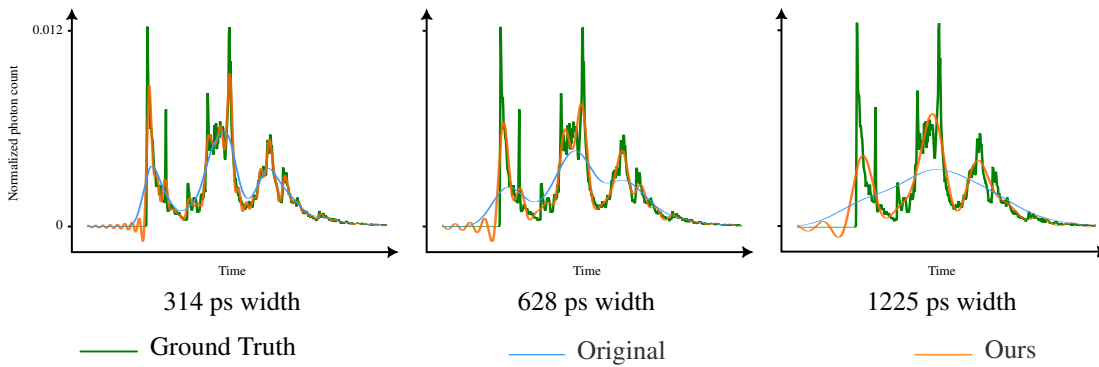


Figure 9.2: Ground truth, original, and ours histograms for an emitter-sensor pair, simulated from the *Dragon* scene under system PSFs of increasing widths: 314 ps, 628 ps, and 1225 ps. Ours (orange) is capable of correctly recovering the main peaks of the histograms from the original data (blue), when compared to the ground truth (green), even in the extreme case of a 1225-ps wide PSF.

	314 ps	628 ps	1225 ps
Original	11.00	8.60	6.30
Ours	17.50	14.09	10.99

Table 9.1: Scale invariant signal to distortion ratio of the original signal versus ours, compared with the ground truth (the higher the better).

In Figure 9.2, we show the histogram comparison of a capture of the *Dragon* scene for each PSF width. We compare three signals: the original capture (blue), our method’s estimation of the ideal capture (orange), and the ground truth (green). Our method correctly recovers the general shape of the ideal capture from the original data if compared to the ground truth. This is visible even in the extreme case of a 1225 ps width pulse, where big uncertainties in the original data result in a single rounded peak, from which we can recover the general shape of the ideal capture. We present a quantitative analysis of the same experiment in Table 9.1, using the scale-invariant signal-to-distortion ratio (SISDR) [63] against our simulated reference. In all cases, our histograms achieve a higher SISDR, meaning ours present more significant data.

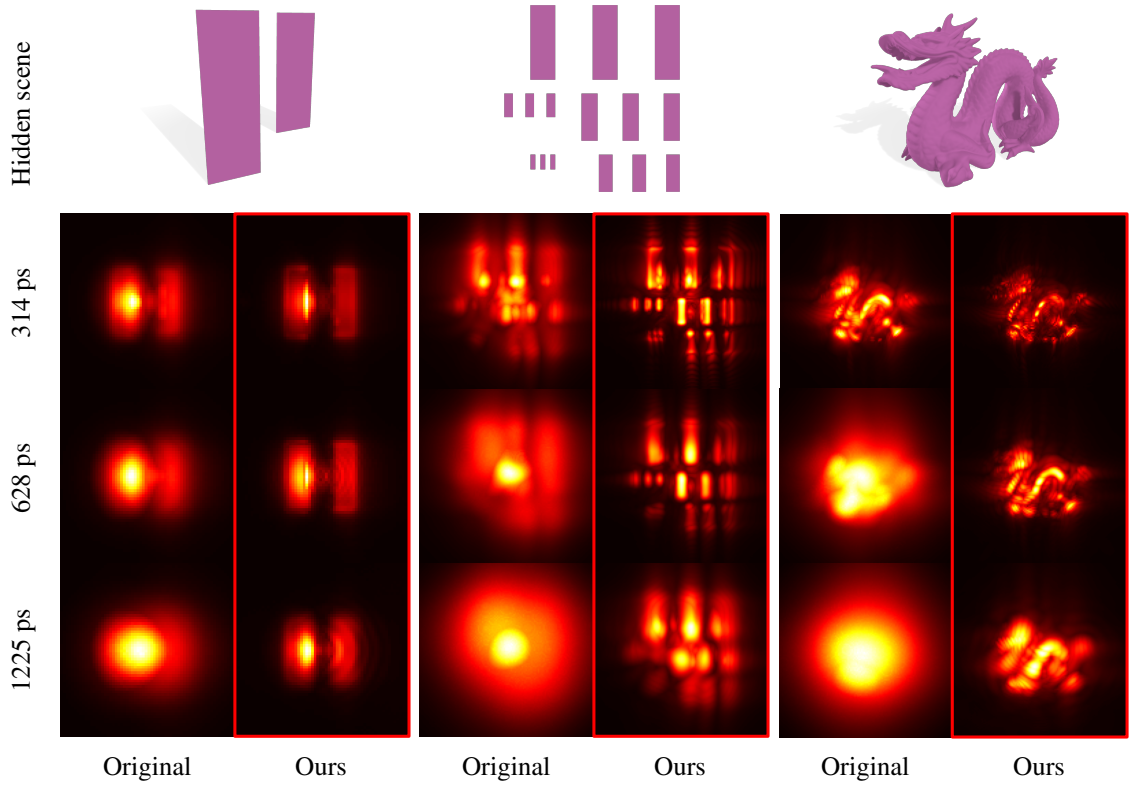


Figure 9.3: Our method’s effect on reconstructions of filtered backprojection. The scenes are simulated using system PSFs of increasing width: 314 ps, 628 ps, and 1225 ps. From left to right: *two planes*, *USAF*, and *Dragon* scenes. Imaging with the original data results in low-lateral resolution on increasing widths of the PSF, whereas our method allows for recovering the hidden-scene features in all cases.

We also present the effect of increasing the PSF width and the impact of our method on NLOS reconstruction methods. Specifically, we employ filtered backprojection in Figure 9.3 under the three widths of the PSF for both the original data and ours. Decreasing temporal resolution associated with an increase in the width of the PSF corresponds to a loss in lateral resolution in the original. Yet, our method retrieves sharper results from the hidden scene in all cases. This is especially noticeable comparing ours with the 628-ps and 1225-ps wide PSF, in the *USAF* and *Dragon* scenes, where the original data yields formless blobs when imaging. Meanwhile, our approach successfully recovers temporal information associated to hidden scene features. It is worth mentioning that the uncertainty derived from the width of the PSF is three orders of magnitude higher than common illumination devices used in NLOS captures, shifting from illumination at a range of picoseconds to nanoseconds.

Finally, as our proposal is general for any NLOS imaging approach, we show the effects of our method using the combined PSF of 628-ps width in fk-migration and Phasor Fields with the *Dragon* in Figure 9.4. Fk-migration requires a confocal configuration, hence we maintain the scene’s setup, but shifting the capture to a confocal configuration of 128x128 points. We observe a clear enhancement of the hidden scene in both cases, reducing the out-of-focus artifacts in Phasor Fields results and increasing lateral resolution for fk-migration.

9.4.2 Including photon-loss noise.

NLOS captures have different sources of noise on top of the PSF of the system. To analyze the effect of our technique, we simulate noise into the previously simulated data. Specifically, we employ a combined PSF of 314 ps width for the same three scenes while introducing photon loss noise. Since photon captures follow a Poisson process [25], we simulate the noise of the captured photons using a Poisson distribution [45]. We

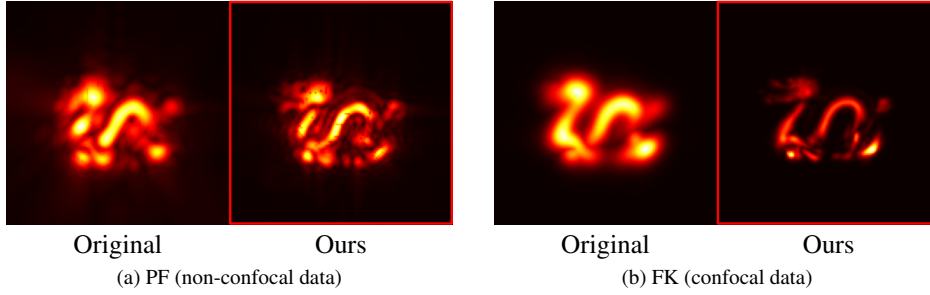


Figure 9.4: Our method’s effect on two different NLOS imaging methods, Phasor Fields and fk-migration, on the *Dragon* scene with a system PSF of 628-ps width. For fk-migration, data is captured in a confocal configuration. The wider pulse appears as out-of-focus artifacts in the Phasor Fields result and as a lower lateral resolution in fk-migration, where ours gains sufficient temporal resolution to improve the sharpness of the results in both cases.

estimate the expected number of events at a specific moment by photons from the clean simulated signal, reduced by the loss percentage of the expected noise. We specifically simulate the 15%, 60%, and 90% loss of photons. This is equivalent to reducing the capture time in a real-world setup [7]. Figure 9.5 shows the histograms of the original capture (blue), and the decoupling using our method (orange), compared to the ground truth (green). Our approach recovers a similar histogram shape to the ground truth, even losing 90% of the photons, keeping the same peak positioning. That is fundamental for proper imaging of the hidden scene. Quantitatively, we obtain similar results in Table 9.2, where we achieve higher SISDR in all cases with our method, compared to the original data.

