Seeing Around Corners Using Time of Flight

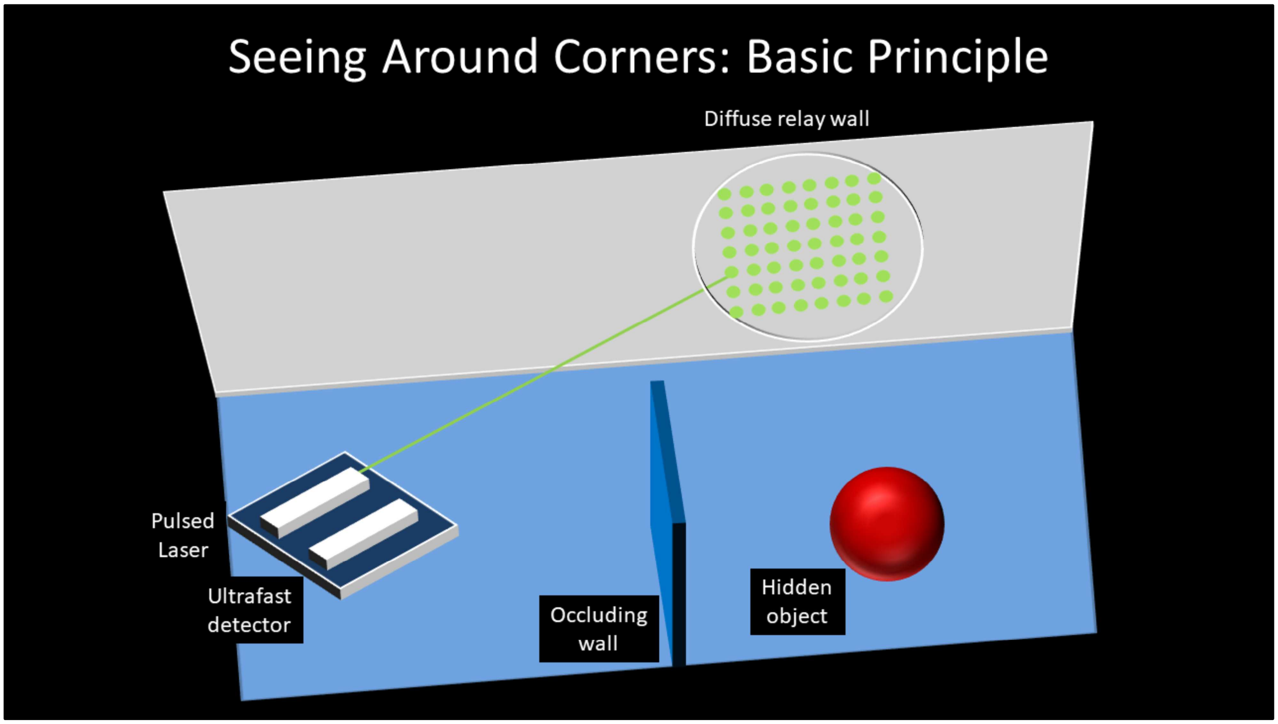
SIGGRAPH Course: Seeing Around Corners Using Time of Flight

Ramesh Raskar, Andreas Velten, Sebastian Bauer, Tristan Swedish

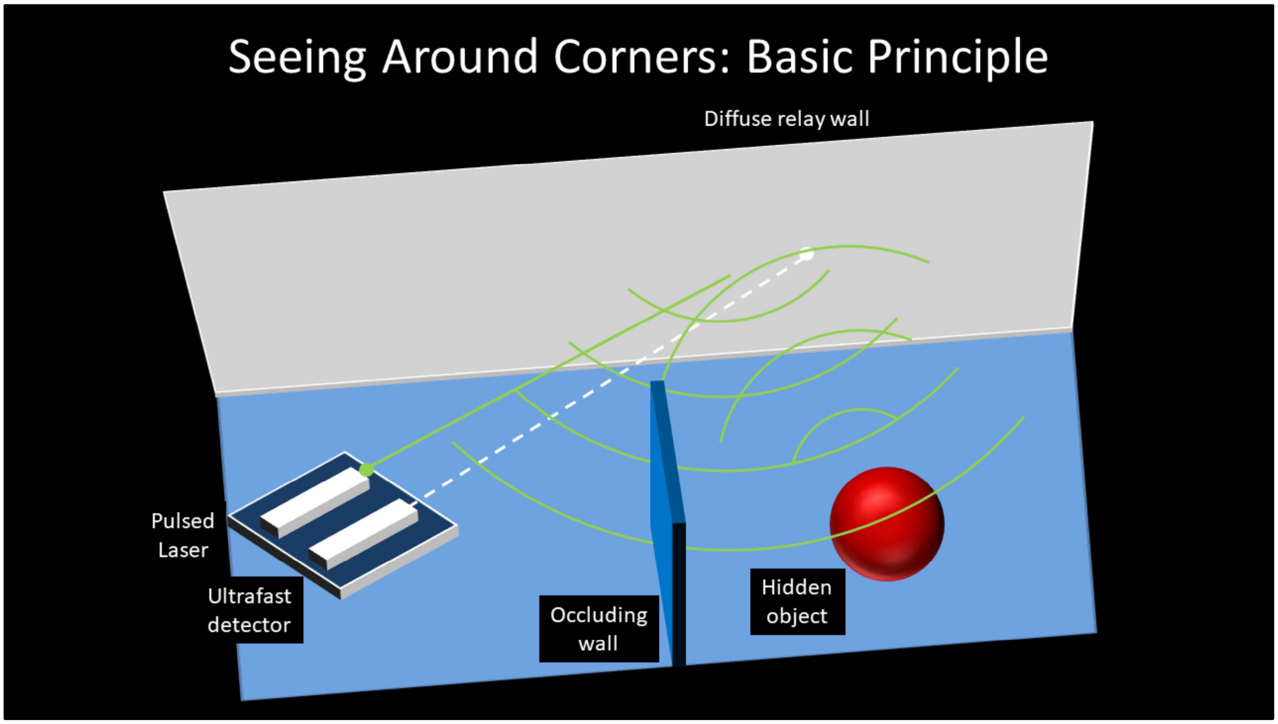
- Description of the basic principle how to see around corners
- Hands-on explanations on
  - How to build your own setup
  - Fast and easy system calibration
- 3D reconstruction algorithms and comparison

SIGGRAPH '20 Courses, August 17, 2020, Virtual Event, USA
ACM 978-1-4503-7972-4/20/08.
10.1145/3388769.3407484

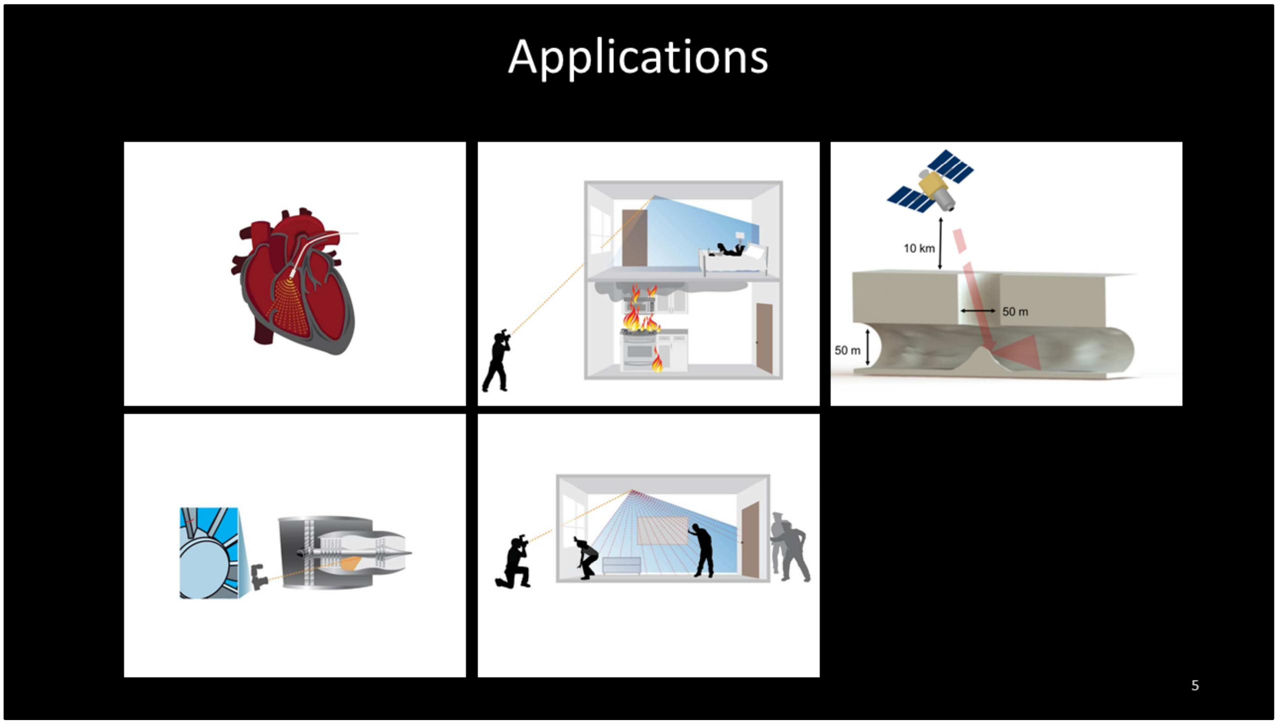
OVERVIEW, BASIC PRINCIPLE AND CHALLENGES



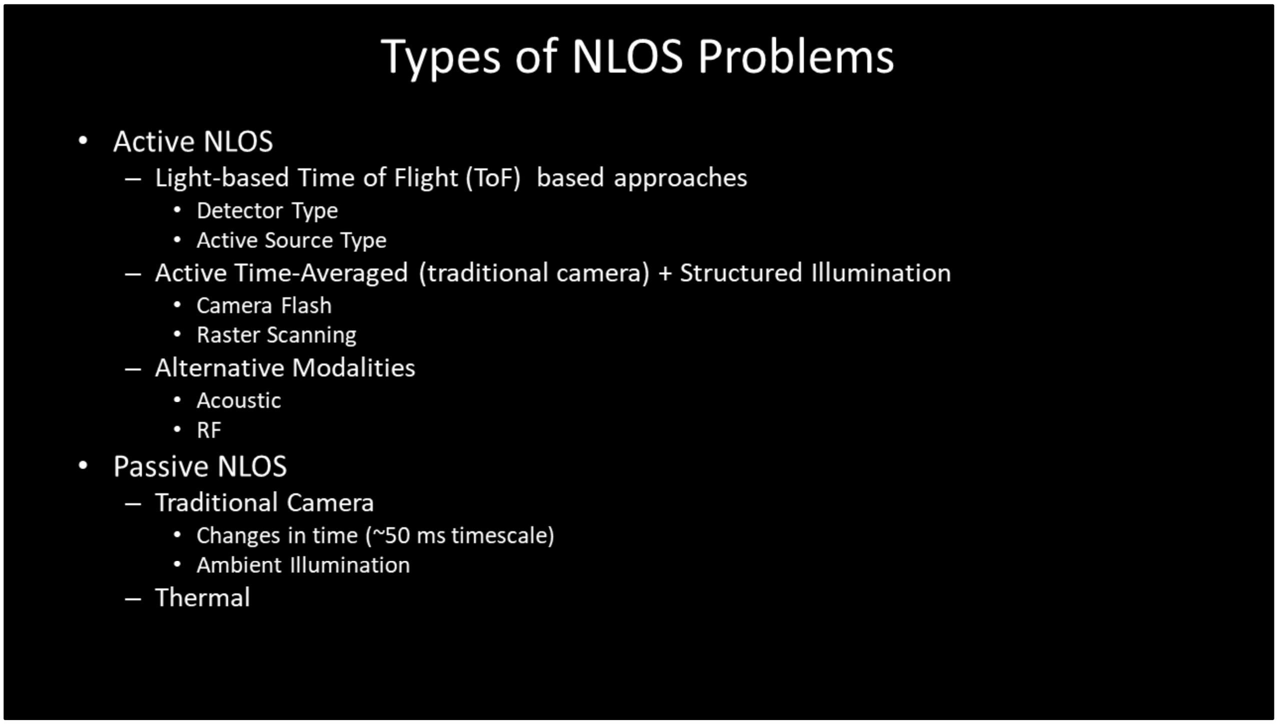
The problem of seeing around corners, often referred in the broader "Non-Line-of-Sight" context, is to use sensed information from directly visible surfaces of an environment to infer properties of the scene not directly visible. For example, the geometry above presents a classic "around the corner" setting, where a flat wall is used as the visible surface, and the hidden scene is occluded by another wall. While many proposed sensing modalities have been proposed, including acoustic and RF signals, most approaches utilize photonic sensors in the visible spectrum due to the availability of hardware, and better temporal and spatial resolution. Approaches range from active time-resolved measurements, time-averaged continuous wave sources, and even to passive exploitation of ambient illumination.



The most common light based Time of Flight (ToF) configurations send a pulsed laser source into the scene, which is then detected by a fast detector (such as a single photon avalanche diode, or SPAD) connected to a time-correlated single photon counter (TCSPC). After many repeated measurements, the time-correlated counts from the TCSPC can be used to build a histogram of backscattered photons in time. This signal is known as a “transient” measurement, as it measures the total time (and therefore distance) each backscattered photon took. Often, such configurations use a collimated active source and detector, such that a single point on the relay wall is illuminated, while another point on the relay wall is used for detection. By raster scanning the source and detector points along a 2D area on the relay wall, a 4D data structure can be constructed, with each spatial configuration producing a transient signal, such that the final measurement can be considered as populating a 5D structure (2D scanning position, 2D detector positions, 1D temporal histogram).

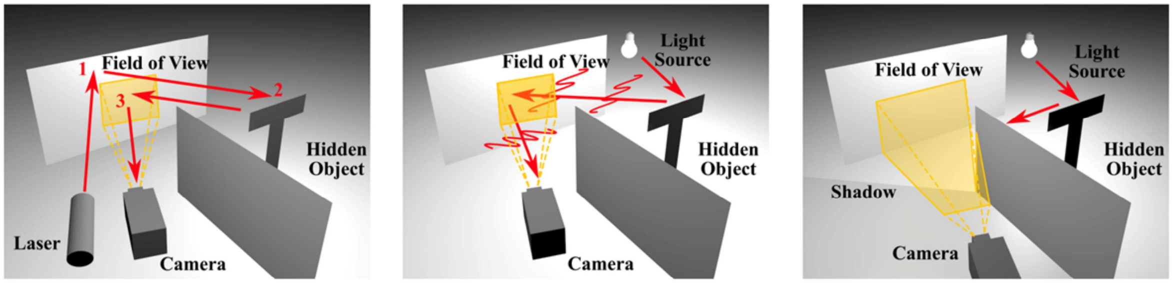


A system that can perform NLOS imaging has many applications, from medicine, industrial inspection, security, search and rescue, or even remote sensing and exploration missions.



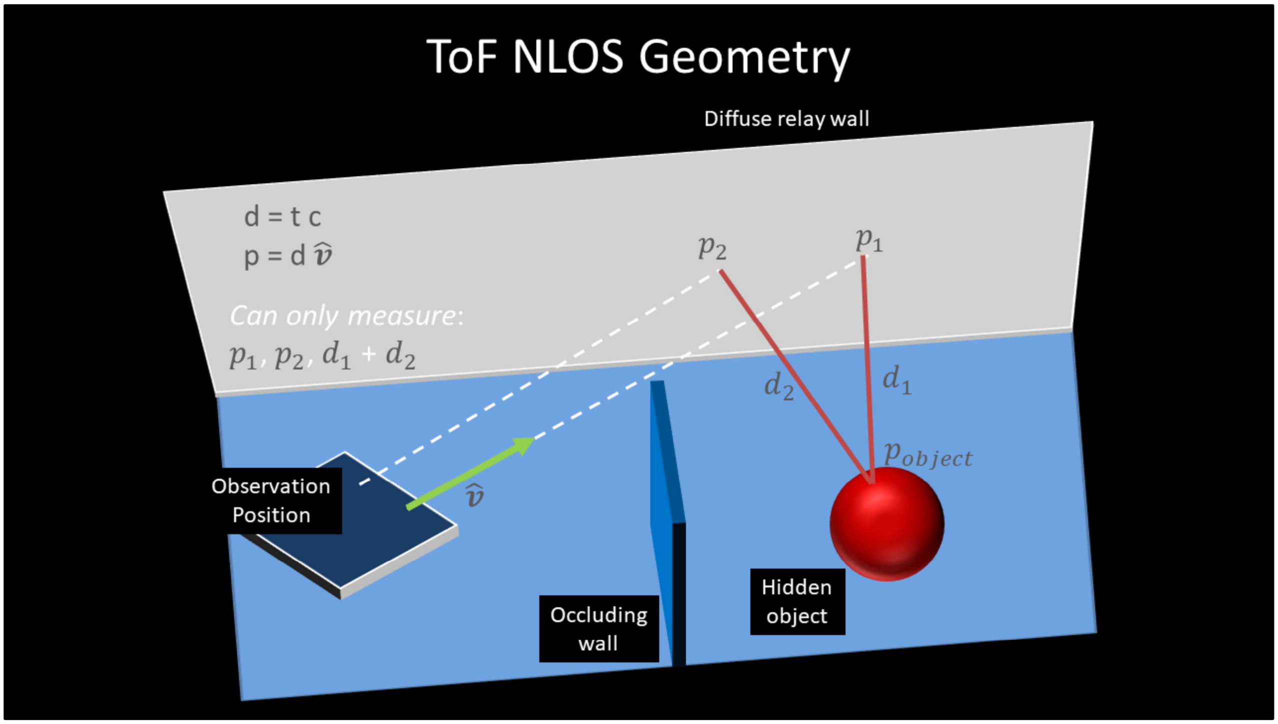
Many NLOS methods sample the full 5D measurement space by fixing either the scanning or illumination points, or scanning them together, reducing the total number of measurements. Typically, this data is then used to reconstruct the albedo within the volume of the hidden scene. Other information can also be extracted from these signals, such as classifying or localizing hidden objects rather than performing full reconstruction. Time-averaged measurements do not utilize a SPAD+TCSPC and instead integrate irradiance in time, using traditional CCD or CMOS camera sensors with an exposure time ~50ms. These methods are less suitable for full 3D reconstruction, but partial reconstruction, localization, and classification have been shown to be possible using time-averaged measurements depending on the properties of the visible scene.

