Efficient Non-Line-of-Sight Imaging
from Transient Sinograms

Mariko Isogawa, Dorian Chan, Ye Yuan, Kris Kitani, and Matthew O'Toole

Carnegie Mellon University, Pittsburgh PA 15213, USA

Abstract. Non-line-of-sight (NLOS) imaging techniques use light that diffusely reflects off of visible surfaces (e.g., walls) to see around corners. One approach involves using pulsed lasers and ultrafast sensors to measure the travel time of multiply scattered light. Unlike existing NLOS techniques that generally require densely raster scanning points across the entirety of a relay wall, we explore a more efficient form of NLOS scanning that reduces both acquisition times and computational requirements. We propose a circular and confocal non-line-of-sight (C^2NLOS) scan that involves illuminating and imaging a common point, and scanning this point in a circular path along a wall. We observe that (1) these C^2NLOS measurements consist of a superposition of sinusoids, which we refer to as a transient sinogram, (2) there exists computationally efficient reconstruction procedures that transform these sinusoidal measurements into 3D positions of hidden scatterers or NLOS images of hidden objects, and (3) despite operating on an order of magnitude fewer measurements than previous approaches, these C^2NLOS scans provide sufficient information about the hidden scene to solve these different NLOS imaging tasks. We show results from both simulated and real C^2NLOS scans.[1]

**Keywords:** Computational imaging, non-line-of-sight imaging

1 Introduction

The ability to image objects hidden outside of a camera's field of view has many potential applications [23], including autonomous driving, search and rescue, and remote imaging. Over the last decade, many different technologies have been used for non-line-of-sight (NLOS) imaging, including transient imaging [2, 6, 7, 10–13, 16, 19, 22, 25, 27–29, 32, 33, 35], conventional cameras [4, 8, 17, 18, 30, 31], WiFi or radio frequency measurements [1, 20], thermal imaging [24], and even audio-based techniques [21]. Transient imaging refers to measuring a scene's temporal response to a pulse of light, and is one of the more successful approaches to reconstructing high-quality 3D shape of hidden scenes.

NLOS imaging techniques are fundamentally dependent on the spatial scanning patterns they utilize. Initially, methods used exhaustive measurements of

---

[1] Project page: https://marikoisogawa.github.io/project/c2nlos
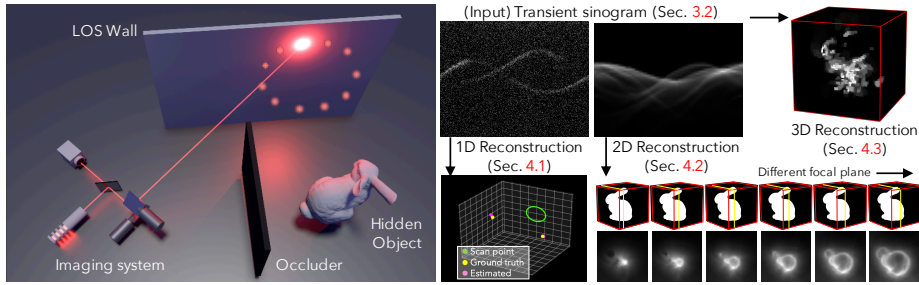
Fig. 1: A circular and confocal non-line-of-sight (C^2NLOS) system scans points along a circular path on a relay wall. Exploiting the sinusoidal properties of C^2NLOS measurements, a circular scan of a wall is sufficient to reconstruct images. C^2NLOS operates on far fewer measurements than existing NLOS techniques.

5D transients [2, 10, 16, 33], requiring explicit scanning of both virtual sources and sensors on a line-of-sight (LOS) wall. To mitigate these issues, alternative approaches proposed co-locating the source and sensor points [27], reducing the dimensionality of the required scanning to just two spatial dimensions which significantly expedites computation. However, even these techniques still require a full raster scan of a wall, which is limited to 2 Hz to 4 Hz for state-of-the-art NLOS systems [22]—too slow for real-time capture.

All of these previous NLOS techniques motivate the following two-fold question. First, what is the dimensionality of the smallest set of measurements that is sufficient for reconstructing a NLOS image? And second, among all measurement sets of this size, which ones lend themselves to efficient reconstruction algorithms? Answering these questions involves many complicated considerations, including the need to define the exact reconstruction problem we are solving.

While we do not provide definitive answers to these questions in this paper, we take first steps towards addressing them. In particular, we identify the subset of measurements produced by a circular and confocal non-line-of-sight (C^2NLOS) scan, which yields powerful properties that facilitate fast reconstructions. As shown in Fig. 1, C^2NLOS scanning involves sampling points that form a circle on a visible surface, reducing the dimensionality of transient measurements under this regime to just 2 dimensions. Our key observation is that NLOS images can be obtained with far fewer measurements than previously expected or demonstrated by existing NLOS systems and reconstruction techniques. With off-the-shelf large beam scanning galvo systems (e.g., Thorlabs GVS012), circular scanning is also fast and potentially supports real-time NLOS tasks at 130 Hz.

In addition to having smaller dimensionality and being efficient to acquire, C^2NLOS measurements satisfy the requirements set out above, sufficiency and computational efficiency, for two important NLOS reconstruction problems. The first problem is localizing a discrete number of small objects ("scatterers"). We show that C^2NLOS measurements are sufficient for this task, and enable recovery

of the unknown locations through a straightforward Hough voting procedure [3]. The second problem is reconstructing a single planar object. For small planar objects, we show that this problem can be reduced to one equivalent to computed tomography, and therefore can be solved using techniques developed for that task such as the inverse Radon transform [15]. Both results rely on a theoretical analysis that shows that the transient measurements from C^2NLOS scanning are a superposition of sinusoids (referred to as a *transient sinogram*) with different amplitudes, phases, and offsets. A one-to-one mapping directly relates the parameters of these sinusoids to the 3D position of the hidden scatterers.

Motivated by the above results, we also empirically investigate two related problems. We show that accurate 2D images of large planar scenes can be obtained by solving a simple linear least squares problem based on C^2NLOS measurements. Furthermore, we demonstrate that approximate 3D reconstructions of the NLOS scene can be efficiently recovered from C^2NLOS measurements.

