

Non-line-of-sight Imaging with Partial Occluders and Surface Normals

FELIX HEIDE, Stanford University¹
MATTHEW O'TOOLE, Stanford University²
KAI ZANG, Stanford University
DAVID B. LINDELL, Stanford University
STEVEN DIAMOND, Stanford University
GORDON WETZSTEIN, Stanford University

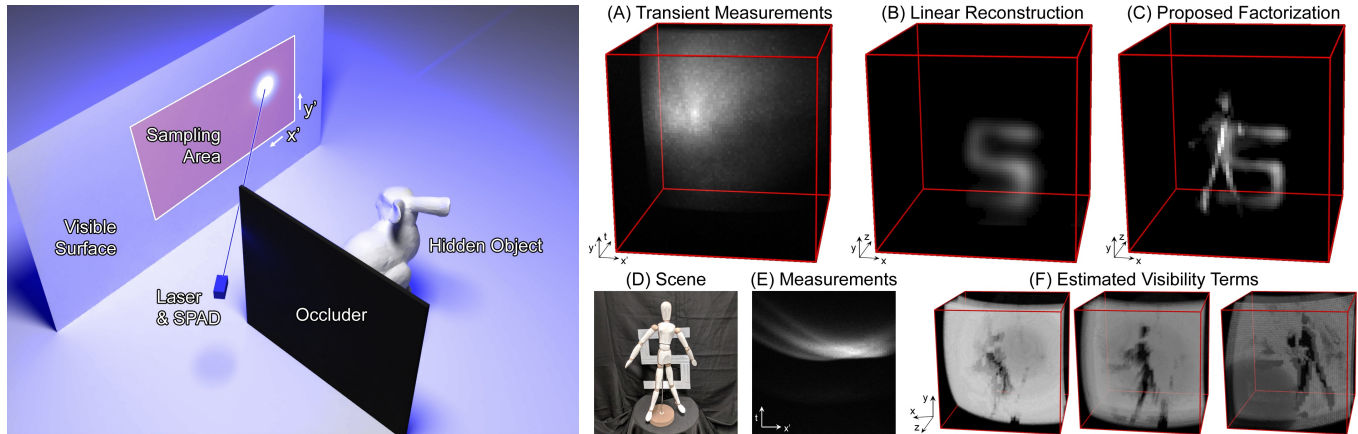


Fig. 1. Non-line-of-sight (NLOS) imaging aims at recovering the shape and albedo of objects hidden from a camera or a light source. Using ultra-fast pulsed illumination and single-photon detectors, the light transport in the scene is sampled on visible surfaces (left). The global illumination components of these time-resolved measurements (A,E) contain sufficient information to estimate the shape of hidden objects (B,C). Using a novel formulation for NLOS light transport that models partial occlusions of hidden objects (D) via visibility terms (F), we demonstrate higher-fidelity reconstructions (C) than previous approaches to NLOS imaging (B).

Imaging objects obscured by occluders is a significant challenge for many applications. A camera that could “see around corners” could help improve navigation and mapping capabilities of autonomous vehicles or make search and rescue missions more effective. Time-resolved single-photon imaging systems have recently been demonstrated to record optical information of a scene that can lead to an estimation of the shape and reflectance of objects hidden from the line of sight of a camera. However, existing non-line-of-sight (NLOS) reconstruction algorithms have been constrained in the types of light transport effects they model for the hidden scene parts. We introduce a factored NLOS light transport representation that accounts for partial occlusions and surface normals. Based on this model, we develop a factorization approach for inverse time-resolved light transport and demonstrate high-fidelity NLOS reconstructions for challenging scenes both in simulation and with an experimental NLOS imaging system.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2019 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2019/3-ART0 \$15.00
https://doi.org/0000001.0000001_2
¹now at Princeton, ²now at CMU

CCS Concepts: • **Computing methodologies** → **Computational photography**; *Reconstruction*; *Visibility*;

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

ACM Reference Format:

Felix Heide, Matthew O’Toole, Kai Zang, David B. Lindell, Steven Diamond, and Gordon Wetzstein. 2019. Non-line-of-sight Imaging with Partial Occluders and Surface Normals. *ACM Trans. Graph.* 38, 1, Article 0 (March 2019), 10 pages. https://doi.org/0000001.0000001_2 ¹now at Princeton, ²now at CMU

1 INTRODUCTION

The capacity of imaging systems must continue to expand to keep pace with rapidly emerging technologies. Autonomous vehicles, for example, would greatly benefit from improved vision in fog, snow, and other scattering media or from being able to see around corners to detect what lies beyond the next bend or another car. Sensing technology offering such non-line-of-sight (NLOS) capabilities could help make self-driving cars safer and unlock unprecedented potential for other machine vision systems.

Two challenges make NLOS imaging with time-resolved detectors difficult. First, the low signal of multiply scattered light places extreme requirements on photon sensitivity of the detectors. Second,

inverse methods that aim at estimating shape, albedo, and other properties of a hidden scene need to model the transient light transport appropriately and devise means to robustly invert it. To address the former issue, we follow recent work proposing acquisition setups with a pulsed picosecond laser and single-photon avalanche diodes (SPADs) for NLOS imaging and tracking [Buttafava et al. 2015; Chan et al. 2017; Garipey et al. 2016; O’Toole et al. 2018]. SPADs are detectors that digitize the time of arrival of individual photons with precision on the order of tens of picoseconds, thus resolving light paths with centimeter resolution.

At the core of our paper is a factored representation of transient light transport that models partial occlusions and surface normals in the hidden scene. These effects have been largely ignored by other NLOS imaging approaches [Buttafava et al. 2015; Gupta et al. 2012; Heide et al. 2014; O’Toole et al. 2018; Velten et al. 2012; Wu et al. 2012]. Moreover, we derive a robust multi-convex reconstruction algorithm that takes a measured transient image as input and factors it into the proposed representation: a volume of hidden albedos and surface normals along with visibility terms that model partial occlusion in the hidden scene (see Fig. 1).

With the presented work, we take first steps towards making NLOS imaging robust and practical for real-world applications. Specifically, we make the following contributions:

- We introduce a factored nonlinear image formation model for non-line-of-sight imaging that accounts for partial occlusions and surface normals in the hidden scene.
- We propose a multi-convex solver for inverse transient light transport and show that it achieves significantly higher reconstruction quality than conventional NLOS imaging methods.
- We implement an experimental NLOS acquisition setup using a single-photon avalanche diode and a picosecond laser.
- We validate the proposed reconstruction algorithms in simulation and with example scenes captured with the prototype.

Overview of Limitations. Although the proposed inverse method improves reconstruction quality for many types of hidden scenes, it is also computationally more expensive than other methods. Specifically, the memory requirements of the proposed factorization method are two orders of magnitude higher than for matrix-free implementations of simpler inverse methods (see Sec. 6 for more details). Similar to other non-line-of-sight methods, we make the assumption that the measured transient light transport contains only first-order and third-order bounces. The first-order bounces correspond to direct illumination that is reflected off a visible wall back to the detector; these contributions are used to estimate surface normals and albedos of the visible wall and can be removed prior to recovering the hidden object from the indirect illumination. The third-order bounces contain indirect illumination that bounced precisely three times before reaching the detector: off the visible wall, then off a hidden object, then off the visible wall again. We used a transient renderer that supports global illumination [Jarabo et al. 2014] to verify that higher-order bounces, which could result from interreflections in the hidden scene, insignificantly contribute to the proposed confocal image formation in our simulated scenes. We thus conclude that these effects can be ignored in the imaging setup described in this work.

2 RELATED WORK

Non-line-of-sight Imaging. Kirmani et al. [2009] first introduced the idea of “looking around corners” by analyzing the feasibility of reconstructing hidden objects from time-resolved light transport. This concept was demonstrated in practice by Velten et al. [2012] with a system capable of resolving the shape of a hidden object. Velten’s hardware setup included a streak camera and a femtosecond laser, which together account for a cost of several hundred thousand dollars. The streak camera provides a theoretical precision of up to 2 ps, which corresponds to a travel distance of 0.6 mm. Correlation-based time-of-flight sensors have also been demonstrated as a low-cost alternative for non-line-of-sight imaging [Heide et al. 2014; Kadambi et al. 2013]. While these systems are about three orders of magnitude less expensive than Velten’s system, they also only offer a very limited temporal resolution, thus limiting the quality of reconstructed NLOS scenes. Recently, single photon avalanche diodes (SPADs) have been proposed for NLOS imaging [Buttafava et al. 2015; O’Toole et al. 2018] as a readily-available hardware platform that offers a good balance between cost and precision.