Types of NLOS Problems

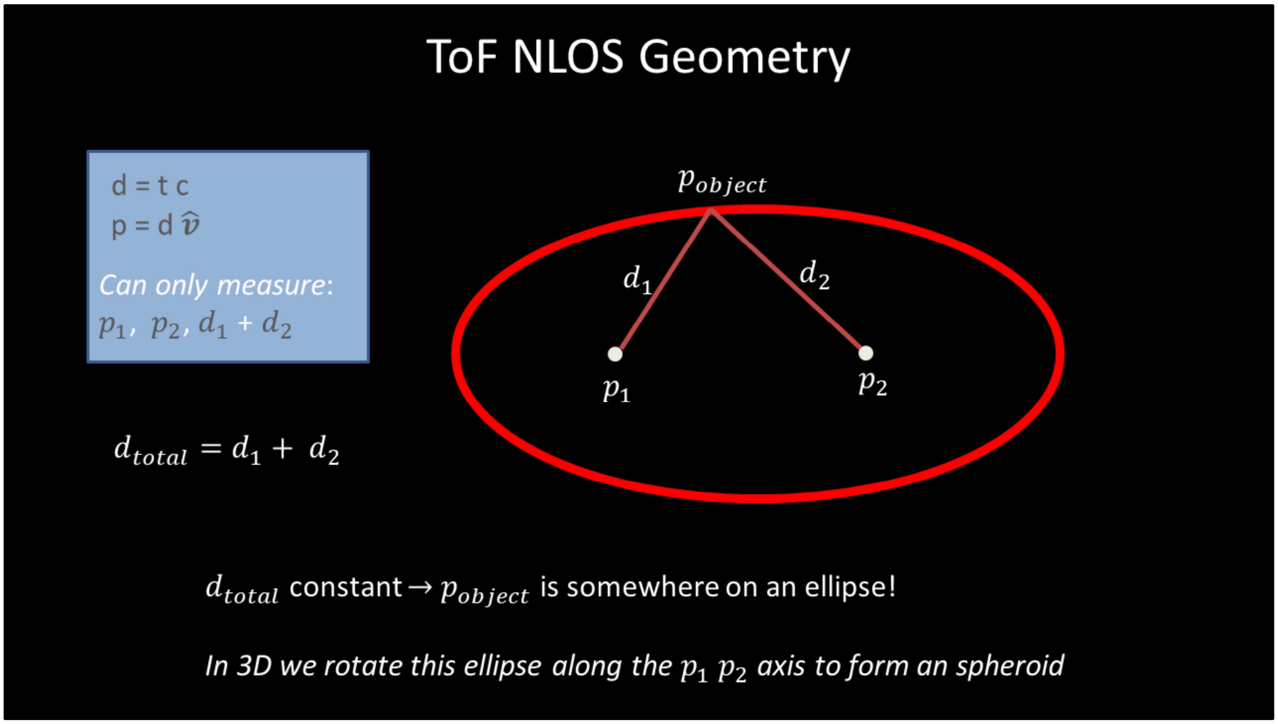


(a) Time-of-flight based methods exploit path-length constraints of photons (b) Coherence-based methods exploit preserved coherence properties (c) Intensity-based methods exploit shadows casted by depth discontinuities

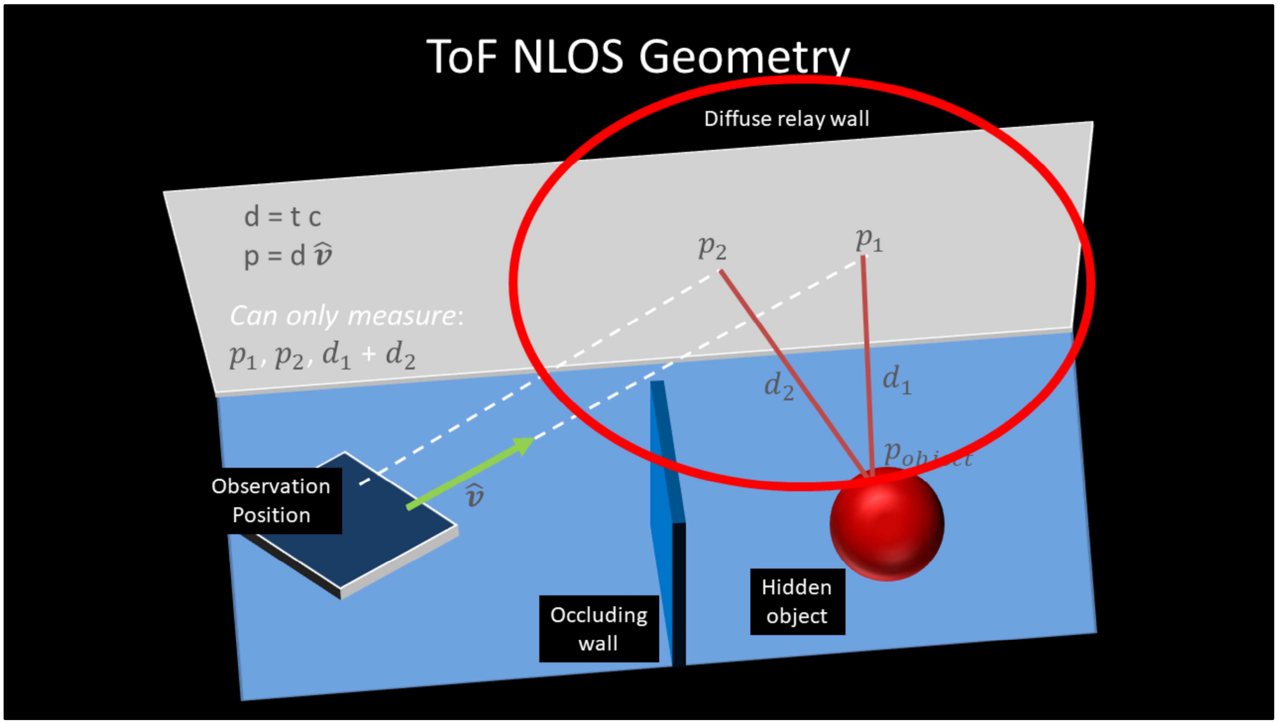
Maeda et al. "Recent Advances in Imaging Around Corners," ArXiv 2019.



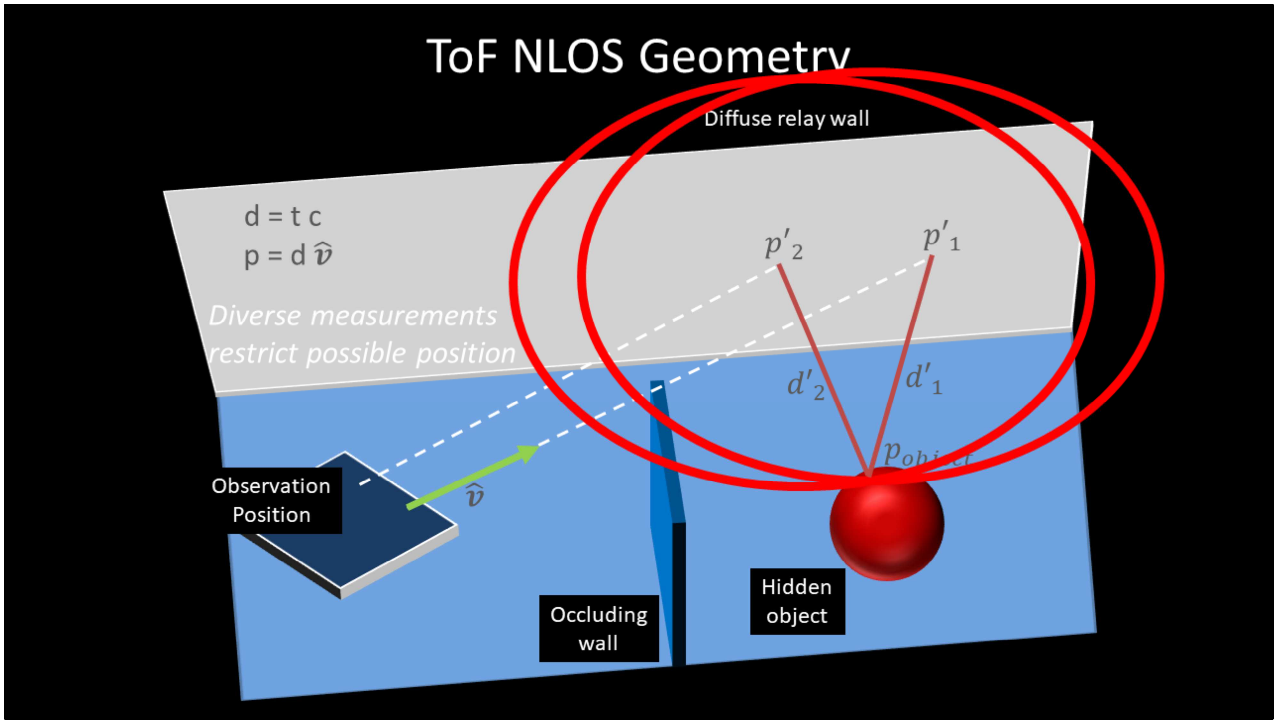
The basic NLOS geometry for a ToF based system consists of a single observation position that projects detection and illumination points on the visible "relay wall." Using the timestamps acquired from the first backscattered photons and initial calibration to determine the scanning direction (v), it's possible to recover the position in 3D of the scanning and illumination points (p1, p2). The total pathlength of multi-bounce photons can be determined via the speed of light (c). Importantly, the distance d1 and d2 in the figure above can not be measured directly, only their sum d1 + d2 is available for measurement.



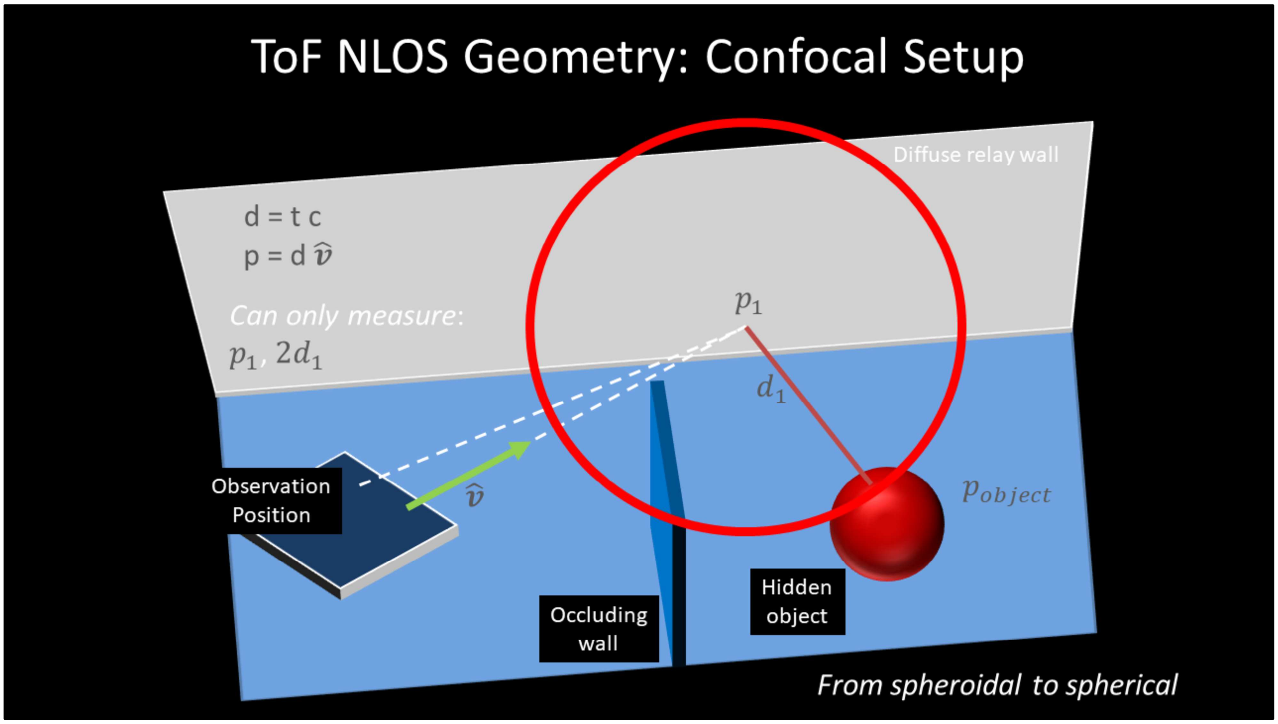
Given d1 + d2 can not be measured directly, the position of any hidden reflectors can at best be restricted to an ellipse in 2D, or a spheroid in 3D.



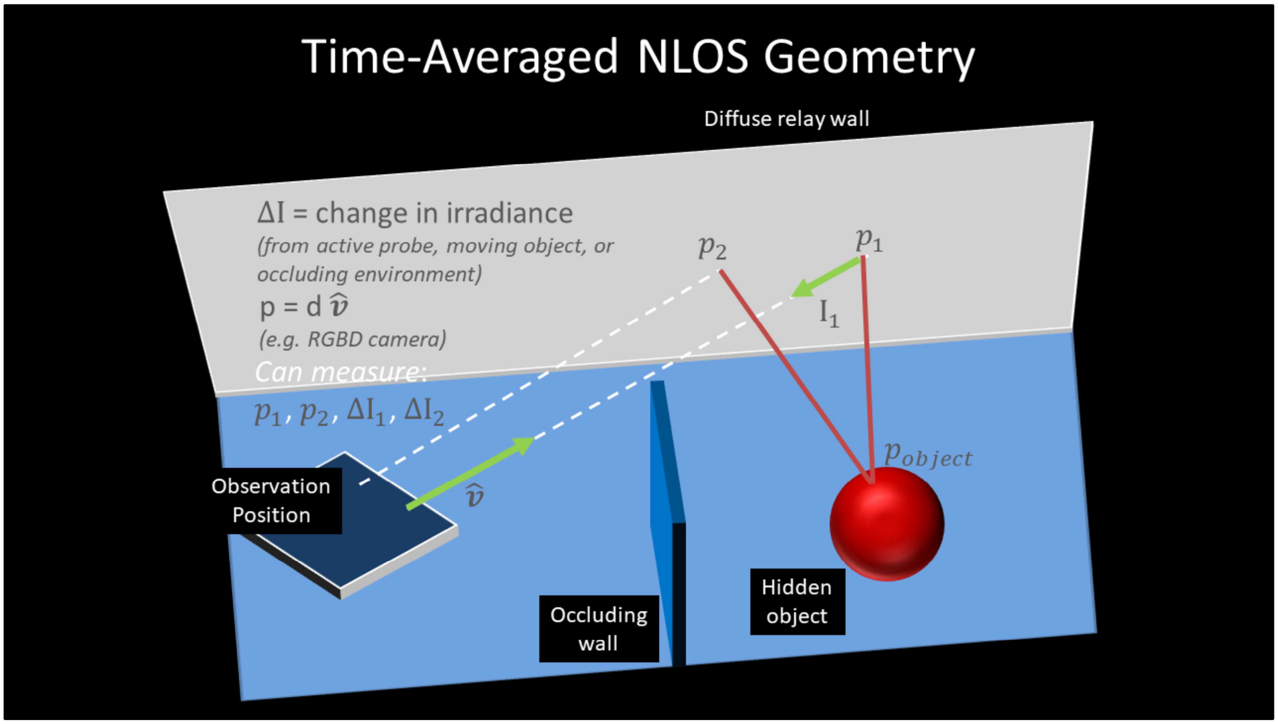
We can draw such an ellipse on the figure and see that our measurements have an ambiguous region that equally explains the position of the hidden reflector.



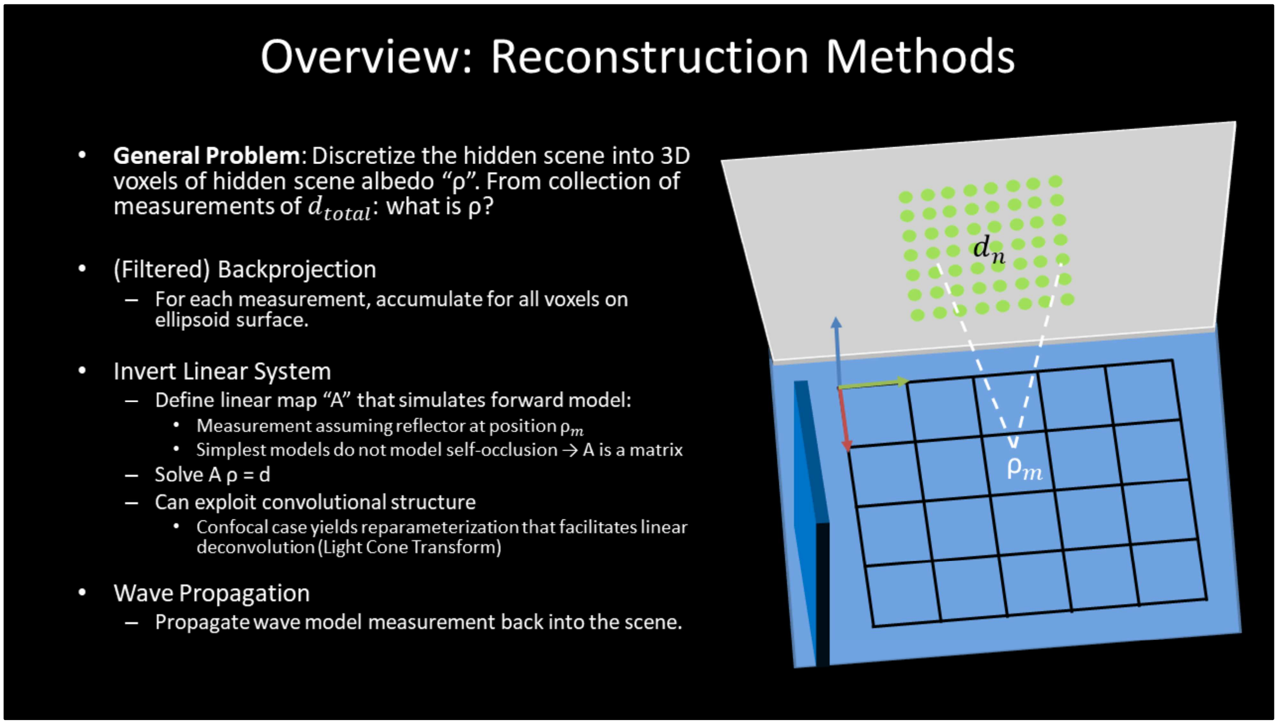
By changing the illumination and detector positions, diverse measurements can restrict the region of ambiguity. This reduces to an intersection of spheroids to localize a single reflector.



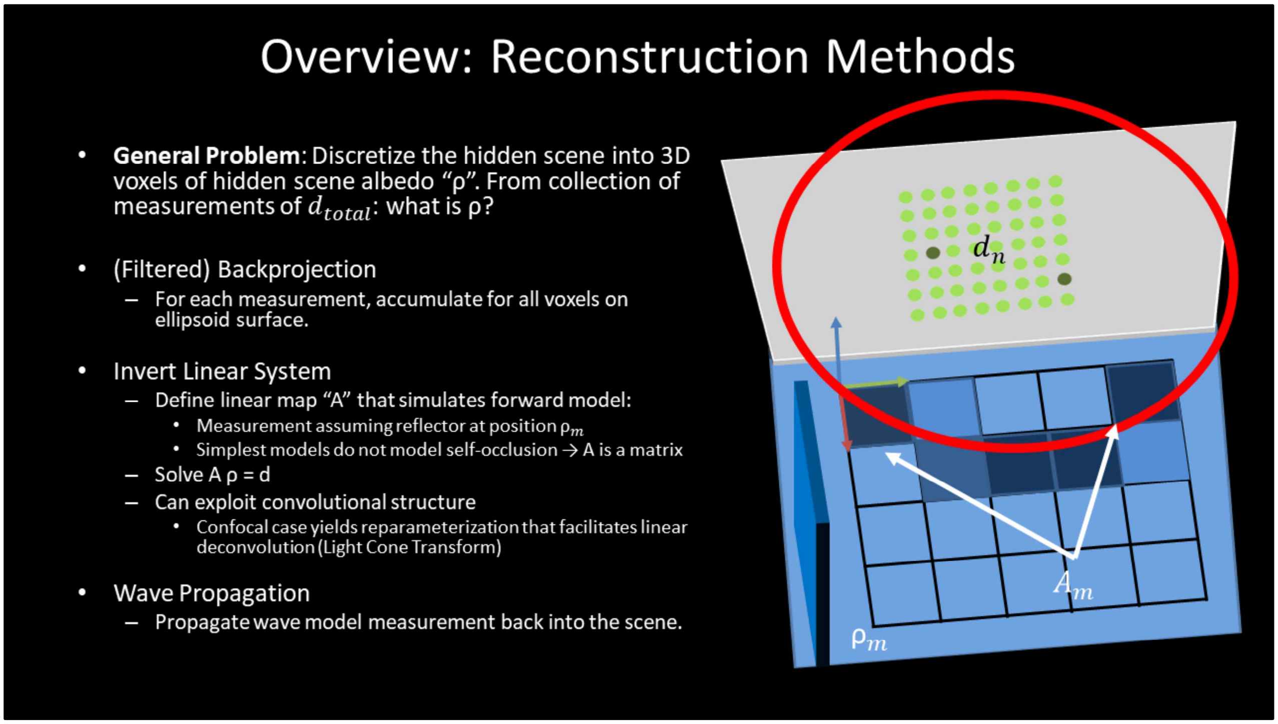
If the scanning and detector points on the relay wall coincide, we obtain a spherical ambiguous region. The symmetry afforded by the resulting sphere facilitates simplified reconstruction algorithms, and can sometimes be advantageous if the hidden object has certain reflection properties (such as retroflection by small particles or road signs).



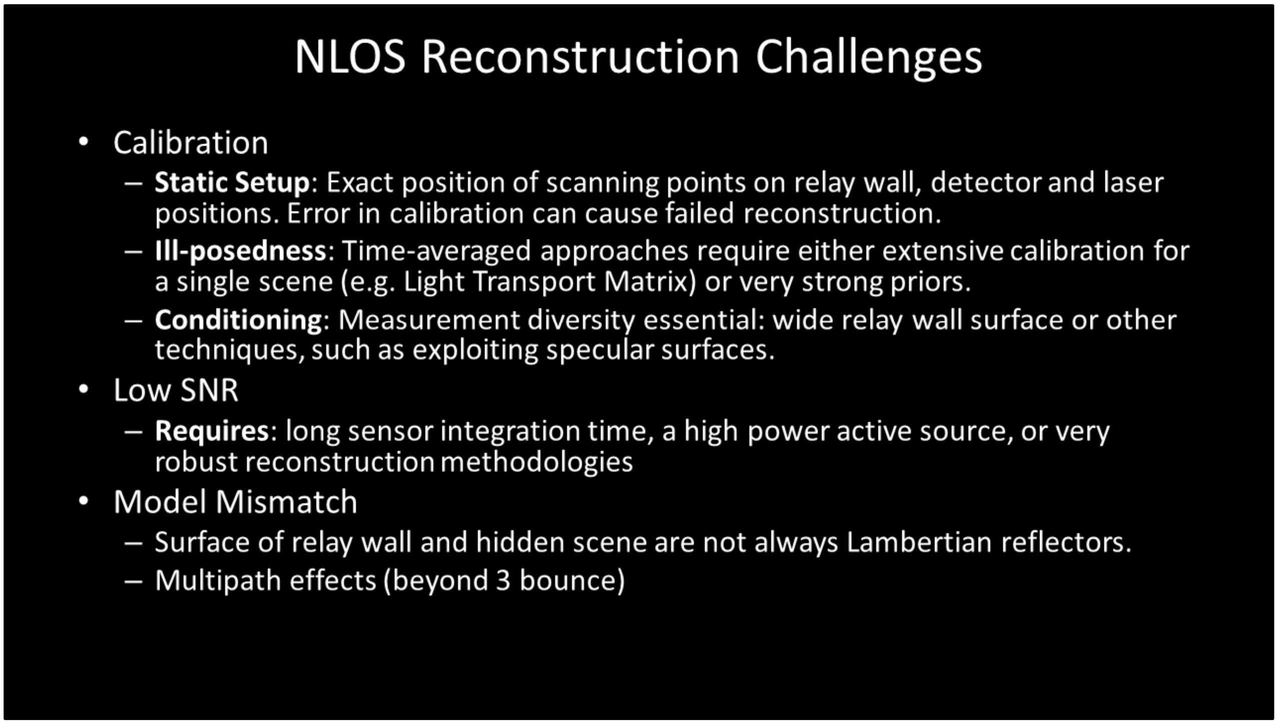
The time-averaged configuration is in principle very similar geometrically. Unfortunately, without time of flight information, only the irradiance at visible scene patches can be observed. Numerous techniques have been proposed that exploit scene structure or moving objects and the resulting change in irradiance.