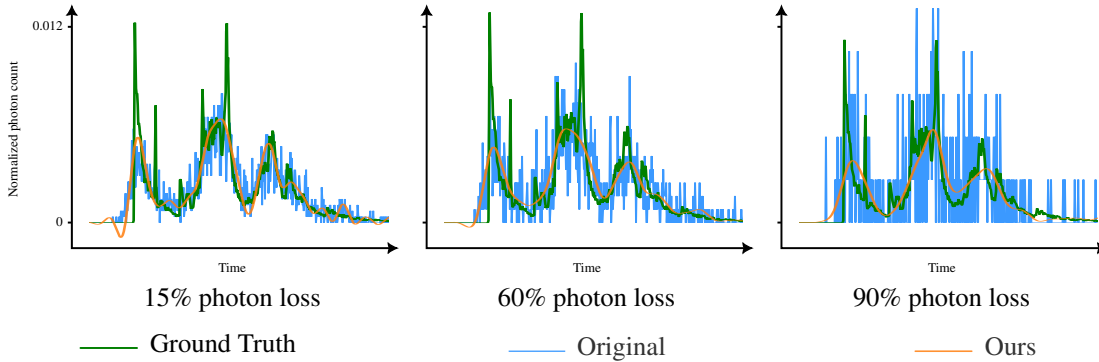


Figure 9.5: Ground truth, original capture, and ours histograms of a single capture of a pair emitter-sensor from the *Dragon*. The system PSF is 314-ps wide, and it presents noise of losing 15%, 40%, and 90% of the photons. While the noise of the original capture (blue) reduces information about impulse response shape, ours (orange) is capable of correctly recovering the histogram main peaks, when compared to the ground truth (green), even in the case of 90% loss.

	15%	60%	90%
Original	9.61	8.50	5.05
Ours	12.02	10.82	10.14

Table 9.2: Scale invariant signal to distortion ratio of the original data versus ours, compared with the ground truth (the higher the better), on a system with a PSF of 314-ps width, and simulating the noise of different photon losses (columns).

The photon-loss effects are also visible when imaging, as Figure 9.6 shows, where we employ filtered backprojection for the reconstruction. In the original data, the signal degrades as photon loss increases, leading to reduced lateral resolution and high-frequency noise in the background. Our method mitigates

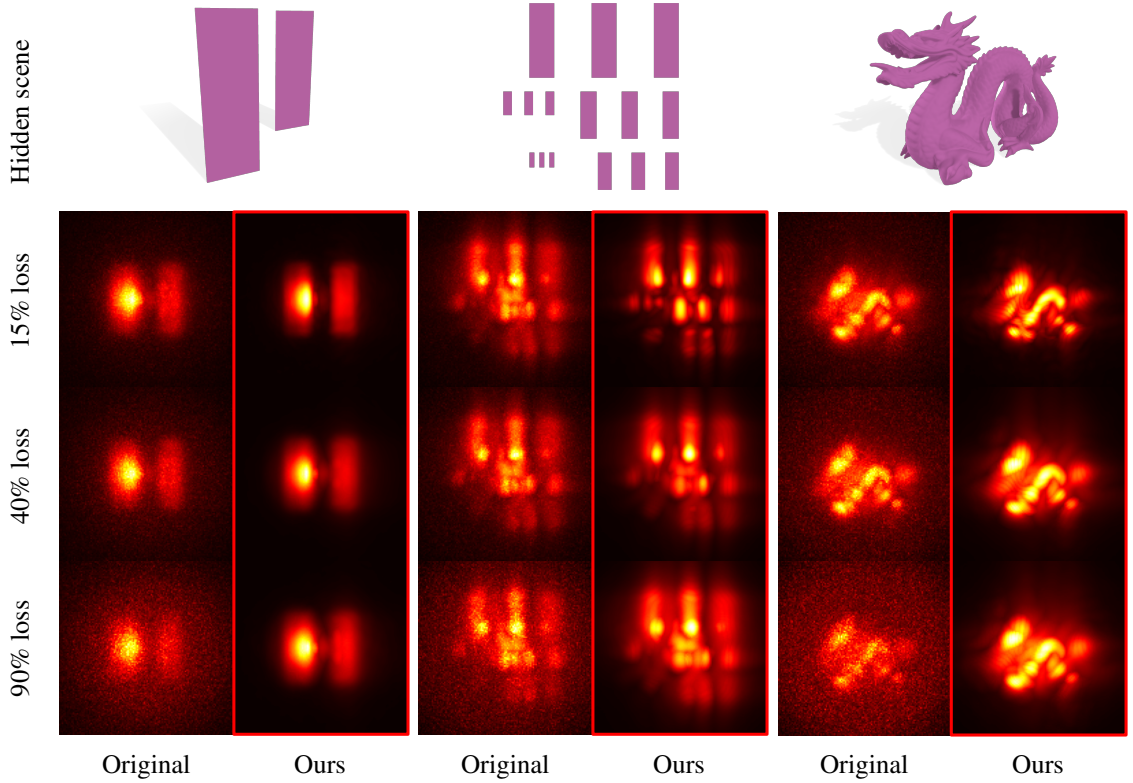


Figure 9.6: Our method’s effect on reconstructions of filtered backprojection with incrementing photon-loss noise. From left to right: *two planes*, *USAF*, and *Dragon* scenes. The system PSF of the capture is 314-ps wide, and we simulated a photon loss of 15%, 40%, and 90% (Poisson noise). The original capture presents high-frequency noise in the reconstructions at increasing loss, consequently affecting the image resolution. Our method mitigates such noise and increases the visible features.

the background noise, blurring some of the edges in the most extreme scenarios, yet increasing the visible features, e.g., some charts of the USAF and the general shape of the dragon. This is also visible on other imaging algorithms, as we show in Figure 9.7 using the *Dragon* scene with a 90% photon loss using Phasor fields and fk-migration. In both cases, high-frequency noise limits the dragon’s visibility, preventing Phasor Fields and fk-migration to image any features from the original data. In contrast, our method allows hidden scene features to appear again.

We additionally apply our method to real-world captured data from the publicly available dataset of the fk-migration work [66]. The data we use comes from a discobulus statue next to a table with a flowerpot, captured outdoors. The data is captured with a confocal configuration, consisting of 512x512 capture and illumination points distributed on a 2x2 meter grid. We compare a long-time capture of 30 minutes and a short one of 10 minutes, the second being even more challenging due to lower SNR. We estimate that this corresponds to photon loss of 76% and 92%, for the long-time and short-time captures, respectively. Furthermore, outdoor ambient lighting increases the capture noise. The system PSF of the capture is 70-ps wide.

We show the impact of our method on fk-migration reconstructions in Figure 9.8 and compare it with the original data. We observe that ours produces cleaner results, diminishing part of the high-frequency noise that appears in the reconstructions of the original captures. The high-frequency noise occludes the hidden scene until it completely hides the features in the case of 92% of photon loss in the original data, whereas our method removes most of the noise while enhancing the visibility of the hidden scene features.

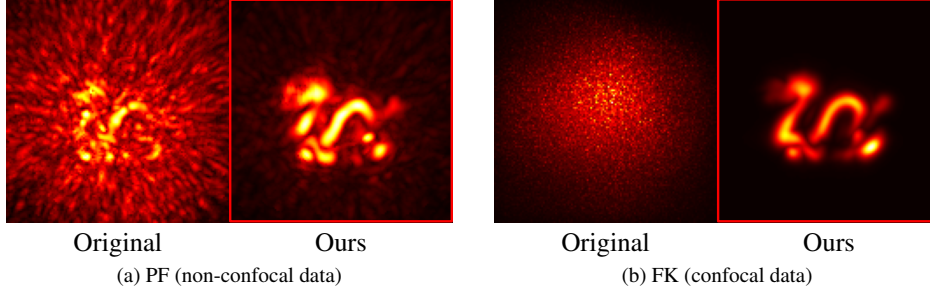


Figure 9.7: Our method’s effect in noisy captures to different NLOS imaging methods: Phasor Fields and fk-migration. We employed the *Dragon* scene with a system with a PSF 314-ps wide, and a 90% of photon loss. Using the raw data, the photon loss (low SNR) yields high-frequency noise, making both algorithms fail at imaging. Our method successfully decouples the ideal signal, enabling both Phasor Fields and fk-migration to image correctly.

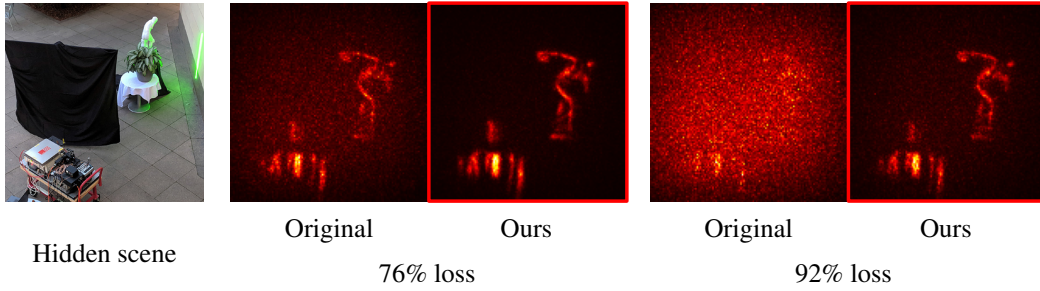


Figure 9.8: Effect of our method on fk-migration reconstructions using publicly available data [66]. This real capture was taken outdoors, and it presents a system PSF of 70-ps width. We employ two cases of 76% and 92% of photon loss (different capture times). Our method allows us to recover the ideal capture, reducing high-frequency noise in the results and enhancing the hidden scene features, even in the challenging case of 92% of photon loss with low SNR.

