

The Geometry of First-Returning Photons for Non-Line-of-Sight Imaging

Chia-Yin Tsai[†], Kiriakos N. Kutulakos[‡], Srinivasa G. Narasimhan[†], and Aswin C. Sankaranarayanan[†]

[†]Carnegie Mellon University

[‡]University of Toronto

Abstract

Non-line-of-sight (NLOS) imaging utilizes the full 5D light transient measurements to reconstruct scenes beyond the camera’s field of view. Mathematically, this requires solving an elliptical tomography problem that unmixes the shape and albedo from spatially-multiplexed measurements of the NLOS scene. In this paper, we propose a new approach for NLOS imaging by studying the properties of first-returning photons from three-bounce light paths. We show that the times of flight of first-returning photons are dependent only on the geometry of the NLOS scene and each observation is almost always generated from a single NLOS scene point. Exploiting these properties, we derive a space carving algorithm for NLOS scenes. In addition, by assuming local planarity, we derive an algorithm to localize NLOS scene points in 3D and estimate their surface normals. Our methods do not require either the full transient measurements or solving the hard elliptical tomography problem. We demonstrate the effectiveness of our methods through simulations as well as real data captured from a SPAD sensor.

1. Introduction

Non-line-of-sight (NLOS) imaging [8, 14] refers to estimation of the shape, texture, and reflectance of scene points that lie *beyond* the field of view of an imaging system. There are numerous approaches for estimating the shape of the scene within the field of view of the imaging system. In contrast, the NLOS shape estimation is a challenging task requiring capture and analysis of photons that have traveled beyond the line of sight (LOS). This is typically achieved by measuring the so-called 5D light transient transport tensor which captures light propagation — from the LOS scene onto the NLOS scene and back — at ultra-high temporal resolutions (typically, in picoseconds).

1.1. NLOS imaging via the 5D light transient

Suppose that the scene which we seek to recover consists of two components — the LOS scene \mathcal{L} that can be illuminated and imaged by the imaging system, and the NLOS

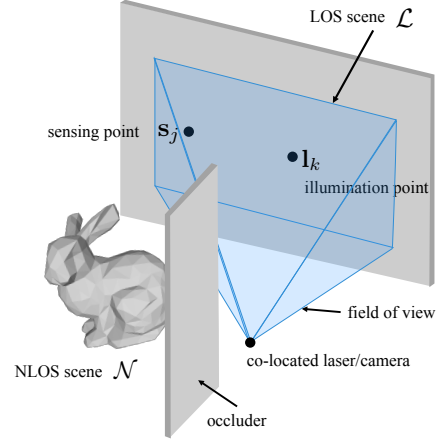


Figure 1. **Non-line-of-sight imaging setup.** NLOS imaging uses properties measured in LOS \mathcal{L} to infer the properties of the NLOS scene \mathcal{N} . In this paper, our goal is to use the first-returning photon associated with illuminating point \mathbf{l}_k and sensing point \mathbf{s}_j to infer shape of the NLOS scene.

scene \mathcal{N} , which is beyond the field of view of the imaging and illumination system (see Fig. 1); here, we assume that \mathcal{L} and \mathcal{N} are simply collections of 3D points.

Consider an imaging system consisting of an ultra-fast laser, capable of emitting an optical pulse with an extremely short duration, and an ultra-fast camera; we will discuss specific approaches to realize such setups in Section 2. Given two points $\mathbf{l}_k, \mathbf{s}_j \in \mathcal{L}$ (see Fig. 1), we define $L(t; \mathbf{l}_k, \mathbf{s}_j)$ as the light transient observed at \mathbf{s}_j while illuminating \mathbf{l}_k with a dirac-delta. We refer to

$$\{L(t; \mathbf{l}_k, \mathbf{s}_j) \mid \forall \mathbf{l}_k, \mathbf{s}_j \in \mathcal{L}\}$$

as the 5D light transient transport tensor [13, 10] since it encodes one degree of freedom along time and two angular degrees, each, for both illumination and sensing.

For simplicity of explanation, let’s assume that the LOS scene \mathcal{L} is convex (no interreflections of photons within LOS). When $\mathbf{l}_k \neq \mathbf{s}_j$, the convexity of \mathcal{L} implies that there are no single- or two-bounce light paths from the laser to the camera; in fact, there are no light paths from the laser to the camera that involve *only* LOS scene points. Hence, a non-zero intensity in the light transient $L(t; \mathbf{l}_k, \mathbf{s}_j)$ en-

codes properties associated with three- and higher-bounce light paths that include LOS as well as NLOS scene points. This forms the basis of NLOS imaging.

We focus solely on three-bounce light paths from the laser to the camera. We assume that the locations of all LOS scene points are known and, for simplicity, that the laser and camera are co-located at the origin. Given this, a non-zero intensity in $L(t; \mathbf{l}_k, \mathbf{s}_j)$ at $t = t_0$, indicates the presence of three-bounce light path(s) whose length is ct_0 , where c is the speed of light, and implies the presence of NLOS scene point(s) \mathbf{p} that satisfy

$$\|\mathbf{p} - \mathbf{l}_k\| + \|\mathbf{p} - \mathbf{s}_j\| = ct_0 - \|\mathbf{l}_k\| - \|\mathbf{s}_j\|, \quad (1)$$

where $\mathbf{p} \in \mathbb{R}^3$. This constrains \mathbf{p} to lie on an ellipsoid whose foci are at \mathbf{l}_k and \mathbf{s}_j [14, 1]. Moreover, the intensity $L(t_0; \mathbf{l}_k, \mathbf{s}_j)$ encodes both the shape and the reflectance of all NLOS scene points on the ellipsoid defined in (1). To further simplify, it is common to assume that the NLOS scene is Lambertian so that the reflectance function is represented by a spatially-varying *scalar* albedo pattern. By sweeping across the entire 5D transient, namely different illumination and sensing points as well as time instants, we can produce a large number of ellipsoidal integral constraints on the NLOS albedo. The albedo is subsequently recovered by solving a complex inverse problem [14].