NLOS imaging requires a model for the light transport of hidden scene parts as well as a large-scale reconstruction framework. Existing proposals for NLOS imaging [Buttafava et al. 2015; Gupta et al. 2012; Heide et al. 2014; Kirmani et al. 2009; O’Toole et al. 2018; Velten et al. 2012; Wu et al. 2012] use an image formation model that makes the following assumptions: (1) light bounces at most three times within the scene; (2) the scene contains no occlusions; (3) light scatters isotropically (i.e., surface normals are ignored). Under these assumptions, the reconstruction becomes a linear inverse problem. Velten et al. [2012], Gupta et al. [2012], Buttafava et al. [2015], and Jarabo et al. [2017] solved this system using variants of the back-projection algorithm. Due to the fact that this tends to emphasize low frequencies and does not actually solve the inverse problem, the reconstruction quality offered by backprojection-type inverse methods tends to be low. Using the same assumptions, Gupta et al. [2012], Wu et al. [2012], Heide et al. [2014], and [O’Toole et al. 2018] proposed to solve the inverse problem via large-scale iterative optimization. While this approach is more accurate than backprojection, the underlying light transport model ignores partial occlusions and surface normals, which we show to be crucial for accurate scene reconstruction. Tsai et al. [2017] recently proposed a space carving algorithm for estimating the convex hull of hidden objects; a full 3D volume of the hidden scene cannot be recovered with this approach. Finally, Pediredla et al. [2017] propose a plane-based parametric model for reconstructing room scenes.

At the core of this paper is a novel image formation model that models NLOS light transport more accurately than existing methods by accounting for partial occlusions and surface normals in the hidden scene; we derive inverse methods tailored to this model.

Single Photon Avalanche Diodes. SPADs are reverse-biased photodiodes that are operated well above their breakdown voltage (see e.g. [Burri et al. 2016]). Every photon incident on a SPAD has some probability of triggering an electron avalanche which is time-stamped. This time-stamping mechanism usually provides an accuracy of tens to hundreds of picoseconds. SPADs and also avalanche photodiodes (APDs) are commonly used for a wide range

of applications in optical telecommunication, fluorescence lifetime imaging, and remote sensing systems (e.g., LIDAR). Often, these imaging modes are referred to as time-correlated single photon counting [O'Connor 2012].

Recently, SPADs were applied to range imaging [Heide et al. 2018; Kirmani et al. 2014; Lindell et al. 2018a; Rapp and Goyal 2017; Shin et al. 2016], transient imaging [Hernandez, Quercus and Gutierrez, Diego and Jarabo, Adrian 2017; Garipey et al. 2015; Lindell et al. 2018b; O'Toole et al. 2017] as well as tracking [Chan et al. 2017; Garipey et al. 2016] and imaging [Buttafava et al. 2015; O'Toole et al. 2018] non-line-of-sight objects. Altmann et al. [2018] also provide a comprehensive review of ultra-fast computational imaging. The works by Buttafava et al. [2015] and O'Toole et al. [2018] are closest to ours, but their reconstruction algorithms ignore hidden surface normals and occlusions. While the approach proposed by O'Toole et al. [2018] achieves computational efficiency by modeling NLOS light transport as a shift-invariant convolution, normals and the visibility term modeling partial occlusions create spatial variation in the image formation and so are incompatible with this representation.

Imaging Through and Around Scene Objects. Other forms of non-line-of-sight imaging have also been demonstrated that do not rely on time-resolved imaging. For example, Sen et al. [2005] proposed a projector-camera system where the viewpoints of camera and projector could be interchanged. This approach allows the scene to be hidden from the camera's perspective, but it must be visible from the projector's perspective. In time-resolved NLOS imaging, the scene is typically not directly observed from either detector or light source. Snapshot NLOS imaging was demonstrated by exploiting correlations that exist in coherent laser speckle [Katz et al. 2014], though this has so far been demonstrated at microscopic scales. Radio and terahertz frequencies were shown to be able to image through objects due to the physical properties of these parts of the electromagnetic spectrum [Adib et al. 2015; Redo-Sanchez et al. 2016]. Recovering and tracking hidden objects was also shown to be possible with intensity measurements of conventional cameras to a limited extent [Bouman et al. 2017; Klein et al. 2016].

In concurrent work, Thrampoulidis et al. [2017] and Xu et al. [2018] developed an alternative model that also includes partial occlusions for NLOS imaging. As opposed to our model, these works assume that the shape of the NLOS scene, including occluders, is known and only surface albedos need to be recovered. We make no such assumptions; we model and recover unknown shape, albedo, visibility and also surface normals.

3 FORWARD AND INVERSE LIGHT TRANSPORT

The non-line-of-sight imaging problem involves estimating 3D shape and albedo of objects outside the line of sight of a detector from third-order bounces of time-resolved global light transport. Specifically, a short light pulse is focused on a visible part of the scene, for example a wall, the light scatters off that surface, reaches a hidden object which scatters some of the light back to the visible surface, where it can be recorded with a time-resolved detector. While several different acquisition setups have been proposed, each warranting a slightly different image formation model, we follow O'Toole et al. [2018] and model a confocal system, where a single time-resolved

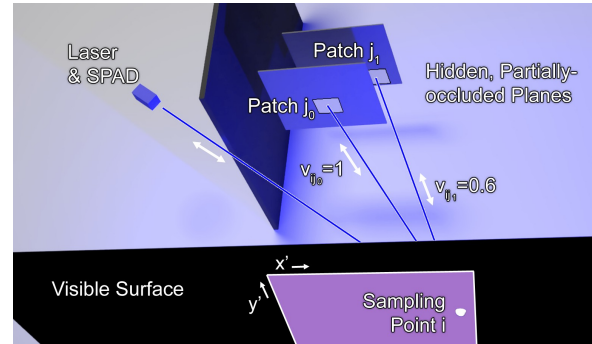


Fig. 2. NLOS scene with partial occlusions. The detector and laser sample the visible wall at point i to record the direct and indirect light transport. The indirect components include contributions from hidden objects, such as two patches j_0 and j_1 . Whereas the optical path between i and j_0 is unoccluded, the path between i and j_1 is partly obscured by another hidden surface. These occlusions are modeled by a visibility term v .

detector is co-axially aligned with a pulsed light source to sample positions x', y' on a visible diffuse wall (see Figs. 1, 2).

3.1 Confocal NLOS Image Formation Model

A time-resolved detector measures the incident photon flux as a function of time relative to an emitted light pulse. Such a detector can be used to record the temporal impulse response of a scene, including direct and global illumination, at sampling positions x', y' on a visible surface, resulting in a 3D space-time volume that is known as a *transient image* τ .

The direct illumination, i.e., light emitted by the source and scattered back to the detector from an object, contains all information necessary to recover the shape and reflectance of visible scene parts. This is commonly done for 3D imaging or light detection and ranging (LIDAR) [Kirmani et al. 2014; McCarthy et al. 2013; Schwarz 2010; Shin et al. 2016]. In the following image formation model, the direct light is not considered because it can be removed from measurements acquired in practice; only the global illumination contains useful information for non-line-of-sight imaging.

The image formation model can be formulated as

$$\tau(x', y', t) = \iiint_{\Omega} \frac{1}{r^4} \rho(x, y, z) \delta\left(2\sqrt{(x' - x)^2 + (y' - y)^2 + z^2} - tc\right) dx dy dz, \quad (1)$$

where the Dirac delta function $\delta(\cdot)$ relates the time of flight t to the distance function $r = \sqrt{(x' - x)^2 + (y' - y)^2 + z^2} = tc/2$. Here, c is the speed of light. This image formation model makes several assumptions on the light transport in the hidden scene: light scatters only once (i.e., back to the visible scene parts), light scatters isotropically (i.e., surface normals are ignored), and no occlusions occur between different scene parts outside the line of sight.