9.4.3 Arbitrary illumination reconstruction: train of delta pulses.

In our previous experiments, we considered Gaussian illumination pulses of varying width for the system PSF, as they are common PSFs for lasers and SPAD sensors and are associated with the excitation and the decay times of the materials. However, our technique can account for any known PSF shape. In this subsection, we demonstrate the decoupling of an ideal capture using our method with an arbitrary illumination pulse, i.e., the PSF shape depends only on the illumination. Specifically, we simulate the *Dragon* scene and show the results of capturing a noiseless scene using a train of 40 delta pulses, each emitted a few ps apart. This could be equivalent to employing a laser with a frequency much higher than the sensor’s, which can increase the SNR or reduce the capture time. We show the results of our technique by applying filtered backprojection to both the original data and our data in Section 9.4.3. Although imaging with the original signal fails, our approach allows us to recover the shape of the hidden scene and image the overall dragon’s features. This result hints at the possibility of imaging systems in which the laser and the sensor are not necessarily synchronized.

9.5 DISCUSSION

Our main contribution is a mathematical model based on the Dirac notation that allows us to remove the artifacts and undesired effects of the suboptimal instrument response in a capture system in non-line-of-sight imaging. Specifically, we decouple the PSF of the optical system, comprising the combination of the PSFs of the illumination and sensor, and the noise caused by the loss of photons with respect

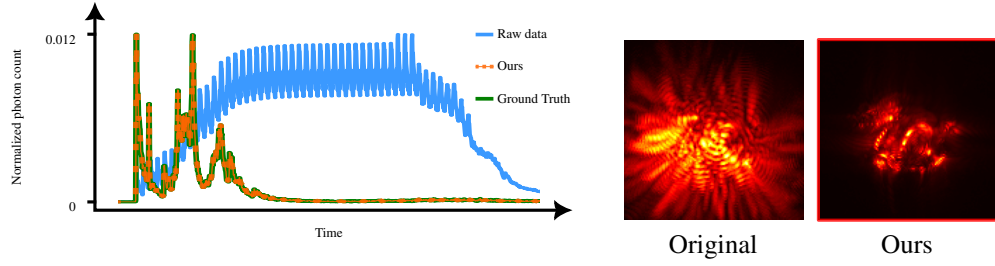


Figure 9.9: *Dragon* scene whose PSF is an illumination formed by a consecutive train of delta functions. Left to right: Histograms of the *Dragon* scene, its effect on the filtered backprojection image, and its effect in the image after using our method. While the train of deltas incorporates temporal uncertainty into the histogram and reconstruction due to de-synchronization, our method recovers the original signal indistinguishable from the ground truth and, therefore, the dragon’s shape in the reconstruction.

to the ideal capture. Our results demonstrate that our method’s solutions enhance the quality of NLOS reconstructions regardless of the imaging method employed. Our method increases the spatial resolution of the reconstructions by increasing the temporal resolution of the capture. We showcase this in our results, where spatial uncertainty in our reconstructions appears in the wide PSFs due to manually accounting for divisions by zero.

Capturing NLOS scenes demands specialized hardware that is often expensive (due to the need to operate in the picosecond temporal range) and that requires precise calibration and long capture times to increase the number of captured photons. Our method supports the modeling of any arbitrary shape of a system PSF, including any illumination pulse shape, and allows reducing the precise calibration and synchronization requirements of any transient NLOS capture. We demonstrate this in Section 9.4.3, showing the reconstruction result using a train of deltas. However, any illumination shape could be employed.

In general, our work facilitates the use of lower-end, less expensive devices and reduces capture times by computationally increasing the temporal resolution and decreasing the effects of the noise from photon loss. We expect future endeavors to focus on novel capture methods, e.g., based on coded illuminations on the relay wall for compressed sensing, which may further reduce capture time, power demands, and shape the pulse and the duration of the illumination, thereby mitigating various noise effects. Future work may focus on different solutions for a general case and further increase the improvements we present in this work. Therefore, it may devise novel and robust processing algorithms for the NLOS captures. Thus, our work is a step forward in applying NLOS imaging in real-world contexts, such as autonomous driving, medical imaging, or search-and-rescue operations.

CONCLUSION.

DISCUSSION AND FUTURE WORK.

10.1 ENGLISH VERSION.

This thesis contains several contributions, described in three complementary directions: (i) relaxing physical assumptions for NLOS reconstructions, (ii) improving the efficiency of NLOS reconstructions, and (iii) enabling NLOS imaging under suboptimal conditions and capture devices using a cohesive formulation. Although these contributions appear separately, they ultimately pursue the practical applicability of NLOS in everyday technology. In that regard, we present a set of physical and computational tools that offer a more robust and faster model for the NLOS reconstruction and the capture process. Each chapter offers a specific discussion on the topic addressed there. Here, we present a holistic summary of the main ideas and outline some of the key challenges that remain in the NLOS imaging field.

A persistent limitation in existing NLOS imaging approaches is a lack of robustness in scenarios common in the real world. Through this thesis, we argue that the causes are simplifying assumptions and ad hoc derivations due to problem tractability and/or computational convenience. Our contributions generalize NLOS to more demanding environments. In Chapters 4 and 5, we surpass the assumptions of light traveling in vacuum and of planar relay walls, respectively. This enables NLOS in dense scattering media such as fog, smoke, or dust. It also allows reconstruction using non-planar relay walls, maintaining temporal efficiency. Moreover, leveraging the Dirac notation from a collaborative work described in Chapter 8, we can employ a compact representation for representing the correct light transport that would be otherwise intractable. We show this in Chapter 9, where we specifically model capture effects such as noise and hardware point-spread functions using the notation. Consequently, our modeling improves the robustness of the capture process and enables the use of lower-end devices in NLOS imaging, as we show. In the future, general models may not be restricted to specific relay wall setups, lighting configurations, or sophisticated hardware, providing an interesting avenue for practical NLOS imaging in the wild.

Another major obstacle on the path to NLOS applicability is scalability. Reconstruction accuracy is closely related to computational resources. Furthermore, the lack of prior knowledge of the hidden scene often forces uniform high-resolution reconstruction across large volumes. In Chapter 6, we address this by proposing a novel reconstruction method that trades off quality and memory consumption, reducing memory usage by up to a factor of 16. Therefore, we enable the use of memory-constrained devices, generating an alternative to reduce computation in empty or irrelevant regions of the scene. Chapter 7 describes another reconstruction algorithm from a different perspective. This algorithm identifies micrometer depth variations, given a centimeter-scale voxelization for reconstruction. Furthermore, if we factor in the reduced capture requirements discussed in Chapter 9, these contributions suggest a path forward to real-time NLOS imaging. In the near future, we may expect substantially lower demands for computational and optical devices for NLOS imaging. Our discussed goal is computational and capture efficiency across these works. Still, our methodology involves understanding the physics of real wave optics to provide insights into the virtual wave optics world. Delving further into physics would provide broader insights, allowing to formulate novel and efficient NLOS imaging techniques. For example, future researchers' endeavors may draw inspiration from fields such as holography to retrieve more features from the hidden scene using the virtual phase. Another alternative may involve the shaping of the virtual illumination, gaining access to information that current methods lose, and diminishing the missing-cone effects.

Overall, these contributions push the boundaries of NLOS imaging in terms of robustness, scalability, modeling fidelity, and practical viability. The outcome paves the way for future researchers to establish more complex and varied setups and applications for NLOS imaging. As the field progresses, we anticipate

several future research directions. First, the Dirac notation proposal could provide a standard in the NLOS field. The notation may represent a learning barrier for the researchers accustomed to previous NLOS notations. However, they may benefit from embracing the generalization and interoperability that our notation offers. Second, current implementations require *a priori* known target resolution and spatial boundaries of the reconstruction, where scene characteristics are unknown by definition. Our Zero-Phase Phasor Fields approach already relaxes these requirements, pointing toward this direction by providing finer grain in detecting the hidden scene details than the voxelization used. In the short term, further work may adapt our approach to non-confocal setups. In the long run, reconstruction algorithms may evolve beyond current voxel grid implementations. Future works may completely forget about voxelization, estimating the whole geometry of the hidden scene in an efficient time, achieving consistently high quality.

A natural extension of this work is the development of adaptive methods, where capture and reconstruction of the hidden scene are jointly optimized. Such methods could focus resolution and, thus, computational resources, on the most relevant parts of the hidden scene while maintaining real-time performance. Implementations would require efficient updates to the reconstruction as additional data becomes available, where intermediate reconstructions guide subsequent acquisition. Our Chapter 5 work may assist in that objective by combining different captures of the same hidden scene, which could be complemented with related work for reconstructing non-regular captures [105]. To guide the use of capture and computational resources, future work may rely on our Zero-Phase Phasor Fields and our simplification based on Fresnel zone plates (Chapters 6 and 7, respectively). These approaches provide efficient solutions to identify areas of interest within the hidden scene, relaxing the computational demands. Furthermore, advances in sensing hardware, such as physically modulating the illumination, may work together with computational tools, as in our Chapter 9 work. Similar to recent works [82], this could improve captures in challenging conditions for NLOS imaging. Altogether, the relationship between hardware and computation may offer approaches for adaptive temporal resolution.

Finally, while this thesis focuses on NLOS scenarios, our contributions may have broader implications. Our proposals involving scattering media in the hidden scene (Chapter 4) and multiple relay-wall propagation (Chapter 5) may inspire parallel imaging ideas in scenarios where the scene is in the line of sight. Examples of this might be imaging through bodies using non-ionizing radiation, or employing higher-order light propagation for material analysis in visible scenes. These applications would employ transient imaging, and they may also benefit from Chapter 9 work and its capture enhancement, as scattering events or the increment of capture surfaces may reduce the affordable time and/or resources for capturing. Furthermore, other applications may draw inspiration from the virtual wave optics abstraction. Applications tailored for the identification of sharp and precise variations may approximate similarly to our virtual-phase usage for Zero-Phase Phasor Fields (Chapter 7).