1.2. First-returning photons

In this paper, we provide a formulation for NLOS shape recovery that avoids solving a complex inverse problem altogether. Specifically, instead of parsing through the entire 5D light transient, we only focus on the path length associated with the first-returning photon, defined as follows:

Definition. The *first-returning photon* at a LOS point \mathbf{s}_j , when we illuminate \mathbf{l}_k , is the photon that traverses the shortest three-bounce light path involving \mathbf{l}_k and \mathbf{s}_j , i.e., it is the first photon that we see at \mathbf{s}_j while illuminating \mathbf{l}_k .

The time of flight (ToF) of the first-returning photon is the smallest time instant t_0 such that $L(t_0; \mathbf{l}_k, \mathbf{s}_j)$ is non-zero. Its path length, $\delta(\mathbf{l}_k, \mathbf{s}_j) = ct_0$, can be derived as

$$\delta(\mathbf{l}_k, \mathbf{s}_j) = \|\mathbf{l}_k\| + \|\mathbf{s}_j\| + \min_{\mathbf{p} \in \mathcal{N}} \|\mathbf{p} - \mathbf{l}_k\| + \|\mathbf{p} - \mathbf{s}_j\|. \quad (2)$$

A key observation is that the shortest path is often unique and determined by the position of a *single* NLOS scene point. This greatly simplifies the shape estimation problem, since we do not need to solve a complex inverse problem.

1.3. Contributions

Our contributions are in the form of constraints induced by the ToF of the first-returning photon:

- **Space carving.** The ToF of the first-returning photon suggests that the closest NLOS scatterers are a certain distance away from the LOS scene points. Based on this

observation, we derive a space carving algorithm that restricts the spatial extent of the NLOS object.

- **Shape from first-returning photons.** We derive a set of constraints that relates the 3D positions of the NLOS scene points to the ToF of first-returning photons. This enables a simple algorithm for estimating the 3D locations of the NLOS scene points.
- **Surface normals from first-returning photons.** Once the scene point is localized in 3D, we show that the surface normal can also be recovered provided the NLOS is locally smooth. The derivation is essentially based on Fermat's principle of shortest path that is associated with the path taken by the first-returning photons.

Advantages. There are numerous advantages to a framework of NLOS shape estimation that relies purely on first-returning photons. First, the ToF of the first-returning photon is purely a function of the shape of the NLOS scene and largely invariant to its reflectance; hence, unlike prior work, our proposed algorithms can be applied to a wide range of reflectances. Second, as we are interested only in the ToF of the first-returning photons, our method does not depend on the radiance measurements, which can help in relaxing the sensing requirements.

Limitations. Shape estimation from first-returning photons has some obvious limitations. First, the paths taken by the first-returning photons depend on the geometry of the NLOS scene and hence, our method has limited control over the sampling of the NLOS scene. Second, the ToF of first-returning photons is a subset of the measurements encoded in the full 5D light transient; having access to the full transient can enable additional robustness especially when imaging complex scenes.

2. Related Work

We discuss several techniques to acquire the 5D transient as well as related work in NLOS imaging.

Acquiring the 5D light transient. Direct acquisition of the 5D light transient requires a time-gated laser to produce a light ray and emit an impulse pulse, and an image sensor that can resolve light at very high temporal resolutions. By orienting the sensor and the laser, independently, to different orientations we can sample the 5D light transient. There are many approaches to realizing this architecture. Velten *et al.* [14] use a streak camera to image at time resolutions of tens of picoseconds. A streak camera uses a time-varying electric field to introduce a time-dependent spatial displacement to incoming photons. Single-photon avalanche diodes (SPADs) provide an alternative, less-expensive approach for measuring the transient. The SPAD has an infinite gain and hence, saturates whenever a photon is incident on the sensor; the time-stamp of this event measures the ToF of the

photon. Buttafava *et al.* [1] show that, when illuminating the scene with an impulse train, the histogram of photon arrivals is a good approximation to the light transient.

An inexpensive approach to measure the light transient is to illuminate the scene with an amplitude-modulated source and measure the phase shifts, in the amplitude wave, caused by light propagation using a photonic mixer device (PMD) [3]. The PMD measurements can be interpreted as the Fourier coefficients of the transient response at the frequency of the illumination; hence, by illuminating the scene with different frequencies, the light transient can be acquired by measuring and inverting the Fourier coefficients [12]. This approach provides temporal resolution in nanoseconds. Another method is to use the wave property of light to find the path length via interferometry [2, 6].

Non-line-of-sight imaging. A common approach to NLOS imaging is to parse the 5D light transient and derive ellipsoidal constraints on the albedo of the NLOS object [14, 1]. Heide *et al.* [4] demonstrate NLOS shape reconstruction directly from a PMD's readout by using a generative model, instead of ellipsoidal constraints. Another common approach is to introduce surface priors to simplify the shape estimation problem. Kadambi *et al.* [7] use signal sparsity to recover distinct scene points. Pediredla *et al.* [11] estimate planar scenes by comparing SPAD readout with rendered results. Klein *et al.* [9] perform NLOS pose estimation using just image intensities. In this case, the number of unknowns is smaller and image intensities are sufficient for estimating the hidden object.

3. Geometry of First-Returning Photons

Problem setup. We assume an imaging system consisting of an ultra-fast laser and camera. We image the NLOS object through a diffuse LOS scene. Given two LOS scene points $\mathbf{l}_k, \mathbf{s}_j \in \mathcal{L}$, we assume that the setup can measure the path length associated with the first-returning photon at \mathbf{s}_j when we illuminate \mathbf{l}_k with an impulse.