To summarize, our contributions are the following: (i) we provide a theoretical analysis of our proposed C^2NLOS scanning procedure which shows that the measurements consist of a superposition of sinusoids, producing a transient sinogram; (ii) we propose efficient reconstruction procedures that build on these sinusoidal properties to localize hidden objects and reconstruct NLOS images; and (iii) we show that the C^2NLOS measurements are sufficient for reconstructing 1D, 2D, and 3D NLOS images from both simulated and real transients, while using far fewer measurements than existing methods.

It should be stated up front that our reconstruction quality is strictly worse than conventional methods that make use of a larger set of transient measurements and longer capture times. In contrast to past NLOS works where the objective is to improve reconstruction quality, our key objective is to show that **a circular scan of transients is sufficient to reconstruct a NLOS image**.

2 Overview of Transient NLOS Imaging

NLOS imaging has received significant attention, with solutions that operate on a wide variety of different principles. However, a common approach to NLOS imaging involves using transient sources and sensors that operate on visible or near-IR light; we refer to these as transient NLOS imaging systems.

C^2NLOS imaging is a transient-based technique which shares many similarities to previous work in confocal NLOS imaging [27]. We therefore review the general NLOS image formation model, followed by confocal NLOS imaging.

General NLOS Imaging [33]. In Fig. 2(a), a laser sends a pulse of light towards a 3D point $\mathbf{x}''$ on a visible wall, and the light diffusely scatters from that point at time $t = 0$. The scattered light illuminates the objects hidden around a corner, and a fraction of that light reflects back towards the wall in response. A transient sensor (e.g., a SPAD) then measures the temporal response at a point $\mathbf{x}'$ on the wall, also known as a transient measurement [26]. The transient measurement, $\tau(\mathbf{x}', \mathbf{x}'', t)$, represents the amount of light detected at point $\mathbf{x}'$ at

(a) General NLOS measurement   (b) Confocal NLOS measurement   (c) Circular Confocal NLOS measurement
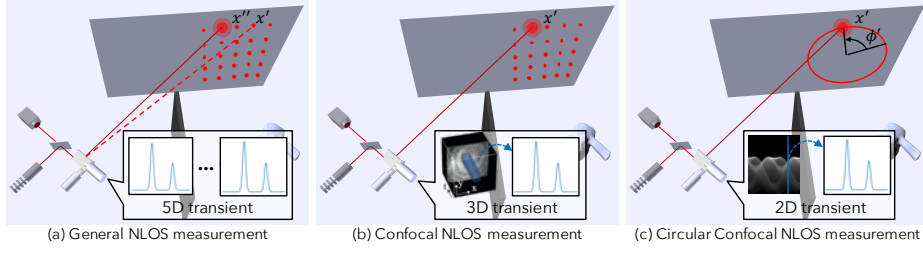
Fig. 2: Illustration of transient NLOS scans. A pulsed laser illuminates a point $\mathbf{x}''$, while a transient sensor images a point $\mathbf{x}'$ on a LOS wall. By changing $\mathbf{x}''$ and $\mathbf{x}'$, we obtain a diverse set of transients that can help identify the position, shape, and appearance of objects hidden from sight. **(a)** Conventional NLOS imaging scans several combinations of light source positions $\mathbf{x}''$ and detector positions $\mathbf{x}'$ on the wall to obtain a 5D transient measurement. **(b)** Confocal NLOS imaging illuminates and images the same point, i.e., $\mathbf{x}'' = \mathbf{x}'$, producing a 3D transient measurement. **(c)** C^{2}NLOS imaging proposes confocally scanning only those points that lie on a circle, yielding a 2D transient measurement. We propose a transformation that reduces this transient image into a sum of sinusoids, called a *transient sinogram*, where the amplitude, phase, and offset of each sinusoid corresponds to the position of a hidden scatterer.

time t, given illumination from point $\mathbf{x}''$ at time $t = 0$. For simplicity, we ignore the travel time between the system to the wall itself, which can be accounted for given the wall's geometry relative to the position of the laser and sensor.

The standard image formation model for NLOS imaging is

$$\tau(\mathbf{x}', \mathbf{x}'', t) = \iiint_{\Omega} \rho(\mathbf{x}) \; \frac{\delta\left(\|\mathbf{x} - \mathbf{x}'\| + \|\mathbf{x} - \mathbf{x}''\| - tc\right)}{\|\mathbf{x} - \mathbf{x}'\|^2 \|\mathbf{x} - \mathbf{x}''\|^2} \; d\mathbf{x}, \qquad (1)$$

where the function $\rho(\mathbf{x})$ represents the albedo of objects at every point $\mathbf{x}$, and $c = 3 \times 10^8$ m/s is the speed of light. The expression inside the Dirac delta $\delta(\cdot)$ relates the distance light travels through the hidden volume to its time of flight. The denominator accounts for the decrease in the intensity of light as a function of distance traveled, as given by the inverse square law.

This image formation model makes three underlying assumptions: (i) it only models three-bounce light paths, (ii) the model ignores the effect of a material's reflectance function and surface orientation, and (iii) the model assumes no occlusions within the hidden volume.

Equation (1) can be discretized into a linear system of equations

$$\boldsymbol{\tau} = \mathbf{A}\boldsymbol{\rho}, \qquad (2)$$

where $\boldsymbol{\tau}$ and $\boldsymbol{\rho}$ are discretized and vectorized representations of the measurements and volume, respectively. Recovering the hidden scene's geometry involves solving the linear system in Equation (2). Unfortunately, the matrix $\mathbf{A}$ can be extremely large in practice. In general, the matrix maps a 3D volume $\boldsymbol{\rho}$ to a 5D

transient measurement $\boldsymbol{\tau}$ (4D spatial + 1D temporal). In this case, the matrix is far too large to construct, store, and invert directly.

As a result, many works have explored different sampling patterns that reduce the size of the measurements and simplify the reconstruction procedure. Certain approaches simply fix the light source and scan the sensor (or vice versa), producing 2D spatial measurements [6, 11]. SNLOS [28] temporally focuses the light reflecting off of a single voxel, by simultaneously illuminating and imaging the wall over an ellipse. To scan a 2D or 3D set of voxels, ellipses of different shapes and sizes are used. Keyhole NLOS imaging [25] illuminates and detects light at a single point on the LOS wall, and relies on the motion of the hidden object to produce measurements for NLOS imaging. Confocal NLOS imaging [22, 27] scans the source and sensor together, and is described next in more detail.