We lift two of these assumptions by augmenting Equation 1 by a visibility factor v as well as surface normals n :

$$\tau(x', y', t) = \left(\frac{2}{tc}\right)^4 \iiint_{\Omega} v(x', y', x, y, z) \rho(x, y, z) (\omega \cdot n(x, y, z)) \delta\left(2\sqrt{(x' - x)^2 + (y' - y)^2 + z^2} - tc\right) dx dy dz, \quad (2)$$

The distance falloff $1/r^4$ (see Eq. 1) is replaced by the factor $(2/tc)^4$, which can be pulled out of the integral because it is space-invariant for confocal scanning setups. The term $v(x', y', x, y, z) \in [0, 1]$ models the visibility of a hidden surface patch at location x, y, z observed from the position x', y' on the visible wall. For partial occlusion in the hidden scene, such a patch may be visible from one sampling point on the wall but it may be occluded by another hidden object from the perspective of another sampling point (see Fig. 2).

In a confocal scanning setup, the bidirectional reflectance distribution function (BRDF) f_r of the hidden scene is only sampled at a subset of all combinations of incident and outgoing light direction: $\omega_i = \omega_o = \omega$, where ω is the normalized direction from a location x, y, z to some sampling point x', y' . Thus, it may be infeasible to recover arbitrary BRDFs from confocal measurements. However, when the BRDF can be modeled as a spatially-varying but directionally constant albedo $\rho(x, y, z) = f_r(x, y, z, \omega, \omega)$, which is the case for diffuse and also retroreflective materials, this albedo can indeed be estimated as shown in previous work.

For a confocal scanning system, retroreflective materials have the benefit of significantly increasing the amount of light reflected from hidden scene parts back to sampling point on the wall. In theory, the distance falloff of $1/r^4$ becomes $1/r^2$ (cf. Eq. 1). In practice, the BRDFs of these materials often exhibit extended retroreflective lobes combined with some amount of Lambertian behavior, making them non-ideal retroreflectors. Even though the proposed confocal scanning scheme *only samples the retroreflective component* of the hidden BRDF, any deviation from ideal retroreflectors leads to a distance falloff term that is somewhere between $1/r^4$ and $1/r^2$. In the Supplemental Material, we experimentally verify that the distance falloff of “diamond grade” retroreflectors is close to ideal while “engineer grade” retroreflectors or retroreflective paints are less than ideal, but still exhibit a falloff that is significantly more favorable than that of purely diffuse objects.

3.2 Factored Image Formation

We discretize Equation 2 by representing the hidden volume as $N \times N \times N$ voxels. Each voxel $j = 1 \dots N^3$ contains an albedo ρ_j and a surface normal \mathbf{n}_j . The discrete transient image is sampled at $N \times N$ locations that coincide with the voxel centers on the visible wall. For notational convenience, we model the transient image with N temporal bins at each spatial location. The image formation model becomes

$$\boldsymbol{\tau} = \mathbf{A}\boldsymbol{\rho} = (\mathbf{T} \circ \mathbf{I}(\mathbf{N} \circ \mathbf{V}))\boldsymbol{\rho}, \quad (3)$$

where $\boldsymbol{\tau} \in \mathbb{R}_+^{N^3}$ is the vectorized transient image and $\boldsymbol{\rho} \in \mathbb{R}_+^{N^3}$ is the vectorized volume of nonnegative hidden albedos. The system matrix $\mathbf{A} \in \mathbb{R}_+^{N^3 \times N^3}$ combines all others terms of the transient

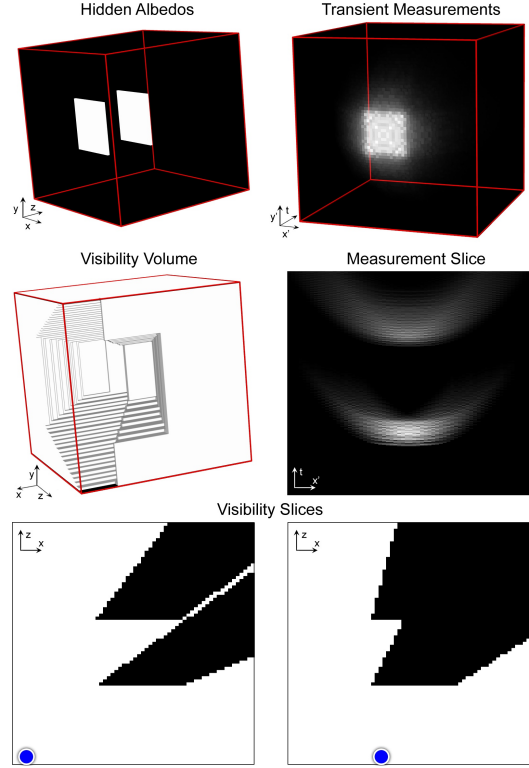


Fig. 3. Illustration of several light transport terms for a simple hidden scene containing two partially-occluding white planes (top left). The transient image of measurements is shown (top right) along with a single xt -slice (center right). A 3D rendering of the ground truth visibility term for one sampling location (center left) and two xz -slices (bottom) for different sampling positions (blue circles) make it intuitive to understand what the visibility terms are.

light transport (cf. Eq. 2). This matrix representation has also been used in most previous approaches to NLOS imaging. We propose to factor the transient light transport matrix \mathbf{A} into several terms, each modeling different aspects of light transport, as discussed in the following and illustrated in Figure 3.

Visibility. The visibility term $\mathbf{V} \in \mathbb{R}_+^{N^2 \times N^3}$ is time-invariant and models how much of the light reflected by voxel j reaches measurement location $i = 1 \dots N^2$ on the visible wall. As shown in Figure 2, when the path between j and i is unoccluded: $\mathbf{V}_{ij} = 1$. When another surface occludes the path between j and i : $\mathbf{V}_{ij} = 0$. We allow for a continuous range of values, i.e. $0 \leq \mathbf{V}_{ij} \leq 1$, to model partial occlusion along a light path.

Normals. The matrix $\mathbf{N} \in \mathbb{R}^{N^2 \times N^3}$ is also time-invariant and models the factor $\omega \cdot \mathbf{n}$, such that $\mathbf{N}_{ij} = \omega_{j \rightarrow i} \cdot \mathbf{n}_j$ where $\omega_{j \rightarrow i}$ is the normalized direction pointing from voxel j to the visible wall location i . For the purpose of this paper, we parameterize the surface normals in spherical coordinates. That is, the normal \mathbf{n}_j at voxel j is represented using two scalars $\mathbf{n}_j^u, \mathbf{n}_j^v$, such that $\mathbf{n}_j = [\cos(\mathbf{n}_j^u) \sin(\mathbf{n}_j^v), \sin(\mathbf{n}_j^u) \sin(\mathbf{n}_j^v), \cos(\mathbf{n}_j^v)]^T$. This representation

enforces unit length on all surface normals and only requires two, instead of three, parameters to be estimated per hidden normal.

Copy Matrix. To account for the fact that neither \mathbf{V} nor \mathbf{N} are time-dependent, but the transport matrix \mathbf{T} is, the matrix $\mathcal{I} \in \mathbb{R}_+^{N^3 \times N^2}$ simply copies the time-independent quantities of the hidden volume projected on the sampling locations to all time bins of the transient image.

Transport Matrix. The matrix $\mathbf{T} \in \mathbb{R}_+^{N^3 \times N^3}$ models the time-dependent aspects of the hidden light transport. Specifically, the j^{th} column of \mathbf{T} is the surface of the hypercone $(x' - x_j)^2 + (y' - y_j)^2 + z_j^2 = (tc/2)^2$ modeling time-resolved propagation in free space from voxel j to the entire transient image. The super-position principle holds, such that the hypercones for each voxel contribute to the transient image in an additive way.

