

Beyond Volumetric Albedo

— A Surface Optimization Framework for Non-Line-of-Sight Imaging

Chia-Yin Tsai, Aswin C. Sankaranarayanan, and Ioannis Gkioulekas
Carnegie Mellon University

Abstract

Non-line-of-sight (NLOS) imaging is the problem of reconstructing properties of scenes occluded from a sensor, using measurements of light that indirectly travels from the occluded scene to the sensor through intermediate diffuse reflections. We introduce an analysis-by-synthesis framework that can reconstruct complex shape and reflectance of an NLOS object. Our framework deviates from prior work on NLOS reconstruction, by directly optimizing for a surface representation of the NLOS object, in place of commonly employed volumetric representations. At the core of our framework is a new rendering formulation that efficiently computes derivatives of radiometric measurements with respect to NLOS geometry and reflectance, while accurately modeling the underlying light transport physics. By coupling this with stochastic optimization and geometry processing techniques, we are able to reconstruct NLOS surface at a level of detail significantly exceeding what is possible with previous volumetric reconstruction methods.

1. Introduction

Non-line-of-sight (NLOS) imaging is an emerging technology that concerns with using higher-order light transport in order to reconstruct properties of a scene that is outside the direct line of sight of a sensor. A common setting is the so-called “looking around the corner” problem [41, 77] (see Figure 1), where information about an NLOS object (geometry, reflectance, motion, class label, and other properties) is extracted from measurements of photons that bounce between a visible wall and the object. This technology has seen rapid advances in the past decade, as several active [68, 56, 40, 74, 13, 73, 81, 36, 42, 63, 48, 61, 50, 80] and passive [8, 66, 6, 10, 5] techniques have been introduced that can operate under progressively more challenging conditions (ambient lighting, real-time capture, and so on).

We are focusing on the problem of shape reconstruction in the looking-around-the-corner setting using active illumination. Typically, active techniques use a controllable

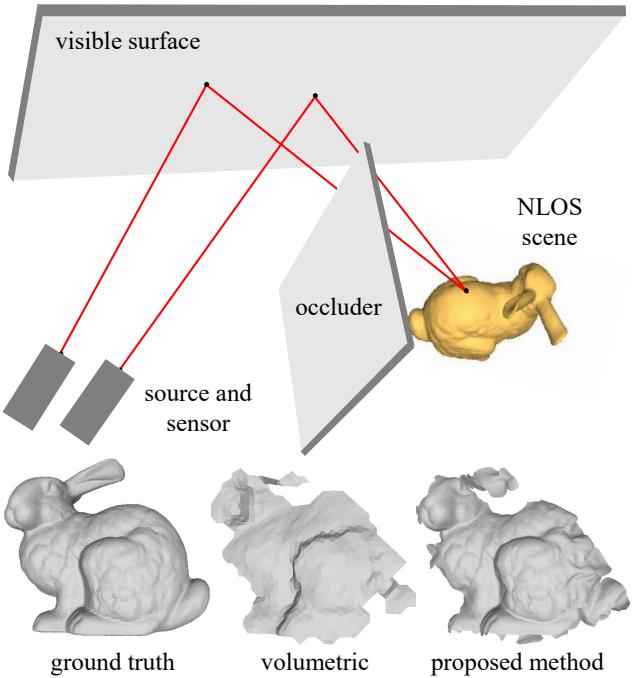


Figure 1. **Looking around the corner:** Non-line-of-sight (NLOS) imaging is the problem of reconstructing parts of a scene occluded from a sensor, by analyzing light that bounces multiple times between visible and occluded surfaces. We develop an inverse rendering pipeline that uses an accurate radiometric image formation model to produce detailed NLOS surface reconstructions.

source, such as a laser beam, to indirectly inject light into the NLOS scene, through a reflection on the visible wall. Then, they use time-resolved, or *transient*, intensity measurements [33] to reconstruct the NLOS scene.

Most existing techniques perform 3D reconstruction using an image formation model introduced by Velten *et al.* [77]. This model represents the NLOS scene as an *albedo volume*, where each voxel is an isotropic reflector with an associated albedo value. This representation allows *approximately* formulating transient light transport in the NLOS scene with only linear algebraic operations. In turn, this allows recovering the unknown albedo volume from the

transient measurements by solving a, potentially regularized, linear least-squares system [31, 29, 56, 30, 3, 48, 50].

This mathematical tractability comes at the cost of physical accuracy: In NLOS scenes consisting of opaque objects, light transport is the result of discrete light-surface interactions at the object interfaces, rather than continuous light-volume interactions. Additionally, these interactions include effects such as normal-dependent shading and non-Lambertian reflectance, which are ignored by the albedo volume. On the other hand, instead of volumetric albedo, given a representation of the NLOS objects' surface and reflectance, light transport can be modeled exactly using the rendering equation [35]. However, unlike the albedo volume model, evaluating this equation is only possible through computationally expensive Monte Carlo rendering operations [75, 64, 18]. This increased computational complexity has so far hindered the adoption of the rendering equation in NLOS reconstruction techniques.

In this paper, we overcome the computational complexity and introduce a computational pipeline that reconstructs NLOS object shape, in the form of a triangular mesh, and complex reflectance, in the form of a microfacet BRDF, while accurately taking into account the underlying light transport physics. At the core of our pipeline is a differentiable formulation of the rendering equation in the NLOS setting. This formulation enables the use of Monte Carlo rendering to efficiently estimate derivatives of radiometric measurements with respect to shape and reflectance parameters. We combine an optimized differentiable rendering implementation with stochastic optimization in an *inverse rendering* framework [53], where we iteratively deform an NLOS surface so as to minimize the difference between measured and rendered light transients. We augment this surface optimization pipeline with geometric processing tools that help improve the quality of the resulting triangular mesh. Through experiments on synthetic and measured data, we show that this pipeline can produce NLOS surface reconstructions at a level of detail comparable to what is achieved by albedo-volume methods using two orders of magnitude more measurements, while additionally recovering non-Lambertian reflectance. We will release our optimized implementation in order to encourage adoption of inverse rendering pipelines in NLOS imaging, either as stand-alone reconstruction tools or in conjunction with albedo-volume methods as post-processing procedures.

2. Related Work

Non-line-of-sight imaging refers to the broad problem of reconstructing properties of scenes that are normally occluded from a sensor. Even though interest in this problem dates back several decades [24], it has recently attracted increased attention within computer vision and graphics, following two seminal papers [41, 77] demonstrating the abil-

ity to reconstruct shape in the *looking around the corner* setting (Figure 1). Most of the NLOS imaging techniques that have been introduced since then use active illuminations, with a few notable exceptions [8, 66, 6, 10, 5].

We can broadly classify active NLOS imaging techniques into three categories. First are coherent illumination techniques, which take advantage of speckle statistics to recover information about the NLOS scene [68, 36, 7, 37]. The second category includes techniques that use incoherent intensity measurements, under laser or flash illumination, to recover NLOS motion information [42], semantic labels [71], or in certain cases even geometry [81, 73].