In conclusion, this thesis addresses key challenges that currently limit the real-world applicability of NLOS imaging. Although the field is still in an early stage, the results presented here provide a set of ideas related to generalization, efficiency, modeling, and adaptability. We hope these contributions inspire future work that makes NLOS imaging more accurate, faster, and more accessible, ultimately enabling its transition from laboratory research to everyday technologies.

In the following Section 10.2, this discussion and future work appears translated to Spanish.

PERSONAL CONCLUSIONS

The next lines treat about considerations that affect me personally regarding this thesis completion. Focusing on a technical aspect, during this thesis, I have acquired the tools to read, study, and research a general topic through specific publications. This information is often complex and requires several layers of abstraction. The task is not trivial, and this maturity and knowledge have allowed me to witness first-hand the advances in cutting-edge innovation. Moreover, I have had the opportunity to advance the innovation with the content described in this thesis. Another set of abilities complements the technical ones, allowing me to express myself. During these past years, I learned that attention to details, wording, and structure are crucial to any communication. I am proud to say that now I am capable of expressing complex ideas, concepts, and results, pitched for a specific audience. In contrast to the beginning of the thesis, I am aware of the importance of these skills: "ideas that only live in your head cost nothing". All this allowed me to contact

researchers around the globe whom I admire, or whom I did not know until that moment. I am trying not to repeat myself as I have already written my acknowledgments, but none of this could have happened without my supervisors and colleagues. Let me finish by saying that I will fondly remember these past four and a half years.

10.2 VERSIÓN EN ESPAÑOL.

En esta sección, se incluyen las mismas conclusiones que las presentadas en el Apartado 10.1, salvo que traducidas al español.

Esta tesis contiene varias contribuciones, descritas en tres direcciones complementarias: (i) la relajación de suposiciones físicas para las reconstrucciones sin línea de visión (NLOS, de sus siglas en inglés), (ii) la mejora de la eficiencia de las mejoras NLOS, y (iii) la habilitación de la imagen NLOS bajo condiciones y dispositivos subóptimos con una formulación cohesiva. Aunque aparecen por separado, estas contribuciones persiguen en última instancia la aplicabilidad práctica de NLOS en la tecnología cotidiana. En ese sentido, presentamos un conjunto de herramientas físicas y computacionales que ofrecen un modelo más robusto y rápido para la reconstrucción NLOS y su proceso de captura. Cada capítulo ofrece una discusión específica sobre el trabajo allí detallado. Aquí, presentamos un resumen holístico de las ideas principales y describimos algunos de los desafíos clave que permanecen dentro del campo de la imagen NLOS.

Una limitación persistente en los enfoques de imagen NLOS es la falta de robustez en escenarios comunes en el mundo real. A lo largo de esta tesis, argumentamos que las causas son asunciones simplificadoras y derivaciones *ad hoc* debidas a la tractabilidad del problema y/o conveniencia computacional. Nuestras contribuciones generalizan NLOS a entornos más exigentes. En Capítulos 4 y 5, superamos las suposiciones de que la luz viaja en el vacío y que las superficies auxiliares son planas, respectivamente. Esto habilita NLOS en medios participativos densos, como niebla, humo o polvo. También, permite la reconstrucción utilizando superficies auxiliares no planares, manteniendo la eficacia temporal. Además, aprovechando la notación de Dirac en un trabajo colaborativo y presentada en Capítulo 8, podemos emplear una representación compacta para una representación correcta transporte de luz que de otra forma sería intratable. Mostramos esto en Capítulo 9, donde modelamos específicamente efectos de captura como el ruido y las funciones temporales propias de efectos del hardware utilizando la notación. En consecuencia, nuestro modelado mejora la robustez del proceso de captura y permite el uso de dispositivos de gama baja en imagen NLOS, como mostramos. En el futuro, los modelos generales podrían no estar restringidos a superficies secundarias específicas, configuraciones de iluminación o hardware sofisticado, proporcionando una vía interesante para la imagen NLOS práctica en entornos reales.

Otro obstáculo mayor en el camino hacia la aplicabilidad de NLOS es su escalabilidad. La precisión de las reconstrucciones está estrechamente relacionada con los recursos computacionales. Además, la falta de conocimiento previo de la escena oculta a menudo fuerza a una reconstrucción uniforme de alta resolución. En Capítulo 6, abordamos esto proponiendo un nuevo método de reconstrucción que ofrece un compromiso entre calidad y consumo de memoria, reduciendo el uso de memoria hasta en un factor de 16. Por lo tanto, habilitamos el uso de dispositivos con memoria limitada, generando una alternativa para reducir la computación en regiones vacías o irrelevantes de la escena. Capítulo 7 describe otro algoritmo de reconstrucción con una perspectiva diferente. Este algoritmo logra la identificación de variaciones micrométricas en la profundidad, dada una voxelización a escala centimétrica para la reconstrucción. Además, si consideramos los requisitos de captura reducidos que discutimos en Capítulo 9, estas contribuciones sugieren un camino hacia la imagen NLOS en tiempo real. En el futuro cercano, podemos esperar demandas sustancialmente menores para los dispositivos computacionales y ópticos para imagen NLOS.

Nuestro objetivo discutido es la eficiencia computacional y de captura a través de estos trabajos. Aun así, nuestra metodología involucra el entendimiento de la física de óptica de ondas para proporcionar contribuciones dentro del mundo de óptica de ondas virtuales. Profundizar más en la física proporcionaría perspectivas más amplias, permitiendo formular técnicas de imagen NLOS más novedosas y eficientes. Por ejemplo, los esfuerzos de futuros investigadores pueden inspirarse en campos como la holografía para

recuperar más detalles de la escena oculta utilizando la fase virtual. Otra alternativa puede implicar el modelado de la iluminación virtual, ganando acceso a información que los métodos actuales pierden, y disminuyendo el efecto del cono faltante.

En su conjunto, estas contribuciones expanden los límites de la imagen NLOS en términos de robustez, escalabilidad, fidelidad de modelado y viabilidad práctica. El resultado allana el camino para que futuros investigadores establezcan configuraciones y aplicaciones más complejas y variadas para la imagen NLOS. A medida que el campo progresa, anticipamos varias direcciones de investigación futura. Primero, la propuesta de la notación de Dirac podría proporcionar un estándar en el campo NLOS. La notación puede representar una barrera de aprendizaje para los investigadores acostumbrados a anotaciones NLOS anteriores. Sin embargo, pueden beneficiarse al adoptar la generalización e interoperabilidad que nuestra notación ofrece. Segundo, las implementaciones actuales requieren conocer la resolución objetivo y límites espaciales de la reconstrucción *a priori*, donde las características de la escena oculta son desconocidas por definición. Nuestro enfoque *Zero-Phase Phasor Fields* Capítulo 7 ya relaja estos requisitos, apuntando hacia esta dirección al proporcionar la detección de detalles de la escena oculta con un grano más fino que la voxelización utilizada. A corto plazo, trabajos adicionales pueden adaptar nuestro enfoque a configuraciones no confocales. A un plazo mayor, los algoritmos de reconstrucción pueden evolucionar más allá de las actuales implementaciones basadas en rejillas de voxelización. Los trabajos futuros pueden olvidar completamente la voxelización, estimando toda la geometría de la escena oculta en un tiempo eficiente, logrando una alta calidad constante.

Una extensión natural de este trabajo es el desarrollo de métodos adaptativos, donde la captura y la reconstrucción de la escena oculta se optimizan conjuntamente. Tales métodos podrían enfocar la resolución y, por ende, recursos computacionales, en las zonas más relevantes de la escena oculta mientras mantienen un rendimiento en tiempo real. Nuestro trabajo de Capítulo 5 puede asistir en dicho objetivo combinando diferentes capturas de la escena oculta, que podría ser complementado con trabajos relacionados consistentes en reconstruir capturas no regulares [105]. Para guiar el uso de captura y recursos computacionales, trabajos futuros podrían emplear nuestros trabajos nuestro *Zero-Phase Phasor Fields* o nuestra simplificación basada en las placas de zona de Fresnel (Capítulos 6 y 7, respectivamente). Estas propuestas proveen soluciones eficientes para identificar áreas de interés dentro de la escena oculta, relajando las demandas computacionales. Además, avances en el hardware de detección, como modular físicamente la iluminación, podría trabajar junto a herramientas computacionales, como nuestro trabajo del Capítulo 9. Similar a trabajos recientes [82], esto podría mejorar las capturas en condiciones desafiantes para la imagen NLOS. En total, la relación entre hardware y computación puede ofrecer soluciones para resolución adaptativa en el tiempo.

Finalmente, aunque esta tesis se enfoca en escenarios NLOS, nuestras contribuciones pueden tener implicaciones más amplias. Nuestras propuestas que involucran medios participativos en la escena oculta (Capítulo 4) y la propagación de múltiples superficies secundarias (Capítulo 5) pueden inspirar ideas de imagen paralela en escenarios donde la escena se encuentra en la línea de visión directa. Ejemplos de esto podrían ser la formación de imágenes a través de cuerpos utilizando radiación no ionizante, o empleando la propagación de luz de orden superior para el análisis de materiales en escenas visibles. Estas aplicaciones emplearían la imagen transitoria, y podrían beneficiarse también del trabajo expuesto en Capítulo 9 y su mejora de captura, ya que los eventos de scattering o el incremento de superficies podrían reducir el tiempo y/o recursos asequibles para la captura. Además, otras aplicaciones pueden inspirarse en la abstracción de ondas virtuales. Aplicaciones adaptadas para la identificación de variaciones nítidas y precisas pueden aproximarse de manera similar a nuestro uso de la fase virtual para *Zero-Phase Phasor Fields* (Capítulo 7).