3.3 Inverse NLOS Light Transport

Several inverse methods for the non-line-of-sight imaging problem have been proposed. We briefly review backprojection-type methods and linear inverse methods before introducing a factorization approach that allows us to recover the unknown visibility terms and surface normals along with the hidden albedos.

3.3.1 Backprojection. The NLOS problem can be reduced to a linear one under certain assumptions. First, the visibility term is ignored (i.e., $V_{ij} = 1, \forall i, j$) and, second, the surface normals are fixed. For the latter, one could either make the assumption that the hidden scene is comprised of isotropic scatterers (i.e., $N_{ij} = 1, \forall i, j$) or assume that the hidden normals are known, but that is typically not the case. Previous work on NLOS imaging has shown that these assumptions lead to the linear image formation model $\boldsymbol{\tau} = \mathbf{A}\boldsymbol{\rho}$, with $\mathbf{A} = \mathbf{T}$ (cf. Eq. 3).

Filtered and unfiltered backprojection methods are standard algorithms for solving many linear inverse problems, particularly in computed tomography [Kak and Slaney 1988]. The beauty of backprojection methods is their simplicity, i.e., both compute time ($O(N^5)$) and memory requirements (matrix-free implementation $O(N^3)$; with sparse matrix $O(N^5)$) are tractable even for large-scale inverse problems. Velten et al. [2012], Gupta et al. [2012], Buttafava et al. [2015], and Arellano et al. [2017] all employ a variant of backprojection by multiplying the measured transient image by the transpose of the system matrix, i.e., $\boldsymbol{\rho} \approx \mathbf{A}^T \boldsymbol{\tau}$, and then optionally applying a sharpening filter, such as a Laplacian, and a thresholding operator [Velten et al. 2012].

Unfortunately, filtered backprojection only solves the linear problem correctly when measurements over the full sphere are available. The acquisition setups of NLOS imaging discussed in the literature resemble that of a limited-baseline tomography problem, for which backprojection only gives a rough estimate of the latent variable, but it does not solve the actual inverse problem.

3.3.2 Linear Inverse Light Transport. Several other NLOS reconstruction algorithms [Gupta et al. 2012; Heide et al. 2014; Wu et al. 2012] solve the system of linear equations directly, but they make the same assumptions on visibility and normals as the backprojection algorithm. The inverse problem of recovering hidden albedos

can be expressed as

$$\underset{\boldsymbol{\rho}}{\text{minimize}} \quad \|\boldsymbol{\tau} - \mathbf{A}\boldsymbol{\rho}\|_2^2 + \Gamma(\boldsymbol{\rho}), \quad \text{s.t. } 0 \leq \boldsymbol{\rho} \quad (4)$$

Although the nonnegativity constraints were not directly enforced by all previous proposals, including it in the reconstruction can improve the estimated solution. An additional prior on the albedos $\Gamma(\boldsymbol{\rho})$ can help further improve the estimated albedos. For example, Heide et al. [2014] used a combination of sparseness and sparse gradients (i.e., total variation). The runtime and memory requirements for an iterative solver are in the same order as those of the backprojection method per iteration.

3.3.3 Factorized Light Transport. Assuming that neither the hidden albedos $\boldsymbol{\rho}$, surface normals \mathbf{n} , or visibility terms \mathbf{V} are known, inverting Equation 3 becomes a nonlinear inverse problem with the cost function

$$\underset{\boldsymbol{\rho}, \mathbf{n}, \mathbf{V}}{\text{minimize}} \quad \|\boldsymbol{\tau} - (\mathbf{T} \circ \mathcal{I}(\mathbf{N} \circ \mathbf{V}))\boldsymbol{\rho}\|_2^2 + \Gamma(\boldsymbol{\rho}). \quad (5)$$

$$\text{s.t. } 0 \leq \mathbf{V} \leq 1, \quad 0 \leq \boldsymbol{\rho}$$

An important insight for solving Equation 5 efficiently is that although the cost function is nonlinear, it is tri-convex when the prior Γ is convex. As is standard practice for multi-convex problems, we use an alternating least-squares (ALS) approach. To this end, Equation 5 is solved in an alternating manner by fixing two of the unknown terms and optimizing for the third. Each of these subproblems is convex; the method is outlined in Algorithm 1.

Algorithm 1 Triconvex Factorization via Alternating Least Squares

- 1: $\mathbf{V}^{(0)} = \mathbf{1}, \mathbf{N}^{(0)} = \mathbf{N}_{\text{iso}}, \boldsymbol{\rho}^{(0)} = \mathbf{0}$
 - 2: **for** $k = 1$ **to** K
 - 3: $\boldsymbol{\rho}^{(k)} \leftarrow \arg \min_{0 \leq \boldsymbol{\rho}} \left\| \boldsymbol{\tau} - (\mathbf{T} \circ \mathcal{I}(\mathbf{N}^{(k-1)} \circ \mathbf{V}^{(k-1)}))\boldsymbol{\rho} \right\|_2^2 + \Gamma_{\boldsymbol{\rho}}(\boldsymbol{\rho})$
 - 4: $\mathbf{V}^{(k)} \leftarrow \arg \min_{0 \leq \mathbf{V} \leq 1} \left\| \boldsymbol{\tau} - (\mathbf{T} \circ \mathcal{I}(\mathbf{N}^{(k-1)} \circ \mathbf{V}))\boldsymbol{\rho}^{(k)} \right\|_2^2$
 - 5: $\mathbf{n}^{(k)} \leftarrow \arg \min_{\mathbf{n}} \left\| \boldsymbol{\tau} - (\mathbf{T} \circ \mathcal{I}(\mathbf{N} \circ \mathbf{V}^{(k)}))\boldsymbol{\rho}^{(k)} \right\|_2^2 + \Gamma_{\mathbf{n}}(\mathbf{n})$
 - 6: **end for**
-

We initialize the unknowns with full visibility, null albedo, and we use isotropic normals \mathbf{N}_{iso} in the first two substeps (until we solve for $\mathbf{n}^{(1)}$). Warmstarting the albedo vector with the backprojection estimate improves the overall runtime of the algorithm.

Updating $\boldsymbol{\rho}$ (Alg. 1, line 3). In this subproblem, the system matrix $\mathbf{A} = \mathbf{T} \circ \mathcal{I}(\mathbf{N}^{(k-1)} \circ \mathbf{V}^{(k-1)})$ is fixed for a given iteration k . The resulting inverse problem is similar to that of Equation 4; we use the alternating direction method of multipliers (ADMM) [Boyd et al. 2011] to solve it.

Updating \mathbf{V} (Alg. 1, line 4). This subproblem is also convex because we could construct a system matrix that absorbs $\mathbf{T}, \mathbf{N}, \mathcal{I}$ and $\boldsymbol{\rho}$ and write the image formation as a matrix-vector multiplication. Unfortunately, the size of this problem is very large – the system matrix would have $N^3 \times N^5$ non-zero elements. Thus, we solve this subproblem using a projected gradient algorithm that minimizes

the objective function $J = \|\tau - (\mathbf{T} \circ \mathcal{I}(\mathbf{N} \circ \mathbf{V})) \rho\|_2^2$ while enforcing $0 \leq \mathbf{V} \leq 1$. For this algorithm, we simply take a step into the direction of the negative gradient of J and clamp the result to the feasible range, i.e., between 0 and 1, using the projection operator Π :

$$\begin{aligned} \mathbf{V} &\leftarrow \Pi(\mathbf{V} - \alpha \nabla_{\mathbf{V}} J), \quad \text{with} \\ \nabla_{\mathbf{V}} J &= -2\mathbf{N} \circ \mathcal{I}^T \left(\left((\tau - (\mathbf{T} \circ \mathcal{I}(\mathbf{N} \circ \mathbf{V})) \rho) \rho^T \right) \circ \mathbf{T} \right) \end{aligned} \quad (6)$$

Here, α is the step length. In practice, this intuitive method can be improved using an adaptive step length which changes per iteration. We derive this in detail in the Supplemental Material.