Confocal NLOS Imaging [27]. Confocal NLOS imaging (Fig. 2(b)) co-locates the source and sensor by setting $\mathbf{x}' = \mathbf{x}''$, and samples a regular 2D grid of points on the wall. This sampling strategy has a number of practical advantages. First, it simplifies the NLOS calibration process, since the shape of the wall is given by direct reflections. Second, confocal scans capture light from retroreflective objects, which helps to enable NLOS imaging at interactive rates [22]. Finally, there exist computationally and memory efficient algorithms for recovering hidden volumes from confocal scans without explicit construction of matrix $\mathbf{A}$.

When co-locating the source and detector, Equation (1) reduces to

$$\tau(\mathbf{x}', t) = \iiint_{\Omega} \rho(\mathbf{x}) \, \frac{\delta\left(2\|\mathbf{x}' - \mathbf{x}\| - tc\right)}{\|\mathbf{x}' - \mathbf{x}\|^4} \, d\mathbf{x}. \tag{3}$$

As discussed by O'Toole et al. [27], a change of variables $v = (tc/2)^2$ produces

$$\tilde{\tau}(\mathbf{x}', v) \equiv v^{\frac{3}{2}} \, \tau(\mathbf{x}', \tfrac{2}{c}\sqrt{v}) = \iiint_{\Omega} \rho(\mathbf{x}) \, \delta\left(\|\mathbf{x}' - \mathbf{x}\|^2 - v\right) \, d\mathbf{x}. \tag{4}$$

When the relay wall is planar (i.e., $z' = 0$), the 3D spatio-temporal response $\tilde{\tau}(\mathbf{x}', v)$ of a scatterer becomes shift-invariant with respect to its 3D position $\mathbf{x}$. Equation (4) can then be expressed as a simple 3D convolution, which can be efficiently evaluated using a fast Fourier transform. The inverse problem involves a simple 3D deconvolution procedure called the light cone transform (LCT).

3   The Geometry of Circular and Confocal Scanning

While previous approaches have successfully reduced both capture and reconstruction times, the scanning paths required by these techniques inherently restrict scanning speeds on current hardware. Typical NLOS imaging systems, such as the one developed by Lindell et al. [22], use a pair of large galvo mirrors to raster scan the wall. The mirrors can only be driven up to a maximum of 65 Hz for a square wave pattern, capping scanning to just 130 lines per second. As a result, even the 2D grids utilized by confocal approaches [22, 27] are limited
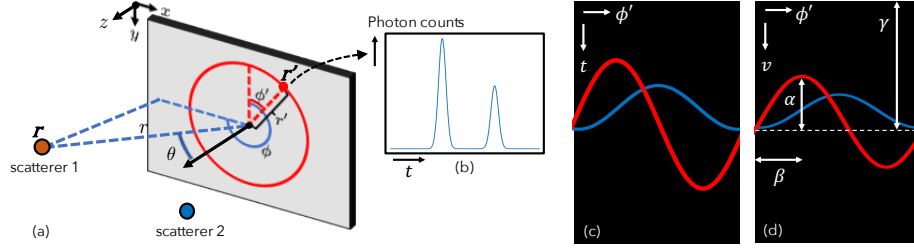
Fig. 3: Geometry of a C^2NLOS scan for individual scatterers. **(a)** The system confocally scans the red circle of radius r' one point at a time to image hidden objects. **(b)** Each point produces a transient, i.e., the temporal response to a pulse of light. This signal represents the travel time from a point on the wall, to the scatterers, and back again. **(c)** Scanning different points on the circle produces a collection of transients. Note that the signals represented here are only approximately sinusoidal. **(d)** By resampling the transients through a change of variables $v = (tc/2)^2$ (as explained in Eq. (4)), we obtain a transient sinogram. Every scatterer produces a unique sinusoid with a specific amplitude α, phase β, and offset γ. The parameters of these sinusoids are directly related to the spherical coordinates of the scatterers; see Eq. (10).

to just a few hertz (e.g., 32×32 at $4\,\text{Hz}$, or 64×64 at $2\,\text{Hz}$), impractical for dynamic scenes. Higher dimensional non-confocal measurements are even slower. Although smaller mirrors enable higher-frequency modulation (e.g., MEMS mirrors operate at kHz rates), this would greatly reduce the light efficiency of the system, lowering the quality of the output measurement.

This fundamental mechanical limitation motivates the following question: *can we further reduce the scanning path to just a single dimension, while still capturing useful information about the hidden scene?* We analyze the case of a circular and confocal scan (see Fig. 2(c)). Such a sinusoidal pattern could easily be captured at 130 Hz under current galvo-mirror systems—a typical NLOS setup can capture an entire 1D circular scan in the time it takes to capture a single row of a 2D grid scan. At the same time, these C^2NLOS scans encode significant information about the hidden scene. We investigate their properties in further detail in the rest of this section.

3.1 Equation (4) in Spherical Coordinates

We start by analyzing the form of Equation (4) when expressed in spherical coordinates. As shown in Fig. 3(a), we express the position of voxels in the hidden scene and scanning positions on the wall as $\mathbf{r} = (r, \theta, \phi)$ and $\mathbf{r}' = (r', \theta', \phi')$ respectively, where

$$x = r\sin(\theta)\cos(\phi), \quad y = r\sin(\theta)\sin(\phi), \quad z = r\cos(\theta), \tag{5}$$

for an azimuth angle $0 \leq \phi \leq 2\pi$, a zenith angle $0 \leq \theta \leq \pi$, and a radius $r \geq 0$.