Most relevant to us is the third category of active techniques, which reconstruct NLOS geometry using *transient* intensity measurements [33]. This has been demonstrated using sensing technologies that include ultrafast photodiodes [41], optical coherence tomography [80], streak cameras [77, 29], continuous-wave time-of-flight cameras [31, 34], and single-photon avalanche diodes (SPADs) [13, 56, 30, 45, 3, 74, 63, 61, 50, 80]. Most of these techniques use a volumetric representation of the NLOS scene and an approximate image formation model introduced by Velten et al. [77]. Our point of departure from this line of work is to show that, using a physically-accurate image formation model based on the rendering equation [35], we can reconstruct surfaces, rather than volumes, for the NLOS scene, at higher geometric detail. Compared to other techniques that reconstruct surface representations from the timestamps of specific events in transients [74, 80], we do so by taking into account the complete transient intensity information, which enables us to additionally reconstruct reflectance.

Surface optimization is a classical approach for 3D reconstruction in computer vision, where it is commonly applied for stereo-based reconstruction [2, 21, 22, 25, 16, 69, 83, 15, 82, 65, 44]. In a related context, surface optimization techniques are used in mesh editing applications for computer graphics [17, 54, 70, 20, 49]. At a high-level, both types of applications operate by first defining an objective function (or energy) as an integral on a surface. Then, they derive expressions for the derivatives of this surface integral with respect to some surface representations. Finally, these derivatives are used to create a *flow* process that progressively deforms some initial surface, until the objective function is minimized. The derivation of derivative expressions typically relies on tools from differential geometry, and has been demonstrated for both implicit (e.g., level sets [55]) and parametric (e.g., triangular meshes [16]) surface representations. Similar surface integrals arise in the context of NLOS imaging through the rendering equation. Therefore, we take advantage of this mathematical machinery to perform surface optimization for NLOS reconstruction.

Differentiable rendering has been introduced as a methodology for recovering physical unknowns from image mea-

surements, which can include direct-only [52] and global illumination effects (e.g., scattering [28, 27, 84, 46, 26, 38], or interreflections [51, 47, 4]). Typically, differentiable rendering is used to perform *analysis-by-synthesis*, also known as *inverse rendering* [53, 60, 59]. This refers to the search for values of physical parameters that, when used to synthesize images, can reproduce input image measurements. To efficiently perform this search through gradient-descent optimization, differentiable rendering is used to estimate derivatives of images with respect to the unknown parameters. In our setting, we devise differentiable rendering algorithms that enable surface optimization, and are tailored to the NLOS image formation model for increased efficiency.

3. Problem setting

We focus on the looking-around-the-corner imaging setting, which we describe in detail in this section. Along the line, we introduce relevant notation, and use this notation to write down expressions for the radiometric measurements captured under this setting as a function of properties of the NLOS scene. These expressions are commonplace in the physics-based rendering literature (see, for instance, [75, 64, 18]), but we describe them in detail as the necessary background for deriving the inverse rendering algorithm of Section 4. To help navigate this section, in Figure 2, we visualize in two dimensions the looking-around-the-corner setting and some of our notation.

We use a pulsed source l_0 and a transient detector s_0 to image a scene that consists of two distinct sets of surfaces: surfaces \mathcal{S}_{LOS} that are visible to both the source and detector, and surfaces $\mathcal{S}_{\text{NLOS}}$ that are occluded from both of them. We assume that there are no surfaces that are neither in \mathcal{S}_{LOS} nor in $\mathcal{S}_{\text{NLOS}}$. We additionally assume that the visible surface \mathcal{S}_{LOS} has Lambertian reflectance.

We use the source to illuminate a point l on the visible surface \mathcal{S}_{LOS} . Likewise, we use the detector to image a point s on \mathcal{S}_{LOS} . We call the points l and s the *virtual source* and *virtual detector* respectively. This terminology stems from the fact that these points effectively act as an isotropic source and detector directly attached to \mathcal{S}_{LOS} , as they redirect light, through a diffuse reflection, from the source to the NLOS scene, and from the NLOS scene to the detector.

3.1. Image formation model

We restrict our attention to light effects from so-called *three-bounce paths* of the form $l_0 \rightarrow l \rightarrow x \rightarrow s \rightarrow s_0$ where $x \in \mathcal{S}_{\text{NLOS}}$; that is, paths that, between the virtual source l and virtual detector s , have a single interaction with the NLOS surface at a point $x \in \mathcal{S}_{\text{NLOS}}$. We make this simplification motivated from previous observations that photons following higher-order paths are difficult to detect with existing sensors [63]. We additionally ignore light following direct paths without interacting with the NLOS surface

$\mathcal{S}_{\text{NLOS}}$, as this light component is typically removed using time-gating mechanisms [13]. Additionally, for each pair of virtual points l and s , we assume that we have calibrated our measurements so that we can neglect the radiometric and pathlength terms for the connections $l_0 \rightarrow l$ and $s \rightarrow s_0$.

Under these assumptions, we can use the *path integral formulation* of light transport [75] to write the intensity measured by the sensor s_0 at time t as

$$I(t; l, s) = \int_{\mathcal{S}_{\text{NLOS}}} \overbrace{W(x; t) f(x, \hat{n}(x))}^{g(x, \hat{n}(x))} \cdot v(x, l) v(x, s) dA(x), \quad (1)$$

where $A(x)$ is the area measure on $\mathcal{S}_{\text{NLOS}}$, \hat{n} is the normal of a surface at a specific point, and $W(x; t)$, $f(x, \hat{n}(x))$, $v(x, s)$ will be discussed below. When considered as a function of all possible times t , virtual sources l , and virtual detectors s , $I(t; l, s)$ is often referred to as the *five-dimensional transient* [13]. We note that, because of the three-bounce assumption, the usual path integral reduces to a single *surface integral* over the NLOS surface $\mathcal{S}_{\text{NLOS}}$.

The *radiometric throughput* f in Equation (1) is the radiance that flows through the path $l \rightarrow x \rightarrow s$,

$$f(x, \hat{n}(x)) = f_s(\hat{n}(x), \hat{\omega}_l(x), \hat{\omega}_s(x)) \cdot \frac{\langle -\hat{\omega}_l(x), \hat{n}(l) \rangle \langle \hat{\omega}_l(x), \hat{n}(x) \rangle}{\|x - l\|^2} \cdot \frac{\langle -\hat{\omega}_s(x), \hat{n}(s) \rangle \langle \hat{\omega}_s(x), \hat{n}(x) \rangle}{\|x - s\|^2}, \quad (2)$$

where f_s is the BRDF of $\mathcal{S}_{\text{NLOS}}$ at point x , $\hat{\omega}_l(x)$ is the normalized vector parallel to $l - x$, and likewise for $\hat{\omega}_s(x)$.