Updating \mathbf{n} (Alg. 1, line 5). Using the spherical coordinate representation of the normals, this subproblem can be solved with an unconstrained nonlinear solver to minimize the objective function J with respect to $\mathbf{n}^u, \mathbf{n}^v$. We chose the L-BFGS algorithm for this task and, using the chain rule, derive the gradient of the objective as

$$\nabla_{\mathbf{n}(u,v)} J = \nabla_{\mathbf{N}} J \cdot \nabla_{\mathbf{n}(u,v)} \mathbf{N}, \quad \text{with} \quad (7)$$

$$\begin{aligned} \nabla_{\mathbf{N}} J &= -2\mathbf{V} \circ \mathcal{I}^T \left(\left((\tau - (\mathbf{T} \circ \mathcal{I}(\mathbf{N} \circ \mathbf{V})) \rho) \rho^T \right) \circ \mathbf{T} \right), \\ \nabla_{\mathbf{n}(u,v)} \mathbf{N}_{ij} &= \begin{bmatrix} -\omega_{j \rightarrow i}^x \sin(\mathbf{n}_j^u) \sin(\mathbf{n}_j^v) + \omega_{j \rightarrow i}^y \cos(\mathbf{n}_j^u) \sin(\mathbf{n}_j^v) \\ \omega_{j \rightarrow i}^x \cos(\mathbf{n}_j^u) \cos(\mathbf{n}_j^v) - \omega_{j \rightarrow i}^y \sin(\mathbf{n}_j^u) \cos(\mathbf{n}_j^v) - \omega_{j \rightarrow i}^z \sin(\mathbf{n}_j^v) \end{bmatrix}^T \end{aligned}$$

4 EVALUATION

In this section, we evaluate the proposed factorization method in simulation and show detailed comparisons to other non-line-of-sight reconstruction methods. Figure 4 shows maximum intensity projections generated with the Chimera volume renderer¹ for one of the scenes we used for this evaluation. Several additional results can be found in the Supplemental Material. This scene contains two hidden objects that partly occlude each other (row 1, column 1). The measured transient image is simulated for 64×64 sampling locations x', y' over an area of $1 \text{ m} \times 1 \text{ m}$ on the visible wall with a temporal bin size of 16 ps (row 1, column 2). Although we simulate time-resolved light transport that includes direct and indirect illumination, we visualize only the indirect components in Figure 4.

Rows 2 and 3 of Figure 4 show two different perspectives of the reconstructions obtained with the following algorithms: backprojection, filtered backprojection as described by Velten et al. [2012], the linear method with *identical regularization parameters* as those used in the proposed factorization method, the proposed factorization method, and a reference solution obtained by fixing the ground truth visibility term and applying the linear method to recover only the hidden albedos. The latter represents an upper bound on the reconstruction quality that can be achieved with the full factorization method, where the visibility is unknown and also needs to be estimated. Whereas previously-proposed algorithms fail in recovering the partially occluded scene parts, our factorization accurately estimates this challenging scene; our solution closely matches the reference solution.

The linear and proposed methods use identical priors and regularization weights. Moreover, in the Supplemental Material, we

¹<http://www.cgl.ucsf.edu/chimera/>

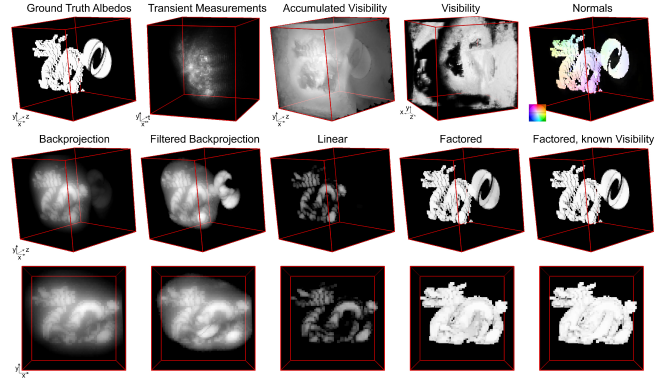


Fig. 4. Evaluation of reconstruction algorithms using the “Dragon & Logo” scene. The hidden scene contains two objects that partially occlude each other. Time-resolved measurements are shown along with reconstructions obtained with several algorithms. Previously-proposed algorithms, including backprojection, filtered backprojection, and the linear method with regularization parameters identical to the ones used in the proposed method fail in adequately recovering partially occluded scene parts, such as the logo in the background. The proposed factorization method accurately estimates this challenging scene via the visibility term and it additionally recovers surface normals. Slight variations in intensity, for example for the logo between backprojection and filtered backprojection, are introduced by the volume rendered used to generate these visualizations.

have included additional results of the linear method with the regularization parameters suggested by Heide et al. [2014], and we also show results of the backprojection method followed by the same regularization used by the proposed method. These supplemental results validate the claim that the improved recovery performance stems from the factorization method and not the choice of priors.

In addition to the hidden albedos, the proposed factorization method also recovers a visibility term (row 1, columns 3–4) and the surface normals of the scene (row 1, column 5). The estimated visibility term contains a full 3D volume of values for each of the 64×64 sampling locations on the wall. We show one of these volumes for a single sampling point (row 1, column 4) as well as the average visibility term for all sampling locations (row 1, column 3). The estimated visibility terms can be interpreted as an intermediate variable that helps improve the estimated albedos in the presence of partial occlusions.

Table 1 shows quantitative comparisons of the estimated albedos. For this purpose, we simulate a set of scenes with isotropic BRDFs without surface normals (Tab. 1, top part) and a set of scenes with Lambertian BRDFs that include surface normals (Tab. 1, bottom part). Again, the filtered backprojection (FBP) method uses the filter described by Velten et al. [2012]. The regularized backprojection (RBP) applies a regularized least-squares projection of the output of the BP method on an identical regularizer as used for the linear (Lin) and the proposed methods. With RBP, regularization can only be applied as a two-step process, i.e. BP followed by regularization, whereas it is jointly solved with the regularized linear method. We list reconstruction fidelity using the peak-signal-to-noise ratio (PSNR) metric. In all cases, the proposed factorization approach

	BP	FBP	RBP	Lin	Factored	Lin with known V
Bunny	15.8	13.3	17.8	26.6	40.8	46.1
Dragon	14.4	8.8	15.9	17.6	19.5	19.8
Dragon & Logo	14.5	10.3	16.5	23.7	38.7	42.3
Dragon & Bunny	14.3	10.3	16.1	22.1	31.6	37.0
Logo	20.9	13.2	24.5	41.1	57.9	62.2
2 Planes	19.7	13.4	22.6	29.9	40.5	40.6
Plane & Logo	19.9	12.1	23.1	33.3	48.2	49.8
Average	17.1	11.6	19.5	27.8	39.6	42.7

	BP	FBP	RBP	Lin	Factored	Lin with known N+V
Bunny	13.5	11.1	15.5	25.2	34.4	41.9
Dragon	13.1	8.3	14.9	17.4	18.7	19.7
Dragon & Logo	13.6	9.4	15.5	24.3	30.9	46.3
Dragon & Bunny	13.3	9.5	15.1	22.3	27.3	40.1
Logo	17.8	9.5	21.3	39.1	52.3	63.9
2 Planes	16.4	11.0	19.1	28.0	37.6	53.1
Plane & Logo	16.1	9.3	19.2	31.5	43.4	60.9
Spheres	14.8	10.1	17.3	24.5	28.5	34.9
Average	14.8	9.8	17.2	26.5	34.1	45.1

Table 1. Quantitative evaluation of estimated albedos for various NLOS reconstruction algorithms: backprojection (BP), filtered backprojection (FBP), regularized backprojection (RBP), the linear estimation method (Lin) with identical parameters to the proposed method, and the proposed factorization method. We compare these algorithms for scenes with isotropic BRDFs (top 7 scenes) and scenes with Lambertian BRDFs, including surface normals (bottom 8 scenes). As a reference solution, we also apply the linear method with fixed ground truth visibility term (top) and visibility as well as normals (bottom); these values represent an upper bound on what quality can be achieved. All values are reported as peak signal-to-noise ratio (PSNR) in dB.

results in the highest PSNR. The solutions Lin w/ V and Lin w/ N+V apply the linear method with ground truth visibility fixed, and with ground truth visibility and normals fixed. Due to the fact that the resulting problems are convex, these values can be interpreted as the reference solution representing an upper bound on what PSNR could be achieved.