Through a change of variables, we can rewrite the confocal NLOS image formation model of Equation (4) as

$$\tilde{\tau}(\mathbf{r}', v) \equiv v^{\frac{3}{2}} \ \tau(\mathbf{r}', \tfrac{2}{c}\sqrt{v}) = \iiint_{\Omega} \rho(\mathbf{r}) \ \delta\left(d(\mathbf{r}', \mathbf{r})^2 - v\right) \ r^2 \sin(\theta) \ dr \ d\theta \ d\phi, \quad (6)$$

where the distance function $d(\cdot, \cdot)$ expressed in spherical coordinates is

$$v(\mathbf{r}') \equiv d(\mathbf{r}', \mathbf{r})^2 = r^2 + r'^2 - 2rr' \left(\sin(\theta)\sin(\theta')\cos(\phi - \phi') + \cos(\theta)\cos(\theta')\right). \quad (7)$$

We restrict scatterers to be on one side of the wall, by setting $\theta \leq \pi/2$.

3.2 Transient Sinograms

We assume that our C^{2}NLOS scans points along a wall where $\theta' = \pi/2$, the points on the wall are on a circle of fixed radius r', and the center of the circle is the origin $\mathbf{0}$. By applying these assumptions to Equation (7), we get

$$v(\phi') = r^2 + r'^2 - 2rr' \sin(\theta)\cos(\phi - \phi') = \gamma - \alpha\cos(\beta - \phi'), \quad (8)$$

where

$$\alpha = 2rr'\sin(\theta), \quad \beta = \phi, \quad \gamma = r^2 + r'^2. \quad (9)$$

Here, α, β, and γ represent the amplitude, phase, and offset of a sinusoid. Therefore, after resampling the transient measurements (Fig. 3(c)) with the substitution $v = (tc/2)^2$, the transient measurement resulting from a C^{2}NLOS scan becomes a 2D image representing a superposition of different sinusoids, where each sinusoid represents a different point in the hidden space (see Fig. 3(d)). We therefore refer to the corresponding measurement as a *transient sinogram* $\boldsymbol{\tau}_{\mathrm{circ}}$.

Consider scatterers of the form $\mathbf{r} = (r, 0, 0)$ for all $r \geq 0$. These scatterers produce a sinusoid with zero amplitude, because all points on the circle are equidistant to the scatterer. As we change the zenith angle $\mathbf{r} = (r, \theta, 0)$ for $0 \leq \theta \leq \pi/2$, the amplitude of the sinusoid also increases up to a maximum of $2rr'$ when scatterers are adjacent to the wall. Finally, introducing an azimuth angle $\mathbf{r} = (r, \theta, \phi)$ for $0 \leq \phi \leq 2\pi$ produces a phase shift of the sinusoid.

After identifying the amplitude, phase, and offset of each scatterer's sinusoid, inverting the expression in Equation (9) recovers the scatterer's position:

$$r = \sqrt{\gamma - r'^2}, \quad \theta = \arcsin\left(\frac{\alpha}{2rr'}\right), \quad \phi = \beta. \quad (10)$$

This expression becomes useful when estimating the positions of one or a handful of scatterers around a corner, as discussed in Section 4.1.

In addition to computing the position of scatterers, another NLOS objective is to reconstruct images of the hidden scene. Consider an object that lies on the surface of a sphere of known radius r and centered about point $\mathbf{0}$ on the wall. This scenario occurs when sufficiently-small planar objects are oriented towards the

origin $\mathbf{0}$. (For planar objects tilted away from the origin, one can scan a different circle centered about another point on the wall.) According to Equation (9), the measurement consists of a combination of sinusoids, all of which have identical offsets. This simplifies the measurement into a conventional sinogram image, the same type of measurement used in computed tomography (CT) [15]. We exploit this property in Section 4.2 to recover 2D images of the hidden scene.

4   Reconstructing Images from Transient Sinograms

### 4.1   1D Reconstruction: Estimating 3D Positions

Given a transient sinogram $\boldsymbol{\tau}_{\mathrm{circ}}$, our first goal is to recover the 3D position of an object located at $\mathbf{x} = (x, y, z)$ by estimating its corresponding sinusoid parameters α, β, and γ, as described in Equation (10). The challenge is to perform this operation both accurately and robustly, e.g., in the presence of sensor noise.

We propose a convolutional approach to the Hough transform for fixed-period sinusoids, based loosely on [36]. First, we generate a 2D Hough kernel representing a sinusoid of a given amplitude. Through a FFT-based convolution between this kernel and the transient sinogram, we obtain a parameter-space image that produces large responses in areas where the kernel aligns well with the sinusoid in the transient sinogram. Second, we repeat this procedure multiple times for kernels representing sinusoids with different amplitudes, producing a three dimensional parameter-space volume. The location of the voxels with the highest values in the volume represent the parameters of the sinusoids in the transient sinogram. Using these sinusoids, we can then recover the scatterers in the hidden scene by applying Equation (10). We illustrate this process in Fig. 6(left).

4.2   2D Reconstructions

Consider the scenario where the hidden scene can be approximately modelled as a single planar object. We propose two ways to reconstruct a 2D image of this object. First, the inverse Radon transform is an integral transform used to solve the CT reconstruction problem; it is therefore possible to directly apply the inverse Radon transform technique on transient sinograms to recover a 2D image. This assumes that the planar patch is small, and tangent to the surface of a sphere of radius r, with the same center as the $\mathrm{C}^2\mathrm{NLOS}$ scan. Second, since it is shown transient sinograms preserve information about the hidden scene and the measurements are much smaller when compared to conventional NLOS scans, it becomes computationally feasible to explicitly construct a matrix $\mathbf{A}$ that directly maps points from a hidden 2D plane to the $\mathrm{C}^2\mathrm{NLOS}$ measurements, and solve a discrete linear system (i.e., Equation (2)) directly.