Since the first-returning photon traverses the shortest three-bounce light path involving \mathbf{l}_k and \mathbf{s}_j , its path length $\delta(\mathbf{l}_k, \mathbf{s}_j)$ is given in (2). Given a known LOS scene, we can deduct from $\delta(\mathbf{l}_k, \mathbf{s}_j)$ the distance of the laser and camera to illumination point \mathbf{l}_k and sensing point \mathbf{s}_j , respectively. This provides us with the length of the shortest path from \mathbf{l}_k to \mathbf{s}_j via the NLOS scene,

$$d(\mathbf{l}_k, \mathbf{s}_j) = \min_{\mathbf{p} \in \mathcal{N}} \|\mathbf{p} - \mathbf{l}_k\| + \|\mathbf{p} - \mathbf{s}_j\|. \quad (3)$$

Our goal is to identify points belonging to the NLOS scene \mathcal{N} given a collection of shortest path lengths:

$$\{d(\mathbf{l}_k, \mathbf{s}_j) \mid \mathbf{l}_k, \mathbf{s}_j \in \mathcal{L}\}.$$

In the following, we will derive constraints on the NLOS scene given the path length of the first-returning photon when illuminating \mathbf{l}_k and sensing at \mathbf{s}_j .

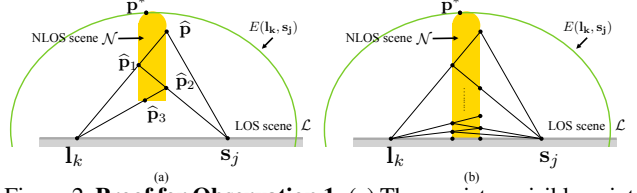


Figure 2. **Proof for Observation 1.** (a) There exists a visible point that creates a shorter light path, which leads to violation of the definition of first-returning photon. (b) Specialized case that needs to be excluded.

Ellipsoidal constraint. Given $d(\mathbf{l}_k, \mathbf{s}_j)$, as in (3), we can deduce that there is at least one NLOS scene point lying on an ellipsoid $E(\mathbf{l}_k, \mathbf{s}_j)$ given as:

$$E(\mathbf{l}_k, \mathbf{s}_j) = \{\mathbf{p} \mid \|\mathbf{p} - \mathbf{l}_k\| + \|\mathbf{p} - \mathbf{s}_j\| = d(\mathbf{l}_k, \mathbf{s}_j)\}.$$

Since the first-returning photon traverses the shortest path associated with \mathbf{l}_k and \mathbf{s}_j , this creates spatial constraint on occupancy of the NLOS object, which can be explained in the following observation.

Observation 1. There are no NLOS scene points in the interior of the ellipsoid $E(\mathbf{l}_k, \mathbf{s}_j)$.

Proof. Suppose that there exists a NLOS scene point $\hat{\mathbf{p}}$ inside the ellipsoid $E(\mathbf{l}_k, \mathbf{s}_j)$.

If $\hat{\mathbf{p}}$ is visible to both \mathbf{l}_k and \mathbf{s}_j , then a three-bounce light path from \mathbf{l}_k to $\hat{\mathbf{p}}$ to \mathbf{s}_j will create a shorter light path, which contradicts the definition of first-returning photon.

Suppose that $\hat{\mathbf{p}}$ is not visible to \mathbf{l}_k . Then there exists an occluder $\hat{\mathbf{p}}_1$ that blocks $\hat{\mathbf{p}}$ from \mathbf{l}_k ; the occluder $\hat{\mathbf{p}}_1$ lies on the line joining $\hat{\mathbf{p}}$ and \mathbf{l}_k and is visible to \mathbf{l}_k (see Fig. 2(a)). If $\hat{\mathbf{p}}_1$ is not visible to \mathbf{s}_j , then we repeat the process to find a point $\hat{\mathbf{p}}_2$ that is visible to \mathbf{s}_j , and so on. We define the path length caused by the occluder $\hat{\mathbf{p}}_i$ to be

$$d_i = \|\hat{\mathbf{p}}_i - \mathbf{l}_k\| + \|\hat{\mathbf{p}}_i - \mathbf{s}_j\|. \quad (4)$$

We observe that $d_1 > d_2 > d_3 > \dots$, that is, the path length decreases each time instant. Given that $\{d_i\}$ is bounded from below by $\|\mathbf{l}_k - \mathbf{s}_j\|$, and is decreasing, it will converge via the monotone convergence theorem.

Let $d^* = \lim_{i \rightarrow \infty} d_i$ be the converged value. When $d^* > \|\mathbf{l}_k - \mathbf{s}_j\|$, it is easily shown that $\hat{\mathbf{p}}_i$ also converges and the converged point will necessarily be visible to both \mathbf{l}_k and \mathbf{s}_j (if not, we repeat the process above and the path length decreases). We can hence find a point visible to both \mathbf{l}_k and \mathbf{s}_j with a shorter light path than the first-returning photon, which is a contradiction.

If $d^* = \|\mathbf{l}_k - \mathbf{s}_j\|$, then $\hat{\mathbf{p}}_i$ lies on the line connecting \mathbf{l}_k and \mathbf{s}_j and can potentially oscillate — this only happens in the scenario of Fig. 2 (b) when \mathbf{l}_k and \mathbf{s}_j are occluded from each other. This scenario is avoided by assuming that the points are visible to each other — a scenario that is entirely consistent with our envisioned operating conditions. \square

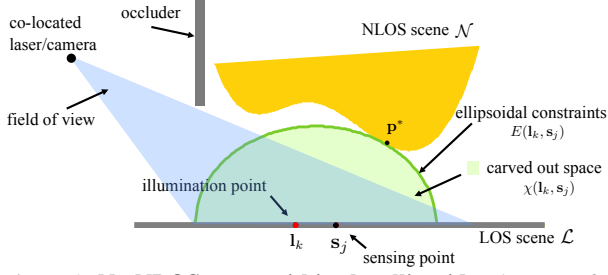


Figure 3. **No NLOS scene within the ellipsoid.** The ToF of the first-returning photon at each illumination and sensing pair creates an ellipsoidal constraint. Since we measure the shortest path between \mathbf{l}_k and \mathbf{s}_j , there should be no other NLOS scene with a shorter light path, and thus we carve out space inside the ellipsoid.