Table 2 shows a quantitative evaluation of the estimated point clouds. For this purpose, all albedo volumes are thresholded by discarding all voxels with an albedo below 0.1. The geometric error between estimated point clouds and ground truth is then calculated as the average Hausdorff distance. This metric can be interpreted as measuring the fidelity of the estimated geometry rather than the albedo. We show visualizations of these point clouds in the Supplemental Material.

5 VALIDATION WITH PROTOTYPE

5.1 SPAD-based Imaging System

Our prototype uses a single photon avalanche diode (SPAD) and a pulsed picosecond laser. We summarize the hardware components, calibration procedure, and acquisition parameters in the following.

5.1.1 Hardware. We use a Micro Photon Devices SPAD with a $100 \times 100 \mu\text{m}$ active area. Light is focused on the detector using a

	BP	FBP	RBP	Lin	Factored	Lin with known V
Bunny	271	263	274	278	23	8
Dragon	413	233	270	243	103	85
Dragon & Logo	295	180	235	298	29	23
Dragon & Bunny	287	412	371	199	54	23
Logo	222	91	136	29	8	8
2 Planes	202	116	135	129	23	23
Plane & Logo	214	139	149	56	23	23
Average	272	205	224	176	38	28

	BP	FBP	RBP	Lin	Factored	Lin with known N+V
Bunny	275	211	222	211	29	15
Dragon	326	211	255	166	99	72
Dragon & Logo	319	199	291	264	42	15
Dragon & Bunny	307	448	417	194	63	8
Logo	222	96	138	29	8	8
2 Planes	217	117	150	129	39	15
Plane & Logo	243	139	165	56	23	15
Spheres	275	159	190	129	56	15
Average	273	198	229	147	45	20

Table 2. Quantitative evaluation of estimated point clouds for various NLOS reconstruction algorithms. We compare the average Hausdorff distance in mm between ground truth and the estimated and thresholded point clouds (lower number is better). This metric represents the per-point depth resolution of the reconstructions. The reported point cloud distances demonstrate that the proposed method outperforms competing approaches in the quality of the recovered geometry, independently of the object albedo.

75 mm achromatic doublet lens (Thorlabs AC254-075-A-ML). Photon arrival times are time-stamped with a PicoHarp 300 Time-Correlated Single Photon Counting (TCSPC) module and stored in histograms of photon counts with a 4 ps bin width. The laser is an ALPHALAS PICOPOWER-LD-670-50, which operates at 670 nm wavelength pulsed with a reported pulse width of 30.6 ps at a 10 MHz repetition rate and 0.11 mW average power. The laser and SPAD are co-axially aligned using a polarizing beam splitter cube (Thorlabs PBS251). The aligned optical path is then scanned over the visible wall using a 2-axis scanning galvanometer (Thorlabs GVS012) from a distance of about 1.5 m from the wall. The combined temporal jitter of SPAD and laser pulse width is measured to be approximately 60 ps. An illustration of the setup and measurements are shown in Figure 5.

5.1.2 Calibration. Aligning the laser and SPAD is done by adjusting the beam splitter position and tilt angle to maximize the recorded photon counts of the light directly reflected off the wall. The SPAD is operated in free-running mode, which creates an effect known as *pileup*. Pileup is basically a masking effect that makes it difficult to see weak signals that occur right after a strong signal in the temporal histogram. To avoid masking the weak indirect reflections with the strong contribution of the direct reflections, we slightly misalign SPAD and laser by moving the beam splitter until we can see both direct and indirect contributions (cf. Fig. 5, lower right). The confocal image formation model is not affected by this procedure. Alternatively, a temporal gating mechanism could

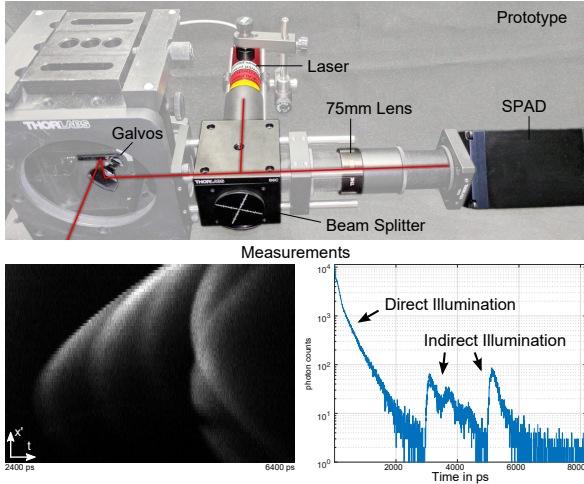


Fig. 5. Prototype and measurements. The prototype (top) consists of a single-photon avalanche diode (SPAD), a pulsed laser, a 2-axis scanning galvanometer, and a beam splitter that combines the optical path of the laser and the SPAD (red line). With this setup, we scan a 2D array of sampling locations capturing a temporal histogram of photon counts (lower right) at each location. A spatio-temporal slice of these measurements is shown (lower left), with only indirect illumination inside the displayed area.

be employed to remove the direct light; we did not have access to hardware with this capability for our experiments.

To account for the differences in path length of different samples on the wall and the imaging system, we align the measured histograms in software such that the peak of the direct light appears at time $t = 0$. Then, we reduce the histograms to a bin size of 16 ps using area downsampling and remove the direct light for further processing by setting the first 600 time bins to 0.

In total we captured six scenes. To keep the acquisition times manageable, five of these scenes contain retroreflective objects and only one scene contains a diffuse object. These scenes are recorded with 64×64 sample points spaced as an equidistant grid on a visible, white planar surface. The sampling points of the retroreflective scenes cover an area of 80×80 cm of the visible wall; the exposure time for each of the samples is 0.1 s; reconstructed volumes of hidden surface albedos have a resolution of $64 \times 64 \times 120$ voxels and cover $80 \times 80 \times 80$ cm. The diffuse scene was sampled over an area of 70×70 cm with an exposure time of 1 s per sample. All methods compared in this work are processed at the same resolution.

For the reconstructions with the linear method, we run 150 iterations with a weight of 0.1 on the sparsity prior and 0.001 on the total variation (TV) prior. The proposed factorization method uses 5 ADMM iterations in total and 20 iterations for the linear method in each of the ADMM iterations. Our source code is implemented as unoptimized MATLAB code and takes about 2 h per scene on a server with an Intel Xeon E5-4620 (2.20 GHz) and 768 GB RAM. Please see the Sec. 6 for a more detailed discussion on potential approaches to improving runtime and memory requirements.

5.2 Experimental Results

Figure 6 shows two of the experimental data sets; all six are shown in the Supplemental Material. We show 3D maximum intensity projections of the acquired measurements and photographs of the imaged objects along with two different perspectives of the reconstructions obtained with several different methods. As expected from our simulations, the backprojection and filtered backprojection methods give a rough idea of the shape of hidden objects but fine geometric detail is missing. The linear method with regularization weights matching the ones of the proposed method achieves significantly better results. The Light Cone Transform (LCT) was recently proposed as a computationally efficient method for NLOS imaging [O’Toole et al. 2018], but it uses similar assumptions on the image formation model as the linear method and thus achieves similar reconstruction quality for low regularization. The proposed factorization approach recovers the scenes exhibiting partial occlusions most accurately among all compared methods. For the “Diffuse S”, our factorization method achieves a better quality than the other methods. The primary benefit of the proposed method for NLOS scenes containing isolated objects without substantial amounts of occlusions is that we can estimate surface normals. As expected for this planar scene, the estimated normals mostly point towards the scanned wall.

For the retroreflective “Sign & S” scene, the proposed factorization method amplifies the shape of the “S” compared to other reconstruction algorithms. This is mostly due to the fact that the image formation model of the other methods does not adequately model partial occlusions in the hidden scenes, but our method does and thus results in a more accurate reconstruction. Finally, there is also ambiguity in the factored light transport representation. An object with very low reflectivity (i.e., small albedo) that is unobstructed could produce the same measurements as the same object with a larger albedo but appropriately down-scaled visibility terms. This ambiguity may boost the brightness of certain objects but could be mitigated by placing additional constraints on the visibility terms (see Sec. 6 for discussion).