Inverse Radon Reconstruction When the hidden object lies on the surface of a sphere of radius r with the same center as the scanning circle (see Fig. 4(a)), each point on the hidden object produces a sinusoid with the same temporal
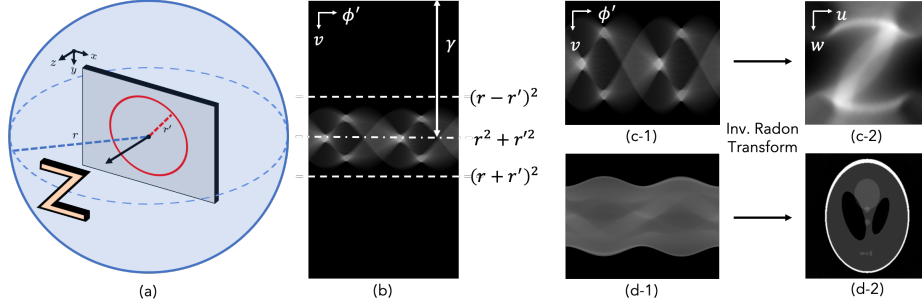
Fig. 4: Inverse Radon reconstruction based 2D imaging. **(a)** Suppose the hidden Z is a planar object that approximately lies on the surface of a sphere of radius r, and the C^{2}NLOS scan radius is r'. **(b)** Each scatterer on the surface of this sphere produces a sinusoidal response with a temporal offset $\gamma = r^2 + r'^2$ within a range $[(r-r')^2, (r+r')^2]$. **(c-1)** The transient sinogram is cropped with this range to recover a 2D image **(c-2)** via the inverse Radon transform. This is inspired by 2D image recovery with X-ray computed tomography (CT), shown in **(d-1)** and **(d-2)**.

offset γ, as shown in Fig. 4(b). We then recover a 2D image of the object from these measurements by using a standard inverse Radon transform procedure.

We choose a value for the radius r, either manually or automatically by computing the mean transient response. All sinusoids from points on this sphere have a corresponding offset $\gamma = r^2 + r'^2$. Because the maximum amplitude of sinusoids is $2rr'$, the transient response is contained within a temporal range $[\gamma - 2rr', \gamma + 2rr'] = [(r - r')^2, (r + r')^2]$. We therefore crop the transient sinogram accordingly, and apply the inverse Radon transform (see Fig. 4(d-1, d-2)) directly to the results to recover a 2D image (see Fig. 4(c-1, c-2)). The pixel coordinate of the recovered image associated for each sinusoid is given by

$$[u, w] = [\alpha \cos(\beta), \alpha \sin(\beta)] \qquad \text{(from Radon transform)} \quad (11)$$
$$= [2rr' \sin(\theta) \cos(\phi), 2rr' \sin(\theta) \sin(\phi)] \quad \text{(from Eq. (9))} \quad (12)$$
$$= 2r'[x, y] \qquad \text{(from Eq. (5))} \quad (13)$$

In other words, the recovered image simply represents a scaled orthographic projection of the hidden object onto the relay wall.

Linear Inversion Our analysis with the Radon Transform demonstrates that significant information about the hidden scene is encoded in a transient sinogram. However, in practice, most common objects are not fully contained within the surface of a sphere, instead consisting of multiple depths and distances from the center of the scanning circle. With that in mind, an important question that arises is whether a C^{2}NLOS measurement contains enough information to reconstruct these more general objects. To explore this question, we empirically

investigate the case of large 2-dimensional planar scenes, which are commonly used by existing NLOS techniques to gauge the accuracy of their reconstructions.

Under this constrained case, efficient recovery of a 2D image $\boldsymbol{\rho}_{2D}$ from C^{2}NLOS measurements involves solving a linear-least squares problem:

$$\boldsymbol{\rho}_{2D} = \arg\min_{\boldsymbol{\rho}} \quad \frac{1}{2}||\boldsymbol{\tau}_{\text{circ}} - \mathbf{A}_d\boldsymbol{\rho}||_2^2 + \frac{\lambda}{2}||\boldsymbol{\rho}||_2^2, \tag{14}$$

where λ controls the weight of the regularization term, and the matrix $\mathbf{A}_d$ represents the mapping from hidden scatterers on a plane d away from the center of the scanning circle, to C^{2}NLOS measurements.

Both our Radon reconstruction and planar inversion algorithms require knowledge of the sphere or plane containing the hidden object. If the object is not contained within the surface of the sphere/plane, the recovered images are blurred; we show an analysis in the supplement. Changing the value for r or d is analogous to a manual refocusing operation that can help produce a clearer image.

### 4.3   3D Reconstruction: 3D Imaging via a Modified LCT

A natural follow-up question that arises is whether a transient sinogram is sufficient for performing a full 3D reconstruction. Empirically, we show that it is feasible to recover full 3D volumes of the hidden scene from C^{2}NLOS measurements. Although this involves solving an underconstrained system due to the limited number of measurements, approximate reconstructions can be achieved by applying non-negativity, sparsity, and total variation priors on the hidden volume, commonly utilized by previous approaches [2, 11, 12, 27].

We propose a modified version of the iterative light cone transform (LCT) procedure used in confocal NLOS imaging [27]. Because a C^{2}NLOS measurement is a subset of a full confocal NLOS measurement, we simply add a sampling term to the iterative LCT procedure, solving the following optimization problem:

$$\boldsymbol{\rho}_{3D} = \arg\min_{\boldsymbol{\rho}} \quad \frac{1}{2}\left\|\boldsymbol{\tau}_{\text{circ}} - \mathbf{MA}\boldsymbol{\rho}\right\|_2^2 + \Gamma(\boldsymbol{\rho}), \tag{15}$$

where the matrix $\mathbf{A}$ maps a 3D volume to a confocal NLOS measurement, the matrix $\mathbf{M}$ subsamples the confocal NLOS measurement to produce a C^{2}NLOS measurement, and $\Gamma(\cdot)$ represents our non-negativity, sparsity, and total variation priors. Because the matrix $\mathbf{A}$ can be modelled as a convolution operation, the above expression can be optimized efficiently without having to construct $\mathbf{A}$ explicitly. We describe our procedure in detail in the supplement.