Typically, the reconstruction problem is to recover the 3D albedo within a discrete grid of hidden scene voxels. High albedos correspond to increased backscatter. The question is: what is the best set of albedos to explain the measurements? Starting with CT-like approach, the first NLOS reconstruction algorithms utilized filtered back projection, which yields a simple, memory efficient, and highly parallelizable implementation. Global methods such as full linear system inversion can in principle produce higher quality results, but also have high computational requirements in general. Simplifications can be found by imposing convolutional structure, through reparameterization or adding some small restrictions on the forward linear operator.



As a case example: the basic idea behind back projection is to fill the region in the hidden volume that corresponds to each measurement. As shown earlier, this corresponds to an ellipsoid that has been rotated around the axis connecting the foci. In the figure above, we represent a single slice of the full 3D hidden volume. For each grid cell, we add some value to the voxels that lie on the ellipse edge. This set of values also corresponds to the values of a row of the full linear forward model "A". With back projection, this linear model does not need to be form explicitly, however, since we can check if the voxel lies on the ellipsoid surface using the implicit equation for the ellipsoid.



There are a few common challenges inherent in NLOS imaging. First, good calibration is very important to achieve high-quality results. In some cases, additional measurements must be performed in order to calibrate the scene (such as determining the Ray-Transport Matrix). Otherwise, these parameters must be selected without sufficient measurements, leading to ill-posed “blind” inverse problems. SNR is a common challenge for NLOS approaches due to the nature of multiply scattered light. Lastly, forward models typically trade computational efficiency for physical accuracy, leading to challenges with model mismatch. In particular, multi-path effects such as interreflections in the hidden volume are not typically modeled.

DATA ACQUISITION AND SENSORS

Data Acquisition and Sensors

- NLOS image resolution depends on the time resolution
  - 50 ps (10^{-12} s) correspond to 1.5 cm in space
- Laser pulses need to be short and sensor has to have a very high time resolution

ToF NLOS imaging resolution is mostly constrained by the time resolution of the light source and the detector. We'll mostly focus on the detector; the reason will become obvious in a few slides

Data Acquisition and Sensors

- Sensors
  - Streak camera (Velten 2012): 2 ps, lateral spatial resolution 1 cm in reconstruction. $150,000, fragile. But: line sensor
  - iCCD camera (Laurenzis 2014): few hundred ps, $80,000, low photon count rate
  - Photonic Mixer Devices (PMDs) (Heide 2014). Compact, less than $500
  - Single-Photon Avalanche Diodes (SPADs): dozens of ps, $20,000, single-photon sensitive, 10 dark counts/s (Buttafava 15)

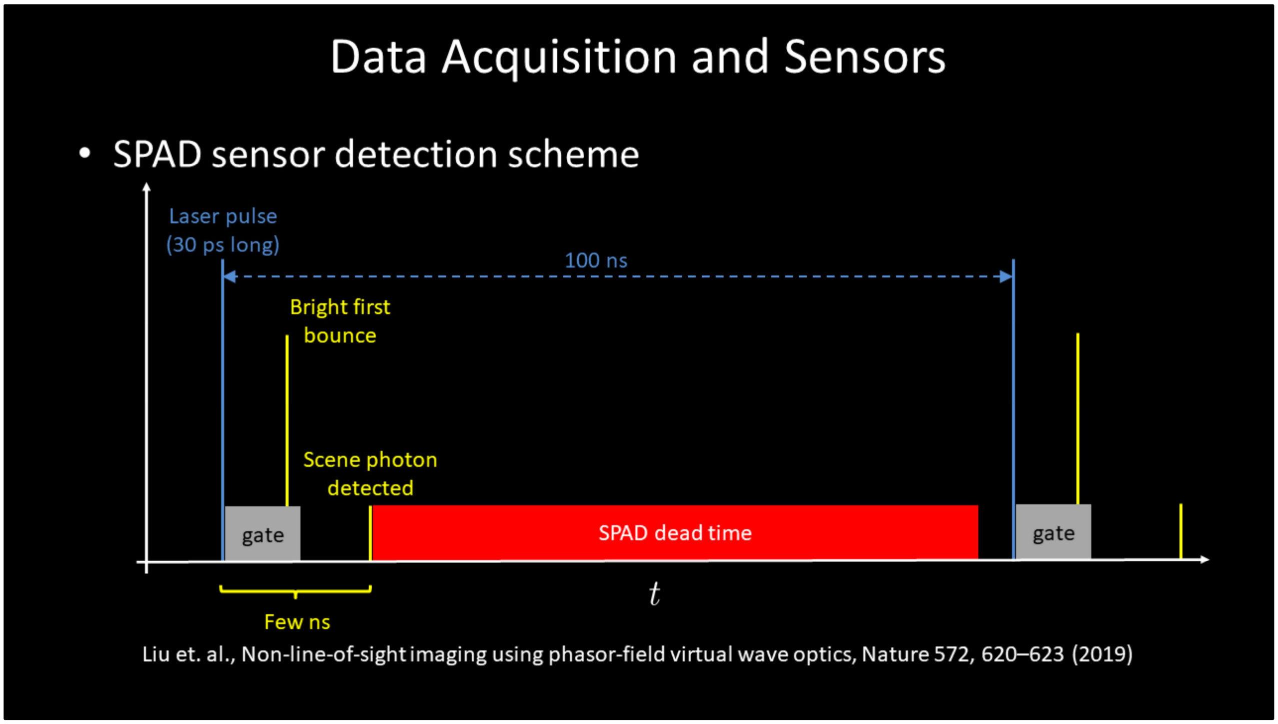
Many different types of detectors have been used for NLOS imaging, with very different properties, and also with very different price tags. While a small price is definitely desirable, the less beneficial properties of cheap sensors that go along with the lower price are a big hurdle for successful NLOS imaging.

Laurenzis, Martin, and Andreas Velten. "Nonline-of-sight laser gated viewing of scattered photons." *Optical Engineering* 53.2 (2014): 023102.
Heide, Felix, et al. "Diffuse mirrors: 3D reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.
Buttafava, M., Zeman, J., Tosi, A., Eliceiri, K., & Velten, A. (2015). Non-line-of-sight imaging using a time-gated single photon avalanche diode. *Optics Express*, *23*(16), 20997-21011.

Data Acquisition and Sensors

- SPAD sensors
  - Reverse biased semiconductor p–n junctions → photon causes electron avalanche that needs to be quenched
  - Extremely high sensitivity (counting single photons)
  - High time resolution (30 – 50 ps)
  - Silicon CMOS
  - After detection of a photon, sensor is blind for 1…100 ns
  - Afterpulsing: trapped carriers after avalanche cause “noise” avalanche
  - Gateable: “blind” for bright first bounce

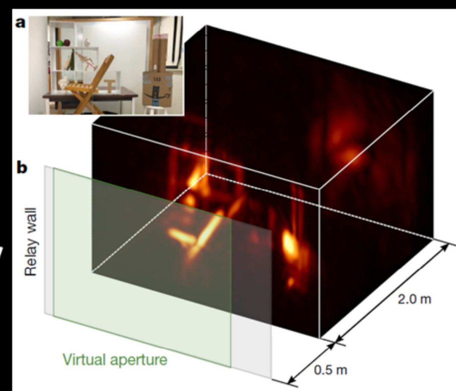
Ignoring the price tag, SPADs definitely have all that is needed for great quality NLOS imaging: they are sensitive to single photons and provide the required time resolution. SPADs are operated at a bias voltage above the breakdown voltage. There is a trade-off between afterpulsing and dead time: if the dead time is reduced, there is a higher likelihood of noise detections because the trapped carriers have not yet been released. Because of the dead time, in our setup 100 ns, a SPAD sensor can detect at most one photon per laser pulse



Time sequence: the laser emits a pulse and the SPAD sensor is "gated" for a specific period of time after the laser pulse, which means it does not capture any photons. This allows to ignore the bright first bounce off the visible relay wall which does not carry information about the hidden scene. Without gating, the SPAD would most likely process first bounce photons and the hidden scene photons arriving during the SPAD dead time would be lost because they arrive during the dead time.

Data Acquisition and Sensors

- About 50 million photons are needed for decent scene reconstruction
- Acquisition time: 240 s
- **Out of ~1 W emitted laser power, only 80 fW ($80\cdot10^{-15}$ W) are captured by the SPAD sensor!**



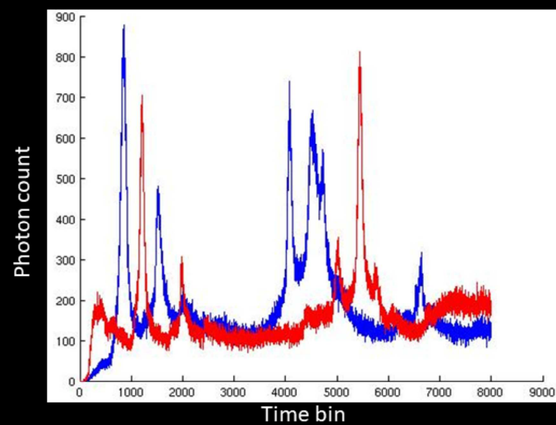
Liu et. al., Non-line-of-sight imaging using phasor-field virtual wave optics, Nature 572, 620–623 (2019)

50 million photons needed for decent scene reconstruction as shown on the right. SPAD only captures light from within a relay wall area of less than one square centimeter.

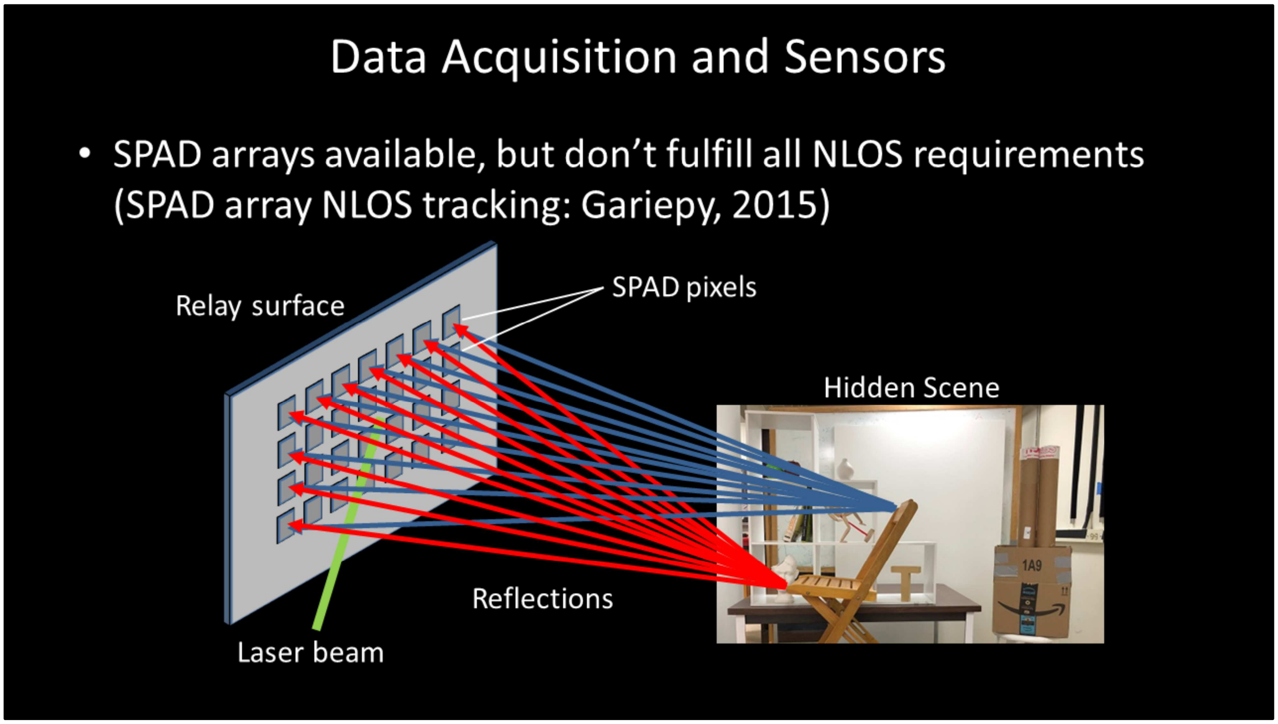
50 million photons acquired in 240 s, times Planck constant, times speed of light, divided by the wavelength: 50e6/240*6.626e-34*3e8/(530e-9)

Data Acquisition and Sensors

- SPAD output connected to TCSPC (Time-correlated single-photon counting) units that create histograms of the arriving photons



Time delay measured with respect to the laser pulse



By requirements in this context, we mainly mean time resolution, total number of photons that can be counted per second (that is both a chip design and a bandwidth issue), dark count, number of pixels, and, very important as well, gateability. High quantum efficiency and fill factor are also desirable.

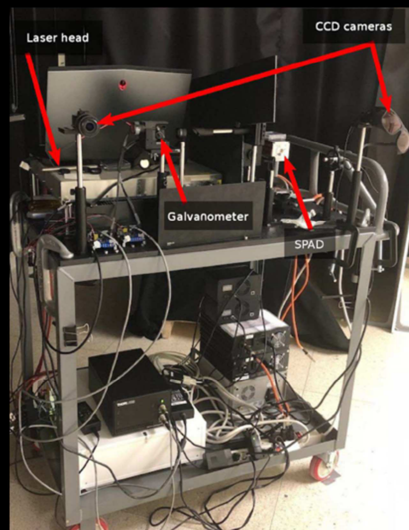
While most SPAD arrays have a low fill factor (~10 %, with microlenses ~50 %), looking at this graphic, it becomes still obvious why their use is beneficial: the scene objects and the relay wall reflect the light in many different directions, and even arrays with low fill factor capture a larger number of photons returning to the relay wall than a single pixel sensor.

With more SPAD pixels observing a larger portion of the wall, the laser power can be reduced proportionally.

Gariepy, G., Tonolini, F., Henderson, R., Leach, J., & Faccio, D. (2015). Tracking hidden objects with a single-photon camera. *arXiv preprint arXiv:1503.01699*.

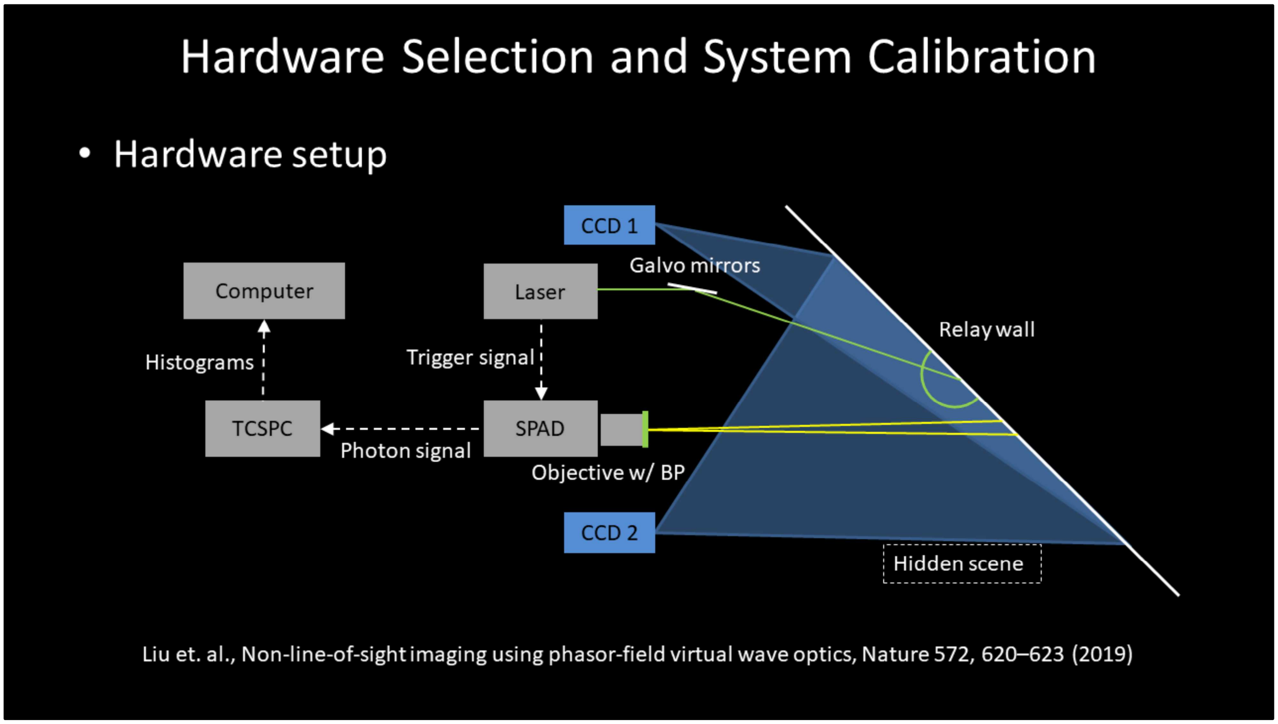
HARDWARE SELECTION AND SYSTEM CALIBRATION

Hardware Selection and System Calibration

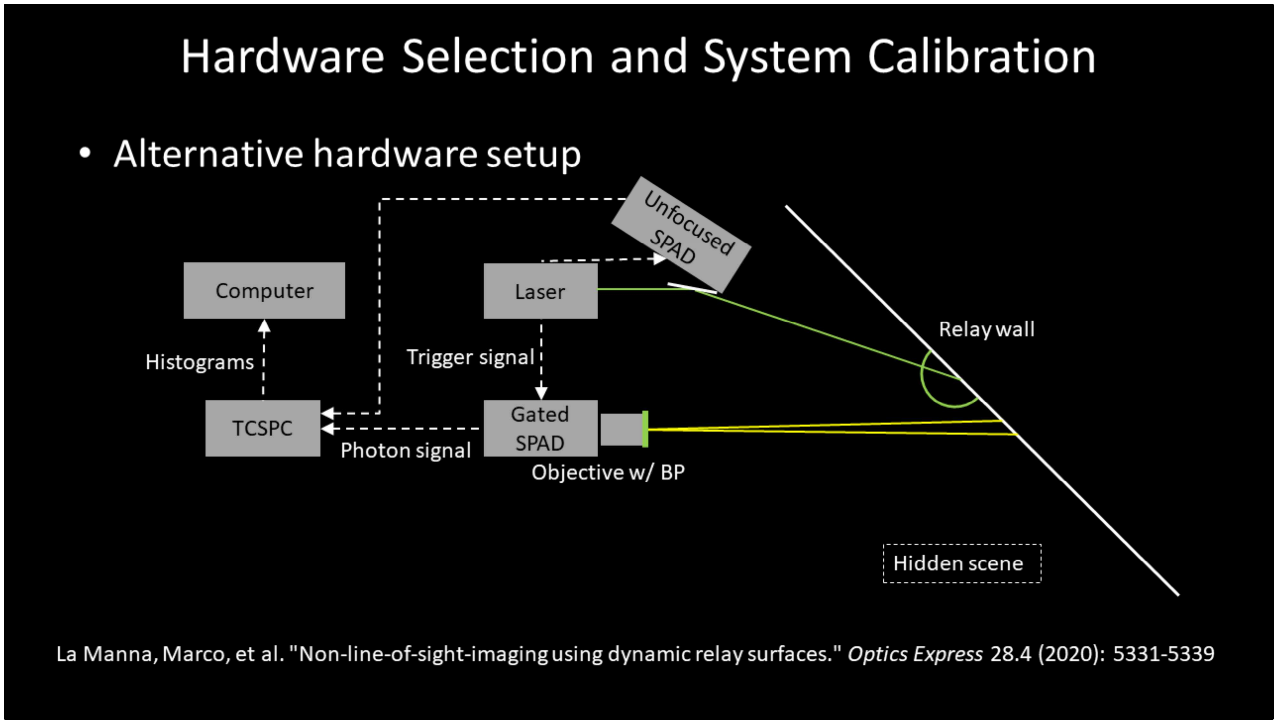


Liu et. al., Non-line-of-sight imaging using phasor-field virtual wave optics, Nature 572, 620–623 (2019)

NLOS setup with laser, galvo for beam scanning, SPAD for photon detection and CCD stereo camera pair for detecting the 3D position of the laser beam on the relay wall.