6 DISCUSSION

In summary, we propose a novel light transport representation for non-line-of-sight imaging along with inverse methods that factor the global illumination components of a transient image of a visible surface into a volume of hidden albedos, surface normals, and visibility terms. The visibility terms model partial occlusions in the hidden scene parts. Our simulations indicate that the proposed factorization approach has the potential to improve the robustness and quality of NLOS imaging for complex scenes that often include partial occlusions. With experimentally-captured data we show similar trends, but also reveal that the non-convex nature of the factored light transport model can result in ambiguities in the reconstructions. Overall, we demonstrate that the reconstruction quality of the proposed method is substantially higher than what is achieved by other methods. In addition, the proposed approach is the first to facilitate surface normal estimation. Whereas normals could alternatively be fitted to the estimated point clouds, the proposed method jointly estimates these two unknowns directly from the measured data.

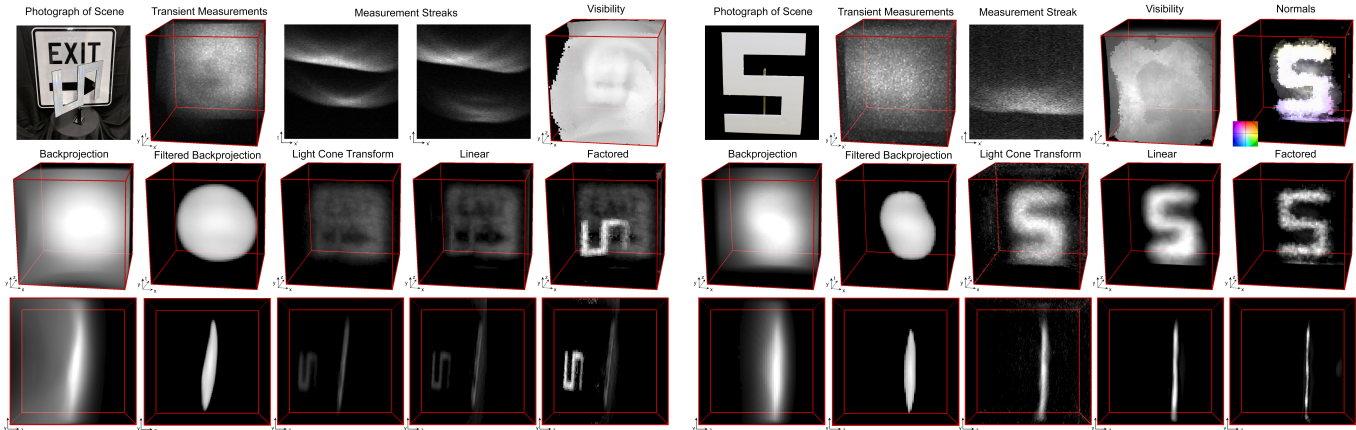


Fig. 6. Experimental results of the “Sign & S” (left) and the “Diffuse S” (right) scenes. The left scene contains two retroreflective objects that partly occlude each other. We show measurements and compare the reconstruction quality of the backprojection method, filtered backprojection, the linear method, the Light Cone Transform, and the proposed factorization method. For this example, the proposed method achieves the highest image quality. In the absence of partial occlusions in the hidden scene, all methods, except for the backprojection, achieve a comparable quality. Unlike other methods, the factorization also estimates surface normals of the hidden scene (top right).

Matrix-free implementations of backprojection-type methods require $O(N^3)$ memory whereas an implementation that uses a sparse matrix representation of the light transport requires $O(N^5)$ memory. The linear method has similar memory requirements but a significantly increased computational cost. The Light Cone Transform provides a closed-form solution to NLOS imaging that achieves a similar quality as the linear method with only $O(N^3 \log N)$ memory. The computational cost of the proposed factorization is slightly higher than the linear method, but it is in the same order of magnitude. However, the fact that the factored light transport representation requires the visibility terms to be stored, which are generally not sparse, results in memory requirements of $O(N^5)$ making our method two orders of magnitude more memory demanding than other matrix-free NLOS algorithms.

6.1 Future Work

In the future, the memory requirements of the proposed method could be mitigated by parameterizing the visibility terms as a separate density volume with N^3 voxels. Inspired by techniques used in the volume rendering community [Max 1995], such a representation would represent the hidden scene using a density value as well as an albedo for each voxel. Thus, memory requirements could be reduced by two orders of magnitude to $O(N^3)$. While such a parametrization is a promising direction to further improve the computational efficiency of the proposed method, we leave this effort to future work.

The runtime of our algorithm can be improved with a parallel implementation on modern graphics processing units (GPUs). We have implemented a preliminary GPU-based solver. This implementation achieves a speedup of $5\times$ compared to our Matlab code at a resolution of $40 \times 40 \times 75$ voxels (matching the resolution reported by Heide et al. [2014]) on a notebook computer (Intel 2.4GHZ, 16GB,

180 s total reconstruction time). However, the GPU solver is fundamentally limited by available memory resources. Scaling this fast solver to operate on the same volume resolutions used throughout this paper would require the memory management to be improved, for example with the approach discussed above.

To improve the prototype hardware, we would like to replace our laser with a more powerful option to reduce acquisition times and improve the quality of hidden objects with Lambertian BRDFs. Developing new inverse methods or learning them with a data-driven approach are additional directions for future research. Finally, similar to other NLOS imaging approaches, we assume that only third-order light bounces contribute to the image formation. This assumption could be lifted in future work to account for diffuse interreflections and higher-order light transport effects in hidden scene parts.

6.2 Conclusion

Non-line-of-sight imaging is a promising technology that has the potential to unlock unprecedented imaging modalities for a variety of applications. Recent advances in single-photon detector technology and large-scale inverse light transport algorithms have demonstrated that NLOS imaging is feasible in certain conditions. With the proposed methods, we lift several important restrictions of previous algorithms and take steps towards making NLOS imaging more robust. Yet, further research and development is needed to enable NLOS imaging “in the wild”, i.e. with strong ambient light, at fast acquisition rates, and for more complex scenarios.

ACKNOWLEDGEMENTS

The authors thank James Harris for fruitful discussions. D.B.L. is supported by a Stanford Graduate Fellowship in Science and Engineering. G.W. is supported by a Terman Faculty Fellowship and

a Sloan Fellowship. Additional funding was generously provided by the National Science Foundation (CAREER Award IIS 1553333), the DARPA REVEAL program, the ARO (Grant W911NF-19-1-0120), and by the KAUST Office of Sponsored Research through the Visual Computing Center CCF grant.