Defining the interior of $E(\mathbf{l}_k, \mathbf{s}_j)$ as

$$\chi(\mathbf{l}_k, \mathbf{s}_j) = \{\mathbf{p} \mid \|\mathbf{p} - \mathbf{l}_k\| + \|\mathbf{p} - \mathbf{s}_j\| < d(\mathbf{l}_k, \mathbf{s}_j)\}, \quad (5)$$

we observe that

$$\chi(\mathbf{l}_k, \mathbf{s}_j) \cap \mathcal{N} = \phi.$$

This observation is visualized in Fig. 3.

Observation 1 constrains space where the NLOS scene \mathcal{N} can exist. We will build our observation how to use local smoothness assumption to estimate the NLOS scene. First, we discuss the assumption that the shortest path is generated by a single NLOS point.

Uniqueness of the shortest path. Given \mathbf{l}_k and \mathbf{s}_j , the shortest three-bounce path between them is assumed to be unique, i.e., we assume that there exists only one NLOS scene point $\mathbf{p}^*(\mathbf{l}_k, \mathbf{s}_j)$ such that

$$d(\mathbf{l}_k, \mathbf{s}_j) = \|\mathbf{p}^*(\mathbf{l}_k, \mathbf{s}_j) - \mathbf{l}_k\| + \|\mathbf{p}^*(\mathbf{l}_k, \mathbf{s}_j) - \mathbf{s}_j\|.$$

For simplicity, we will denote $\mathbf{p}^*(\mathbf{l}_k, \mathbf{s}_j)$ simply as \mathbf{p}^* .

Recall from Section 1.1 that given (1), the NLOS scene point(s) that contribute to the three-bounce light transient $L(t_0; \mathbf{l}_k, \mathbf{s}_j)$ lie on an ellipsoid; further, when t_0 is increased, the ellipsoid increases in size. Consider the point(s) of contact between the NLOS scene and the ellipsoid when t_0 is gradually increased; the shortest path between \mathbf{l}_k and \mathbf{s}_j is unique only when the contact is at a single location. This is always the case when the NLOS scene is a convex shape.

There are instances of non-convex NLOS scenes that violate this uniqueness assumption. In fact, given \mathbf{l}_k and \mathbf{s}_j , it is not hard to construct a NLOS scene such that there are multiple points of contact to the ellipsoid. However, this is not a generic condition in that it requires very careful design of the NLOS scene for a given LOS scene pair; hence, if we perturb \mathbf{l}_k and \mathbf{s}_j , the symmetry of the scene is invariably broken and we recover uniqueness of the shortest path. In practice, this implies that non-unique shortest path scenarios occur, at best, for a tiny subset of LOS scene pairs and

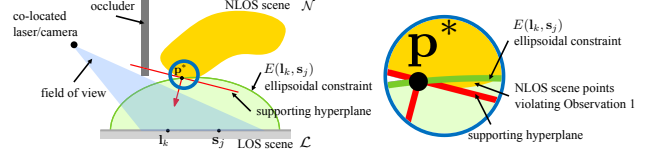


Figure 4. **When Observation 2 is violated, there exist NLOS scene points violating Observation 1.** If the supporting hyperplane at \mathbf{p}^* is not tangential to the ellipsoid $E(\mathbf{l}_k, \mathbf{s}_j)$, there will exist NLOS scene points belonging to the interior of the ellipsoid.

can be handled as outliers. We elaborate specific examples in the supplemental material.

With local smoothness assumption, we can characterize the local neighborhood of the unique scene point \mathbf{p}^* .

Observation 2. Suppose that NLOS scene is locally smooth at \mathbf{p}^* . Then, the (unique) supporting hyperplane at \mathbf{p}^* is tangential to the ellipsoid $E(\mathbf{l}_k, \mathbf{s}_j)$.

Proof. Suppose that the supporting hyperplane to the NLOS scene at \mathbf{p}^* is not tangential to the ellipsoid $E(\mathbf{l}_k, \mathbf{s}_j)$. Then, given local smoothness, we can show that there are NLOS scene points in an infinitesimal neighborhood of \mathbf{p}^* that belong to the interior of the ellipsoid, $\chi(\mathbf{l}_k, \mathbf{s}_j)$. This contradicts Observation 1. \square

We visualize this scenario in Fig. 4 and provide a detailed algebraic proof in the supplemental material. It is also worth noting that when local smoothness is violated — for example, at corners — the supporting hyperplane to the NLOS object is not unique.

Observation 2 also implies that the supporting hyperplanes at \mathbf{p}^* to both the NLOS scene and the ellipsoid are identical. This provides us with an explicit expression for the surface normal of the NLOS object at \mathbf{p}^* .

Observation 3. Under local smoothness of the NLOS scene at \mathbf{p}^* , the surface normal $\mathbf{n}(\mathbf{p}^*)$ is the angular bisector of the vectors from \mathbf{p}^* to the illumination spot \mathbf{l}_k and sensing spot \mathbf{s}_j , respectively; that is,

$$\mathbf{n}(\mathbf{p}^*) \propto \frac{\mathbf{l}_k - \mathbf{p}^*}{\|\mathbf{l}_k - \mathbf{p}^*\|} + \frac{\mathbf{s}_j - \mathbf{p}^*}{\|\mathbf{s}_j - \mathbf{p}^*\|}. \quad (6)$$

Observation 3 directly follows the property of ellipsoids. The following geometric interpretation of Observation 3 is useful for the shape recovery algorithms in Section 4. Given the supporting hyperplane H at \mathbf{p}^* , we find the mirror image of the illumination point \mathbf{l}_k with respect to H ; we denote the mirror image as \mathbf{l}'_k . When the surface normal \mathbf{n} satisfies (6), then it can be shown that the straight line joining the mirror image \mathbf{l}'_k to \mathbf{s}_j passes through \mathbf{p}^* . A brief proof of this is presented in Fig. 5.