The *temporal importance* W models the mechanism by which the sensor selects paths of length within some specific range for each measurement $I(t; l, s)$ ¹,

$$W(x, t) = \text{rect}\left(\frac{\tau(x) - t}{T}\right), \quad (3)$$

where rect is the unit rectangular function, T is the sensor's temporal resolution, and τ is the length of path $l \rightarrow x \rightarrow s$,

$$\tau(x) = \|x - l\| + \|x - s\|. \quad (4)$$

Finally, the *visibility function* v is a binary indicator of occlusion between two points,

$$v(x, y) = \begin{cases} 1, & \text{if } x, y \text{ are visible to each other,} \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

Comparison to albedo volume model. It is instructive to compare the surface integral formulation of Equation (1) with the albedo volume model of Velten et al. [77].

¹We treat geometric pathlength and time of flight as equivalent, with the understanding that they relate to each other through the speed of light.

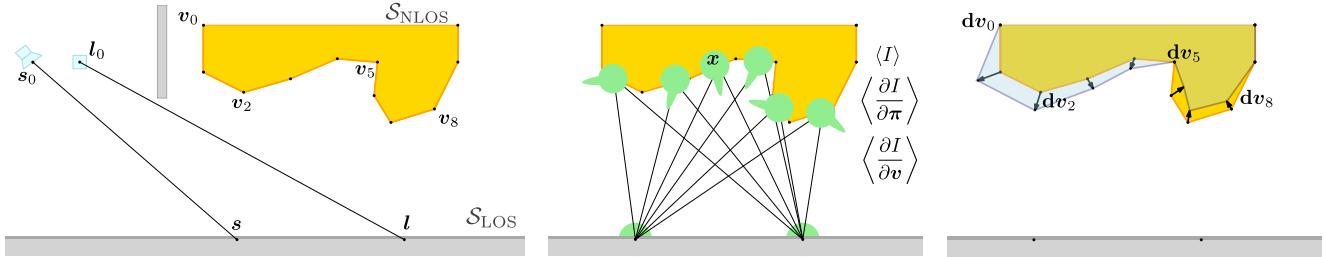


Figure 2. **Pipeline overview:** (Left) 2D visualization and notation. (Middle) We sample points x to estimate the transient and its gradients with respect to reflectance π and vertices v of a triangular mesh. (Right) We use the computed gradients to evolve the surface.

This model represents the NLOS scene as an albedo function $\rho(x)$, defined on all points of a continuous three-dimensional volume, $x \in \mathcal{V}_{\text{NLOS}}$. Then, transient measurements are expressed as a *volume integral*,

$$I(t; l, s) = \int_{\mathcal{V}_{\text{NLOS}}} \frac{W(x; t) \rho(x)}{\|x - l\|^2 \|x - s\|^2} dV(x), \quad (6)$$

where $V(x)$ is the standard measure on $\mathcal{V}_{\text{NLOS}}$. Compared to Equation (1), we note that the integrand of Equation (6) constrains the reflectance function f_s to be Lambertian, and omits the normal-related shading terms and the visibility terms v . Recent extensions incorporate normal and visibility effects through additional volumetric functions defined everywhere on $\mathcal{V}_{\text{NLOS}}$ [30]. However, the albedo, normal, and visibility volumetric functions are treated as independent of each other, even though they are in fact intertwined as functions of the underlying NLOS surface $\mathcal{S}_{\text{NLOS}}$.

Despite the lack of physical accuracy, the albedo volume model is attractive because of its mathematical convenience: Through a straightforward discretization of the volume integral of Equation (6), forward evaluations of the model become simple matrix-vector multiplication operations. Consequently, inverting the model to reconstruct the NLOS scene can be posed as a linear least-squares problem. By contrast, forward evaluations of the surface integral of Equation (1) rely on involved surface quadrature methods, or Monte Carlo rendering. In turn, this makes inverting the model for NLOS reconstruction non-trivial. We defer discussion of Monte Carlo rendering until Section 4.2, after we first develop our approach for performing this inversion.

4. Analysis-by-synthesis optimization

We can now formulate the NLOS reconstruction problem. We are given a set of calibrated transient measurements $\{\tilde{I}_m(t), m = 1, \dots, M\}$, corresponding to pairs of virtual points $\{(l_m, s_m), m = 1, \dots, M\}$. We additionally adopt parametric forms $\mathcal{S}_{\text{NLOS}}[v]$ and $f_s[\pi]$ for the NLOS surface and reflectance, respectively. Then, we recover the unknown parameters from the measurements through *analysis-by-synthesis*, also known as *inverse rendering*: We

search for the parameter values that can be used to simulate transients that best match our measurements. Formally, we minimize the following loss function,

$$E(v, \pi) = \frac{1}{2} \sum_{m,t} \left\| \tilde{I}_m(t) - I[v, \pi](t; l_m, s_m) \right\|^2. \quad (7)$$

We use the notation $I[v, \pi](t; l, s)$ to indicate that a rendered transient is a function, through Equation (1), of the surface and reflectance parameters v and π . While we use the L_2 loss for convenience, our technique can be used to minimize arbitrary losses differentiable with respect to $I[v, \pi](t; l, s)$, including losses derived from the noise model of the underlying transient sensors [32].

We aim to use gradient-descent optimization, in order to efficiently minimize the analysis-by-synthesis objective and recover the NLOS surface and reflectance parameters. Differentiating the loss function $E(v, \pi)$ of Equation (7) with respect to surface and reflectance parameters, we obtain

$$\frac{\partial E}{\partial y} = - \sum_{m,t} \left(\tilde{I}_m(t) - I(t; l_m, s_m) \right) \frac{\partial I(t; l_m, s_m)}{\partial y}, \quad (8)$$

where y can be either v or π . Evaluating the derivatives requires computing not only the transients I , but also their derivatives $\partial I / \partial \pi$ and $\partial I / \partial v$ with respect to reflectance and surface parameters. This is challenging because I is not an analytical function of these parameters, but is only related to them through the surface integral of Equation (1).

We overcome this obstacle using an approach based on *differentiable rendering*. We prove that the derivatives $\partial I / \partial \pi$ and $\partial I / \partial v$ can be expressed as surface integrals analogous to that of Equation (1). This allows us to derive efficient Monte Carlo rendering algorithms for *stochastically* approximating the reflectance and surface derivatives. We can, then, combine these stochastic estimates with stochastic gradient descent optimization [39] to minimize Equation (7). In the rest of this section, we first describe our choices for NLOS surface and reflectance parameterization, then provide an overview of our approach differentiable rendering approach, deferring details to the supplement.

4.1. Differentiating transients

Surface parameterization. We represent the NLOS surface $\mathcal{S}_{\text{NLOS}}$ as a triangular mesh with boundary, which we represent using two matrices: First, a $3 \times V$ *geometry matrix* \mathbf{V} providing the three-dimensional coordinates of its V vertices. Second, a $3 \times T$ *topology matrix* \mathbf{T} providing the integer vertex indices of its T triangles. We do not assign any normal or texture parameters to the vertices, and at every point on the mesh, we assume that the surface normal is equal to the corresponding triangle’s face normal. We use meshes instead of an implicit representation (e.g., level sets [55] or signed distance functions [14] to facilitate efficient Monte Carlo rendering (see Section 4.2). On the other hand, this complicates optimization due to the need to handle the discrete topology matrix \mathbf{T} . As is common in mesh optimization, we use differentiable rendering to minimize Equation (7) only with respect to mesh vertices. During this iterative minimization, we use standard geometry processing tools to improve the mesh topology (Section 5).