En conclusión, esta tesis aborda desafíos clave que limitan actualmente la aplicación de la imagen NLOS en el mundo real. Aunque el campo está todavía en una etapa temprana, los resultados aquí presentados proporcionan un conjunto de ideas relacionadas con la generalización, eficiencia, modelado y adaptabilidad. Esperamos que estas contribuciones inspiren trabajo futuro que hagan la imagen NLOS más precisa, más rápida, y más accesible, habilitando en última instancia su transición de la investigación de laboratorio a las tecnologías cotidianas.

APPENDIX

PROPAGATION OPERATORS ARE UNITARY.

We will demonstrate that any propagation operator $\hat{U}$ is a unitary operator. This property is key in demonstrating that backprojection and wave-based imaging methods are equivalent, which is demonstrated in Section 8.7. A unitary operator is a bounded linear operator on a Hilbert space that satisfies:

$$\hat{U}\hat{U}^\dagger = \mathbb{I} \quad (\text{A.1})$$

where $\hat{U}^\dagger$ is the adjoint operator of $\hat{U}$ and $\mathbb{I}$ is the identity operator. In other words, a unitary operator is an operator such that the adjoint operator is also the inverse operator, as it follows from Equation (A.1):

$$\hat{U}\hat{U}^\dagger = \mathbb{I} \quad \Rightarrow \quad \hat{U}^\dagger = \hat{U}^{-1} \quad (\text{A.2})$$

We will demonstrate that point-to-point propagation operators in Equation (8.11) are unitary operators. The propagation of light from and illumination point n in virtual light source surface $\mathcal{L}$ to a sensing point m in virtual camera $\mathcal{S}$ is defined as:

$$\hat{T}_{m,n} = |s_m\rangle \langle l_n| \quad (\text{A.3})$$

where $|s\rangle$ represent the state of light at sensing point m and $\langle l|$ represent the state of light at illumination point n .

The adjoint of Equation (A.3) is:

$$\hat{T}_{m,n}^\dagger = (|s_m\rangle \langle l_n|)^\dagger = |l_n\rangle \langle s_m| \quad (\text{A.4})$$

In quantum mechanics, an observable represents and experimental measurement of operator $\hat{O}$. Observable O corresponding to operator $\hat{O}$ is calculated according to Born's rule, which states that the probability density of finding a system in a given state is proportional to the square of the amplitude of a system's wavefunction. In practical terms, the observable is the expected value of operator $\hat{O}$ using a suitable basis functions $\{|b_a\rangle\}$ representing the system's state:

$$O = \sum_a \langle b_a | \hat{O} | b_a \rangle \quad (\text{A.5})$$

The basis functions must behave in a way such that $\langle b_{k'} | b_k \rangle = \delta_{k',k}$. Notice that the *observable* $O \in \mathbb{R}$ whereas $\hat{O} \in \mathbb{C}$ (symbol $\hat{\cdot}$ denotes the operator).

The observable for Equation (A.3) is computed as:

$$T_{m,n} = \sum_a \langle b_a | \hat{T}_{m,n} | b_a \rangle \quad (\text{A.6})$$

Similarly, the observable for Equation (A.4) is computed as:

$$T_{m,n}^\dagger = \sum_a \langle b_a | \hat{T}_{m,n}^\dagger | b_a \rangle \quad (\text{A.7})$$

Notice that both Equation (A.6) and Equation (A.7) represent the experimental measurements of the propagation of light between two points, which are Real numbers (as opposed to Complex numbers) and reversible in time due to the Helmholtz reciprocity principle. Therefore:

$$T_{m,n} = T_{m,n}^\dagger \quad (\text{A.8})$$

which implies that the corresponding operators must be equal too:

$$\hat{T}_{m,n} = \hat{T}_{m,n}^\dagger \quad (\text{A.9})$$

The propagation of light from an illumination point n to a sensing point m is defined as:

$$|s_m\rangle = \hat{U}_{m,n} |l_n\rangle \quad (\text{A.10})$$

and the inverse propagation from a sensing point m to an illumination point n is defined as:

$$|l_n\rangle = \hat{V}_{n,m} |s_m\rangle \quad (\text{A.11})$$

where $V_{n,m}$ is the inverse propagation operator.

The *bra* corresponding to Equation (A.11) is:

$$\langle l_n| = \langle s_m| \hat{V}_{n,m}^\dagger \quad (\text{A.12})$$

Using Equation (A.12) in Equation (A.3) yields:

$$\hat{T}_{m,n} = |s_m\rangle \langle l_n| = |s_m\rangle \langle s_m| \hat{V}_{n,m}^\dagger \quad (\text{A.13})$$

and using now Equation (A.10) in Equation (A.13) yields:

$$\hat{T}_{m,n} = |s_m\rangle \langle l_n| = |s_m\rangle \langle s_m| \hat{V}_{n,m}^\dagger = \hat{U}_{m,n} |l_n\rangle \langle s_m| \hat{V}_{n,m}^\dagger = \hat{U}_{m,n} \hat{V}_{n,m}^\dagger |l_n\rangle \langle s_m| = \hat{T}_{m,n}^\dagger \quad (\text{A.14})$$

We can commute $|l_n\rangle \langle s_m| \hat{V}_{n,m}^\dagger = \hat{V}_{n,m}^\dagger |l_n\rangle \langle s_m|$ because the outer product is a linear operation that is commutative.

The only way Equation (A.14) is equal to Equation (A.9) is if:

$$\boxed{\hat{U}_{m,n} \hat{V}_{n,m}^\dagger = \mathbb{I} \quad \Rightarrow \quad \hat{V}_{n,m}^\dagger = \hat{U}_{m,n}^\dagger \quad \Rightarrow \quad \hat{U}_{m,n} \hat{U}_{m,n}^\dagger = \mathbb{I}} \quad (\text{A.15})$$

which proves that a propagation operator $\hat{U}_{m,n}$ is unitary.

IDEAL SIGNAL DERIVATION IN THE FREQUENCY DOMAIN.

We can split Equation (9.13) in terms, the first one being $\langle l_p | l \rangle \langle s | l_p \rangle$, the second one $\langle l_p | \hat{N} | l_p \rangle \langle l_p | s \rangle$, and the third one $\langle l_p | s_{p\eta} \rangle$. If we now factor Equation (9.14) and Equation (9.15) in Equation (9.13), so the first term yields

$$\begin{aligned} \langle l_p | l \rangle \langle s | l_p \rangle &= \iiint d\omega d\omega_\delta d\omega_s L_p^*(\omega) L_0 S(\omega_s) L_p(\omega) \langle \omega | \omega_\delta \rangle \langle \omega_s | \omega \rangle = \\ &= L_0 \iiint d\omega d\omega_\delta d\omega_s S(\omega_s) |L_p(\omega)|^2 \delta(\omega - \omega_\delta) \delta(\omega_s - \omega) = \\ &= L_0 \int d\omega S(\omega) |L_p(\omega)|^2 \equiv \boxed{\int d\omega S(\omega) |L_p(\omega)|^2}, \end{aligned} \quad (\text{B.1})$$

where ω with any or none subindex is the frequency basis, and we simplify $L_0 = 1$ as it only affects as a scale factor, representing the amplitude of our searched delta illumination.

We set $\langle l_p | \hat{N} | l_p \rangle$ as the mean of the spectral density. Thus, we define a unitless kernel for noise $N(\omega, \omega')$ that corresponds to the operator $\hat{N}$, so

$$\langle l_p | \hat{N} | l_p \rangle = \int d\omega_l L_p^*(\omega_l) \hat{N}(\omega, \omega_l) L_p(\omega_l) = \int d\omega_l |L_p(\omega_l)|^2 \hat{N}(\omega, \omega_l) \equiv N(\omega). \quad (\text{B.2})$$

Now, we can apply a similar definition from Equation (B.1) to the second term from Equation (9.13), thus

$$\langle l_p | \hat{N} | l_p \rangle \langle l_p | s_\delta \rangle = \int d\omega N(\omega) L_p^*(\omega) S_\delta(\omega) \langle \omega | \omega \rangle = \boxed{\int d\omega N(\omega) L_p^*(\omega) S_\delta(\omega)}. \quad (\text{B.3})$$

Finally, we expand the third term from Equation (9.13) similar to the other two, defining

$$\langle l_p | s_{p\eta} \rangle = \int d\omega L_p^*(\omega) S_{p\eta}(\omega) \langle \omega | \omega \rangle = \boxed{\int d\omega L_p^*(\omega) S_{p\eta}(\omega)}. \quad (\text{B.4})$$

Plugging all three expanded terms from Equation (B.1), Equation (B.3) and Equation (B.4), we now define Equation (9.13) as

$$\begin{aligned} \langle l_p | s \rangle \langle l | l_p \rangle + \langle l_p | \hat{N} | l_p \rangle \langle l_p | s \rangle - \langle l_p | s_{p\eta} \rangle &= \\ &= \int d\omega S(\omega) |L_p(\omega)|^2 + \int d\omega N(\omega) L_p^*(\omega) S(\omega) - \int d\omega L_p^*(\omega) S_{p\eta}(\omega) = \\ &= \boxed{\int d\omega S(\omega) [|L_p(\omega)|^2 + N(\omega) L_p^*(\omega)] - L_p^*(\omega) S_{p\eta}(\omega)}. \end{aligned} \quad (\text{B.5})$$