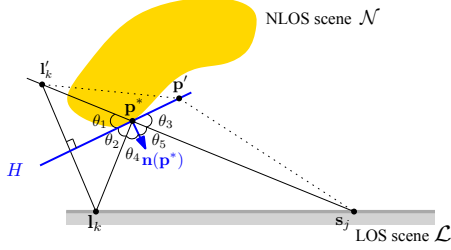


Figure 5. **Geometric interpretation of the normal being the angle bisector.** l'_k is the mirror image of l_k w.r.t H . Then, $\theta_1 = \theta_2$ and the distance from any point on H to l_k is the same as the distance to l'_k . Since the shortest path from l'_k to s_j is the straight line connecting them, p^* lies on this straight line; otherwise, there exists an alternate path that is shorter. Therefore, $\theta_3 = \theta_1$. We can conclude that $\theta_4 = \pi/2 - \theta_2 = \pi/2 - \theta_1 = \pi/2 - \theta_3 = \theta_5$. The normal $\mathbf{n}(p^*)$ is the angular bisector.

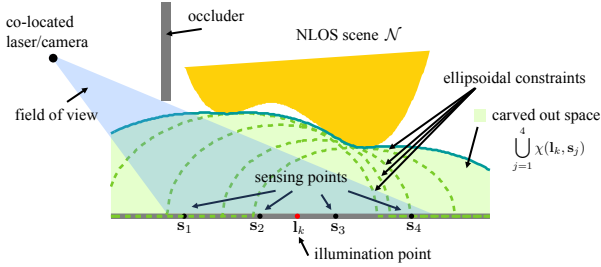


Figure 6. **Space carving for NLOS imaging.** From Observation 1, we carve out space inside each ellipsoid. The space carved region will contain no NLOS scene points.

4. Shape from First-Returning Photons

We propose two algorithms utilizing the finding in the three observations in Section 3 — the first for carving out the space the NLOS object cannot occupy and the second for recovering the location and surface normals of NLOS scene points that generate the first-returning photons.

Algorithm 1 — Space carving for NLOS imaging. We can extend Observation 1 by incorporating space carving constraints from all pairs of illumination and sensor points. Specifically, the NLOS scene \mathcal{N} cannot lie within the union of the individual ellipsoids, i.e.,

$$\bigcup_{l_k, s_j \in \mathcal{L}} \chi(l_k, s_j) \cap \mathcal{N} = \emptyset,$$

where $\chi(l_k, s_j)$ is defined in (5). We illustrate space carving from multiple first-returning photons in Fig. 6, when illuminating a single LOS spot.

Algorithm 2 — NLOS shape recovery under local planarity assumption. We assume that a small neighborhood of LOS sensing spots $\{s_j \in \Omega\}$ receive first-returning photons from a locally-planar NLOS scene patch. Therefore, we can combine the measured ToFs to infer the location and orientation of the locally-planar patch. We achieve

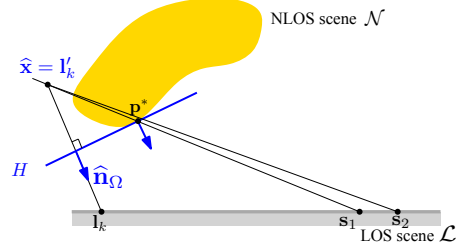


Figure 7. **NLOS shape and normal under local planarity.** If we assume the first-returning photons from neighboring sensors are from the same local planar patch, then they share the same mirrored light source location l'_k . Thus, finding l'_k can subsequently lead to the recovery of NLOS scene point and surface normal.

this by estimating the location of the mirror image of the illumination point (see Fig. 7).

Recall that the length of the shortest path from l_k to s_j is equal to the distance between the mirror image l'_k and s_j , $\|l'_k - s_j\|$; hence, $d(l_k, s_j) = \|l'_k - s_j\|$. Given the collection $\{d(l_k, s_j) \mid s_j \in \Omega\}$, we can solve for the location of the mirror image as

$$\min_{\mathbf{x}} \sum_{s_j \in \Omega} (d(l_k, s_j) - \|\mathbf{x} - s_j\|)^2. \quad (7)$$

The optimization problem is non-convex; we solve it using gradient descent techniques and, thus, the result depends heavily on the initialization. We initialize with the algebraic minimizer of the objective function, which we detail in the supplemental material. Once we have an estimate for the mirror image $\hat{\mathbf{x}}$, the estimate of the surface normal to the planar patch is given as

$$\hat{\mathbf{n}}_\Omega = \frac{(l_k - \hat{\mathbf{x}})}{\|l_k - \hat{\mathbf{x}}\|}.$$

We can also identify $(l_k + \hat{\mathbf{x}})/2$ as a point on the supporting hyperplane H , which gives us the equation of the planar patch. We now identify points on the plane that produce the first-returning photons by intersecting this plane with the straight lines from $\hat{\mathbf{x}}$ to each of the points $s_j \in \Omega$.

5. Experiments

We demonstrate the effectiveness of our two proposed algorithms. We show how space carving can help reduce the free space where the NLOS object resides. Also, we use local planarity to localize scene points. Unless otherwise noted, the NLOS scene is Lambertian and the neighborhood size used for (7) is 15.

5.1. Simulated Results

Rendering setup. We used the code base from [5] to render the 5D light transient. Rendering three-bounce light paths is a time consuming process, therefore, to improve the efficiency of rendering, we modify the imaging setup by

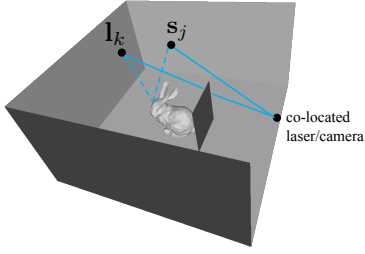


Figure 8. **Scene setup for demonstrating the space carving algorithm.** We place a bunny behind an occluder and orient the laser and SPAD toward the wall to image the NLOS bunny. For ease of visualization, we omit the top wall. We show the space carving result of using 100 pairs of illumination and sensing spots.