Reflectance parameterization. We assume that the NLOS surface has a spatially-uniform BRDF, which we represent using the widely-adopted GGX microfacet BRDF, as described by Walter et al. [79]. For completeness, we provide in the supplement the full expression f_s for GGX.

Derivatives as surface integrals. We now state the main technical result of the paper, which allows us to derive expressions for the derivatives of the image formation model (1) with respect to surface geometry and reflectance.

Proposition 1 *The derivatives of a transient $I(t; \mathbf{l}, \mathbf{s})$ with respect to reflectance and mesh vertices can be written as:*

$$\frac{\partial I}{\partial \mathbf{v}} = \int_{\mathcal{S}_{\text{NLOS}}} g_s(\mathbf{x}, \hat{\mathbf{n}}(\mathbf{x})) v(\mathbf{x}, \mathbf{l}) v(\mathbf{x}, \mathbf{s}) dA(\mathbf{x}), \quad (9)$$

$$\frac{\partial I}{\partial \pi} = \int_{\mathcal{S}_{\text{NLOS}}} g_r(\mathbf{x}, \hat{\mathbf{n}}(\mathbf{x})) v(\mathbf{x}, \mathbf{l}) v(\mathbf{x}, \mathbf{s}) dA(\mathbf{x}), \quad (10)$$

for appropriate functions g_s and g_r .

We provide the proof and detailed expressions for g_s and g_r in the supplement. In the case of reflectance, this simply involves changing the order of differentiation and integration. However, in the case of mesh vertices, differentiating Equation (1) is complicated by the fact that the integration over surface is also a function of the mesh vertices. We tackle this by using recent results on analytically expressing gradient flows from mesh functionals as surface integrals [20, 19]. These results have also been used by Delaunoy and Prados [16] for surface optimization in line-of-sight reconstruction problems (e.g., shape from shading).

Our proof makes the approximation that the visibility terms v are independent of the mesh geometry. This approximation is justified by the fact that the visibility terms have non-zero derivatives only on a zero-measure part of the

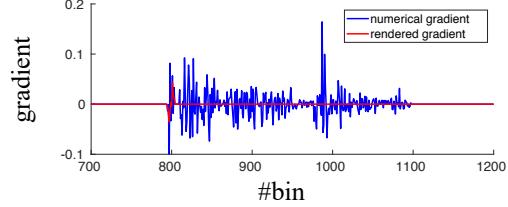


Figure 3. Comparison of gradient estimation techniques: We estimate the derivative of a transient with respect to one coordinate of one NLOS surface vertex. We plot two estimates, one computed using the rendering algorithm of Section 4.2, and another using finite differences and antithetic variates [23]. We observe that the numerical gradient is significantly noisier than the rendered one, despite being computed using twice as many path samples.

surface (that is, on occluding contours [43]), and is common in computer vision and graphics problems [2, 67, 49]. Delaunoy and Prados [16] show that differentiation is possible even without approximating the visibility terms v as constant. As we discuss in the supplement, we can similarly extend Proposition 1 to account for visibility. However, in practice we found that this complicates Monte Carlo rendering without significantly improving the optimization results.

Surface regularization. We note that, when optimizing geometry, we follow Delaunoy and Prados [16] and augment the loss function $E(\mathbf{v}, \pi)$ with a normal smoothing regularization term. We discuss this in the supplement.

4.2. Stochastic estimation and optimization

We can now describe our two core computational tools for efficiently minimizing the loss function of Equation (7).

Monte Carlo rendering. The surface integrals of Equations (1), (9), and (10) can be approximated using Monte Carlo integration: We first use any probability distribution μ on $\mathcal{S}_{\text{NLOS}}$ to sample a set of points $\{\mathbf{x}_j \in \mathcal{S}_{\text{NLOS}}, j = 1, \dots, J\}$. Then, we can form the respective *unbiased and consistent* estimates [23]:

$$\langle I \rangle = \sum_{j=1}^J \frac{g(\mathbf{x}_j, \hat{\mathbf{n}}(\mathbf{x}_j)) v(\mathbf{x}_j, \mathbf{l}) v(\mathbf{x}_j, \mathbf{s})}{\mu(\mathbf{x}_j)}, \quad (11)$$

$$\left\langle \frac{\partial I}{\partial \mathbf{v}} \right\rangle = \sum_{j=1}^J \frac{g_s(\mathbf{x}_j, \hat{\mathbf{n}}(\mathbf{x}_j)) v(\mathbf{x}_j, \mathbf{l}) v(\mathbf{x}_j, \mathbf{s})}{\mu(\mathbf{x}_j)}, \quad (12)$$

$$\left\langle \frac{\partial I}{\partial \pi} \right\rangle = \sum_{j=1}^J \frac{g_r(\mathbf{x}_j, \hat{\mathbf{n}}(\mathbf{x}_j)) v(\mathbf{x}_j, \mathbf{l}) v(\mathbf{x}_j, \mathbf{s})}{\mu(\mathbf{x}_j)}. \quad (13)$$

In the supplement, we describe a *stratified* area sampling procedure, which greatly accelerates rendering.

Stochastic gradient descent. Using these Monte Carlo estimates, we can *approximately* compute the derivatives of Equation (8). We can combine these stochastic deriva-

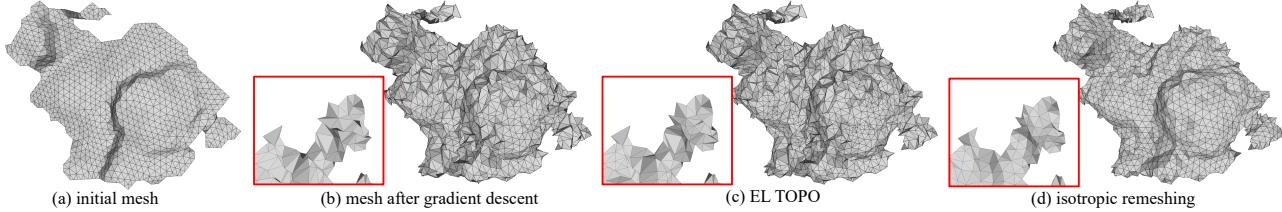


Figure 4. **Geometry processing:** We show an example of the geometric processing operations we use to improve mesh topology. (a) Initial mesh. (b) Mesh after gradient descent steps, with self-intersections at several places (see inset). (c) Mesh evolution using El Topo, which helps reduce self-intersections. (d) Mesh after isotropic remeshing, which increases mesh detail while decreasing high frequency artifacts.

tive estimates with *stochastic gradient descent* (SGD) algorithms to perform the analysis-by-synthesis optimization of Equation (7). We use Adam [39] to alternately optimize for reflectance and surface, and provide more details about our optimization procedure in the supplement.