BIBLIOGRAPHY

- [1] ABRAMSON, N. Light-in-flight recording by holography. *Optics letters* 3, 4 (1978), 121–123.
- [2] ARELLANO, V., GUTIERREZ, D., AND JARABO, A. Fast back-projection for non-line of sight reconstruction. *Optics express* 25, 10 (2017), 11574–11583.
- [3] BARRETT, H. H. Fresnel zone plate imaging in nuclear medicine. *Journal of nuclear medicine* 13, 6 (1972), 382–385.
- [4] BECKER, W. *Advanced time-correlated single photon counting applications*, vol. 111. Springer, 2015.
- [5] BOUMAN, K. L., YE, V., YEDIDIA, A. B., DURAND, F., WORNELL, G. W., TORRALBA, A., AND FREEMAN, W. T. Turning corners into cameras: Principles and methods. In *Proceedings of the IEEE International Conference on Computer Vision* (2017), pp. 2270–2278.
- [6] BRUECK, S. R. Optical and interferometric lithography-nanotechnology enablers. *Proceedings of the IEEE* 93, 10 (2005), 1704–1721.
- [7] BUTTAFAVA, M., ZEMAN, J., TOSI, A., ELICEIRI, K., AND VELTEN, A. Non-line-of-sight imaging using a time-gated single photon avalanche diode. *Opt. Express* 23, 16 (2015), 20997–21011.
- [8] CAO, R., GOUMOENS, F. D., BLOCHET, B., XU, J., AND YANG, C. High-resolution non-line-of-sight imaging employing active focusing. *Nature Photonics* 16, 6 (2022), 462–468.
- [9] CARAMAZZA, P., BOCCOLINI, A., BUSCHEK, D., HULLIN, M., HIGHAM, C. F., HENDERSON, R., MURRAY-SMITH, R., AND FACCIO, D. Neural network identification of people hidden from view with a single-pixel, single-photon detector. *Scientific reports* 8, 1 (2018), 11945.
- [10] CHEN, W., DANEAU, S., MANNAN, F., AND HEIDE, F. Steady-state non-line-of-sight imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2019), pp. 6790–6799.
- [11] CHEN, W., WEI, F., KUTULAKOS, K. N., RUSINKIEWICZ, S., AND HEIDE, F. Learned feature embeddings for non-line-of-sight imaging and recognition. *ACM Transactions on Graphics (ToG)* 39, 6 (2020), 1–18.
- [12] CONWAY, J. B. *A course in functional analysis*. Springer, 2019.
- [13] CRONIN, A. D., SCHMIEDMAYER, J., AND PRITCHARD, D. E. Optics and interferometry with atoms and molecules. *Reviews of Modern Physics* 81, 3 (2009), 1051–1129.
- [14] DEMOS, S., AND ALFANO, R. Optical polarization imaging. *Appl. Opt.* 36, 1 (1997), 150–155.
- [15] DIRAC, P. A. M. A new notation for quantum mechanics. *Mathematical Proceedings of the Cambridge Philosophical Society* 35, 3 (1939), 416–418.
- [16] DIRAC, P. A. M. *The Principles of Quantum Mechanics*. Clarendon Press, 1981.
- [17] DOVE, J., AND SHAPIRO, J. H. Paraxial theory of phasor-field imaging. *Optics Express* 27, 13 (June 2019), 18016–18037.
- [18] DOVE, J., AND SHAPIRO, J. H. Nonparaxial phasor-field propagation. *Optics Express* 28, 20 (Sept. 2020), 29212–29229. Publisher: Optical Society of America.

- [19] DOVE, J., AND SHAPIRO, J. H. Paraxial phasor-field physical optics. *Optics Express* 28, 14 (July 2020), 21095–21109. Publisher: Optical Society of America.
- [20] EISBERG, R. M., AND RESNICK, R. *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles*. Wiley, 1985.
- [21] FACCIO, D., VELTEN, A., AND WETZSTEIN, G. Non-line-of-sight imaging. *Nature Reviews Physics* 2, 6 (June 2020), 318–327. Number: 6 Publisher: Nature Publishing Group.
- [22] FADE, J., PANIGRAHI, S., CARRÉ, A., FREIN, L., HAMEL, C., BRETENAKER, F., RAMACHANDRAN, H., AND ALOUINI, M. Long-range polarimetric imaging through fog. *Applied optics* 53, 18 (2014), 3854–3865.
- [23] FANG, Y., WU, J., ZHONG, S., ZHANG, X., AN, Y., WANG, X., SU, B., AND WANG, K. Non-line-of-sight imaging via scalable scattering mapping using tof cameras. *Photonics Research* 13, 8 (2025), 2172–2183.
- [24] FEYNMAN, R. P., LEIGHTON, R. B., AND SANDS, M. *The Feynman Lectures on Physics, Vol. III*. Basic Books, 2015.
- [25] FISHBURN, M. W. *Fundamentals of CMOS single-photon avalanche diodes*. fishburn, 2012.
- [26] FRENCH, A. P., AND TAYLOR, E. F. *An Introduction to Quantum Physics*. Norton, 1978.
- [27] GARIEPY, G., KRSTAJIĆ, N., HENDERSON, R., LI, C., THOMSON, R. R., BULLER, G. S., HESHMAT, B., RASKAR, R., LEACH, J., AND FACCIO, D. Single-photon sensitive light-in-flight imaging. *Nature communications* 6, 1 (2015), 6021.
- [28] GARIEPY, G., TONOLINI, F., HENDERSON, R., LEACH, J., AND FACCIO, D. Detection and tracking of moving objects hidden from view. *Nature Photonics* 10, 1 (2016), 23–26.
- [29] GASBARRI, G., BELENCHIA, A., CARLESSO, M., DONADI, S., BASSI, A., KALTENBAEK, R., PATERNOSTRO, M., AND ULBRICHT, H. Testing the foundation of quantum physics in space via interferometric and non-interferometric experiments with mesoscopic nanoparticles. *Communications Physics* 4, 1 (2021), 155.
- [30] GENG, R., HU, Y., AND CHEN, Y. Recent advances on non-line-of-sight imaging: Conventional physical models, deep learning, and new scenes. *APSIPA Transactions on Signal and Information Processing* 11, 1 (2022), e1.
- [31] GIBSON, A., HEBDEN, J., AND ARRIDGE, S. R. Recent advances in diffuse optical imaging. *Phys. Med. Biol.* 50, 4 (2005), R1.
- [32] GKIOULEKAS, I., LEVIN, A., AND ZICKLER, T. An evaluation of computational imaging techniques for heterogeneous inverse scattering. In *European Conference on Computer Vision* (2016), Springer, pp. 685–701.
- [33] GRIFFITHS, D. J., AND SCHROETER, D. F. *Introduction to Quantum Mechanics*. Cambridge University Press, 2018.
- [34] GU, C., SULTAN, T., MASUMNIA-BISHEH, K., WALLER, L., AND VELTEN, A. Fast non-line-of-sight imaging with non-planar relay surfaces. In *2023 IEEE International Conference on Computational Photography (ICCP)* (2023), pp. 1–12.
- [35] GUILLÉN, I., LIU, X., VELTEN, A., GUTIERREZ, D., AND JARABO, A. On the effect of reflectance on phasor field non-line-of-sight imaging. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2020), IEEE, pp. 9269–9273.
- [36] GUPTA, O., WILLWACHER, T., VELTEN, A., VEERARAGHAVAN, A., AND RASKAR, R. Reconstruction of hidden 3d shapes using diffuse reflections. *Opt. Express* 20, 17 (2012), 19096–19108.

- [37] HAN, P., CHO, G., AND ZHANG, X.-C. Time-domain transillumination of biological tissues with terahertz pulses. *Opt. Lett.* 25, 4 (2000), 242–244.
- [38] HANNAH, M. J. *Computer matching of areas in stereo images*. Stanford University, 1974.
- [39] HE, J., WU, S., WEI, R., AND ZHANG, Y. Non-line-of-sight imaging and tracking of moving objects based on deep learning. *Optics Express* 30, 10 (2022), 16758–16772.
- [40] HEIDE, F., HULLIN, M. B., GREGSON, J., AND HEIDRICH, W. Low-budget transient imaging using photonic mixer devices. *ACM Transactions on Graphics (ToG)* 32, 4 (2013), 1–10.
- [41] HEIDE, F., O’TOOLE, M., ZANG, K., LINDELL, D. B., DIAMOND, S., AND WETZSTEIN, G. Non-line-of-sight imaging with partial occluders and surface normals. *ACM Trans. Graph.* 38, 3 (2019), 22.
- [42] HEIDE, F., XIAO, L., HEIDRICH, W., AND HULLIN, M. B. Diffuse mirrors: 3d reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2014), pp. 3222–3229.
- [43] HEIDE, F., XIAO, L., KOLB, A., HULLIN, M. B., AND HEIDRICH, W. Imaging in scattering media using correlation image sensors and sparse convolutional coding. *Opt. Express* 22, 21 (2014), 26338–26350.
- [44] HENYEV, L., AND GREENSTEIN, J. Diffuse radiation in the galaxy. *The Astrophysical Journal* 93 (1941), 70–83.
- [45] HERNANDEZ, Q., GUTIERREZ, D., AND JARABO, A. A computational model of a single-photon avalanche diode sensor for transient imaging, 2017.
- [46] HIRSCHMULLER, H. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on pattern analysis and machine intelligence* 30, 2 (2007), 328–341.
- [47] HUANG, D., SWANSON, E. A., LIN, C. P., SCHUMAN, J. S., STINSON, W. G., CHANG, W., HEE, M. R., FLOTTE, T., GREGORY, K., PULIAFITO, C. A., ET AL. Optical coherence tomography. *Science* 254, 5035 (1991), 1178–1181.
- [48] ISERINGHAUSEN, J., AND HULLIN, M. B. Non-line-of-sight reconstruction using efficient transient rendering. *ACM Transactions on Graphics (ToG)* 39, 1 (2020), 1–14.
- [49] ISOGAWA, M., YUAN, Y., O’TOOLE, M., AND KITANI, K. M. Optical non-line-of-sight physics-based 3d human pose estimation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2020), pp. 7013–7022.
- [50] JARABO, A., MARCO, J., MUÑOZ, A., BUISAN, R., JAROSZ, W., AND GUTIERREZ, D. A framework for transient rendering. *ACM Trans. Graph.* 33, 6 (2014).
- [51] JARABO, A., MASIA, B., MARCO, J., AND GUTIERREZ, D. Recent advances in transient imaging: A computer graphics and vision perspective. *Visual Informatics* 1, 1 (2017).
- [52] JIANG, D., LIU, X., LUO, J., LIAO, Z., VELTEN, A., AND LOU, X. Ring and radius sampling based phasor field diffraction algorithm for non-line-of-sight reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 11 (2022), 7841–7853.
- [53] JIN, C., XIE, J., ZHANG, S., ZHANG, Z., AND ZHAO, Y. Richardson–lucy deconvolution of time histograms for high-resolution non-line-of-sight imaging based on a back-projection method. *Optics Letters* 43, 23 (2018), 5885–5888.
- [54] KENDALL, A., MARTIROSYAN, H., DASGUPTA, S., HENRY, P., KENNEDY, R., BACHRACH, A., AND BRY, A. End-to-end learning of geometry and context for deep stereo regression. In *Proceedings of the IEEE international conference on computer vision* (2017), pp. 66–75.