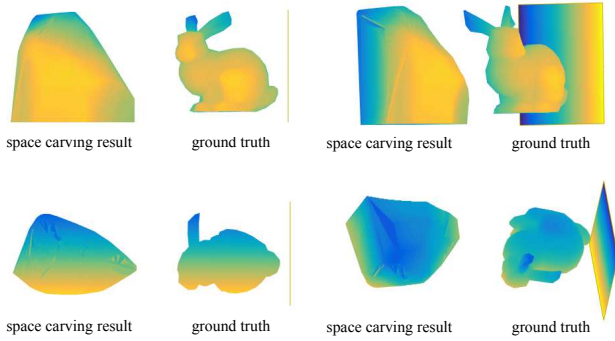


Figure 9. **Space carving algorithm result.** We present the space carving result alongside with the ground truth in four views. For best visualization, please refer to the supplementary material.

placing an omnidirectional light source and an omnidirectional sensor on the LOS scene and rendering single-bounce light paths. This setup is equivalent to NLOS imaging.

Space carving for NLOS imaging. We demonstrate NLOS space carving for the scene, shown in Fig. 8, consisting of a bunny in a room occluded from the laser/camera location. The result of the space carving is visualized in Fig. 9. We observe that the bunny is outside the carved region. Further, the volume of the entire NLOS scene is 5.56 cubic meters. With space carving, we identify free space in the NLOS scene and hence, the resulting NLOS scene volume is reduced to 0.44 cubic meters, which is only 7.86% of the original space. The volume that the bunny occupies is 0.20 cubic meters. Our method successfully decreases the space of possible NLOS scene. However, due to the complex shape of the bunny, some space will not be carved out.

Shape recovery with different reflectances. The ToF is solely a function of the geometry of the NLOS scene. Thus, our proposed method can work for NLOS scene with highly specular reflectance as long as some diffuse component is present. We verify this claim by rendering the 5D light transient of the NLOS scene with different reflectances, from

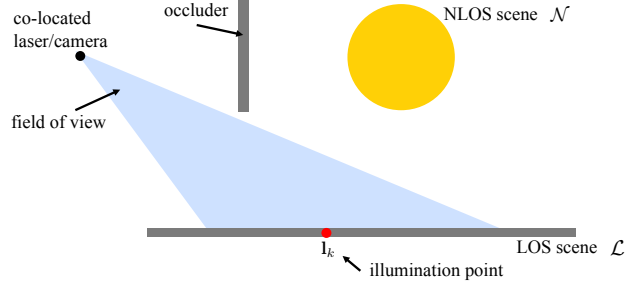


Figure 10. **Shape recovery scene setup.** Here we show the side view of the setup. The NLOS scene consists of a sphere. We image the sphere through 1 illumination point and 957 sensing points.

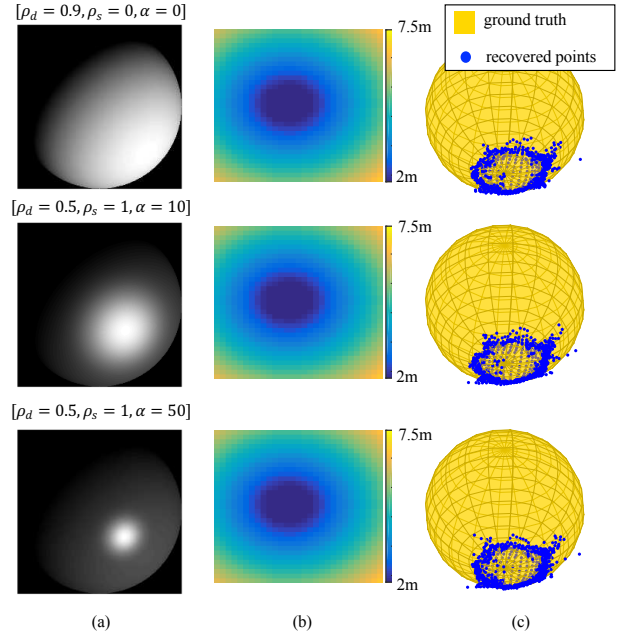


Figure 11. **Shape reconstruction under varying NLOS scene reflectance.** Here we show results of 3 different material reflectances. (a) Reflectance visualization, shown by illuminating the sphere with a point light source. (b) The rendered ToF at each sensing location. (c) The recovered NLOS scene points.

purely Lambertian to highly specular. We control the reflectance via the parameters of the Blinn-Phong model. We show the scene setup in Fig. 10. The NLOS object is a sphere and we illuminate one point and sense at 957 points. As shown in Fig. 11, the rendered ToF and hence, the recovered points, are largely invariant to the NLOS reflectances.

Shape recovery with different noise levels. The scene contains a sphere where the location of the center \mathbf{o} and the radius r are known. We compute the recovery error by finding the average distance between the recovered point to the surface of the sphere.

$$E_{\mathbf{p}} = \frac{1}{m} \sum_{i=1}^m |||\mathbf{p}_i - \mathbf{o}| - r|,$$

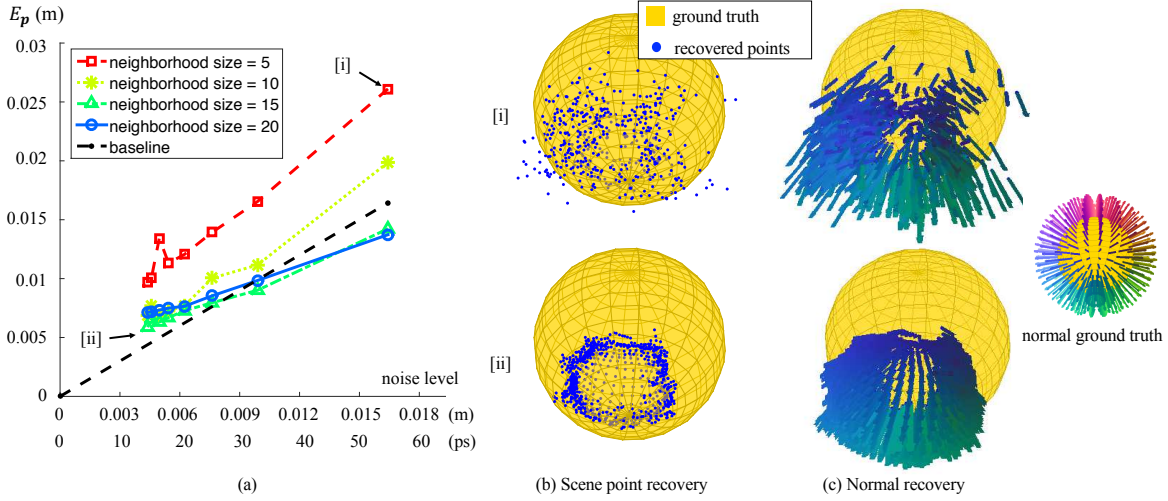


Figure 12. **Shape recovery with different noise levels and different neighborhood sizes.** We are able to vary the noise level by varying the number of photons used in rendering. (a) The shape recovery error with respect to different noise levels. [i] Recovery result of noise level at 0.0164 m with neighborhood size = 5. [ii] Recovery result of noise level at 0.0044 m with neighborhood size = 15.