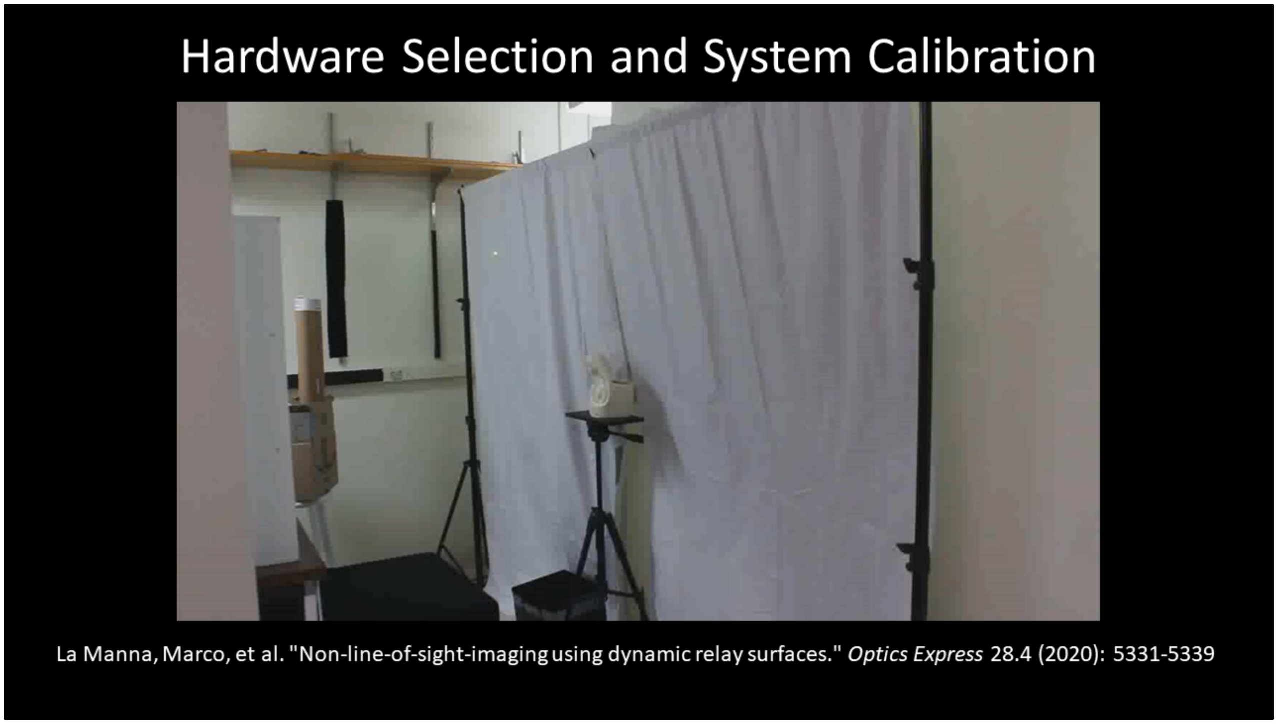


The SPAD objective focuses the SPAD on the relay wall and has a bandpass in front of it to suppress ambient light

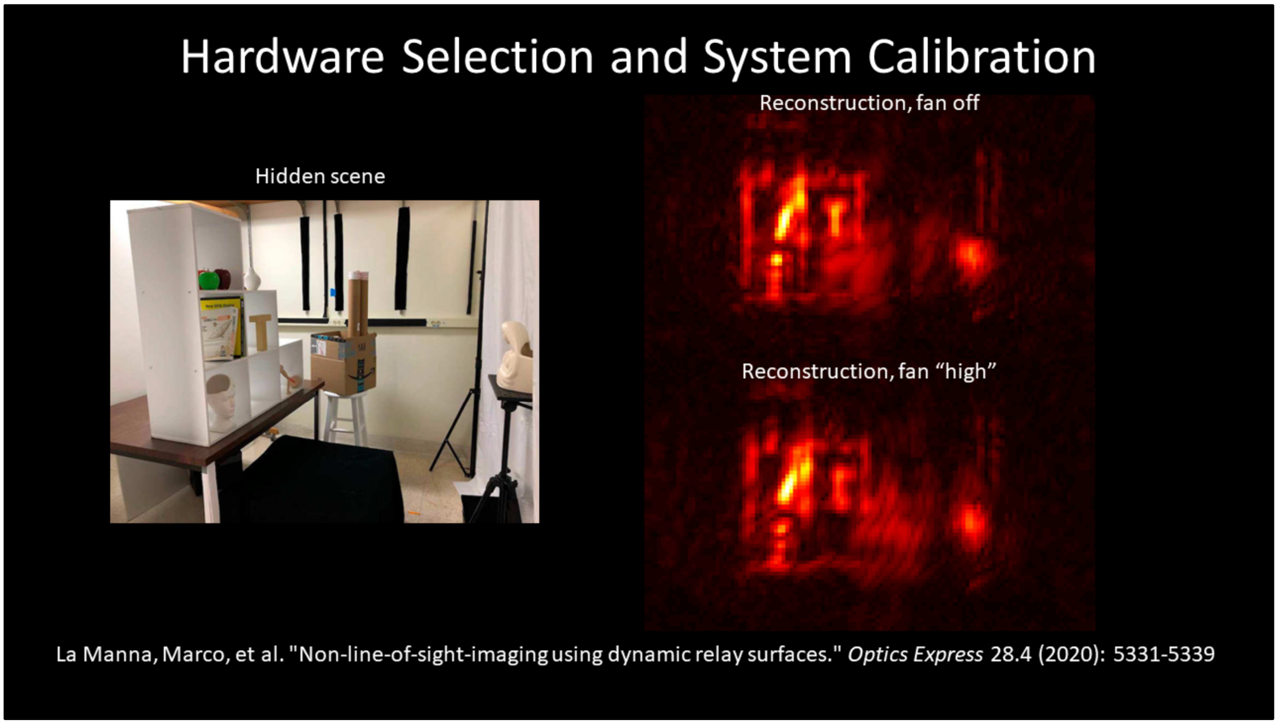


Alternatively, for determining the 3D beam position, the stereo cameras are removed and an ungated, unfocused SPAD is placed next to the galvo. Its job is to provide the arrival time of the laser pulse at the relay wall. Together with the laser pulse direction, we then have the depth along the laser ray and are able to uniquely determine the location where the laser pulse hits the relay wall, which in this case can be dynamic because of the fast acquisition of the required data.

Also, the TCSPC this time does not provide a full histogram, but is run in *time-tagged time-resolved* (TTTR) mode, where it only provides the time stamps (i.e., the delay with respect to the laser pulse) for each individual photon

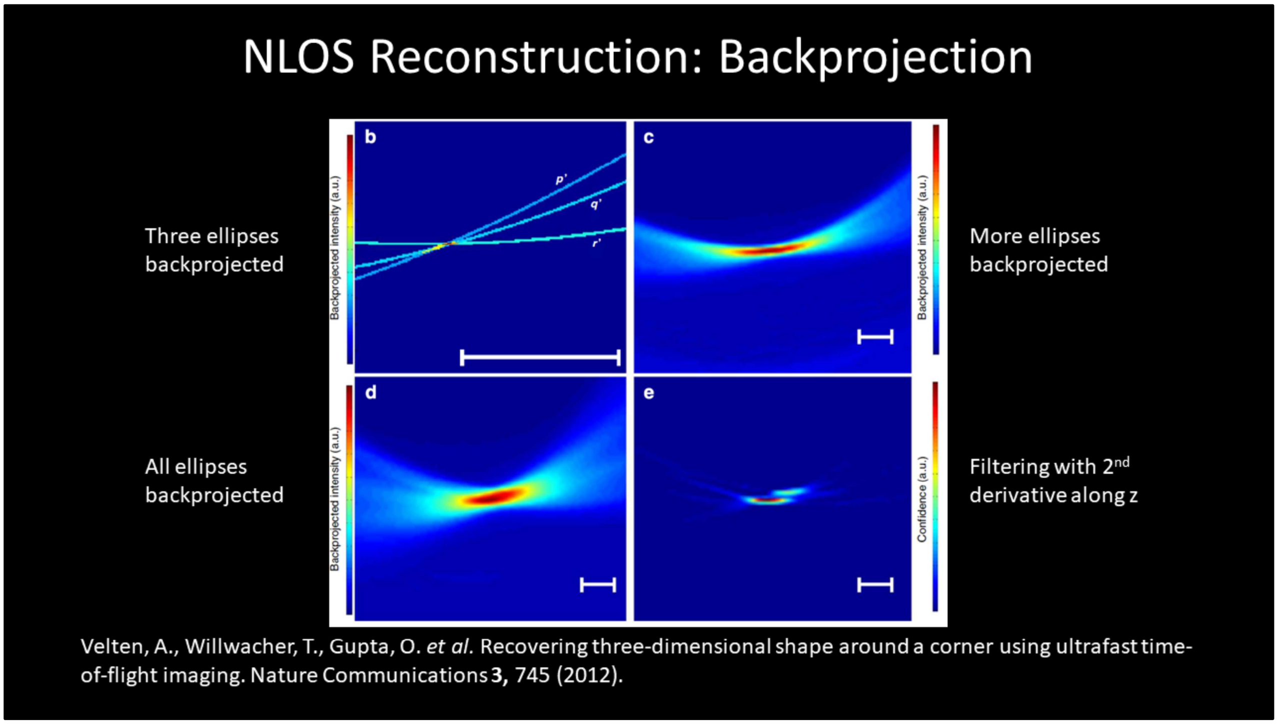


Here, the planar relay wall has been replaced by a curtain, and in addition, there is a fan behind the curtain that makes it move. Obviously, the goal still is to reconstruct the hidden scene. In practice, it does not matter if the sensor or the relay wall is moving; only their position with respect to each other needs to be known.

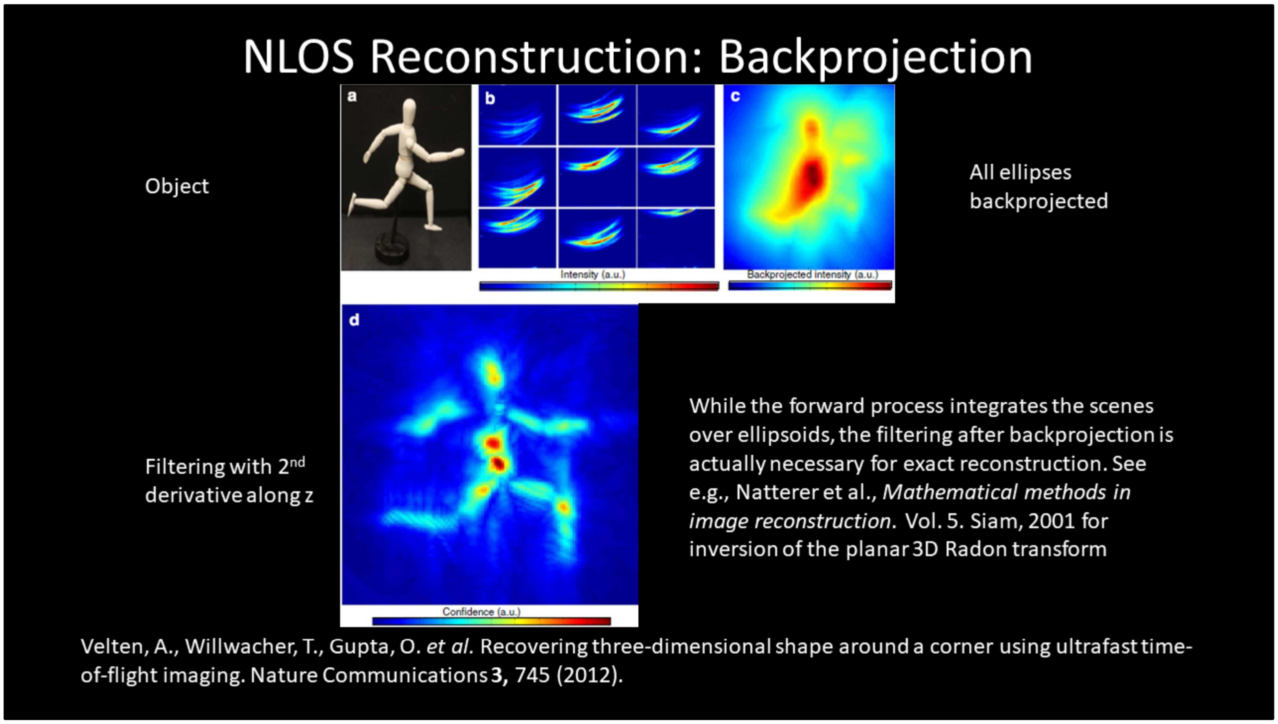


The moving curtain does not lead to significant reconstruction degradation.

RECONSTRUCTION METHODS – BACKPROJECTION

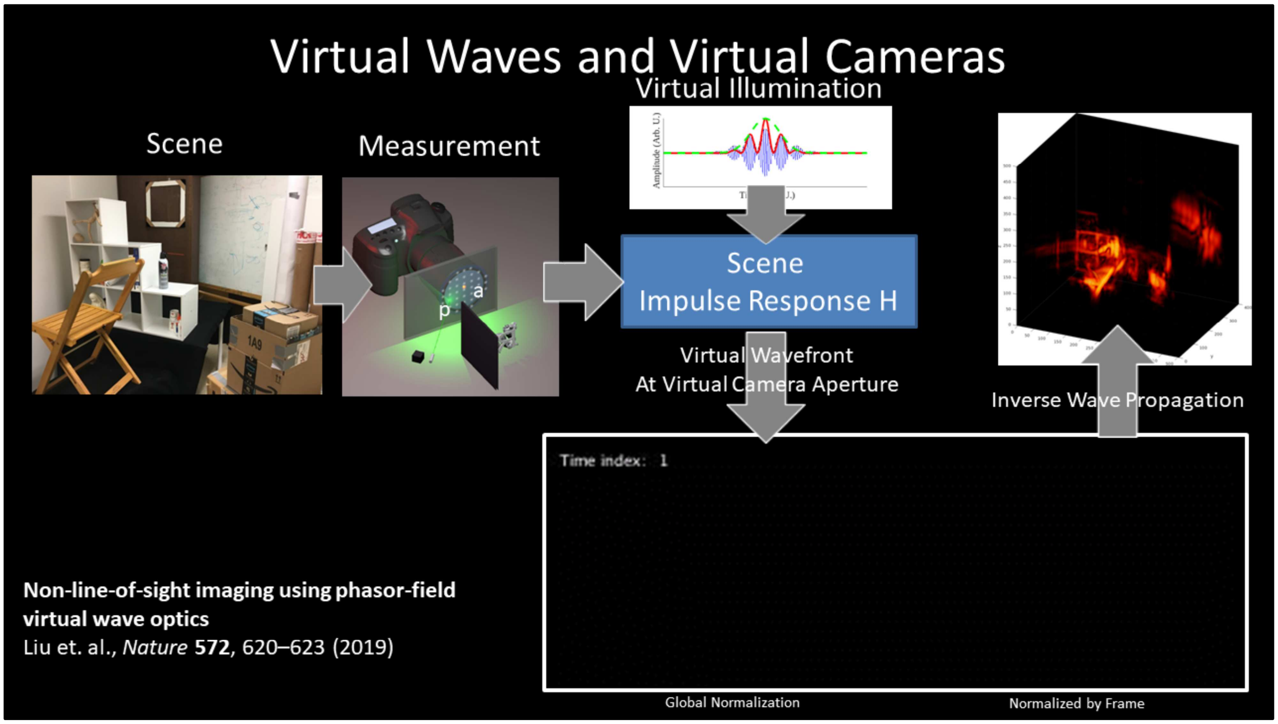


The first paper showing practical reconstruction of objects hidden around a corner used backprojection. This procedure is well know in computational tomography (CT): the acquired data are backprojected into the scene to reconstruct it. In the context of NLOS imaging, each respective spheroid is drawn in the reconstruction space with the magnitude of the corresponding time bin. Subsequently, the result is filtered by a Laplacian (second derivative).

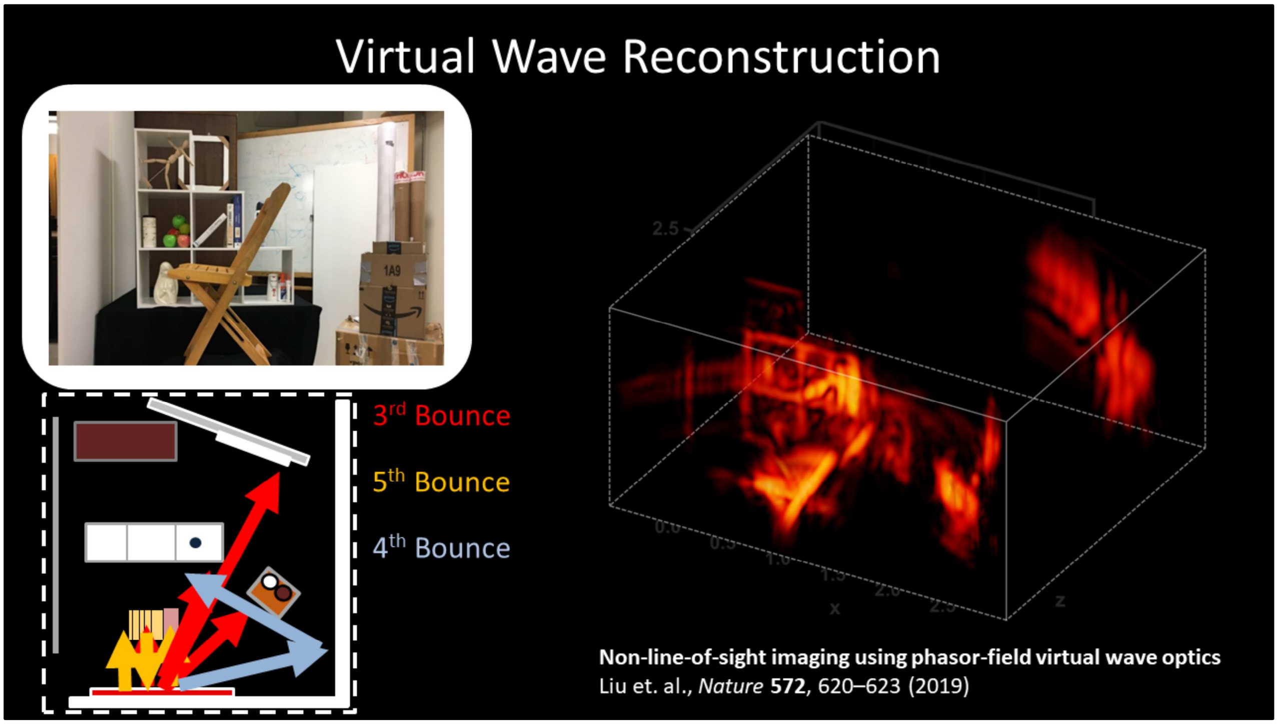


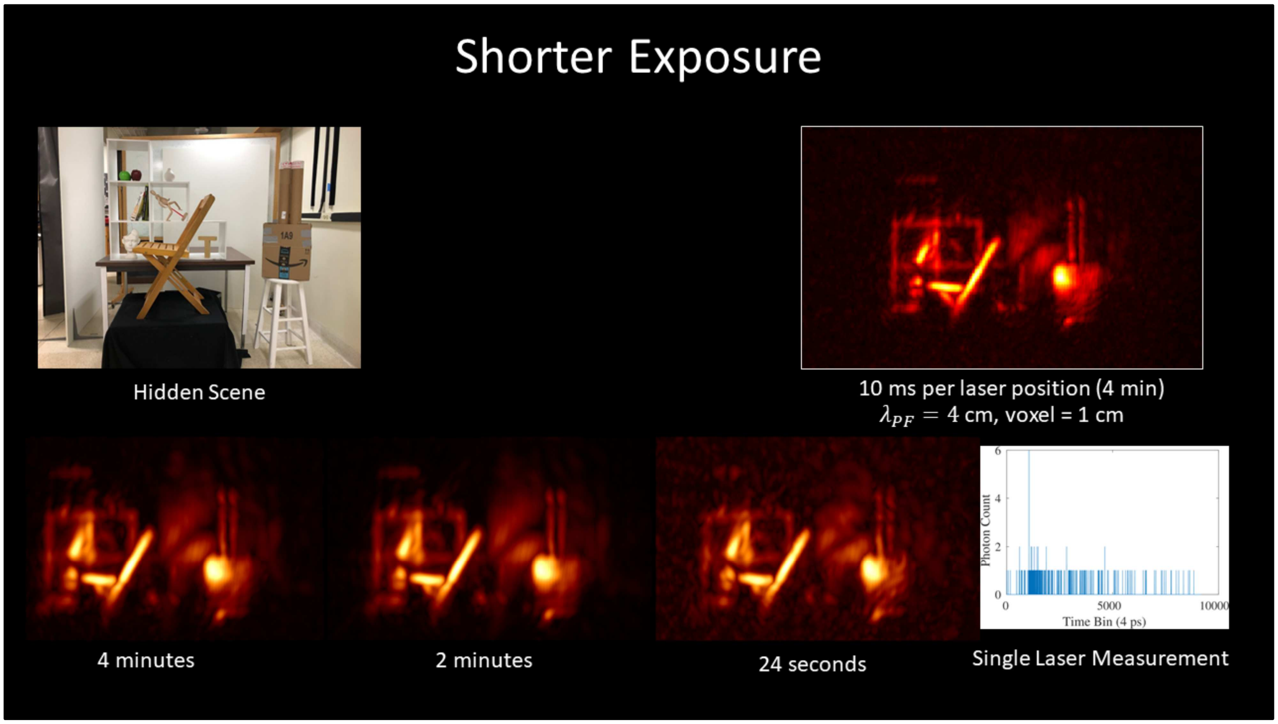
The Laplacian is actually required to (theoretically) exactly reconstruct the scene. This is known from the 3D Radon transform integrating the scene over all possible planes. As the second derivative decreases the SNR, the Laplacian of Gaussian (LoG) can be used for noise reduction.