Comparison with numerical differentiation. The performance of SGD optimization critically depends on the ability to compute *unbiased* gradient estimates of *low variance*. To highlight the importance of our differentiable rendering formulation in facilitating this optimization, we compare in Figure 3 rendered gradient estimates with estimates computed using a finite-difference approximation. We observe that the numerical gradients have significantly higher variance; therefore, using them with SGD would greatly slow down convergence. Alternatively, we could eliminate variance in finite-difference estimation, by using a quadrature technique (e.g., finite elements) to compute the forward integral (1). However, this could introduce strong bias, and therefore affect the physical accuracy of the result.

5. Geometry processing operations

As discussed in Section 4.1, we use stochastic gradient descent to optimize only the vertices of the mesh $\mathcal{S}_{\text{NLOS}}$, and not its topology. We instead improve the mesh topology by using, at various times during the analysis-by-synthesis optimization, geometry processing tools, as shown in Figure 4.

Robust surface evolution. As the mesh $\mathcal{S}_{\text{NLOS}}$ evolves over multiple SGD iterations, triangle quality typically degrades. Motivated by other optimization-driven mesh editing algorithms [49], we instead evolve $\mathcal{S}_{\text{NLOS}}$ using the pipeline of Brochu and Bridson [12], implemented in El Topo [11]. Given initial vertices $\{v_i, i = 1, \dots, V\}$, and associated displacements $\{dv_i, i = 1, \dots, V\}$, El Topo performs two types of operations: First, it alters the displacement vectors and mesh topology, to produce a non-intersecting mesh. Second, it uses local topology operations to improve overall mesh quality. In our implementation, we accumulate displacement vectors dv_i over multiple gradient-descent iterations, then use El Topo to evolve the mesh.

Progressive refinement and isotropic remeshing. As an additional means of regularization, we optimize the NLOS

surface $\mathcal{S}_{\text{NLOS}}$ in a coarse-to-fine fashion. We start with a mesh of a relatively small number of vertices V and triangles T . Then, during the gradient-descent optimization of $\mathcal{S}_{\text{NLOS}}$, we progressively increase the number of vertices and triangles. We implement mesh refinement by performing isotropic remeshing operations [9] with increasing target number of vertices. In addition to increasing the mesh detail, isotropic remeshing improves mesh quality and filters out high-frequency artifacts on the mesh surface.

6. Experiments

Implementation. Our framework has three major components: differentiable rendering, geometry processing, and stochastic gradient descent. For rendering, we have developed a C++ implementation based on Embree [78] for fast CPU execution. For geometry processing, our C++ implementation is built using the El Topo [11], CGAL [72] and libigl [57] libraries. Finally, both the rendering and geometry components are interfaced with Pytorch [58], which we use for stochastic gradient descent optimization with Adam [39]. Our implementation can scale up to optimization of meshes with more than 100,000 vertices, using 4096 transient measurements of 1200 temporal bins each. We run experiments on a 72-core Amazon EC2 c5.18xlarge instance, with a runtime of around two hours per scene. Our implementation and data are available online [1].

Scanning configuration. In all our experiments, we use a confocal scanning procedure, $l = s$ [56]. The scanning points are on a 64×64 regular grid on the visible surface.

Initialization. Except where specified otherwise, we initialize using the light cone transform algorithm of O’Toole et al. [56]. We convert the resulting albedo volume to a surface by first computing the maximum albedo voxel along the depth axis, then pruning albedo values below a threshold, and finally triangulating the remaining points.

6.1. Synthetic experiments

We use synthetic data to evaluate the ability of our method to reconstruct NLOS surface shape and reflectance. In our synthetic experiments, NLOS objects are placed at a distance of 0.4 m from a visible wall of size $0.5m \times 0.5m$.



Figure 5. **Surface reconstruction examples:** (Top) Ground truth. (Middle) Reconstructions using the light cone transform [56]. (Bottom) Reconstructions from our method. We can reconstruct shapes with different surface characteristics, including strong non-convexities, large depth variations, and bas-relief details.

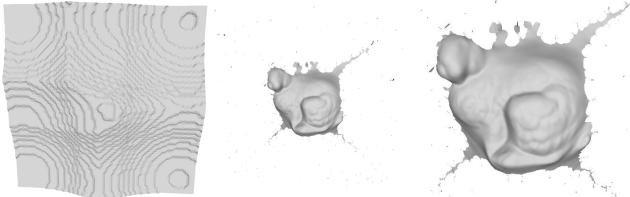


Figure 6. **Alternative initialization:** (Left) Reconstruction using space carving [74]. (Middle, right) Reconstruction from our method, shown in same scale (middle) and zoomed-in (right).

We use Monte Carlo transient rendering [62] to synthesize data, to which we add noise [32]. In the supplement, we show additional simulations evaluating performance for different numbers of measurements and amounts of noise.

Shape reconstruction. Figures 1 and 5 show reconstructions for a variety of NLOS shapes with known Lambertian reflectance. Our method reconstructs surface details that are completely missing from the initial volumetric reconstruction. A notable result is the soap bar, where we can reconstruct the relief letters (depth 2 transient bins). In the

supplement, we use surface distance metrics to quantify the reconstruction improvements.

Alternative initialization. Figure 6 shows an example where we initialize our optimization using the space carving algorithm of Tsai et al. [74]. We observe that, despite the very crude initialization, our method still produces a reconstruction of comparable detail to that produced from the more accurate volumetric initialization in Figure 1.

Simultaneous shape and reflectance reconstruction. Figure 7 shows simulated experiments for reconstructing both shape and reflectance. We experiment with a range of GGX α values, going from very smooth to very rough specular reflectance (Figure 7(b)-(c)). We observe that our algorithm successfully reconstructs a rough estimate of both shape and reflectance in all cases, but the reconstruction quality deteriorates as the surface becomes more specular.

6.2. Experiments with measured data

We perform experiments using datasets from three real NLOS scenes, captured with SPAD-based transient imaging systems. The first dataset is the diffuse 'S' shape ob-

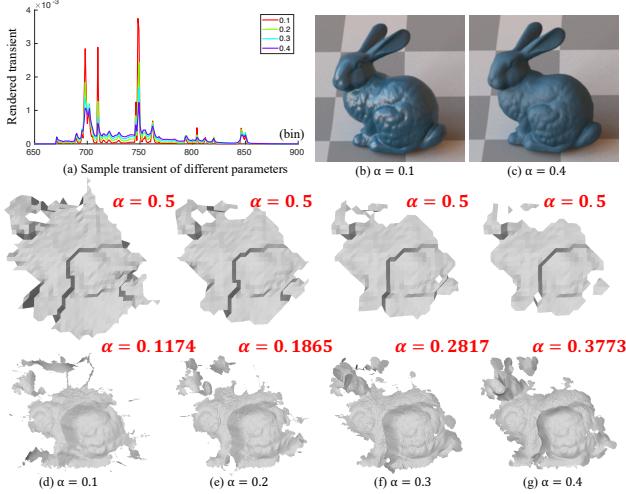


Figure 7. Reconstruction of both shape and reflectance: (a) We render transients for different GGX α values to visualize the effect of reflectance on NLOS measurements. (b, c) We also visualize the reflectance by rendering the scene under ambient light. (d - g) We show optimization results for different α values, with the initial shape and α at the top, and the optimized results at the bottom.

ject from [56]. As shown in Figure 8, our recovered shape closely resembles the ground truth geometry and is overall flatter than the volumetric reconstruction.