There are some drawbacks to this formulation. In its current form, it makes no explicit usage of the sinusoidal properties of C^{2}NLOS measurements that we utilized for object detection and 2D imaging, that could simplify the reconstruction. At the same time, the matrix $\mathbf{M}$ complicates a frequency analysis of the LCT, making it much more unclear which parts of the hidden scene can and cannot be reconstructed. We plan to investigate these phenomena in the future.
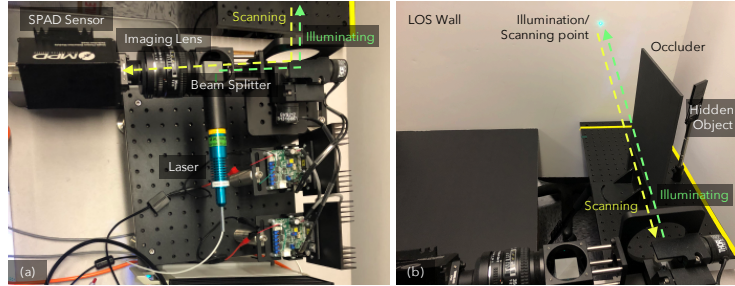
Fig. 5: (a) Our hardware prototype. **(b)** A hidden scene with a single NLOS object (a retroreflector) used for our object localization experiments.
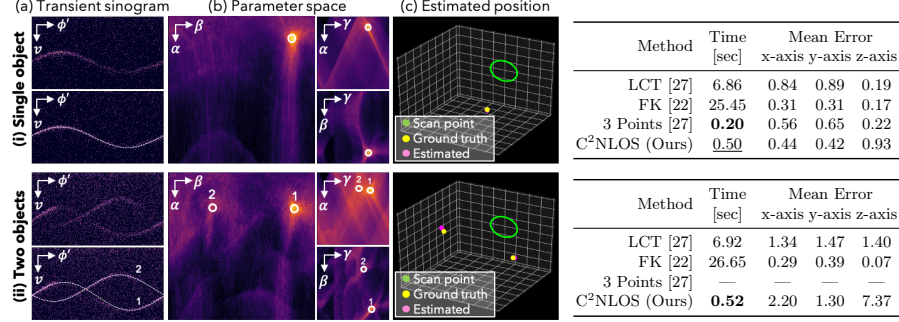
5 Experiments

Baseline algorithms. No existing algorithms operate on C^{2}NLOS scans, or even just 1D scans. Thus, we compare our method with two volume reconstruction approaches that rely on full 2D confocal scans: LCT [27] and FK [22]. We identify the peak and compute a maximum intensity projection from each of the output volumes to generate 1D and 2D reconstructions, respectively. To estimate scatterer positions, we also test a trilateration-based approach ("3 Points") that uses only three scanning points [27], which we describe in the supplement.

Hardware. Our prototype C^{2}NLOS system (Fig. 5) is based on the system proposed in O'Toole et al. [27]. Please refer to the supplemental material for more details. To estimate the computational efficiency of each algorithm, we ran each reconstruction algorithm on a 2017 Macbook Pro (2.5 GHz Intel Core i7).

Transient Measurement Data. For object localization, we use real captured data from our C^{2}NLOS acquisition system (see Fig. 6(a)), as well as simulated data. In our hardware acquisition system, we captured transients of size $1024(\phi') \times 4096(t)$ from a circular scanning pattern of diameter 1.0m. To qualitatively evaluate our single-object localization, we synthesized transients of size $64 \times 64 \times 2048$ from 100 randomly generated NLOS scenes, where we placed a single scatterer at a random location in a 1.0m×1.0m×1.0m volume 2.0m away from the LOS wall. We used 200 scenes in the two-object case. For 2D imaging, we used simulated transient data from the Z-NLOS Dataset [9, 14], which we resized to $64 \times 64 \times 2048$. For 3D imaging, we test our algorithm on real captured data provided by O'Toole et al. [27], as well as simulated data from the Z-NLOS Dataset, all rescaled to $64 \times 64 \times 512$. In all cases, to synthesize C^{2}NLOS data, we sampled 360 angles along the inscribed circle of the confocal grid.

On a typical NLOS hardware setup, the 64×64 grid data used by FK/LCT would only be captured at roughly 2 Hz. In contrast, C^{2}NLOS measurements can be captured at 130 Hz, corresponding to just 1.6% of the capture time.

Method	Time [sec]	Mean Error x-axis y-axis z-axis		
LCT [27]	6.86	0.84	0.89	0.19
FK [22]	25.45	0.31	0.31	0.17
3 Points [27]	**0.20**	0.56	0.65	0.22
C²NLOS (Ours)	0.50	0.44	0.42	0.93

Method	Time [sec]	Mean Error x-axis y-axis z-axis		
LCT [27]	6.92	1.34	1.47	1.40
FK [22]	26.65	0.29	0.39	0.07
3 Points [27]	—	—	—	—
C²NLOS (Ours)	**0.52**	2.20	1.30	7.37

Fig. 6: Estimating the position of (i) one and (ii) two scatterers. **Left:** Given a transient measurement in (a) top, we generate a sinusoid parameter space (b). The sinusoid parameters that best fit the transient sinogram (see (a) bottom) are obtained by finding its peak (see annotations on (b)). The 3D position of the object is reconstructed with the estimated parameters. **Right:** Quantitative evaluation with mean estimation error and computational time. Despite only requiring roughly 1.6% of the capture time of LCT or FK, C²NLOS estimated the position within almost the same order of accuracy as the other methods. Our approach was also faster than LCT and FK.

5.1   1D Reconstruction: Object Localization

Fig. 6(left) demonstrates our methodology for estimating a scatterer's position with real captured transient data, using the approach described in Section 4.1. Please note that the proposed method is applicable for more than two objects without loss of generality. Due to the robustness of Hough voting, our method detects the position of the hidden object(s) even though the transient measurements are quite noisy. See supplemental materials for more results.

For quantitative validation, Fig. 6(right) compares our method to the baseline approaches using the experimental setup outlined in the previous section. Despite the much smaller number of spatial samples, C²NLOS was able to achieve similar accuracy to LCT or FK, both of which require an order-of-magnitude more measurements. Computationally, our method was also faster than LCT and FK, but slower than the 3-Points method in the single object case. However, note that the 3-Points method does not generalize beyond a single object.

### 5.2   2D Reconstruction: 2D Plane Imaging

Fig. 7 shows the qualitative and quantitative 2D imaging results on large planar scenes. Despite requiring just 1.6% of the capture time, the linear inversion method was able to visualize the hidden image plane. At the same time, even though our inverse Radon reconstruction approach was not designed for large planar objects, it still recovers an approximate reconstruction of the 2D scene.