RECONSTRUCTION METHODS – WAVE BASED

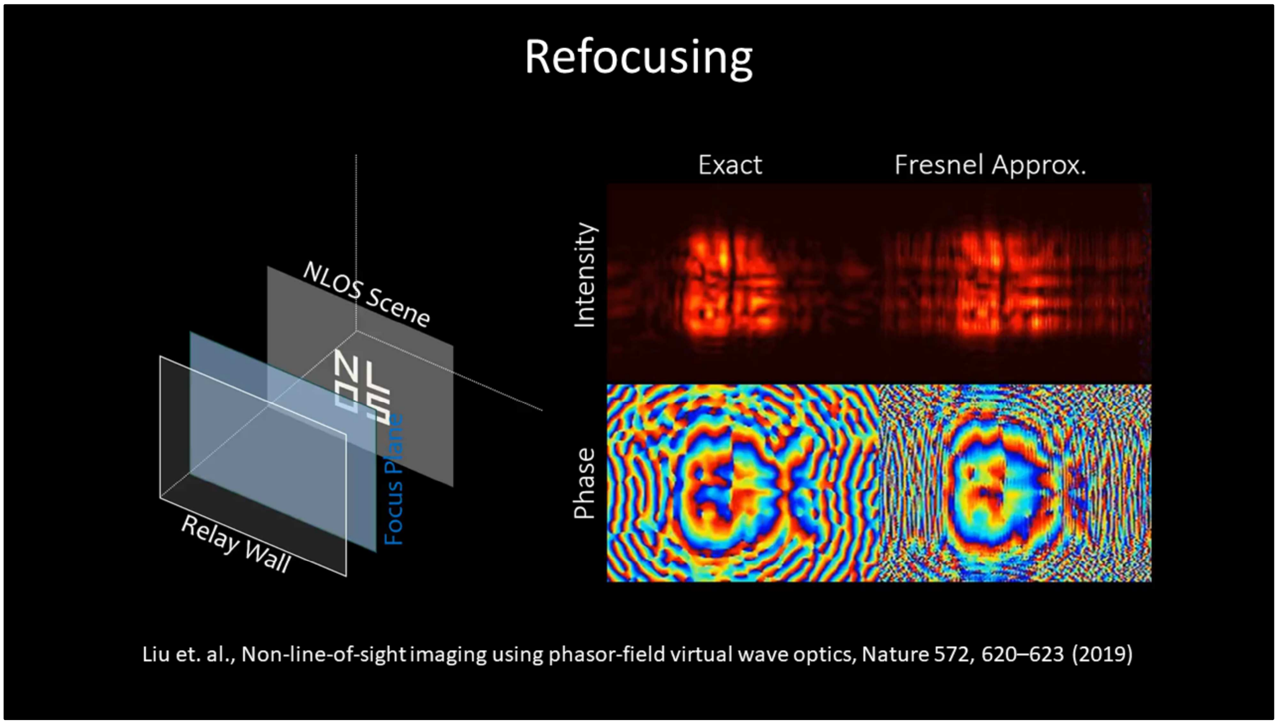


The acquired time responses can be interpreted as the impulse responses of the scene at different relay wall locations. Therefore, we can calculate the response of the scene to any input function. Using a sinusoidal modulation multiplied by a Gaussian therefore provides the scene response to this virtual wave pulse. This allows to formulate NLOS imaging as an inverse wave propagation problem; the acquired wavefronts are propagated back into the scene and at the object location, these wavefronts take the form of the physical objects. This allows to interpret the relay wall as the aperture of a virtual camera operating at the wavelength of the modulation sinusoid.

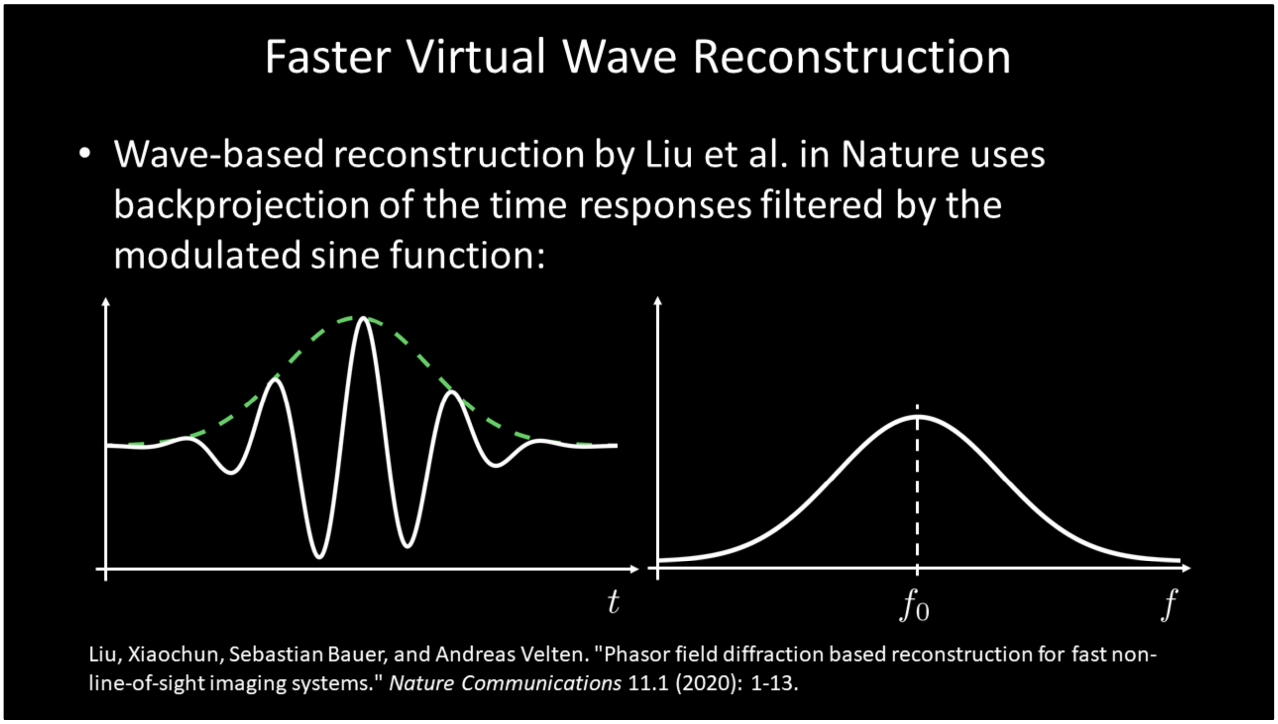

Virtual Wave Reconstruction
3rd Bounce
5th Bounce
4th Bounce
2.5
x
z
Non-line-of-sight imaging using phasor-field virtual wave optics
Liu et. al., Nature 572, 620–623 (2019)



Bottom row: 6 cm modulation wavelength



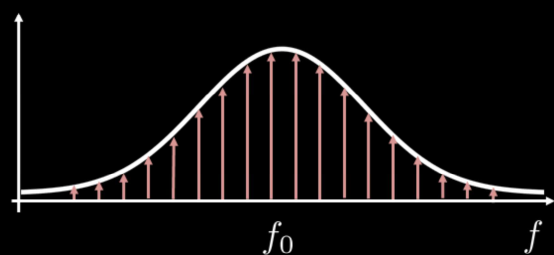
Refocusing
Exact
Fresnel Approx.
Intensity
Phase
NLOS Scene
Focus Plane
Relay Wall
Liu et. al., Non-line-of-sight imaging using phasor-field virtual wave optics, Nature 572, 620–623 (2019)



For the Nature paper, the results have been calculated by backprojecting not the original impulse response, but the filtered one. Another approach is to consider the frequency domain representation of the modulation pulse which is just a Gaussian around a center frequency f_0

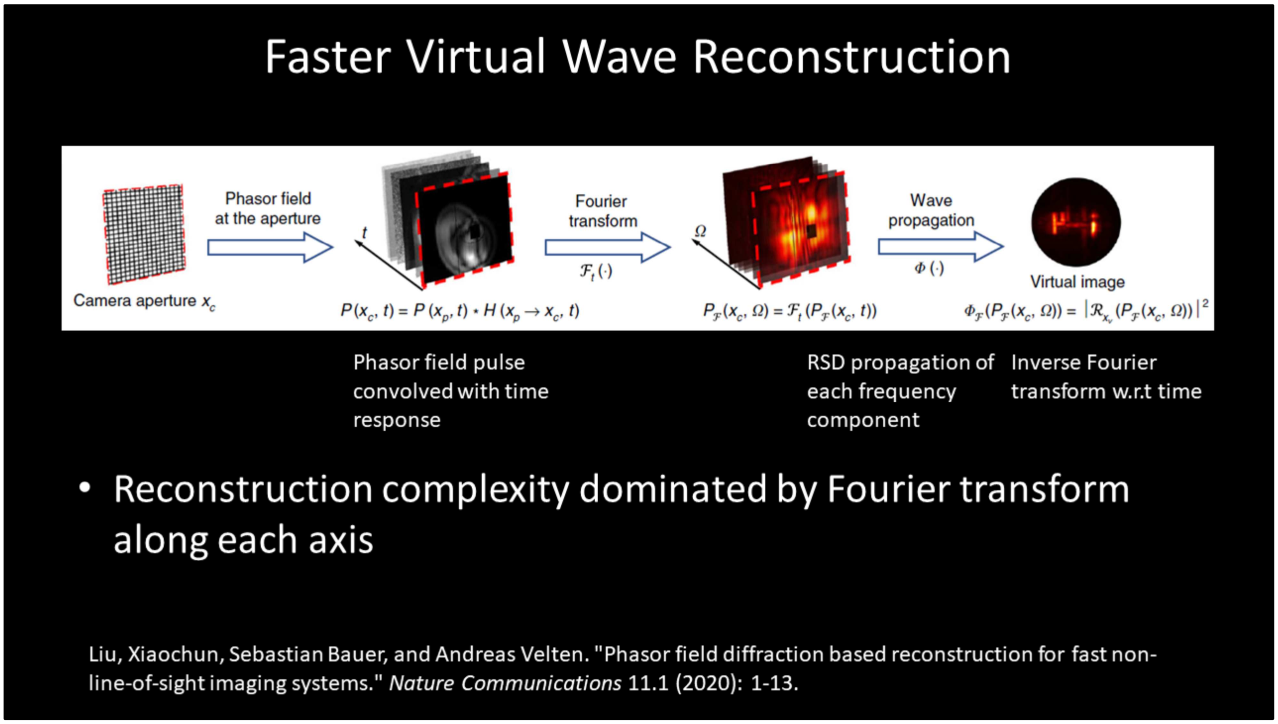
Faster Virtual Wave Reconstruction

- Rayleigh-Sommerfeld Diffraction (RSD) operator calculates scalar wave propagation for a single frequency
- Propagation from relay wall plane to a parallel plane: 2D spatial convolution



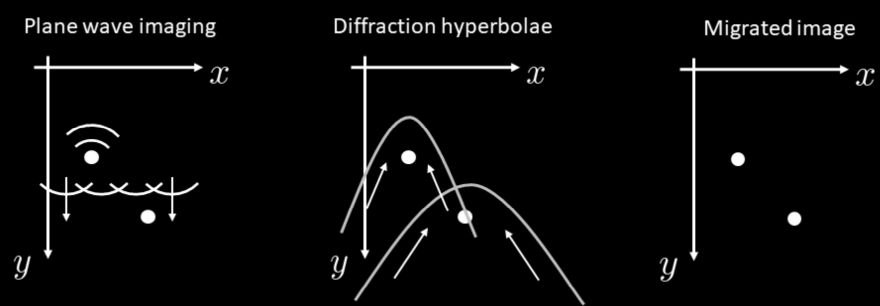
Liu, Xiaochun, Sebastian Bauer, and Andreas Velten. "Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems." *Nature Communications* 11.1 (2020): 1-13.

It is therefore possible to just select a number of frequencies and propagate these frequencies separately with the RSD.



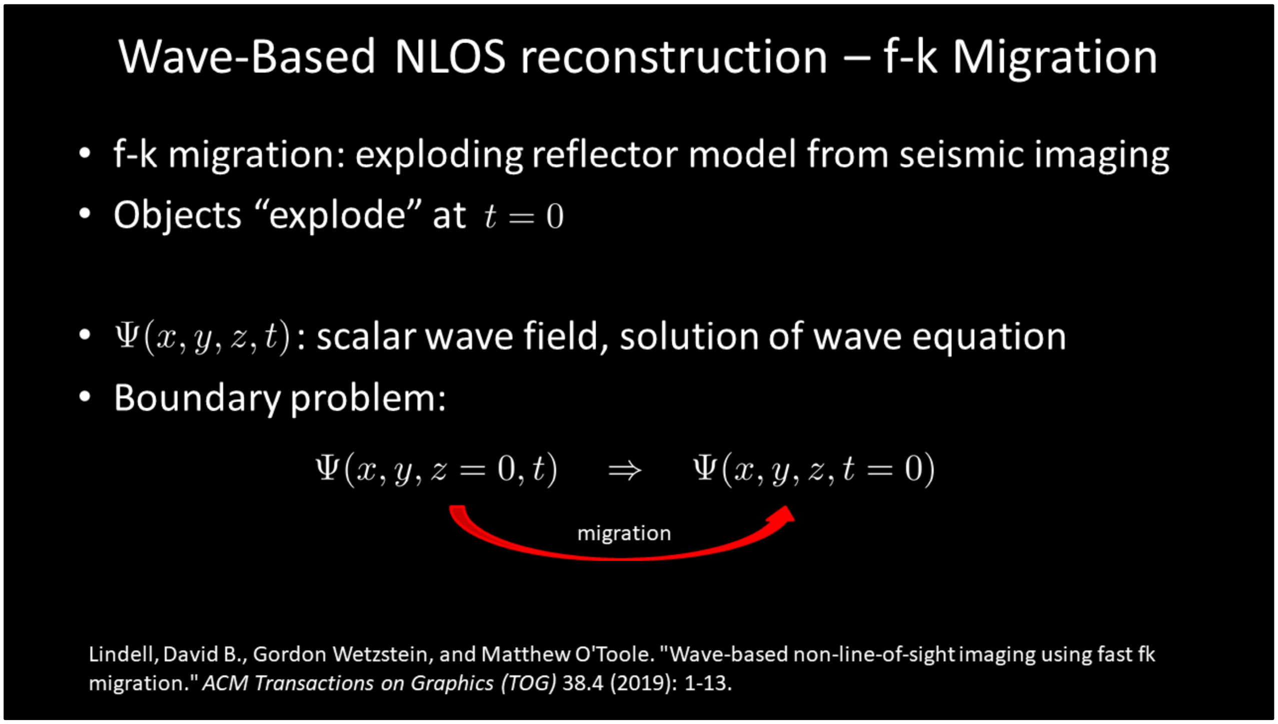
In total, the measured time responses are filtered with the virtual pulse, and after Fourier transform with respect to time, each wavelength is propagated back into the scene. Inverse Fourier transform with respect to time/frequency and taking the absolute value squared is the reconstruction. The RSD for propagating the waves from the relay wall plane to a parallel one is just a 2D convolution, which can be calculated efficiently in the 2D spatial frequency domain.

Wave-Based NLOS reconstruction – f-k Migration

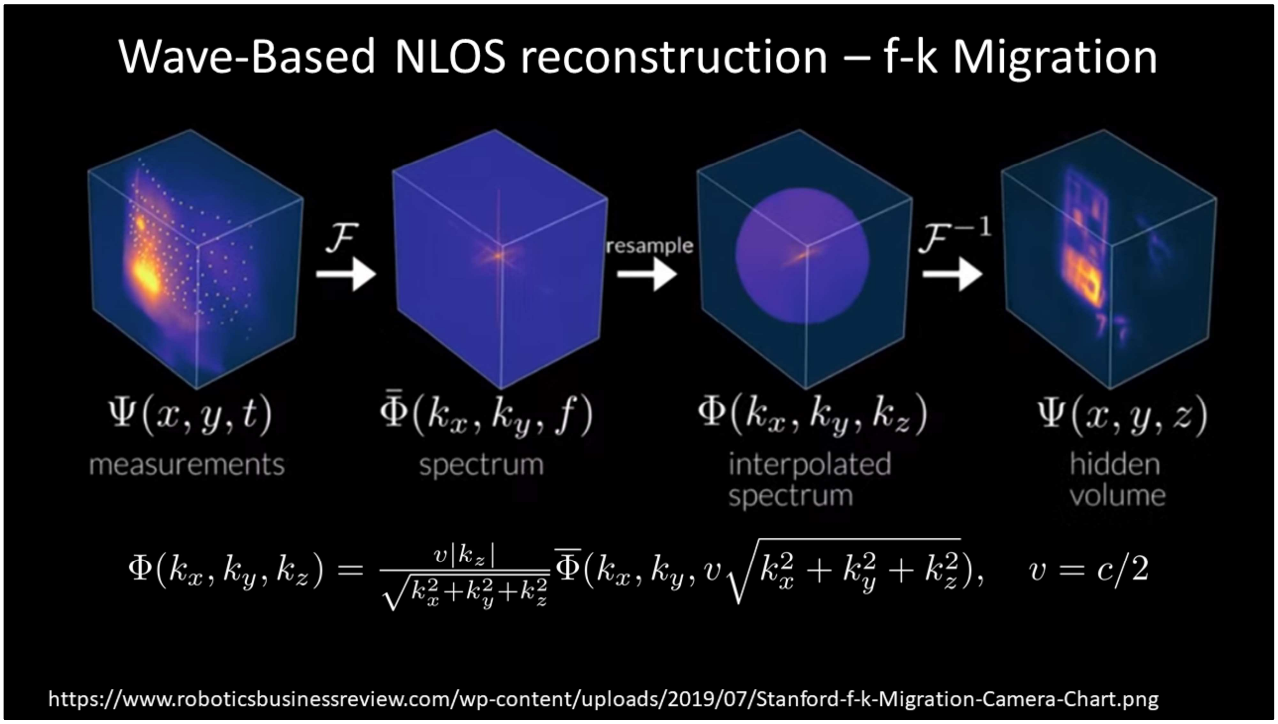
- f-k migration: exploding reflector model from seismic imaging
- Objects "explode" at $t = 0$
- 

Lindell, David B., Gordon Wetzstein, and Matthew O'Toole. "Wave-based non-line-of-sight imaging using fast fk migration." *ACM Transactions on Graphics (TOG)* 38.4 (2019): 1-13.

Migration process for plane wave imaging. In seismic imaging, small explosions are initiated at the earth's surface, and the resulting sound waves travel through the different layers of rock. The scatterers reached by a plane wave become secondary sources that emit upward spherical waves which show up as diffraction hyperbolas (the deeper the scatterer, the flatter the hyperbola) in the backscattered signals. Migration allows one to recover the scatterers' position. In the exploding reflector model, all scatterers are thought to "explode" at the same time, ignoring the time it takes the wave to travel downwards; this is accounted for by dividing the wave speed by 2.



Mathematically, we know the scalar wave field at the plane z=0, the relay wall plane, and want to know its shape in space at t=0, i.e., when all scene objects “exploded” simultaneously.



The process is then quite straightforward: the 3D Fourier transform of the acquired data cube is calculated, the frequency coefficients are resampled/interpolated, and the inverse 3D Fourier transform yields the scene reconstruction.