We additionally show reconstructions for two datasets captured with our own implementation of the SPAD setup of O’Toole et al. [56], for two NLOS objects of greater surface complexity. The first object is a diffuse horse bust with fine geometric details. As shown in Figure 8, our recovered result reproduces the flat and curved surface areas better. The second object is a planar scene with a 6mm tall (5 transient bins) relief in the shape of two digits. Our recovered result better differentiates the digits from the background surface.

7. Discussion

We discuss some limitations of our NLOS surface optimization framework. Since we do reconstruction by optimizing a very non-linear loss function, our final result can be strongly dependent on the initialization. Our experiments indicate that the quality of the initialization strongly affects the extent of the NLOS object that is recovered, but has a small impact on reconstruction detail. We hope to address the former issue by incorporating boundary evolution techniques into our optimization pipeline. Additionally, our pipeline performs worse as the reflectance of the NLOS object becomes more specular. We believe this is primarily caused by the area sampling procedure we use for rendering, which becomes very inefficient for highly-specular reflectance. We can potentially improve performance in such cases by considering multiple importance sampling techniques [76]. Finally, our results show a difference in per-

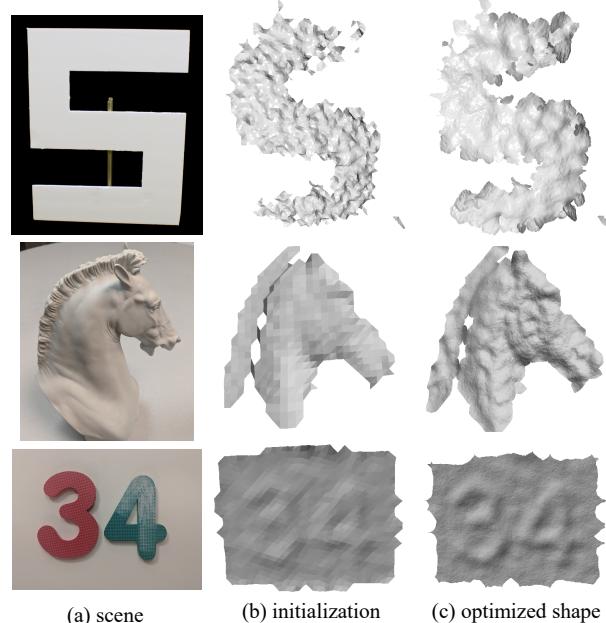


Figure 8. NLOS surface reconstruction using SPAD measurements: (Top) Diffuse object from [56]. (Middle) A diffuse horse statue. (Bottom) Digit relief on a planar object. In our experiment, we cover the digits with white paper, to increase SNR.

formance between synthetic and real data. The noise sensitivity experiments in the supplement indicate that the difference is primarily due to inaccurate modeling of SPAD sensors (Poisson noise, pile-up, jitter [32]). We expect that we can close the performance gap by changing the loss function of Equation (7) to account for these effects.

Despite these limitations, our experiments demonstrate that our surface optimization framework significantly improves the quality of reconstruction possible in NLOS settings. Because it is based on the rendering equation, our framework can be used to process not only transients, but all types of radiometric measurements: steady-state, continuous-wave time-of-flight, and so on. Therefore, we hope it can serve as a platform for exploring NLOS imaging schemes that use, independently or in combinations, alternative radiometric sensors, transient or otherwise. Additionally, our framework can be used to empirically investigate fundamental resolution limits inherent in each of these sensing modalities, without concern about information loss from approximations to the image formation model. Such empirical investigations can complement existing theoretical results [34] on resolution limits, or even provide insights that will eventually lead to such results.

Acknowledgments. This work was supported by DARPA REVEAL (HR0011-16-C-0025, HR0011-16-C-0028) and NSF Expeditions (CCF-1730147) grants. CYT gratefully acknowledges support from the Bertucci Graduate Fellowship and the Google PhD Fellowship.