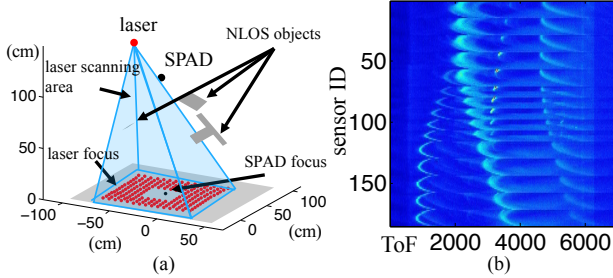


Figure 13. **NLOS imaging setup in [1].** (a) The laser scans 185 locations on the wall while SPAD focuses on one spot on the wall. (b) The 5D light transient collected by [1]. We extract first-returning photons from (b) to demonstrate the effectiveness of our proposed algorithms.

where \mathbf{p}_i is the i -th recovered point. As for the normal recovery error, we find the average angular error between the recovered normal and the normal of the projected point on the sphere.

$$E_n = \frac{1}{m} \sum_{i=1}^m \cos^{-1} \left(\mathbf{n}_i^\top \frac{\mathbf{p}_i - \mathbf{o}}{\|\mathbf{p}_i - \mathbf{o}\|} \right),$$

where \mathbf{n}_i is the normal estimation of recovered point \mathbf{p}_i .

As shown in Fig. 12(a), we observe that the reconstruction error is roughly linear to noise level. We also show the recovered shapes in Fig. 12(b) for visualization. We show recovered normals in Fig. 12(c) with color coded normal; E_n is less than 0.1° in all configurations.

Shape recovery with different neighborhood sizes. In (7), we use a local neighborhood to solve the mirrored light source location. By choosing larger area, we include more measurements, thus the effect of noise can be alleviated. However, for very large neighborhoods, the locally planar

assumption can be violated, causing large model misfit errors. We use the setup in Fig. 10 and compare the reconstruction results with respect to different neighborhood sizes. We observe in Fig. 12 that the error first reduces with increasing neighborhood size and subsequently, increases due to violation of the local planarity assumption.

5.2. Real Scene

Buttafava *et al.* [1] collect three-bounce light paths using 185 illumination points and sensing at a single spot on the wall (see Fig. 13). There are three NLOS objects, a T shape, a larger square behind the T shape, and a smaller square. There is an additional surface due to optical setup.

From the 5D light transient, we find the ToF of first-returning photon by finding the first time instant that exceeds a set threshold (14 in our experiment). Because of Helmholtz reciprocity, we can switch the role of illumination and sensing, then apply the shape recovery algorithm. We use a neighborhood size of 5 for shape recovery using local planarity assumption (Algorithm 2) and filter out points that satisfy the space carving constraint (Algorithm 1). The recovered result is shown in Fig. 14(a). Notice that we can only recover the two parts closest to the wall since the square behind the T shape cannot create first-returning photon observations.

By observing the 5D light transient (see Fig. 13(b)), we can clearly see multiple peaks in the light transient. Those are first-returning photons generated by other surfaces that are farther from the wall, thus they will not be picked out by finding the first time instant when the photon count exceeds the threshold. Therefore, we use signal processing method to find the peaks in the light transient then assign labels by clustering. As seen in Fig. 14(b), we are able to recover more scene points.

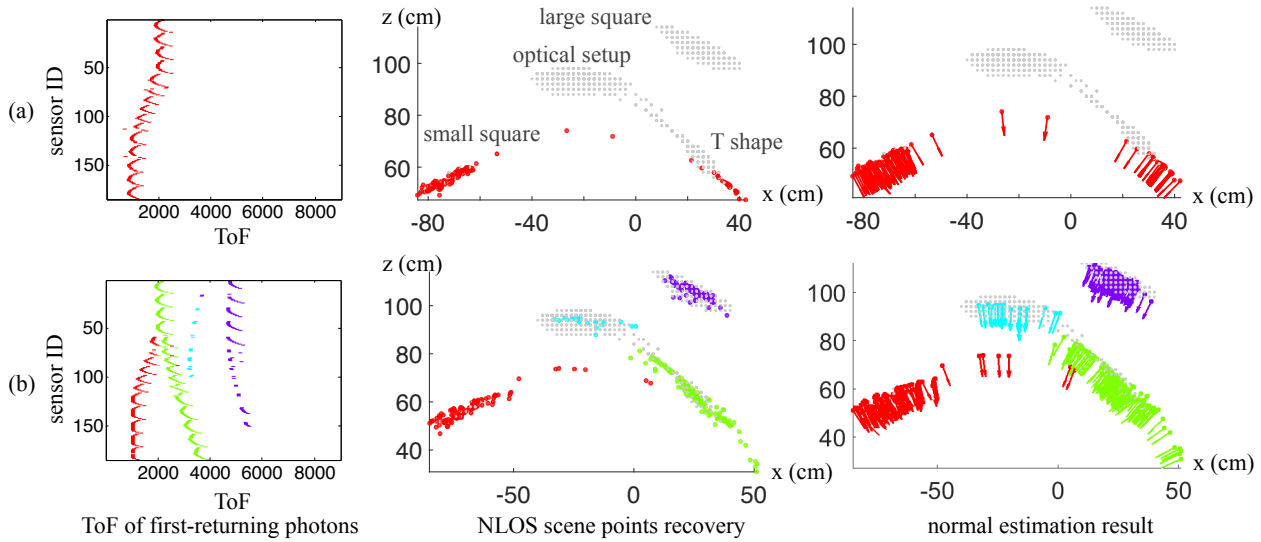


Figure 14. **Shape estimation from first-returning photons collected by SPAD [1].** We compare our proposed method using solely first-returning photons (red and other color dots) with elliptical tomography method using full 5D light transient (grey dots). For visualization, we only show the side view of the reconstruction, please refer to our supplementary video for complete 3D view. (a) First-returning photons by thresholding the photon count in the light transient. We show that we can recover points closer to the wall and the recovered points align with the result using elliptical tomography method. (b) We use signal process to extract the first-returning photon from different NLOS surfaces. For each surface, we repeat the process in (a). We can see that by finding first-returning photons from different surfaces in the 5D light transient, we can recover most of the NLOS scene.