Wave-Based NLOS reconstruction – f-k Migration

- Fast calculation
- Method designed for confocal acquisition, preprocessing step provided for approximating non-confocal data as confocal
- Resampling requires memory intense interpolation (unoptimized Matlab)

The paper uses a computer with 256 GB RAM for a scene of 512^3 voxels

RECONSTRUCTION METHODS – LIGHT CONE TRANSFORM

Light-Cone Transform

- Confocal acquisition:

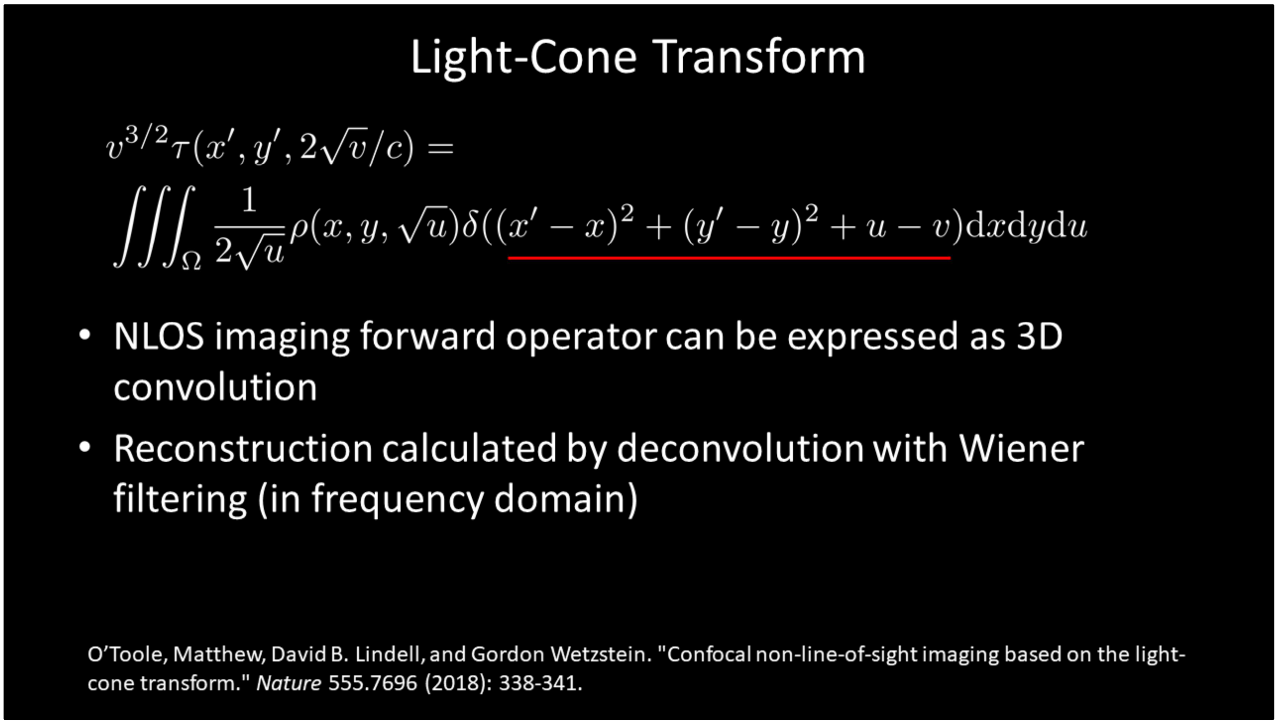
$$\tau(x', y', t) =$$
$$\iiint_{\Omega} \frac{1}{r^4} \rho(x, y, z) \delta(2\sqrt{(x' - x)^2 + (y' - y)^2 + z^2} - tc) \mathrm{d}x\mathrm{d}y\mathrm{d}z$$

- substitute $z = \sqrt{u},\ dz/du = 1/(2\sqrt{u}),\ v = (tc/2)^2$

$$v^{3/2}\tau(x', y', 2\sqrt{v}/c) =$$
$$\iiint_{\Omega} \frac{1}{2\sqrt{u}} \rho(x, y, \sqrt{u}) \delta((x' - x)^2 + (y' - y)^2 + u - v) \mathrm{d}x\mathrm{d}y\mathrm{d}u$$

O'Toole, Matthew, David B. Lindell, and Gordon Wetzstein. "Confocal non-line-of-sight imaging based on the light-cone transform." *Nature* 555.7696 (2018): 338-341.

Light cone: from relativity theory. The authors of the paper proposed confocal acquisition, where the laser spot and the SPAD observation position on the relay wall are collocated



In the frequency domain, convolution corresponds to multiplication, so calculating the reconstruction of the scene can be performed by dividing the measured data by the Fourier transformed delta function. As this leads to poor SNR, Wiener filtering is used which minimizes the mean squared error between the reconstructed volume and the ground truth.

Light-Cone Transform

- Fast calculation
- Resampling requires memory intense interpolation (unoptimized Matlab)

O'Toole, Matthew, David B. Lindell, and Gordon Wetzstein. "Confocal non-line-of-sight imaging based on the light-cone transform." *Nature* 555.7696 (2018): 338-341.

Comparison

	Compute	Memory
f-k migration	$O(n^3 \log n)$	$O(n^3)$
Light-cone transform	$O(n^3 \log n)$	$O(n^3)$
Wave-based, RSD	$O(n^3 \log n)$	$O(n^3)$
Backprojection (original or wave-based)	$O(n^5)$ *	$O(n^3)$

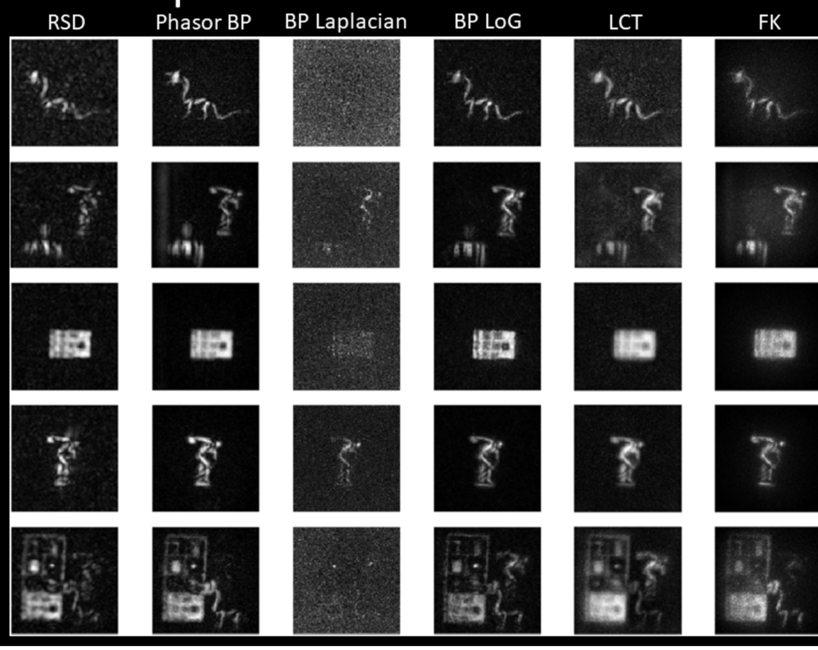
* Backprojection can actually be performed with lower complexity by using intermediate results similar to FFT, Natterer, F., & Wübbeling, F. (2001). *Mathematical methods in image reconstruction* (Vol. 5). Siam.

We want to point out that despite the relatively high complexity, backprojection is still one of the standard algorithms for computed tomography (CT) reconstruction and widely used. Fast implementations are based on GPUs

LCT/FK: approximation of non-confocal data as described in the FK paper which effectively shrinks the aperture size (small images); for the large images, the aperture size was increased by zero padding to effectively have the same aperture size as in the non-confocal case

Comparison

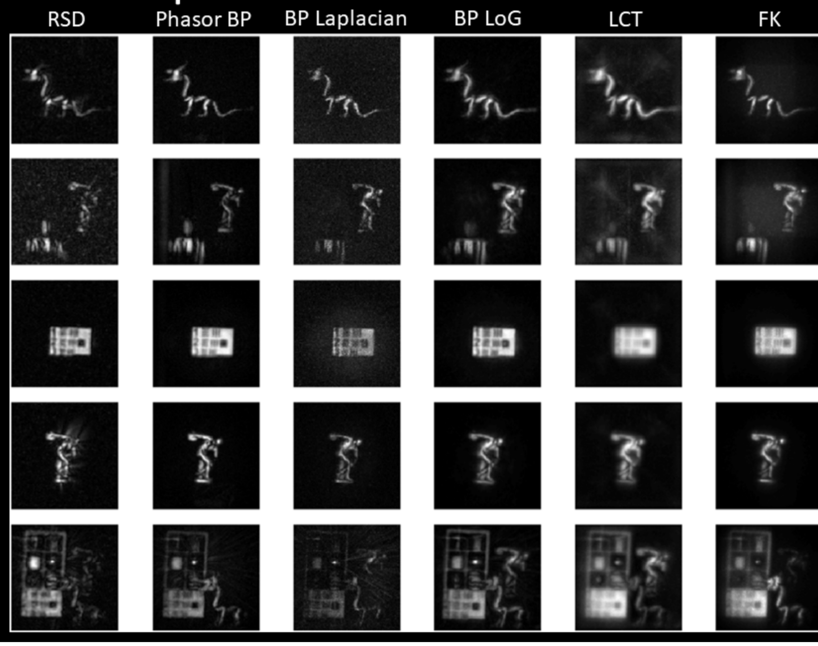
- Confocal data, short exposure



Liu, Xiaochun, Sebastian Bauer, and Andreas Velten. "Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems." *Nature Communications* 11.1 (2020): 1-13.

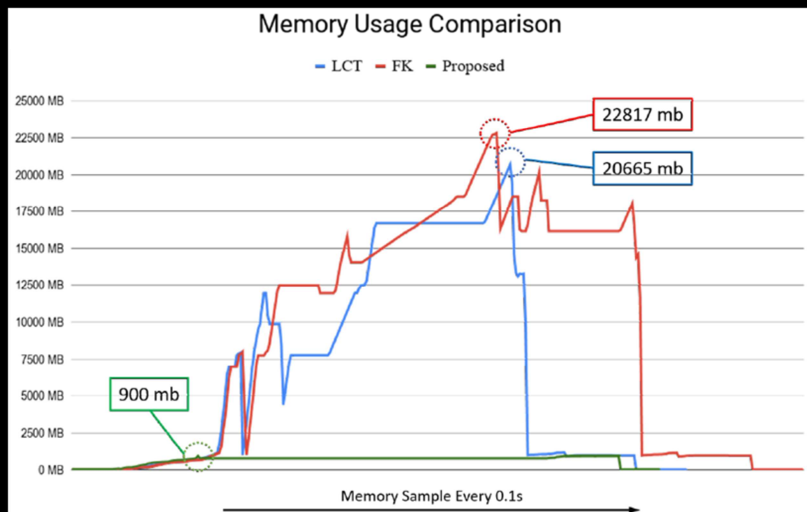
Comparison

- Confocal data, long exposure



Liu, Xiaochun, Sebastian Bauer, and Andreas Velten. "Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems." *Nature Communications* 11.1 (2020): 1-13.

Comparison



Liu, Xiaochun, Sebastian Bauer, and Andreas Velten. "Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems." *Nature Communications* 11.1 (2020): 1-13.

This is a comparison based on the respective current (unoptimized) Matlab implementations. The bottleneck for LCT and FK is the interpolation, which could be performed more efficiently in optimized implementations.

RECONSTRUCTION METHODS – DATA DRIVEN

Reconstruction with Neural Networks

- Typical reconstruction methods rely on a forward model, whose parameters are provided through calibration
  - Understanding details of the light transport matrix may require knowledge of the hidden scene
  - Can we learn prior knowledge from representative datasets?
- Difficult to define forward models (e.g. classification) are a clear win for data driven methods
- Forward model may be poorly conditioned, or the problem may be ill-posed.
  - Spatial structure and regularization can be imposed using neural networks without a dataset (e.g. Deep Image Priors)

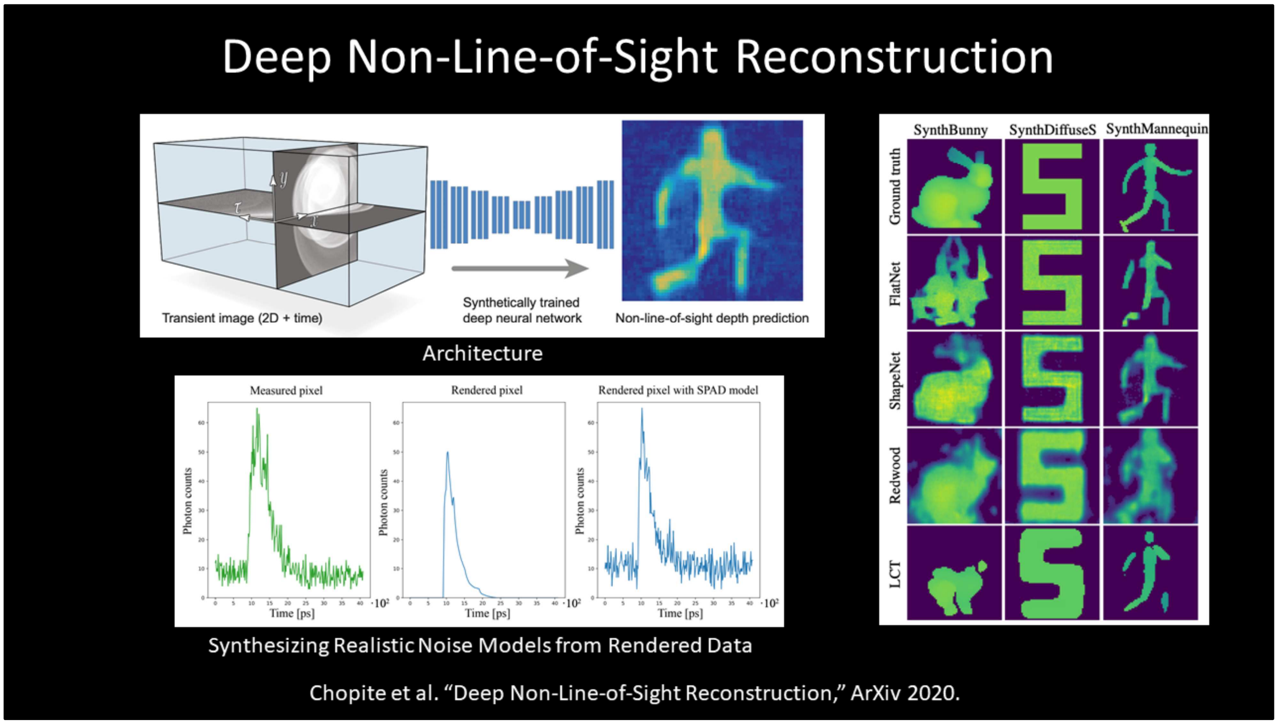
Challenges when using Data-Driven Methods

- Lack of available data
  - Non-traditional sensors (e.g. SPAD)
    - Difficult to characterize and model
  - Reliance on simulation and shape datasets (e.g. ShapeNet)
- Labeling when generating a dataset using real-world data
  - Requires additional experimental setup to capture labels in addition to measurements
- Generalization
  - Ensuring any learned model works outside the distribution of the training data is a (very) hard problem

Transient Rendering

- Essential for creating training data
  - Sensor noise characteristics still a difficult problem
  - Can use Mitsuba (https://www.mitsuba-renderer.org/) for time averaged
    - ToF version: https://github.com/cmu-ci-lab/MitsubaToFRenderer
- Transient Rendering Literature (non-exhaustive):

Arellano et al. “Fast back-projection for non-line of sight reconstruction,” Optics Express 2017.
Hernandez et al. “A computational model of a single-photon avalanche diode sensor for transient imaging,” ArXiv 2017.
Iseringhausen et al. “Non-line-of-sight reconstruction using efficient transient rendering,” ArXiv 2018.
Jarabo et al. “A framework for transient rendering,” ACM TOG 2014.
Marco et al. “Transient photon beams,” CEIG 2017.
Smith et al. “Transient Rendering,” Tech Report UCSC 2008.



A recent line of work has been to explore if deep learning can be used to learn an approximate mapping from measurements directly to the hidden albedo volume. Effectively, this paper learns a function mapping transient measurements directly to a depth map of the hidden scene (rather than full voxel reconstruction). In order to create the dataset, this paper utilized an efficient rendering system for creating simulated data. The paper explores the importance of the dataset to for producing good reconstructions.

Comparing NLOS Methods

	Illumination	Sensor	Cost	Sensitivity to Ambient Light	Need of Priors on Geometry
Pulsed ToF	Pulsed Laser	Streak Camera SPAD	High	Robust	Not Required
AMCW ToF	Modulated Laser, LED	Correlation Camera	Medium	Sensitive to Strong Light	Not Required
Passive Coherence	None	Traditional Camera	Low	Sensitive	Scene Geometry
Passive Spatial Coherence	None	Dual Phase Sagnac Interferometer	Medium	Sensitive	Scene Geometry
Active Coherence	Coherent Source	Traditional Camera	Low	Sensitive	Scene Geometry
Passive Intensity	None	Traditional Camera	Low	Sensitive	Scene Geometry Occlusion Geometry
Active Intensity	Flashlight Laser	Traditional Camera	Low	Sensitive	Scene Geometry

Maeda et al. "Recent Advances in Imaging Around Corners," ArXiv 2019.
(See for full reference list)

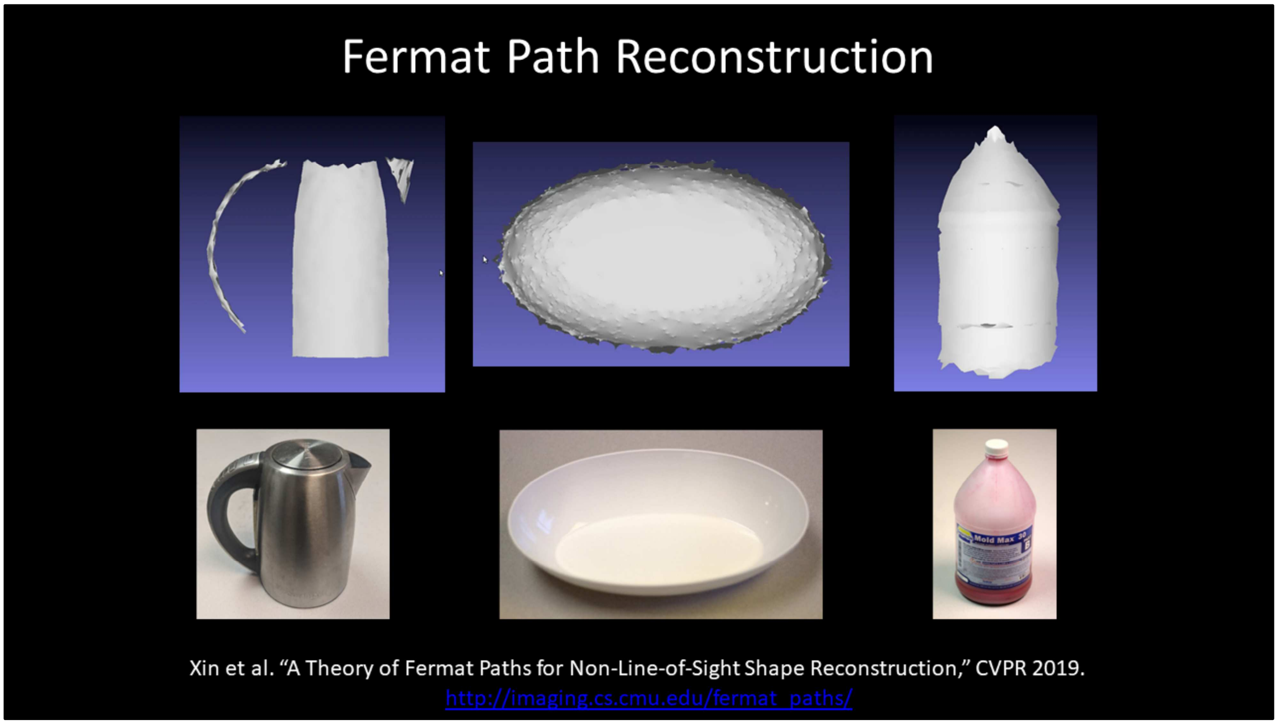
In addition to the comparison in the table above, an important consideration when comparing the NLOS sensing approaches is their suitability for long-range operation. In principle, ToF based approaches can work at long range (100s of meters or more), while other proposed methods using coherence (e.g. speckle) have been demonstrated on smaller scenes (less than 1 meter).