As we mentioned in Section 4.2, changing the value for the radius r of the sphere or distance d is analogous to a manual refocusing operation. Fig. 8 shows the results with the proposed Radon reconstruction-based method, in which the
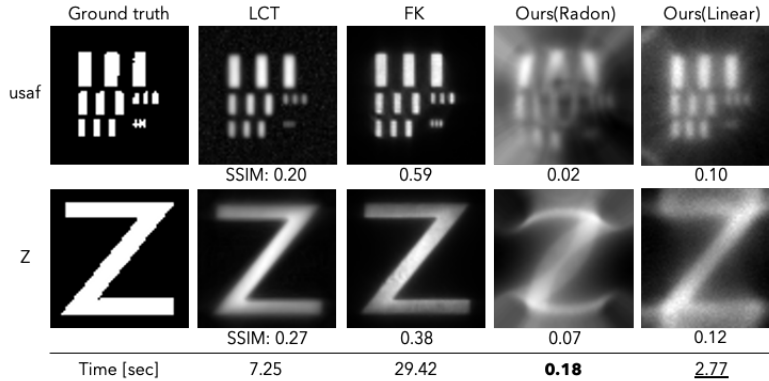
Fig. 7: Quantitative and qualitative results on 2D imaging on large planar scenes. Despite sampling far fewer measurements than LCT and FK, both of our inverse Radon reconstruction/linear inversion-based methods reconstructed images that were similar in quality. The SSIM scores for the proposed methods were slightly worse than LCT or FK. However, our methods were much more computationally efficient (e.g., inverse Radon reconstruction yielded a 50x speedup over LCT).
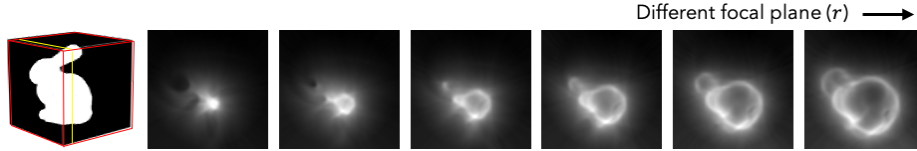

Fig. 8: Our inverse Radon reconstruction-based 2D imaging with different focus planes (right focuses towards larger depth). Even for non-planar objects like a bunny, C^2NLOS measurements contain sufficient information about the hidden object for an inverse Radon transform to approximately reconstruct its visual appearance.

results with larger r (farther from the wall) are shown to the right. Our method was able to visualize not only flat objects, but also objects with a wider range of depths, as shown by the *bunny* scene.

For quantitative validation, we used SSIM [34] as a metric. As shown in Fig. 7, both of our 2D imaging methods yield slightly worse results compared to LCT and FK. However, both our Radon-reconstruction and linear-inversion procedures are significantly faster than LCT and FK.

5.3   3D Reconstruction: 3D Volume Imaging

We use the Alternating Direction Method of Multipliers (ADMM) [5] to minimize the optimization problem from Equation (15). We show a full derivation in the supplement. Because of the iterative optimization process required by our strong priors, our reconstruction operator is inherently slower than both FK and LCT. However, our method yields similar runtimes to the iterative versions of LCT and the Gram operator [2], both of which use a similar optimization formulation.
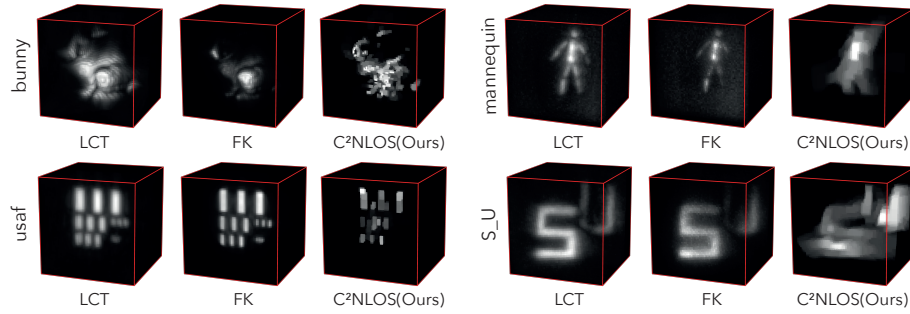
Fig. 9: 3D volume reconstruction results. Even though a C^2NLOS scan requires just 1.6% of the capture time needed by LCT and FK, our approach still generates an approximate reconstruction. More results in the supplement.

In order to evaluate our 3D reconstructions, we test LCT, FK, and our modified LCT procedure on a variety of different scenes in Fig. 9. In general, while FK and LCT demonstrate much higher reconstruction quality, our approach still captures important features of the hidden scene, like the presence and depth of multiple planar objects in the S_U scene and the overall pose in the *mannequin* scene. Empirically, this shows that significant volumetric information of the hidden scene can be recovered from a single transient sinogram.

6   Conclusion

We show that a transient sinogram acquired through C^2NLOS scanning is sufficient for solving a number of imaging tasks, even though the dimensionality of the measurement is smaller than those captured by existing NLOS methods. Through an analysis of the image formation model, we explain how the measurements are fundamentally sinusoidal and lend themselves to efficient reconstruction algorithms, including a Hough voting procedure for estimating the 3D position of scatterers and an inverse Radon technique for recovering 2D images of hidden objects. We empirically demonstrate that the measurements can also be applied to recover full 3D volumes. We believe these contributions mark a significant step in our understanding of *efficient* imaging techniques for revealing objects hidden just around a corner.