Notice that we recover each scene point independently. The recovered scene points lie on different planar scene objects, which are consistent to the setup described in [1]. Also, the recovered normals (see Fig. 14) are roughly perpendicular to the point collection, which means our normal estimation is meaningful.

6. Conclusion and Discussions

This paper studies the geometry of first-returning photons in NLOS imaging and identifies novel constraints that arise from the study of shortest paths. To our knowledge, ours is the first technique that directly resolves NLOS shape and normals without solving complex inverse problems. The method proposed in this paper is computationally lightweight, hence is suitable for initializing and accelerating more complex methods that utilize the full light transient. To this end, we believe that this is an important step towards simpler techniques for NLOS imaging. In the following, we discuss some aspects of our approach.

Sampling of the NLOS scene points. A drawback of using first-returning photon is that the NLOS scene points that produce the first-returning photon is a function of the scene geometry, i.e., we have limited control over the points that end up generating the first-returning photons. As a consequence, it is entirely possible that some scene points will never create first-returning photons. One such example is the surface on the top right corner of Fig. 14. Only if we use signal processing techniques to extract more first-returning

photons from the light transient can we recover the shape.

NLOS reflectances, smoothness, and convexity From the Observation 3, the light path of first-returning photon follows the mirror direction. This means that specular BRDF is actually most advantageous for smooth convex object, since all photons will be first-returning photons. Lambertian reflectance, on the contrary, creates light paths that belongs to the tail of the light transient. This makes determining the ToF of first-returning photon harder.

Specular BRDF is not always favorable, especially for smooth non-convex objects, when there exist inter-reflections on the NLOS object. This makes separating three-bounce light path from higher-bounce light paths harder. However, in this case, for Lambertian reflectance, higher order bounce light paths attenuate faster and it is easier to identify three-bounce light paths. Specular BRDF is also not suitable for non-smooth object since some lighting/sensing pairs cannot receive any photon.

Acknowledgments

We thank Anonymous Reviewer 2 for valuable suggesting corrections to the proof of Observation 1. We thank Andreas Velten at Wisconsin-Madison for sharing the SPAD dataset and Ashok Veeraraghavan at Rice for valuable discussions. All authors were supported, in part, by the DARPA REVEAL program. CYT gratefully acknowledges support from the Bertucci Graduate Fellowship and the Google PhD Fellowship. ACS was partially supported by the NSF CAREER award CCF-1652569.

References

- [1] M. Buttafava, J. Zeman, A. Tosi, K. Eliceiri, and A. Velten. Non-line-of-sight imaging using a time-gated single photon avalanche diode. *Optics Express*, 23(16):20997–21011, 2015. 2, 3, 7, 8
- [2] I. Gkioulekas, A. Levin, F. Durand, and T. Zickler. Micron-scale light transport decomposition using interferometry. *ACM Trans. Graphics*, 34(4):37:1–37:14, 2015. 3
- [3] F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich. Low-budget transient imaging using photonic mixer devices. *ACM Trans. Graphics*, 32(4):45:1–45:10, 2013. 3
- [4] F. Heide, L. Xiao, W. Heidrich, and M. Hullin. Diffuse mirrors: 3d reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors. In *CVPR*, 2014. 3
- [5] A. Jarabo, J. Marco, A. Muñoz, R. Buisan, W. Jarosz, and D. Gutierrez. A framework for transient rendering. *ACM Trans. Graphics*, 33(6):177:1–177:10, 2014. 5
- [6] A. Kadambi, J. Schiel, and R. Raskar. Macroscopic interferometry: Rethinking depth estimation with frequency-domain time-of-flight. In *CVPR*, 2016. 3
- [7] A. Kadambi, H. Zhao, B. Shi, and R. Raskar. Occluded imaging with time-of-flight sensors. *ACM Trans. Graphics*, 35(2):15:1–15:12, 2016. 3
- [8] A. Kirmani, T. Hutchison, J. Davis, and R. Raskar. Looking around the corner using transient imaging. In *ICCV*, 2009. 1
- [9] J. Klein, C. Peters, J. Martín, M. Laurenzis, and M. B. Hullin. Tracking objects outside the line of sight using 2d intensity images. *Scientific Reports*, 6:32491, 2016. 3
- [10] M. O’Toole, F. Heide, L. Xiao, M. B. Hullin, W. Heidrich, and K. N. Kutulakos. Temporal frequency probing for 5d transient analysis of global light transport. *ACM Trans. Graphics*, 33(4):87:1–87:11, 2014. 1
- [11] A. K. Pediredla, M. Buttafava, A. Tosi, O. Cossairt, and A. Veeraraghavan. Reconstructing rooms using photon echoes: A plane based model and reconstruction algorithm for looking around the corner. In *ICCP*, 2017. 3
- [12] C. Peters, J. Klein, M. B. Hullin, and R. Klein. Solving trigonometric moment problems for fast transient imaging. *ACM Trans. Graphics*, 34(6):220:1–220:11, 2015. 3
- [13] R. Ramesh and J. Davis. 5d time-light transport matrix: What can we reason about scene properties? *Tech. rep., MIT*, 2008. 1
- [14] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Comm.*, 3:745 – 758, 2012. 1, 2, 3