Non-Line-of-Sight Imaging Conclusion

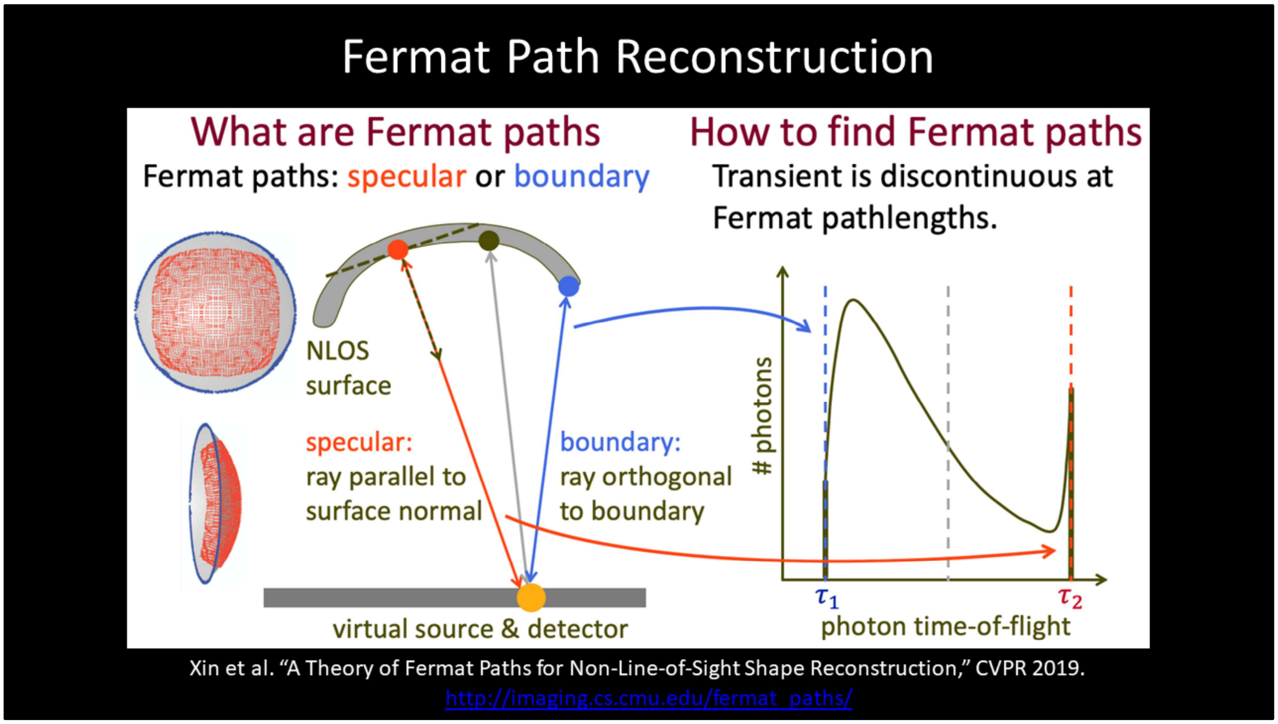
- There are numerous modalities used to extract NLOS signals
  - ToF with SPAD sensors have been leading state of the art, but other methods have unique advantages and tradeoffs
- Future Outlook and Challenges
  - Low SNR capture, Realtime Reconstruction
  - Compact and mobile systems
    - Dynamic Calibration
    - Relay Wall Surface Materials (e.g. heterogeneous)
    - Complex environments and clutter
  - Demonstration of scientific and consumer applications

ADDITIONAL MATERIAL

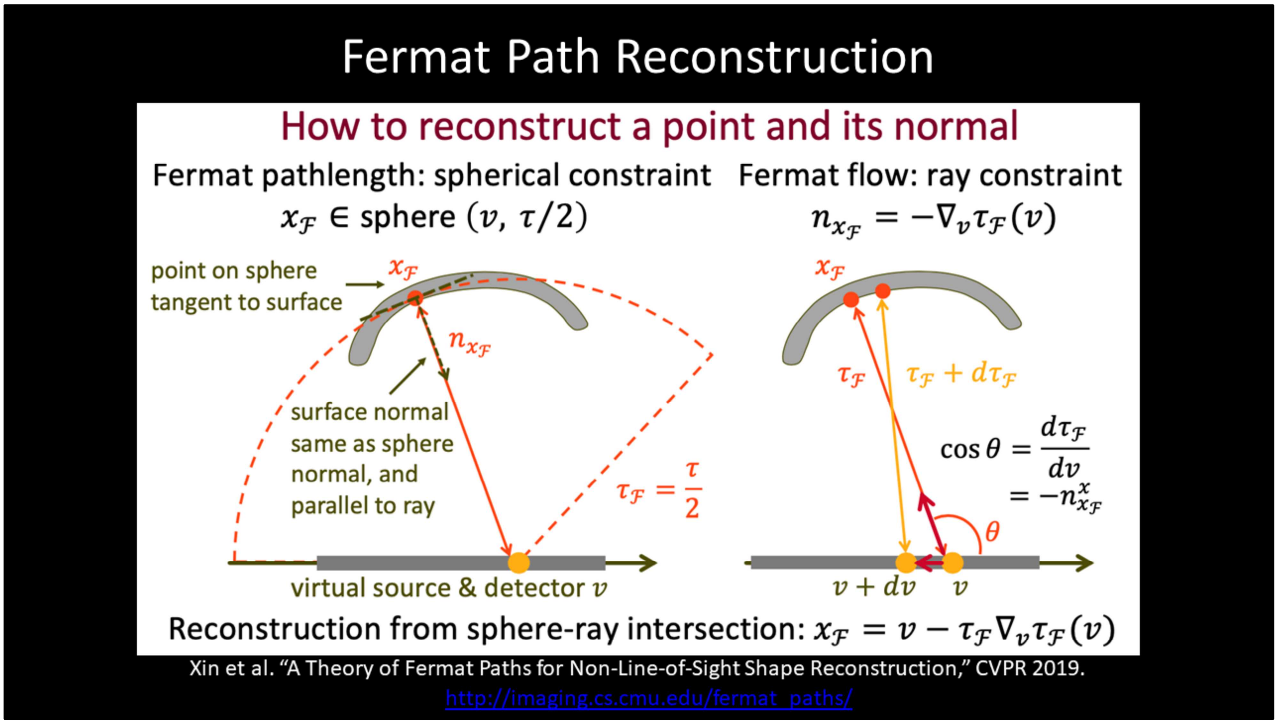
RECONSTRUCTION METHODS – FERMAT PATHS



Fermat Paths are particular photon trajectories that have well defined signatures in the transient signal. These correspond largely to specular paths or paths at hidden object edges. An interesting consequence to using Fermat paths is that the algorithm is suitable for recovering mesh-like representations.

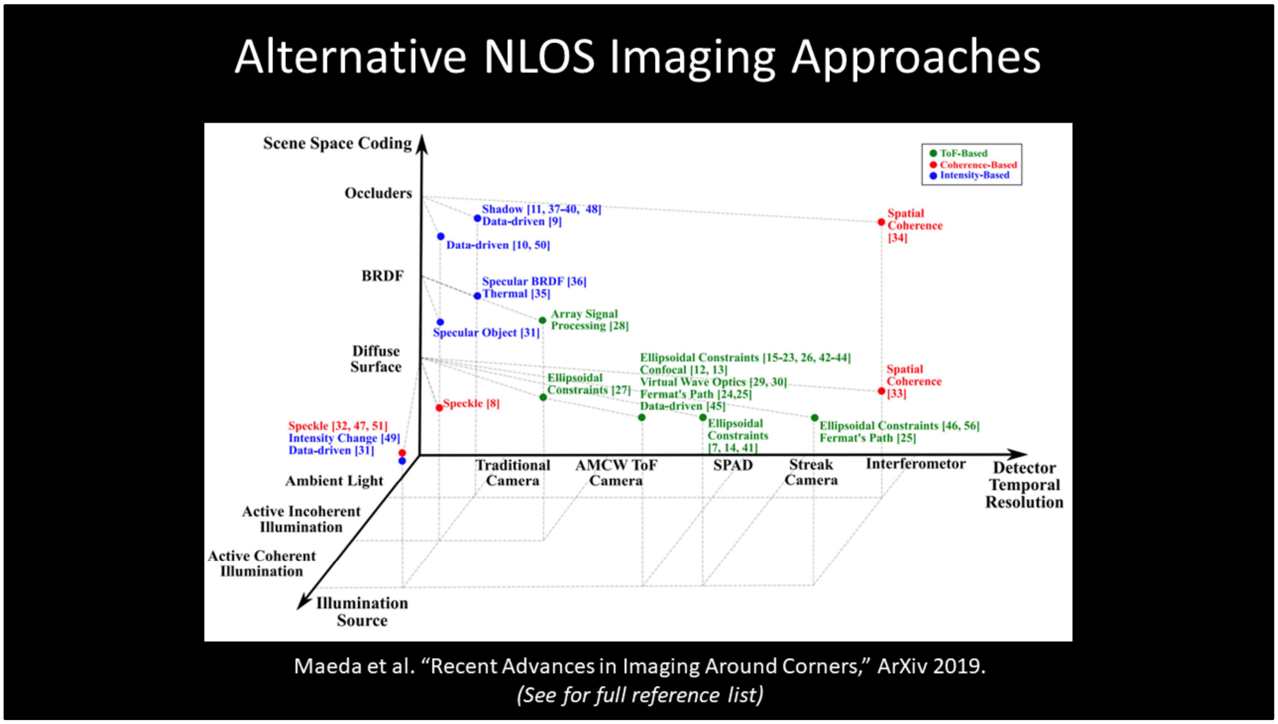


After identifying signatures of Fermat paths in the transient response, the paper shows how the hidden surface can be recovered.



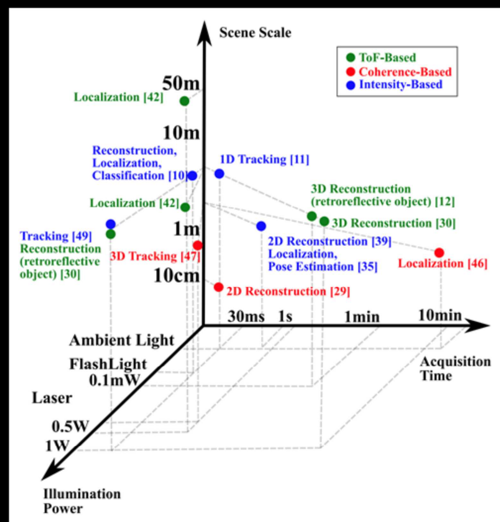
Note that the location on the transient of the Fermat path restricts the location of the recovered point to an ellipse, but correspondence between these transients for adjacent scan points enables recovery of the normal of the hidden surface. Therefore a single point can be found in the hidden region from these Fermat paths, and a point cloud can be formed.

OTHER METHODS FOR SEEING AROUND CORNERS



Numerous capture methodologies have been proposed in the literature. Broadly, we can distinguish them by their illumination source, the temporal resolution of the detector, and what kind of structure exists in the visible scene. Without physical constraints, we would ideally like a system that operates on any surface, under ambient light, and with poor detector resolution. However, the laws of physics impose limitations, forcing different approaches to make a tradeoffs between these capabilities. We notice that different modalities tend to cluster together, making different kinds of tradeoffs. In short, ToF based approaches tend to work on diffuse relay walls, with little prior knowledge of the visible scene required, but with added hardware complexity. Coherence-based approaches tend to require either complex sensing or illumination hardware. Intensity-based (or time-averaged) approaches require more scene structure but greatly simplify the necessary hardware complexity, but often do not perform as detailed reconstructions.

Demonstrated Capability of NLOS Imaging Approaches



Maeda et al. "Recent Advances in Imaging Around Corners," ArXiv 2019.
(See for full reference list)

NLOS imaging has been demonstrated in a variety of conditions. Full reconstruction typically requires better SNR than other tasks such as localization, and as such must either operate on smaller scenes, with longer integration times, or with high illumination power.

Alternative NLOS Imaging Approaches

	ToF	Coherence	Intensity
3D Reconstruction	µm-cm Resolution [7, 12–30]	None	Planar/Specular Object [31]
2D Reconstruction	Included in 3D	Diffraction Limited Resolution [8, 32] cm Resolution [33, 34]	Coarse Resolution [31], Thermal [35] Occlusion Dependent [9, 10, 36–40]
Localization/ Tracking	cm Precision [41–46]	1D Distance Recovery [33] 3D Tracking [47]	Occlusion Dependent [10, 11, 48] 6D Tracking [49], 3D localization [35, 50]
Classification	Human Identification [45]	MNIST, Human Pose Classification [51]	Object Type Classification [10, 50], Human Pose Detection [35]

Maeda et al. "Recent Advances in Imaging Around Corners," ArXiv 2019.
(See for full reference list)

The different capture methodologies have demonstrated various capability across tasks. ToF based approaches are most mature, and have been shown state of the art performance in 3D and 2D reconstruction, localization, and classification tasks. The capture setups are more complex, however, and simplified measurement setups have been shown to work well for "simpler" tasks such as Classification and Localization, but reconstruction remains a major challenge.

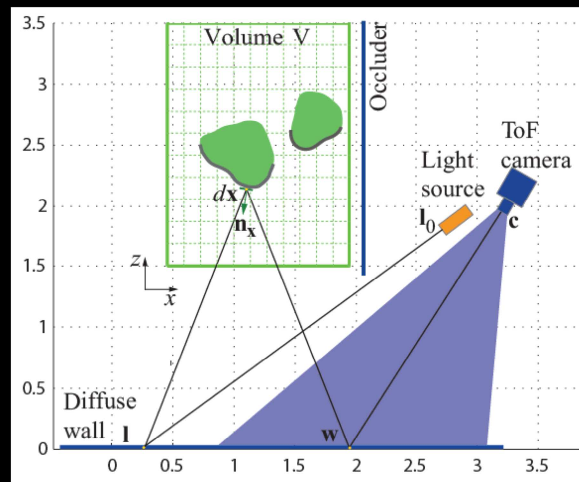
OTHER METHODS FOR SEEING AROUND CORNERS – CONTINUOUS WAVE

Continuous Wave NLOS Imaging

- ToF hardware is expensive, Continuous Wave (CW) hardware relatively cheap
- 2 approaches by Heide et al. and Kadambi et al.

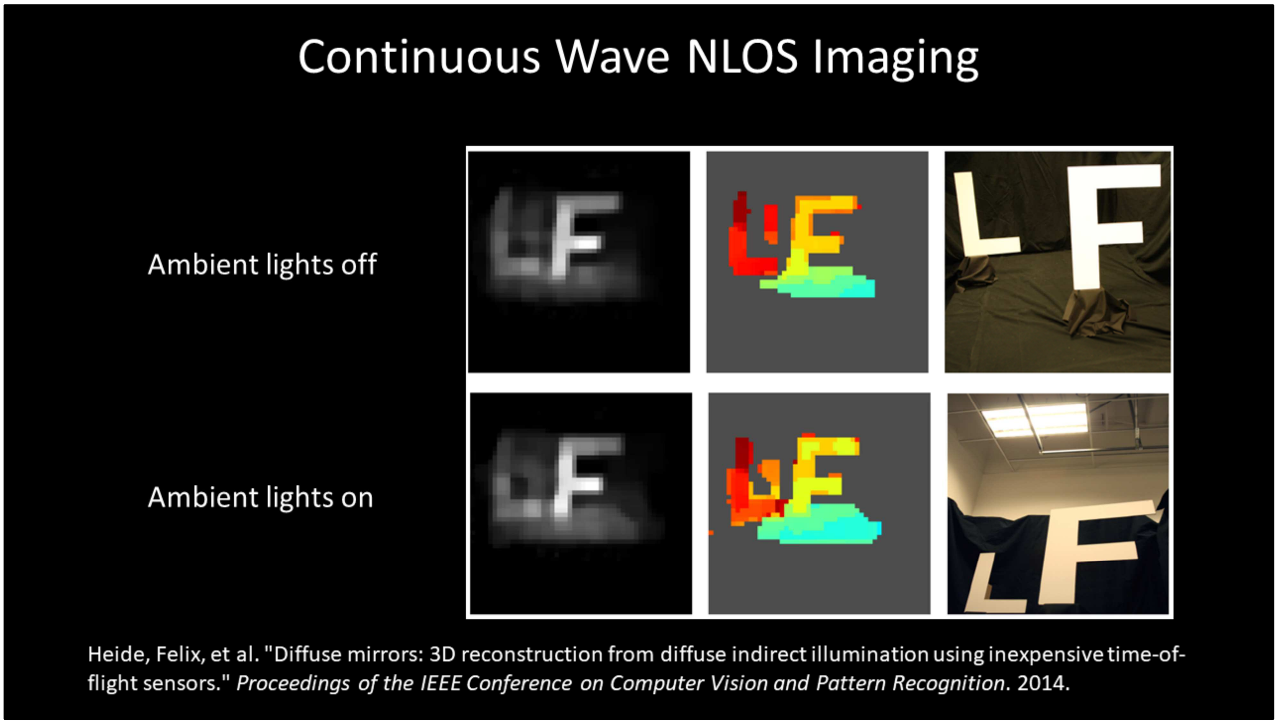
Continuous Wave NLOS Imaging

- 6 250 mW laser diodes focused on one spot, 2-3 ns long pulses, ~ 1 % duty cycle
- PMD (Photonic Mixer Device) camera measuring phase shift
- Total cost about $1,500



Heide, Felix, et al. "Diffuse mirrors: 3D reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.

We are unsure about the actually used light levels and modulation frequencies. In their preceding non-NLOS paper (**Low-budget Transient Imaging using Photonic Mixer Devices)**, the authors say that they use multiple frequencies, but in this paper, they don't say anything about that. 2-3 ns pulses with variable repetition rate does not help a lot, because in the previous paper they said they can use frequencies up to 180 Mhz, stably only up to 110 MHz, which means a period length of about 9 ns. From the previous paper: "In order to complete a full measurement over a frequency range from 10 to 120MHz in steps of 0.5 MHz, our capture system takes about 90 seconds. Note that we operate the setup at a duty cycle of less than 1%, in order to avoid overheating of the critical components (signal buffer and laser driver ICs, laser diodes, PMD sensor) that are not equipped with heatsinks or active cooling. We therefore estimate that with proper thermal management, another significant speedup will be achieved, reducing the overall acquisition time to only a few seconds." also: "Our light source is a bank of six 650 nm laser diodes with a total average output power of 2.4W", but it is uncertain if that includes the 1 % duty factor or not; most likely, that would be the full 100 % duty cycle power

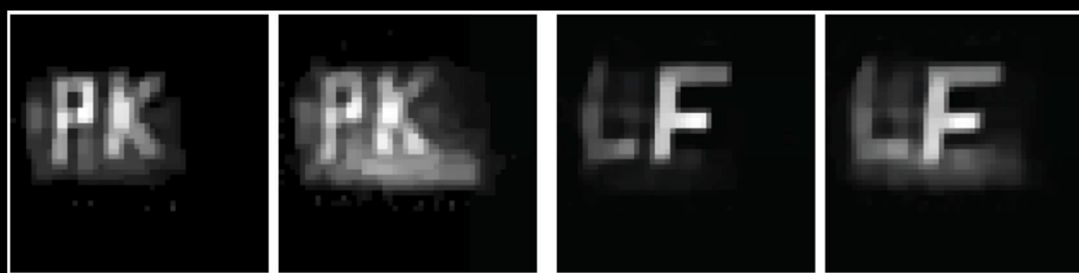


Reconstruction calculated with regularization, 3 terms, 2 of them enforce sparsity

Continuous Wave NLOS Imaging

- Reconstruction calculated by optimization using sparsity constraints

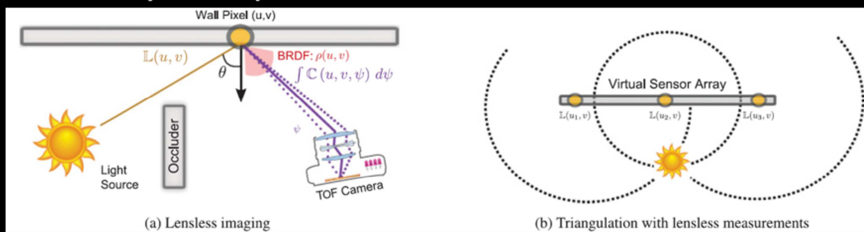
500 images averaged (200 minutes) | 10 images averaged (4 minutes) | 500 images averaged (200 minutes) | 10 images averaged (4 minutes)



Heide, Felix, et al. "Diffuse mirrors: 3D reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2014.

Continuous Wave NLOS Imaging

- ToF hardware is expensive, Continuous Wave (CW) hardware relatively cheap



(a) Lensless imaging (b) Triangulation with lensless measurements

- $c(u,v,t) = A(u,v)\sin(2\pi f_{\mathrm{M}} + \phi(u,v)) + \xi(u,v)$
- (u,v) relay wall coordinates, f_{M} modulation frequency
- Phasor formulation: $\mathbb{M}(u,v) = A_{\mathbb{M}}(u,v)e^{j\phi_{\mathbb{M}}(u,v)}$

Kadambi, Achuta, Hang Zhao, Boxin Shi, Ramesh Raskar. "Occluded imaging with time-of-flight sensors." *ACM Transactions on Graphics (ToG)* 35.2 (2016): 1-12.