References

1. Adib, F., Hsu, C.Y., Mao, H., Katabi, D., Durand, F.: Capturing the human figure through a wall. ACM Transactions on Graphics (TOG) **34**(6), 1–13 (2015)
2. Ahn, B., Dave, A., Veeraraghavan, A., Gkioulekas, I., Sankaranarayanan, A.C.: Convolutional approximations to the general non-line-of-sight imaging operator. In: IEEE International Conference on Computer Vision (ICCV). pp. 7888–7898 (2019)
3. Ballard, D.H.: Generalizing the Hough transform to fetect arbitrary shapes. In: Readings in Computer Vision, pp. 714 – 725 (1987)
4. Bouman, K.L., Ye, V., Yedidia, A.B., Durand, F., Wornell, G.W., Torralba, A., Freeman, W.T.: Turning corners into cameras: Principles and methods. In: IEEE International Conference on Computer Vision (ICCV). pp. 2289–2297 (2017)
5. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning **3**(1), 1–122 (2011)
6. Buttafava, M., Zeman, J., Tosi, A., Eliceiri, K., Velten, A.: Non-line-of-sight imaging using a time-gated single photon avalanche diode. Optics Express **23**(16), 20997–21011 (2015)
7. Chan, S., Warburton, R.E., Gariepy, G., Leach, J., Faccio, D.: Non-line-of-sight tracking of people at long range. Optics Express **25**(9), 10109–10117 (2017)
8. Chandran, S., Jayasuriya, S.: Adaptive lighting for data-driven non-line-of-sight 3d localization and object identification. In: The British Machine Vision Conference (BMVC) (2019)
9. Galindo, M., Marco, J., O'Toole, M., Wetzstein, G., Gutierrez, D., Jarabo, A.: A dataset for benchmarking time-resolved non-line-of-sight imaging. In: ACM SIGGRAPH 2019 Posters, pp. 1–2 (2019)
10. Gupta, Otkrist; Willwacher, Thomas; Velten, Andreas; Veeraraghavan, Ashok; Raskar, R.: Reconstruction of hidden 3d shapes using diffuse reflections. Optics Express (20(17):19096–19108) (2012)
11. Heide, F., Heidrich, W., Hullin, M.B.: Diffuse mirrors: 3d reconstruction from diffuse indirect illumination using inexpensive time-of-flight sens. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3222–3229 (2014)
12. Heide, F., O'Toole, M., Zang, K., Lindell, D.B., Diamond, S., Wetzstein, G.: Non-line-of-sight imaging with partial occluders and surface normals. ACM Transactions on Graphics **38**(3) (2019)
13. Isogawa, M., Yuan, Y., O'Toole, M., Kitani, K.M.: Optical non-line-of-sight physics-based 3d human pose estimation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 7013–7022 (2020)
14. Jarabo, A., Marco, J., Muñoz, A., Buisan, R., Jarosz, W., Gutierrez, D.: A framework for transient rendering. ACM Transactions on Graphics (ToG) **33**(6), 1–10 (2014)
15. Kak, A.C., Slaney, M., Wang, G.: Principles of computerized tomographic imaging. Medical Physics **29**(1), 107–107 (2002)
16. Kirmani, A., Hutchison, T., Davis, J., Raskar, R.: Looking around the corner using transient imaging. In: IEEE International Conference on Computer Vision (ICCV). pp. 159–166 (2009)
17. Klein, J., Peters, C., Laurenzis, M., Hullin, M.: Tracking objects outside the line of sight using 2d intensity images. Scientific Reports **6**(32491), 1–9 (2016)

18. Klein, J., Peters, C., Laurenzis, M., Hullin, M.: Non-line-of-sight mocap. In: ACM SIGGRAPH Emerging Technologies. pp. 18:1–18:2 (2017)
19. La Manna, M., Kine, F., Breitbach, E., Jackson, J., Sultan, T., Velten, A.: Error backprojection algorithms for non-line-of-sight imaging. IEEE Transactions on Pattern Analysis and Machine Intelligence **41**(7), 1615–1626 (2019)
20. Li, T., Fan, L., Zhao, M., Liu, Y., Katabi, D.: Making the invisible visible: Action recognition through walls and occlusions. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 872–881 (2019)
21. Lindell, D.B., Wetzstein, G., Koltun, V.: Acoustic non-line-of-sight imaging. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 6780–6789 (2019)
22. Lindell, D.B., Wetzstein, G., O'Toole, M.: Wave-based non-line-of-sight imaging using fast f-k migration. ACM Transactions on Graphics (TOG) **38**(4), 116 (2019)
23. Maeda, T., Satat, G., Swedish, T., Sinha, L., Raskar, R.: Recent advances in imaging around corners. In: arXiv (2019)
24. Maeda, T., Wang, Y., Raskar, R., Kadambi, A.: Thermal non-line-of-sight imaging. In: IEEE International Conference on Computational Photography (ICCP). pp. 1–11 (2019)
25. Metzler, C.A., Lindell, D.B., Wetzstein, G.: Keyhole imaging: Non-line-of-sight imaging and tracking of moving objects along a single optical path at long standoff distances. In: arXiv (2019)
26. O'Toole, M., Heide, F., Lindell, D.B., Zang, K., Diamond, S., Wetzstein, G.: Reconstructing transient images from single-photon sensors. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2289–2297 (2017)
27. O'Toole, M., Lindell, D.B., Wetzstein, G.: Confocal non-line-of-sight imaging based on the light-cone transform. Nature **555**(7696), 338 (2018)
28. Pediredla, A., Dave, A., Veeraraghavan, A.: SNLOS: Non-line-of-sight scanning through temporal focusing. In: IEEE International Conference on Computational Photography (ICCP). pp. 1–13 (2019)
29. Redo-Sanchez, A., Heshmat, B., Aghasi, A., Naqvi, S., Zhang, M., Romberg, J., Raskar, R.: Terahertz time-gated spectral imaging for content extraction through layered structures. Nature Communications **7**, 12665 (2016)
30. Saunders, C., Murray-Bruce, J., Goyal, V.K.: Computational periscopy with an ordinary digital camera. Nature **565**(7740), 472–475 (2019)
31. Tancik, M., Satat, G., Raskar, R.: Flash photography for data-driven hidden scene recovery. In: arXiv (2018)
32. Tsai, C.Y., Sankaranarayanan, A.C., Gkioulekas, I.: Beyond volumetric albedo – a surface optimization framework for non-line-of-sight imaging. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
33. Velten, A., Willwacher, T., Gupta, O., Veeraraghavan, A., Bawendi, M.G., Raskar, R.: Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. Nature Communications **3**, 745 (2012)
34. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing **13**(4), 600–612 (2004)
35. Xin, S., Nousias, S., Kutulakos, K.N., Sankaranarayanan, A.C., Narasimhan, S.G., Gkioulekas, I.: A theory of Fermat paths for non-line-of-sight shape reconstruction. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2019)
36. Zou, C.C., Ge, S.: A Hough transform-based method for fast detection of fixed period sinusoidal curves in images. In: International Conference on Signal Processing (ICSP). pp. 909 – 912 vol.1 (2002)