References

- [1] NLOS surface optimization with differentiable rendering, 2019. https://github.com/cmu-ci-lab/nlos_surface_optimization. 6
- [2] Ben Appleton and Hugues Talbot. Globally minimal surfaces by continuous maximal flows. *IEEE transactions on pattern analysis and machine intelligence*, 28(1):106–118, 2006. 2, 5
- [3] Victor Arellano, Diego Gutierrez, and Adrian Jarabo. Fast back-projection for non-line of sight reconstruction. *Optics Express*, 25(10):11574–11583, 2017. 2
- [4] Dejan Azinović, Tzu-Mao Li, Anton Kaplanyan, and Matthias Nießner. Inverse path tracing for joint material and lighting estimation. *CVPR*, 2019. 3
- [5] Manel Baradad, Vickie Ye, Adam B Yedidia, Frédéric Durand, William T Freeman, Gregory W Wornell, and Antonio Torralba. Inferring light fields from shadows. In *CVPR*, 2018. 1, 2
- [6] Mufeed Batarseh, Sergey V Sukhov, Zhiqin Shen, Heath Gemar, Reza Rezvani, and Aristide Dogariu. Passive sensing around the corner using spatial coherence. *Nature communications*, 9(1):3629, 2018. 1, 2
- [7] Jacopo Bertolotti, Elbert G van Putten, Christian Blum, Ad Lagendijk, Willem L Vos, and Allard P Mosk. Non-invasive imaging through opaque scattering layers. *Nature*, 491(7423):232, 2012. 2
- [8] Jeremy Boger-Lombard and Ori Katz. Non line-of-sight localization by passive optical time-of-flight. *ArXiv e-prints*, Aug. 2018. 1, 2
- [9] Mario Botsch and Leif Kobbelt. A remeshing approach to multiresolution modeling. In *Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 185–192, 2004. 6
- [10] Katherine L Bouman, Vickie Ye, Adam B Yedidia, Frédéric Durand, Gregory W Wornell, Antonio Torralba, and William T Freeman. Turning corners into cameras: Principles and methods. In *ICCV*, 2017. 1, 2
- [11] Tyson Brochu and Robert Bridson. El topo, 2009. <https://www.cs.ubc.ca/labs/imager/tr/2009/eltopo/eltopo.html>. 6
- [12] Tyson Brochu and Robert Bridson. Robust topological operations for dynamic explicit surfaces. *SIAM Journal on Scientific Computing*, 31(4):2472–2493, 2009. 6
- [13] Mauro Buttafava, Jessica Zeman, Alberto Tosi, Kevin Eliceiri, and Andreas Velten. Non-line-of-sight imaging using a time-gated single photon avalanche diode. *Optics Express*, 23(16):20997–21011, 2015. 1, 2, 3
- [14] Brian Curless and Marc Levoy. A volumetric method for building complex models from range images. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 303–312, 1996. 5
- [15] Amaël Delaunoy and Marc Pollefeys. Photometric bundle adjustment for dense multi-view 3d modeling. In *CVPR*, 2014. 2
- [16] Amaël Delaunoy and Emmanuel Prados. Gradient flows for optimizing triangular mesh-based surfaces: Applications to 3d reconstruction problems dealing with visibility. *IJCV*, 95(2):100–123, 2011. 2, 5
- [17] Mathieu Desbrun, Mark Meyer, Peter Schröder, and Alan H Barr. Implicit fairing of irregular meshes using diffusion and curvature flow. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 317–324. ACM Press/Addison-Wesley Publishing Co., 1999. 2
- [18] Philip Dutré, Kavita Bala, and Philippe Bekaert. *Advanced global illumination*. AK Peters, Ltd., 2006. 2, 3
- [19] Gerhard Dziuk and Charles M Elliott. Finite elements on evolving surfaces. *IMA journal of numerical analysis*, 27(2):262–292, 2007. 5
- [20] Ilya Eckstein, Jean-Philippe Pons, Yiying Tong, Chung Chieh Jay Kuo, and Mathieu Desbrun. Generalized surface flows for mesh processing. In *Proceedings of the fifth Eurographics symposium on Geometry processing*, pages 183–192. Eurographics Association, 2007. 2, 5
- [21] Olivier Faugeras and Renaud Keriven. Complete dense stereovision using level set methods. In *European conference on computer vision*, pages 379–393. Springer, 1998. 2
- [22] Olivier Faugeras and Renaud Keriven. *Variational principles, surface evolution, PDE's, level set methods and the stereo problem*. IEEE, 2002. 2
- [23] George Fishman. *Monte Carlo: concepts, algorithms, and applications*. Springer Science & Business Media, 1996. 5
- [24] Isaac Freund. Looking through walls and around corners. *Physica A: Statistical Mechanics and its Applications*, 168(1):49–65, 1990. 2
- [25] Pau Gargallo, Emmanuel Prados, and Peter Sturm. Minimizing the reprojection error in surface reconstruction from images. In *ICCV*, 2007. 2
- [26] Adam Geva, Yoav Y Schechner, Yonatan Chernyak, and Rajaiv Gupta. X-ray computed tomography through scatter. In *ECCV*, 2018. 3
- [27] Ioannis Gkioulekas, Anat Levin, and Todd Zickler. An evaluation of computational imaging techniques for heterogeneous inverse scattering. *ECCV*, 2016. 3
- [28] Ioannis Gkioulekas, Shuang Zhao, Kavita Bala, Todd Zickler, and Todd Levin. Inverse volume rendering with material dictionaries. *SIGGRAPH Asia*, 2013. 3
- [29] Otkrist Gupta, Thomas Willwacher, Andreas Velten, Ashok Veeraraghavan, and Ramesh Raskar. Reconstruction of hidden 3d shapes using diffuse reflections. *Optics express*, 20(17):19096–19108, 2012. 2
- [30] Felix Heide, Matthew O’Toole, Kai Zhang, David Lindell, Steven Diamond, and Gordon Wetzstein. Non-line-of-sight imaging with partial occluders and surface normals. *ACM Transactions on Graphics (ToG)*, 2019. 2, 4
- [31] Felix Heide, Lei Xiao, Wolfgang Heidrich, and Matthias Hullin. Diffuse mirrors: 3d reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors. In *CVPR*, 2014. 2
- [32] Quercus Hernandez, Diego Gutierrez, and Adrian Jarabo. A computational model of a single-photon avalanche diode sensor for transient imaging. *arXiv*, preprint arXiv:1703.02635, 2017. 4, 7, 8

[33] Adrian Jarabo, Belen Masia, Julio Marco, and Diego Gutierrez. Recent advances in transient imaging: A computer graphics and vision perspective. *Visual Informatics*, 1(1):65–79, 2017. 1, 2

[34] Achuta Kadambi, Hang Zhao, Boxin Shi, and Ramesh Raskar. Occluded imaging with time-of-flight sensors. *ACM TOG*, 35(2):15, 2016. 2, 8

[35] James T Kajiya. The rendering equation. In *ACM Siggraph Computer Graphics*, volume 20, pages 143–150, 1986. 2

[36] Ori Katz, Pierre Heidmann, Mathias Fink, and Sylvain Gigan. Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations. *Nature photonics*, 8(10):784, 2014. 1, 2

[37] Ori Katz, Eran Small, and Yaron Silberberg. Looking around corners and through thin turbid layers in real time with scattered incoherent light. *Nature photonics*, 6(8):549–553, 2012. 2

[38] Pramook Khungurn, Daniel Schroeder, Shuang Zhao, Kavita Bala, and Steve Marschner. Matching real fabrics with micro-appearance models. *ACM TOG*, 35(1):1:1–1:26, 2015. 3

[39] Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *ICLR*, 2015. 4, 6

[40] Ahmed Kirmani, Tyler Hutchison, James Davis, and Ramesh Raskar. Looking around the corner using transient imaging. In *ICCV*, 2009. 1

[41] Ahmed Kirmani, Tyler Hutchison, James Davis, and Ramesh Raskar. Looking around the corner using ultrafast transient imaging. *IJCV*, 95(1):13–28, 2011. 1, 2

[42] Jonathan Klein, Christoph Peters, Jaime Martín, Martin Laurenzis, and Matthias B Hullin. Tracking objects outside the line of sight using 2d intensity images. *Scientific reports*, 6, 2016. 1, 2

[43] Jan J Koenderink. What does the occluding contour tell us about solid shape? *Perception*, 13(3):321–330, 1984. 5

[44] Kalin Kolev and Daniel Cremers. Integration of multiview stereo and silhouettes via convex functionals on convex domains. In *ECCV*, 2008. 2

[45] Marco La Manna, Fiona Kine, Eric Breitbach, Jonathan Jackson, Talha Sultan, and Andreas Velten. Error back-projection algorithms for non-line-of-sight imaging. *IEEE transactions on pattern analysis and machine intelligence*, 2018. 2

[46] Aviad Levis, Yoav Y Schechner, Amit Aides, and Anthony B. Davis. Airborne three-dimensional cloud tomography. *ICCV*, 2015. 3

[47] Tzu-Mao Li, Miika Aittala, Frédo Durand, and Jaakko Lehtinen. Differentiable monte carlo ray tracing through edge sampling. In *SIGGRAPH Asia 2018 Technical Papers*, page 222. ACM, 2018. 3

[48] David Lindell, Gordon Wetzstein, and Vladlen Koltun. Acoustic non-line-of-sight imaging. In *CVPR*, 2019. 1, 2

[49] Hsueh-Ti Derek Liu, Michael Tao, and Alec Jacobson. Paparazzi: Surface editing by way of multi-view image processing. *ACM TOG*, 2018. 2, 5, 6

[50] Xiaochun Liu, Sebastian Bauer, and Andreas Velten. Analysis of feature visibility in non-line-of-sight measurements. In *CVPR*, 2019. 1, 2