REFERENCES

- Fadel Adib, Chen-Yu Hsu, Hongzi Mao, Dina Katabi, and Frédo Durand. 2015. Capturing the Human Figure Through a Wall. *ACM Trans. Graph. (SIGGRAPH Asia)* 34, 6 (2015).
- Yoann Altmann, Stephen McLaughlin, Miles J. Padgett, Vivek K Goyal, Alfred O. Hero, and Daniele Faccio. 2018. Quantum-inspired computational imaging. *Science* 361 (2018). Issue 6403.
- Victor Arellano, Diego Gutierrez, and Adrian Jarabo. 2017. Fast back-projection for non-line of sight reconstruction. *Opt. Express* 25, 10 (2017), 11574–11583.
- Katherine Bouman, Vickie Ye, Adam Yedidia, Fredo Durand, Gregory W Wornell, Antonio Torralba, and William T Freeman. 2017. Turning Corners into Cameras: Principles and Methods. In *Proc. CVPR*. 2270–2278.
- Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. 2011. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Found. Trends Mach. Learn.* 3, 1 (2011), 1–122.
- Samuel Burri, Harald Homulle, Claudio Bruschini, and Edoardo Charbon. 2016. LinoSPAD: a time-resolved 256x1 CMOS SPAD line sensor system featuring 64 FPGA-based TDC channels running at up to 8.5 giga-events per second. *Proc. SPIE* 9899 (2016), 98990D–10.
- Mauro Buttafava, Jessica Zeman, Alberto Tosi, Kevin Eliceiri, and Andreas Velten. 2015. Non-line-of-sight imaging using a time-gated single photon avalanche diode. *Opt. Express* 23, 16 (2015).
- Susan Chan, Ryan E. Warburton, Genevieve Garipey, Jonathan Leach, and Daniele Faccio. 2017. Non-line-of-sight tracking of people at long range. *Opt. Express* 25, 9 (2017), 10109–10117.
- Hernandez, Quercus and Gutierrez, Diego and Jarabo, Adrian. 2017. A Computational Model of a Single-Photon Avalanche Diode Sensor for Transient Imaging. *arXiv preprint arXiv:1703.02635* (2017).
- Pediredla, Adithya Kumar and Buttafava, Mauro and Tosi, Alberto and Cossairt, Oliver and Veeraraghavan, Ashok. 2017. Reconstructing rooms using photon echoes: A plane based model and reconstruction algorithm for looking around the corner. In *Proc. ICCP*. 1–12.
- Genevieve Garipey, Nikola Krstajic, Robert Henderson, Chunyong Li, Robert R. Thomson, Gerald S. Buller, Barmak Heshmat, Ramesh Raskar, Jonathan Leach, and Daniele Faccio. 2015. Single-photon sensitive light-in-flight imaging. *Nature Communications* 6, 6021 (2015).
- Genevieve Garipey, Francesco Tonolini, Robert Henderson, Jonathan Leach, and Daniele Faccio. 2016. Detection and tracking of moving objects hidden from view. *Nature Photonics* 10 (2016).
- Otkrist Gupta, Thomas Willwacher, Andreas Velten, Ashok Veeraraghavan, and Ramesh Raskar. 2012. Reconstruction of hidden 3D shapes using diffuse reflections. *OSA Opt. Express* 20, 17 (2012).
- Felix Heide, Steven Diamond, David B. Lindell, and Gordon Wetzstein. 2018. Sub-picosecond photon-efficient 3D imaging using single-photon sensors. *Scientific Reports* 8 (2018).
- Felix Heide, Lei Xiao, Wolfgang Heidrich, and Matthias B. Hullin. 2014. Diffuse Mirrors: 3D Reconstruction from Diffuse Indirect Illumination Using Inexpensive Time-of-Flight Sensors. In *Proc. CVPR*.
- Adrian Jarabo, Julio Marco, Adolfo Muñoz, Raul Buisan, Wojciech Jarosz, and Diego Gutierrez. 2014. A Framework for Transient Rendering. *ACM Trans. Graph.* 33, 6, Article 177 (Nov. 2014), 10 pages. <https://doi.org/10.1145/2661229.2661251>
- Achuta Kadambi, Rafael Whyte, Ayush Bhandari, Lee Streeter, Christopher Barsi, Adrian Dorrington, and Ramesh Raskar. 2013. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Trans. Graph. (SIGGRAPH Asia)* 32, 6 (2013).
- Avinash Kak and Malcolm Slaney. 1988. *Principles of Computerized Tomographic Imaging*. IEEE Press.
- Ori Katz, Pierre Heidmann, Mathias Fink, and Sylvain Gigan. 2014. Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations. *Nature Photonics* 8, 10 (2014).
- Ahmed Kirmani, Tyler Hutchison, James Davis, and Ramesh Raskar. 2009. Looking around the corner using transient imaging. *Proc. ICCV* (2009), 159–166.
- Ahmed Kirmani, Dheera Venkatraman, Donggeek Shin, Andrea Colaco, Franco N. C. Wong, Jeffrey H. Shapiro, and Vivek K Goyal. 2014. First-Photon Imaging. *Science* 343, 6166 (2014), 58–61.
- Jonathan Klein, Christoph Peters, Jaime Martin, Martin Laurenzis, and Matthias B. Hullin. 2016. Tracking objects outside the line of sight using 2D intensity images. *Scientific Reports* (2016), 177–188.
- David B. Lindell, Matthew O’Toole, and Gordon Wetzstein. 2018a. Single-Photon 3D Imaging with Deep Sensor Fusion. *ACM Trans. Graph. (SIGGRAPH)* 4 (2018). Issue 37.
- David B. Lindell, Matthew O’Toole, and Gordon Wetzstein. 2018b. Towards Transient Imaging at Interactive Rates with Single-Photon Detectors. *Proc. IEEE ICCP* (2018).
- Nelson Max. 1995. Optical models for direct volume rendering. *IEEE TVCG* 1, 2 (1995), 99–108.
- Aongus McCarthy, Nils J. Krichel, Nathan R. Gemmell, Ximing Ren, Michael G. Tanner, Sander N. Dorenbos, Val Zwiller, Robert H. Hadfield, and Gerald S. Buller. 2013. Kilometer-range, high resolution depth imaging via 1560 nm wavelength single-photon detection. *Opt. Express* 21, 7 (2013), 8904–8915.
- Desmond O’Connor. 2012. *Time-correlated single photon counting*. Academic Press.
- Matthew O’Toole, Felix Heide, David Lindell, Kai Zang, Stephen Diamond, and Gordon Wetzstein. 2017. Reconstructing Transient Images from Single-Photon Sensors. *Proc. CVPR* (2017).
- Matthew O’Toole, David B. Lindell, and Gordon Wetzstein. 2018. Confocal Non-line-of-sight imaging based on the light cone transform. *Nature* (2018), 338–341. Issue 555.
- Joshua Rapp and Vivek K. Goyal. 2017. A Few Photons Among Many: Unmixing Signal and Noise for Photon-Efficient Active Imaging. *IEEE Trans. Computational Imaging* 3, 3 (2017), 445–459.
- Albert Redo-Sanchez, Barmak Heshmat, Alireza Aghasi, Salman Naqvi, Mingjie Zhang, Justin Romberg, and Ramesh Raskar. 2016. Terahertz time-gated spectral imaging for content extraction through layered structures. *Nature Communications* 7 (2016).
- Brent Schwarz. 2010. LIDAR: Mapping the world in 3D. *Nat. Photonics* 4, 7 (2010), 1749–4885.
- Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik P. A. Lensch. 2005. Dual Photography. *ACM Trans. Graph. (SIGGRAPH)* 24, 3 (2005), 745–755.
- Donggeek Shin, Feihu Xu, Dheera Venkatraman, Rudi Lussana, Federica Villa, Franco Zappa, Vivek K. Goyal, Franco N. C. Wong, and Jeffrey H. Shapiro. 2016. Photon-efficient imaging with a single-photon camera. *Nature Communications* 7, 12046 (2016), 1–7.
- Christos Thrampoulidis, Gal Shulkind, Feihu Xu, William T. Freeman, Jeffrey H. Shapiro, Antonio Torralba, Franco N. C. Wong, and Gregory W. Wornell. 2017. Exploiting Occlusion in Non-Line-of-Sight Active Imaging. (2017). [arXiv:1711.06297](https://arxiv.org/abs/1711.06297)
- Chia-Yin Tsai, Kiriakos N. Kutulakos, Srinivasa G. Narasimhan, and Aswin C. Sankaranarayanan. 2017. The Geometry of First-Returning Photons for Non-Line-of-Sight Imaging. In *Proc. CVPR*. 1–9.
- Andreas Velten, Thomas Willwacher, Otkrist Gupta, Ashok Veeraraghavan, Mounsi G. Bawendi, and Ramesh Raskar. 2012. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Communications* 3 (2012).
- Di Wu, Gordon Wetzstein, Christopher Barsi, Thomas Willwacher, Matthew O’Toole, Nikhil Naik, Qionghai Dai, Kyros Kutulakos, and Ramesh Raskar. 2012. Frequency analysis of transient light transport with applications in bare sensor imaging. In *Proc. ECCV*.
- Feihu Xu, Gal Shulkind, Christos Thrampoulidis, Jeffrey H. Shapiro, Antonio Torralba, Franco N. C. Wong, and Gregory W. Wornell. 2018. Revealing hidden scenes by photon-efficient occlusion-based opportunistic active imaging. *OSA Opt. Express* 26, 8 (2018), 9945–9962.