- [55] KIRMANI, A., HUTCHISON, T., DAVIS, J., AND RASKAR, R. Looking around the corner using transient imaging. In *2009 IEEE 12th International Conference on Computer Vision (2009)*, IEEE, pp. 159–166.
- [56] KIRMANI, A., HUTCHISON, T., DAVIS, J., AND RASKAR, R. Looking around the corner using ultrafast transient imaging. *International journal of computer vision* 95, 1 (2011), 13–28.
- [57] KLEIN, J., PETERS, C., MARTÍN, J., LAURENZIS, M., AND HULLIN, M. B. Tracking objects outside the line of sight using 2D intensity images. *Scientific Reports* 6, 1 (2016), 32491.
- [58] KRSKA, W., SEIDEL, S. W., SAUNDERS, C., CZAJKOWSKI, R., YU, C., MURRAY-BRUCE, J., AND GOYAL, V. Double your corners, double your fun: the doorway camera. In *2022 IEEE International Conference on Computational Photography (ICCP) (2022)*, IEEE, pp. 1–12.
- [59] LA MANNA, M., KINE, F., BREITBACH, E., JACKSON, J., SULTAN, T., AND VELTEN, A. Error backprojection algorithms for non-line-of-sight imaging. *IEEE transactions on pattern analysis and machine intelligence* 41, 7 (2018), 1615–1626.
- [60] LA MANNA, M., NAM, J.-H., REZA, S. A., AND VELTEN, A. Non-line-of-sight-imaging using dynamic relay surfaces. *Optics express* 28, 4 (2020), 5331–5339.
- [61] LAURENZIS, M., AND VELTEN, A. Feature selection and back-projection algorithms for nonlinear-of-sight laser-gated viewing. *Journal of Electronic Imaging* 23, 6 (2014), 063003–063003.
- [62] LAURENZIS, M., AND VELTEN, A. Nonline-of-sight laser gated viewing of scattered photons. *Opt. Eng.* 53, 2 (2014), 023102.
- [63] LE ROUX, J., WISDOM, S., ERDOGAN, H., AND HERSHEY, J. R. Sdr-half-baked or well done? In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2019)*, IEEE, pp. 626–630.
- [64] LIAO, Z., JIANG, D., LIU, X., VELTEN, A., HA, Y., AND LOU, X. FPGA Accelerator for Real-Time Non-Line-of-Sight Imaging. *IEEE Transactions on Circuits and Systems I: Regular Papers* 69, 2 (Feb. 2022), 721–734. Conference Name: IEEE Transactions on Circuits and Systems I: Regular Papers.
- [65] LINDELL, D. B., AND WETZSTEIN, G. Three-dimensional imaging through scattering media based on confocal diffuse tomography. *Nat. Commun.* 11, 1 (2020), 1–8.
- [66] LINDELL, D. B., WETZSTEIN, G., AND O'TOOLE, M. Wave-based non-line-of-sight imaging using fast fk migration. *ACM Transactions on Graphics (ToG)* 38, 4 (2019), 1–13.
- [67] LIU, X., BAUER, S., AND VELTEN, A. Analysis of feature visibility in non-line-of-sight measurements. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2019)*, pp. 10140–10148.
- [68] LIU, X., BAUER, S., AND VELTEN, A. Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems. *Nat. Commun.* 11, 1 (2020), 1–13.
- [69] LIU, X., GUILLÉN, I., LA MANNA, M., NAM, J. H., REZA, S. A., HUU LE, T., JARABO, A., GUTIERREZ, D., AND VELTEN, A. Non-line-of-sight imaging using phasor-field virtual wave optics. *Nature* 572, 7771 (2019), 620–623.
- [70] LIU, X., AND VELTEN, A. 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.
- [71] LUESIA, P., CRESPO, M., JARABO, A., AND REDO-SANCHEZ, A. Non-line-of-sight imaging in the presence of scattering media using phasor fields. *Optics Letters* 47, 15 (2022), 3796–3799.

- [72] LUESIA-LAHOZ, P., CARTIEL, S., AND MUÑOZ, A. Stereo non-line-of-sight imaging. *The Visual Computer* 42, 2 (Jan 2026), 148.
- [73] LUESIA-LAHOZ, P., GUTIERREZ, D., AND MUÑOZ, A. Zone plate virtual lenses for memory-constrained nlos imaging. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2023), IEEE, pp. 1–5.
- [74] LUESIA-LAHOZ, P., SULTAN, T., PETERSON, F. B., VELTEN, A., GUTIERREZ, D., AND MUÑOZ, A. Zero-phase phasor fields for non-line-of-sight imaging. In *2025 IEEE International Conference on Computational Photography (ICCP)* (2025), pp. 1–9.
- [75] LYONS, A., TONOLINI, F., BOCCOLINI, A., REPETTI, A., HENDERSON, R., WIAUX, Y., AND FACCIO, D. Computational time-of-flight diffuse optical tomography. *Nature Photonics* 13, 8 (2019), 575–579.
- [76] MAEDA, T., SATAT, G., SWEDISH, T., SINHA, L., AND RASKAR, R. Recent advances in imaging around corners. *arXiv preprint arXiv:1910.05613* (2019).
- [77] MAEDA, T., WANG, Y., RASKAR, R., AND KADAMBI, A. Thermal non-line-of-sight imaging. In *2019 IEEE International Conference on Computational Photography (ICCP)* (2019), IEEE, pp. 1–11.
- [78] MARCO, J., HERNANDEZ, Q., MUNOZ, A., DONG, Y., JARABO, A., KIM, M. H., TONG, X., AND GUTIERREZ, D. Deeptof: off-the-shelf real-time correction of multipath interference in time-of-flight imaging. *ACM Transactions on Graphics (ToG)* 36, 6 (2017), 1–12.
- [79] MARCO, J., JARABO, A., NAM, J. H., LIU, X., COSCULLUELA, M. A., VELTEN, A., AND GUTIERREZ, D. Virtual light transport matrices for non-line-of-sight imaging. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (October 2021), pp. 2440–2449.
- [80] MIAO, J., GUO, E., SHI, Y., CAI, F., BAI, L., AND HAN, J. Super-resolution non-line-of-sight imaging based on temporal encoding. *Optics Express* 31, 24 (2023), 40235–40248.
- [81] NAM, J. H., BRANDT, E., BAUER, S., LIU, X., RENNA, M., TOSI, A., SIFAKIS, E., AND VELTEN, A. Low-latency time-of-flight non-line-of-sight imaging at 5 frames per second. *Nature communications* 12, 1 (2021), 6526.
- [82] NOUSIAS, S., WEI, M., XIAO, H., WU, M., ATHAR, S., WANG, K. J., MALIK, A., BARMHERZIG, D. A., LINDELL, D. B., AND KUTULAKOS, K. N. Opportunistic single-photon time of flight. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2025), pp. 15852–15862.
- [83] O’LEARY, M. A. *Imaging with diffuse photon density waves*. University of Pennsylvania, 1996.
- [84] O’TOOLE, M., HEIDE, F., LINDELL, D. B., ZANG, K., DIAMOND, S., AND WETZSTEIN, G. Reconstructing transient images from single-photon sensors. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017), pp. 1539–1547.
- [85] OUTER, P. N. D., LAGENDIJK, A., AND NIEUWENHUIZEN, T. M. Location of objects in multiple-scattering media. *J. Opt. Soc. Am. A* 10, 6 (1993), 1209.
- [86] O’TOOLE, M., LINDELL, D. B., AND WETZSTEIN, G. Confocal non-line-of-sight imaging based on the light-cone transform. *Nature* 555, 7696 (2018), 338–341.
- [87] PEI, C., ZHANG, A., DENG, Y., XU, F., WU, J., LI, D. U.-L., QIAO, H., FANG, L., AND DAI, Q. Dynamic non-line-of-sight imaging system based on the optimization of point spread functions. *Optics Express* 29, 20 (2021), 32349.