[51] Stephen Lombardi and Ko Nishino. Radiometric scene decomposition: Scene reflectance, illumination, and geometry from rgb-d images. In *3D Vision (3DV), 2016 Fourth International Conference on*, pages 305–313. IEEE, 2016. 3

[52] Matthew M Loper and Michael J Black. OpenDr: An approximate differentiable renderer. In *ECCV*, 2014. 3

[53] Stephen R Marschner. Inverse rendering in computer graphics, phd thesis. *Ithaca, NY, USA*, 1998. 2, 3

[54] Mark Meyer, Mathieu Desbrun, Peter Schröder, and Alan H Barr. Discrete differential-geometry operators for triangulated 2-manifolds. In *Visualization and mathematics III*, pages 35–57. Springer, 2003. 2

[55] Stanley Osher and Ronald Fedkiw. *Level set methods and dynamic implicit surfaces*, volume 153. Springer Science & Business Media, 2006. 2, 5

[56] Matthew O’Toole, David B Lindell, and Gordon Wetzstein. Confocal non-line-of-sight imaging based on the light-cone transform. *Nature*, 555(7696):338, 2018. 1, 2, 6, 7, 8

[57] Daniele Panozzo and Alec Jacobson. *LIBIGL: A C++ library for geometry processing without a mesh data structure*. SGP 2014 Graduate School, 2014. 6

[58] Adam Paszke, Soumith Chintala, Ronan Collobert, Koray Kavukcuoglu, Clement Farabet, Samy Bengio, Iain Melvin, Jason Weston, and Johnny Mariethoz. Pytorch: Tensors and dynamic neural networks in python with strong gpu acceleration, may 2017. 6

[59] Gustavo Patow and Xavier Pueyo. A survey of inverse rendering problems. In *Computer graphics forum*, volume 22, pages 663–687. Wiley Online Library, 2003. 3

[60] Gustavo Patow and Xavier Pueyo. A survey of inverse surface design from light transport behavior specification. In *Computer Graphics Forum*, volume 24, pages 773–789. Wiley Online Library, 2005. 3

[61] Adithya Pediredla, Akshat Dave, and Ashok Veeraraghavan. Snlos: Non-line-of-sight scanning through temporal focusing. *International Conference on Computational Photography*, 2019. 1, 2

[62] Adithya Pediredla, Ashok Veeraraghavan, and Ioannis Gkioulekas. Elliptic path sampling for time-gated rendering. *ACM Transactions on Graphics (TOG)*, 2019. 7

[63] Adithya K. Pediredla, Mauro Buttafava, Alberto Tosi, Oliver Cossairt, and Ashok Veeraraghavan. Reconstructing rooms using photon echoes: A plane based model and reconstruction algorithm for looking around the corner. In *ICCP*, 2017. 1, 2, 3

[64] Matt Pharr, Wenzel Jakob, and Greg Humphreys. *Physically based rendering: From theory to implementation*. Morgan Kaufmann, 2016. 2, 3

[65] Jean-Philippe Pons, Renaud Keriven, and Olivier Faugeras. Modelling dynamic scenes by registering multi-view image sequences. In *CVPR*, 2005. 2

[66] Charles Saunders, John Murray-Bruce, and Vivek K Goyal. Computational periscopy with an ordinary digital camera. *Nature*, 565(7740):472, 2019. 1, 2

[67] Steven M Seitz, Brian Curless, James Diebel, Daniel Scharstein, and Richard Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *CVPR*, 2006. 5

[68] Brandon M. Smith, Matthew O’Toole, and Mohit Gupta. Tracking multiple objects outside the line of sight using speckle imaging. In *CVPR*, 2018. 1, 2

[69] Jan Erik Solem and Niels Chr. A geometric formulation of gradient descent for variational problems with moving surfaces. In *Scale-Space*, pages 419–430. Springer, 2005. 2

[70] Olga Sorkine, Daniel Cohen-Or, Yaron Lipman, Marc Alexa, Christian Rössl, and H-P Seidel. Laplacian surface editing. In *Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 175–184. ACM, 2004. 2

[71] Matthew Tancik, Guy Satat, and Ramesh Raskar. Flash photography for data-driven hidden scene recovery. *arXiv preprint arXiv:1810.11710*, 2018. 2

[72] The CGAL Project. *CGAL User and Reference Manual*. CGAL Editorial Board, 4.13 edition, 2018. 6

[73] Christos Thrampoulidis, Gal Shulkin, Fei Xu, William T Freeman, Jeffrey H Shapiro, Antonio Torralba, Franco N C Wong, and Gregory W Wornell. Exploiting occlusion in non-line-of-sight active imaging. *IEEE Transactions on Computational Imaging*, 4(3):419–431, 2018. 1, 2

[74] Chia-Yin Tsai, Kiriakos N. Kutulakos, Srinivasa G. Narasimhan, and Aswin C. Sankaranarayanan. The geometry of first-returning photons for non-line-of-sight imaging. In *CVPR*, 2017. 1, 2, 7

[75] Eric Veach. *Robust monte carlo methods for light transport simulation*. Stanford University Stanford, 1998. 2, 3

[76] Eric Veach and Leonidas J Guibas. Optimally combining sampling techniques for monte carlo rendering. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 419–428. ACM, 1995. 8

[77] Andreas Velten, Thomas Willwacher, Otkrist Gupta, Ashok Veeraraghavan, Mouni G Bawendi, and Ramesh Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Comm.*, 3:745, 2012. 1, 2, 3

[78] Ingo Wald, Sven Woop, Carsten Benthin, Gregory S Johnson, and Manfred Ernst. Embree: a kernel framework for efficient cpu ray tracing. *ACM TOG*, 33(4):143, 2014. 6

[79] Bruce Walter, Stephen R Marschner, Hongsong Li, and Kenneth E Torrance. Microfacet models for refraction through rough surfaces. In *Proceedings of the 18th Eurographics conference on Rendering Techniques*, pages 195–206. Eurographics Association, 2007. 5

[80] Shumian Xin, Sotiris Nousias, Kyriakos N. Kutulakos, Aswin C. Sankaranarayanan, Srinivasa G. Narasimhan, and Ioannis Gkioulekas. A theory of fermat paths for non-line-of-sight shape reconstruction. In *CVPR*, 2019. 1, 2

[81] Feihu Xu, Gal Shulkin, Christos Thrampoulidis, Jeffrey H Shapiro, Antonio Torralba, Franco NC Wong, and Gregory W Wornell. Revealing hidden scenes by photon-efficient occlusion-based opportunistic active imaging. *Optics express*, 26(8):9945–9962, 2018. 1, 2

[82] Anthony Yezzi and Stefano Soatto. Stereoscopic segmentation. *IJCV*, 53(1):31–43, 2003. 2

[83] Kuk-Jin Yoon, Emmanuel Prados, and Peter Sturm. Joint estimation of shape and reflectance using multiple images with known illumination conditions. *IJCV*, 86(2):192–210, 2010. 2

[84] Shuang Zhao, Lifan Wu, Frédo Durand, and Ravi Ramamoorthi. Downsampling scattering parameters for rendering anisotropic media. *ACM TOG*, 35(6):166:1–166:11, 2016. 3