- [88] PLACK, M., CALLENBERG, C., SCHNEIDER, M., AND HULLIN, M. B. Fast differentiable transient rendering for non-line-of-sight reconstruction. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (2023), pp. 3067–3076.
- [89] PUEYO-CIUTAD, O., MARCO, J., SCHERTZER, S., CHRISTNACHER, F., LAURENZIS, M., GUTIERREZ, D., AND REDO-SANCHEZ, A. Time-gated polarization for active non-line-of-sight imaging. In *SIGGRAPH Asia 2024 Conference Papers* (2024), pp. 1–11.
- [90] RAMACHANDRAN, H. Imaging through turbid media. *Current Science* (1999), 1334–1340.
- [91] REDO-SANCHEZ, A., LUESIA-LAHOZ, P., GUTIERREZ, D., AND MUÑOZ, A. Cohesive framework for non-line-of-sight imaging based on Dirac notation. *Optics Express* 32, 6 (2024), 10505–10526.
- [92] REZA, S. A., LA MANNA, M., BAUER, S., AND VELTEN, A. Phasor field waves: A Huygens-like light transport model for non-line-of-sight imaging applications. *Optics express* 27, 20 (2019), 29380–29400.
- [93] REZA, S. A., LA MANNA, M., BAUER, S., AND VELTEN, A. Phasor field waves: experimental demonstrations of wave-like properties. *Optics Express* 27, 22 (Oct. 2019), 32587.
- [94] REZA, S. A., LA MANNA, M., AND VELTEN, A. A Physical Light Transport Model for Non-Line-of-Sight Imaging Applications. *arXiv:1802.01823 [physics]* (Feb. 2018). arXiv: 1802.01823.
- [95] ROYO, D., CRESPO, M., AND GARCIA-PUEYO, J. mitransient. <https://github.com/diegoroyo/mitransient>, 2024.
- [96] ROYO, D., GARCÍA, J., MUÑOZ, A., AND JARABO, A. Non-line-of-sight transient rendering. *Computers & Graphics* 107 (2022), 84–92.
- [97] ROYO, D., HUANG, Z., LIANG, Y., SONG, B., NOZ, A. M., GUTIERREZ, D., AND MARCO, J. Structure-aware parametric representations for time-resolved light transport. *Opt. Lett.* 47, 19 (Oct 2022), 5212–5215.
- [98] ROYO, D., SULTAN, T., MUÑOZ, A., MASUMNIA-BISHEH, K., BRANDT, E., GUTIERREZ, D., VELTEN, A., AND MARCO, J. Virtual mirrors: Non-line-of-sight imaging beyond the third bounce. *ACM Transactions on Graphics* 42, 4 (2023).
- [99] RUDIN, W. *Functional Analysis*. McGraw-Hill, 1991.
- [100] SAKURAI, J. J., AND NAPOLITANO, J. *Modern Quantum Mechanics*. Cambridge University Press, 2021.
- [101] SAUNDERS, C., MURRAY-BRUCE, J., AND GOYAL, V. K. Computational periscopy with an ordinary digital camera. *Nature* 565, 7740 (2019), 472–475.
- [102] SAVUSKAN, V., BROUK, I., JAVITT, M., AND NEMIROVSKY, Y. An estimation of single photon avalanche diode (spad) photon detection efficiency (pde) nonuniformity. *IEEE Sensors Journal* 13, 5 (2013), 1637–1640.
- [103] SCHARSTEIN, D., AND SZELISKI, R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International journal of computer vision* 47 (2002), 7–42.
- [104] SMITH, B. M., O’TOOLE, M., AND GUPTA, M. Tracking Multiple Objects Outside the Line of Sight Using Speckle Imaging. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2018), 6258–6266.
- [105] SULTAN, T., BOCCHIERI, A., GU, C., LIU, X., POLYNKIN, P., AND VELTEN, A. Optimized sampling for non-line-of-sight imaging using modified fast Fourier transforms, 2025.

- [106] SULTAN, T., BRANDT, E., MASUMNIA-BISHEH, K., RICCARDO, S., POLYNKIN, P., TOSI, A., AND VELTEN, A. Iterating the transient light transport matrix for non-line-of-sight imaging. *arXiv preprint arXiv:2412.10300* (2024).
- [107] SULTAN, T., REZA, S. A., AND VELTEN, A. Towards a more accurate light transport model for non-line-of-sight imaging. *Optics Express* 32, 5 (Feb. 2024), 7731–7761. Publisher: Optica Publishing Group.
- [108] SZELISKI, R. *Computer vision: algorithms and applications*. Springer Nature, 2022.
- [109] TANAKA, K., MUKAIGAWA, Y., AND KADAMBI, A. Polarized Non-Line-of-Sight Imaging. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 00* (2020), 2133–2142.
- [110] TEICHMAN, J. A. Phasor field waves: a mathematical treatment. *Optics Express* 27, 20 (Sept. 2019), 27500–27506. Publisher: Optical Society of America.
- [111] TIAN, X., YU, J., QIAO, K., TANG, M., ZHANG, S., AND JIN, C. Non-line-of-sight virtual modulated range migration imaging based on super-resolution histograms. *Optics Letters* 50, 2 (2025), 519–522.
- [112] TOBIN, R., HALIMI, A., MCCARTHY, A., LAURENZIS, M., CHRISTNACHER, F., AND BULLER, G. S. Three-dimensional single-photon imaging through obscurants. *Optics express* 27, 4 (2019), 4590–4611.
- [113] TSAI, C.-Y., SANKARANARAYANAN, A. C., AND GKIOULEKAS, I. 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, June 2019), IEEE, pp. 1545–1555.
- [114] VELLEKOOP, I. M., AND MOSK, A. Focusing coherent light through opaque strongly scattering media. *Opt. Lett.* 32, 16 (2007), 2309–2311.
- [115] VELTEN, A., WILLWACHER, T., GUPTA, O., VEERARAGHAVAN, A., BAWENDI, M. G., AND RASKAR, R. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature communications* 3, 1 (2012), 1–8.
- [116] VELTEN, A., WU, D., JARABO, A., MASIA, B., BARSİ, C., JOSHI, C., LAWSON, E., BAWENDI, M., GUTIERREZ, D., AND RASKAR, R. Femto-photography: capturing and visualizing the propagation of light. *ACM Transactions on Graphics (ToG)* 32, 4 (2013), 1–8.
- [117] WANG, B., ZHENG, M.-Y., HAN, J.-J., HUANG, X., XIE, X.-P., XU, F., ZHANG, Q., AND PAN, J.-W. Non-line-of-sight imaging with picosecond temporal resolution. *Physical Review Letters* 127, 5 (2021), 053602.
- [118] WANG, D., HAO, W., TIAN, Y., XU, W., TIAN, Y., CHENG, H., CHEN, S., ZHANG, N., ZHU, W., AND SU, X. Enhancing the spatial resolution of time-of-flight based non-line-of-sight imaging via instrument response function deconvolution. *Optics Express* 32, 7 (2024), 12303–12317.
- [119] WEBB, G., MININ, I., AND MININ, O. Variable reference phase in diffractive antennas: Review, applications, new results. *IEEE Antennas and Propagation Magazine* 53, 2 (2011), 77–94.
- [120] WILLOMITZER, F., RANGARAJAN, P. V., LI, F., BALAJI, M. M., CHRISTENSEN, M. P., AND COSSAIRT, O. Fast non-line-of-sight imaging with high-resolution and wide field of view using synthetic wavelength holography. *Nature Communications* 12, 1 (2021), 6647.
- [121] WU, C., LIU, J., HUANG, X., LI, Z.-P., YU, C., YE, J.-T., ZHANG, J., ZHANG, Q., DOU, X., GOYAL, V. K., ET AL. Non-line-of-sight imaging over 1.43 km. *Proceedings of the National Academy of Sciences* 118, 10 (2021), e2024468118.

- [122] WU, D., VELTEN, A., O'TOOLE, M., MASIA, B., AGRAWAL, A., DAI, Q., AND RASKAR, R. Decomposing global light transport using time of flight imaging. *International journal of computer vision* 107, 2 (2014), 123–138.
- [123] WU, R., JARABO, A., SUO, J., DAI, F., ZHANG, Y., DAI, Q., AND GUTIERREZ, D. Adaptive polarization-difference transient imaging for depth estimation in scattering media. *Opt. Lett.* 43, 6 (2018).
- [124] XIN, S., NOUSIAS, S., KUTULAKOS, K. N., SANKARANARAYANAN, A. C., NARASIMHAN, S. G., AND GKIOULEKAS, I. A theory of Fermat paths for non-line-of-sight shape reconstruction. In *IEEE CVPR* (2019), pp. 6800–6809.
- [125] XU, W., CHEN, S., TIAN, Y., WANG, D., AND SU, X. Fast non-line-of-sight imaging based on product-convolution expansions. *Optics Letters* 47, 18 (2022), 4680–4683.
- [126] YE, J., HONG, Y., SU, X., YUAN, X., AND XU, F. Plug-and-play algorithms for dynamic non-line-of-sight imaging. *ACM Transactions on Graphics* 43, 5 (2024), 1–12.
- [127] YI, S., KIM, D., CHOI, K., JARABO, A., GUTIERREZ, D., AND KIM, M. H. Differentiable transient rendering. *ACM Transactions on Graphics (TOG)* 40, 6 (2021), 1–11.
- [128] YODH, A., AND CHANCE, B. Spectroscopy and imaging with diffusing light. *Physics today* 48, 3 (1995), 34–41.
- [129] YOUNG, S. I., LINDELL, D. B., GIROD, B., TAUBMAN, D., AND WETZSTEIN, G. 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.

This document was typeset using the typographical look-and-feel classicthesis developed by André Miede. The style was inspired by Robert Bringhurst’s seminal book on typography “*The Elements of Typographic Style*”.

Final Version as of June 30, 2026 (version 1.0).