Continuous Wave NLOS Imaging

- Forward problem is phasor multiplication then:

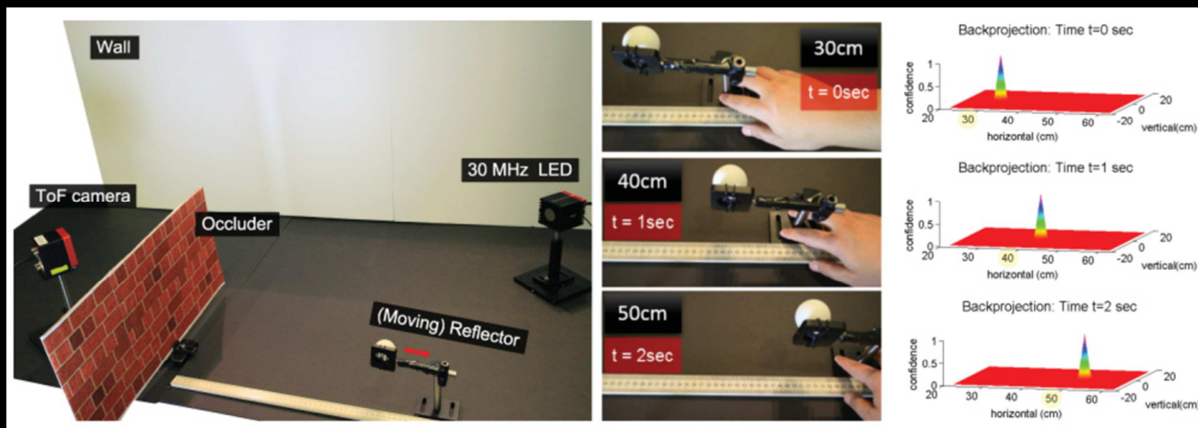
$$\mathbb{M}(u,v,\psi) = A_0 \underbrace{\mathbb{L}(u,v)}_{\text{Transport phasor object} \rightarrow \text{wall}} \overbrace{\mathbb{C}(u,v,\psi)}^{\text{Transport phasor wall} \rightarrow \text{camera}}$$

- Reconstruction performed using sparse solver applied to dictionary containing all possible target locations

Kadambi, Achuta, Hang Zhao, Boxin Shi, Ramesh Raskar. "Occluded imaging with time-of-flight sensors." *ACM Transactions on Graphics (ToG)* 35.2 (2016): 1-12.

Light source position must be known

Continuous Wave NLOS Imaging



(a) Moving object around the corner (b) Video of object (c) Object tracking

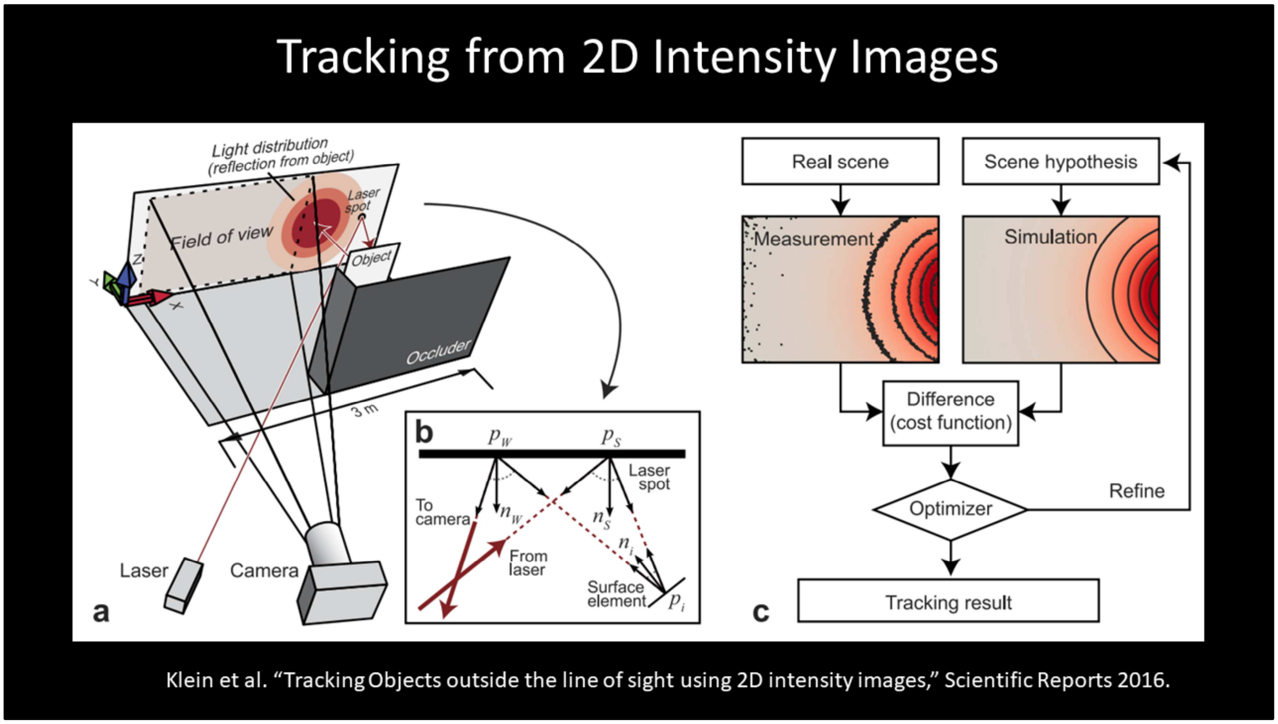
Kadambi, Achuta, Hang Zhao, Boxin Shi, Ramesh Raskar. "Occluded imaging with time-of-flight sensors." *ACM Transactions on Graphics (ToG)* 35.2 (2016): 1-12.

Note that the light source is placed on the hidden scene side

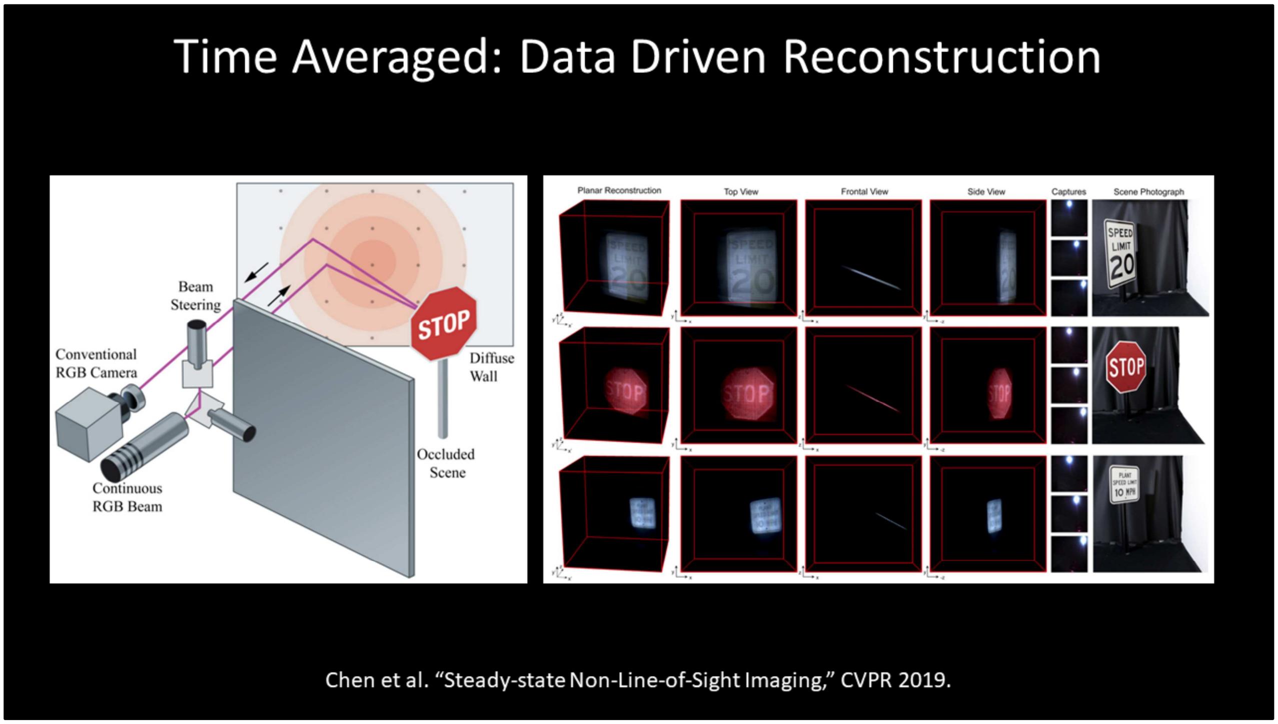
OTHER METHODS FOR SEEING AROUND CORNERS – COMPLETELY PASSIVE

Time-averaged and Passive NLOS

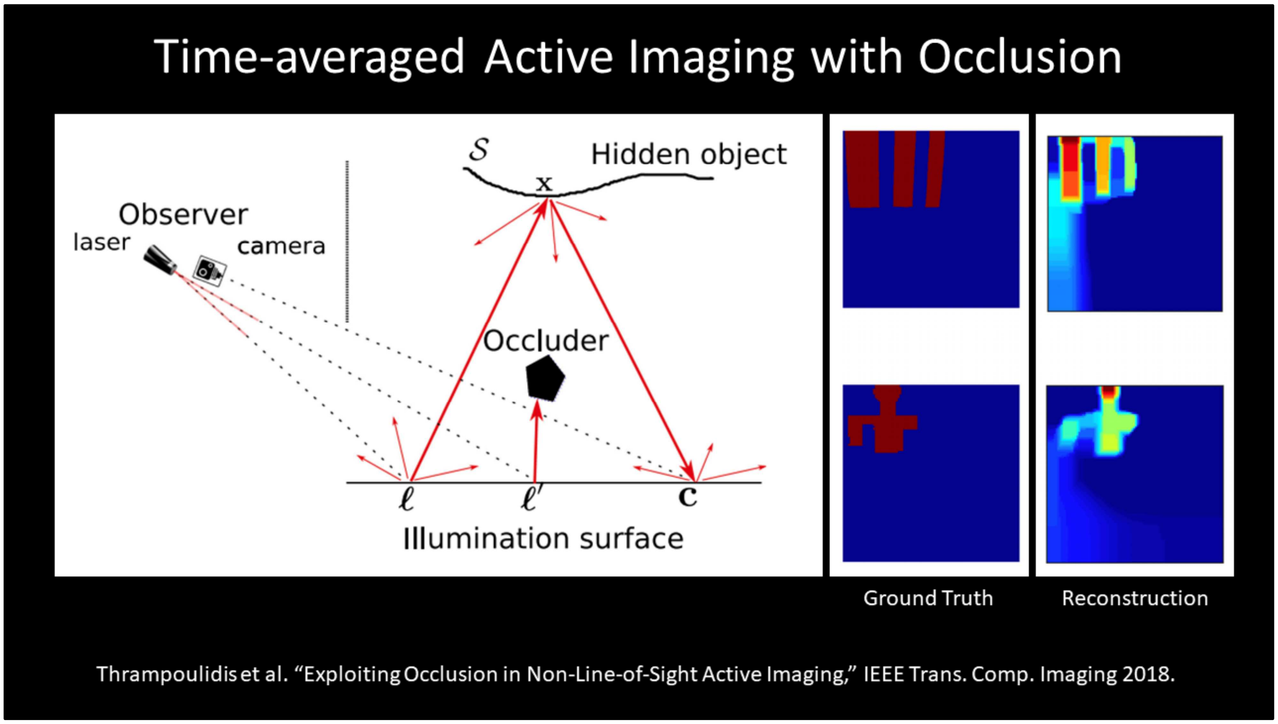
- Can a ”normal” video camera be used for NLOS imaging?
  - Full albedo reconstruction is challenging, but recent results have demonstrated localization, tracking, classification, and partial reconstruction
  - Due to ill-posed nature of the problem, many approaches use data driven methods to implicitly develop priors
  - High sensitivity required by capturing device unless scene contains additional structure (e.g. occlusions or specular paths)
- Some of these methods also can include NLOS imaging using ambient illumination, known as *passive* NLOS



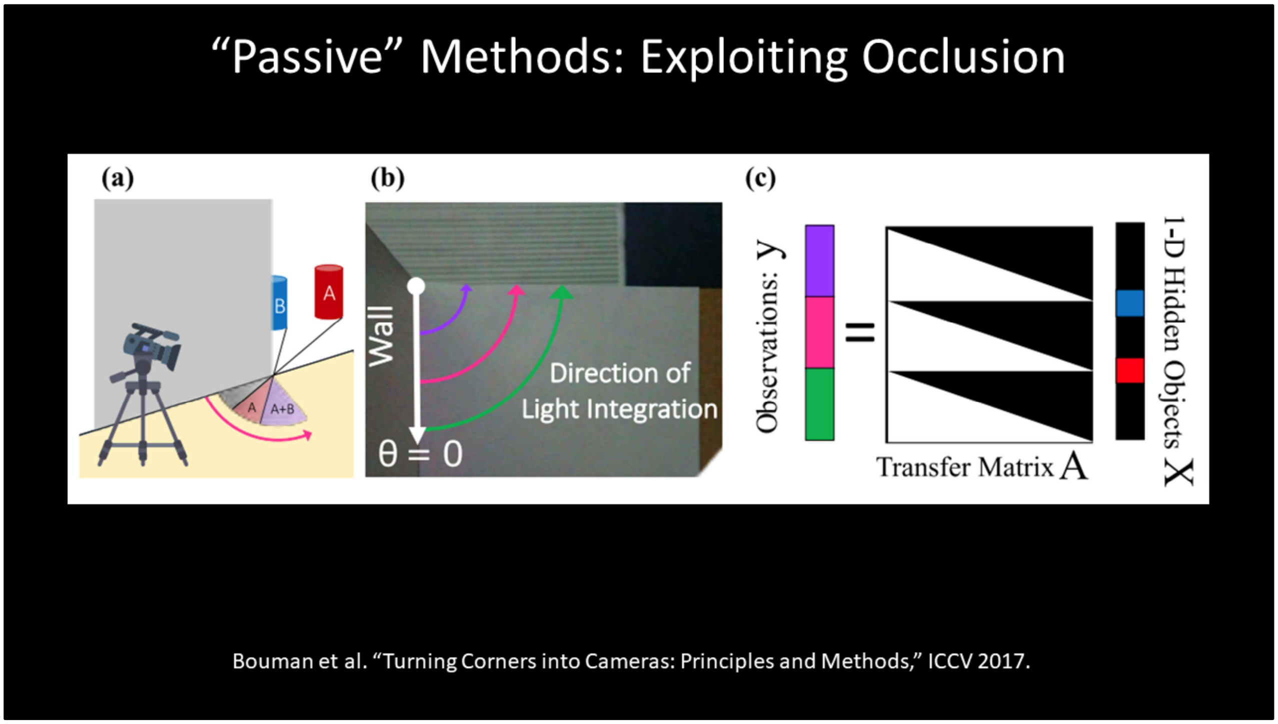
Using a inverse-rendering approach, Klein et al. were one of the first to demonstrate tracking of objects from time-averaged measurements.



Active imaging was also used by Chen et al. to capture a set of time averaged measurements for a few scanned locations of the active source. The paper proposes a method for reconstructing planar scenes, as well as the addition of a CNN to learn how to reconstruct orthogonal projections of more complex hidden scenes.

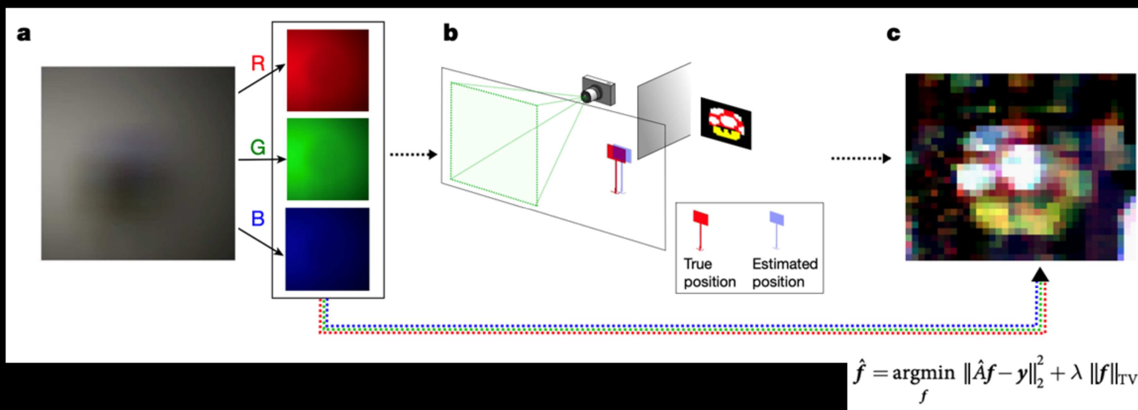


One interesting line of research is the use of occlusion in this hidden scene to better condition the underlying light transport matrix, making inversion more feasible. Thrampoulidis et al. demonstrated reconstruction using time-averaged measurements, an active illumination source, and occlusion in the hidden scene.



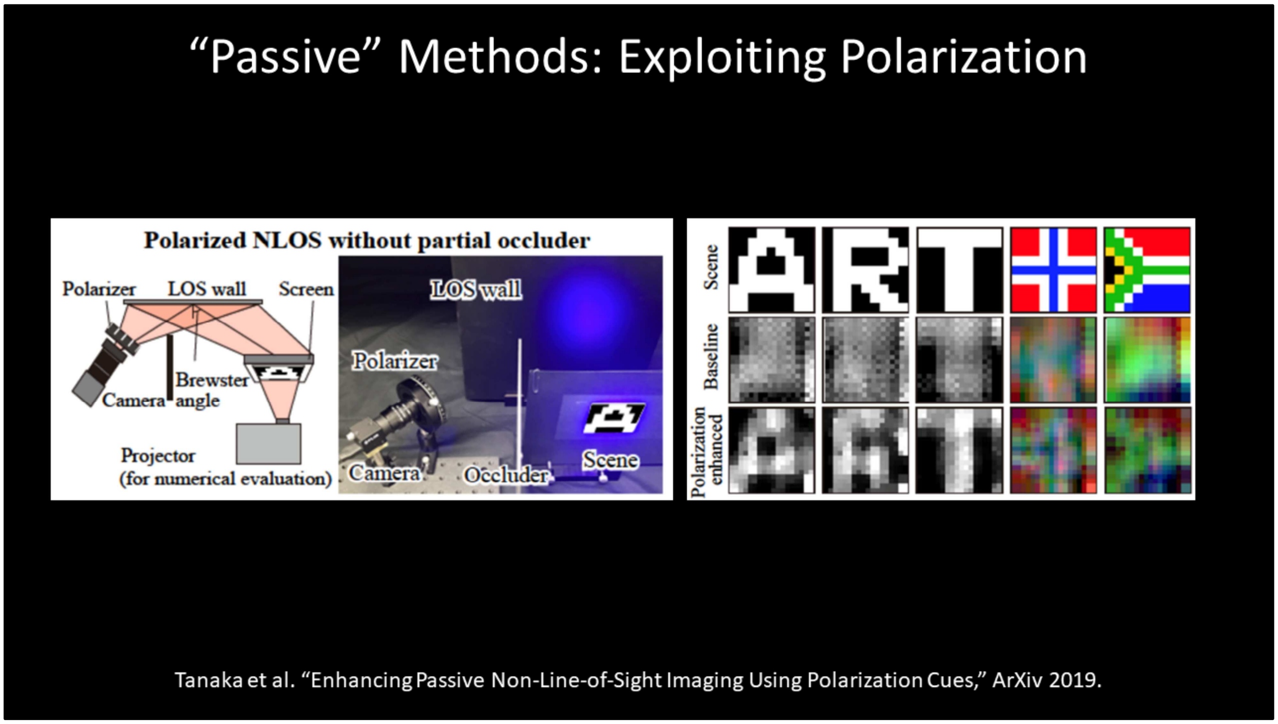
Bouman et al demonstrated that the occluding wall used in many experiments creates signal in the visible scene that can be used to recover a 1D projection of the hidden scene. This method works under a variety of ambient lighting conditions, and can operate using ambient illumination in favorable conditions.

“Passive” Methods: Exploiting Occlusion



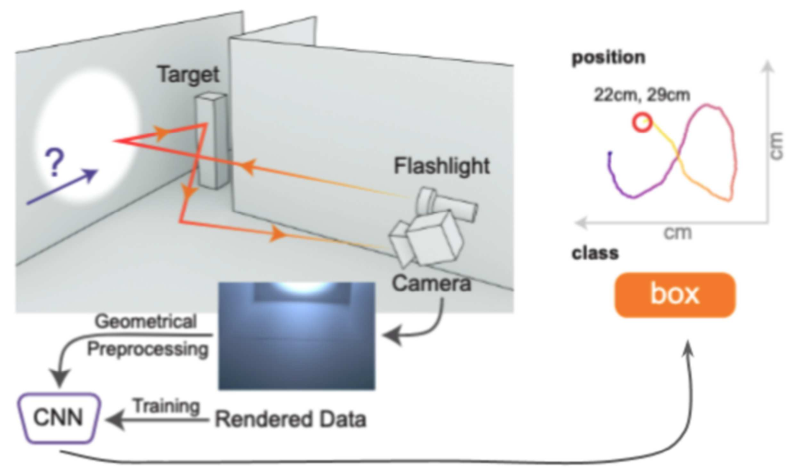
Saunders et al. “Computational periscopy with an ordinary digital camera,” Nature 2019.
See also:
Aittala et al. “Computational Mirrors: Blind inverse light transport by deep matrix factorization,” NeurIPS 2019.
Baradad et al. “Inferring light fields from shadows,” CVPR 2018.
Yedidia et al. “Using unknown occluders to recover hidden scenes,” CVPR 2019.

Occlusion is a powerful feature that can modulate light paths as they travel through the scene, effectively making the ray transport matrix better conditioned. A recent line of work has considered various blind inversion scenarios, where partial information about occlusion can be used to simultaneously recover the ray transport matrix and projections of the hidden scene.



Other ways of better conditioning the ray transport matrix have been demonstrated, including the use of polarization cues.

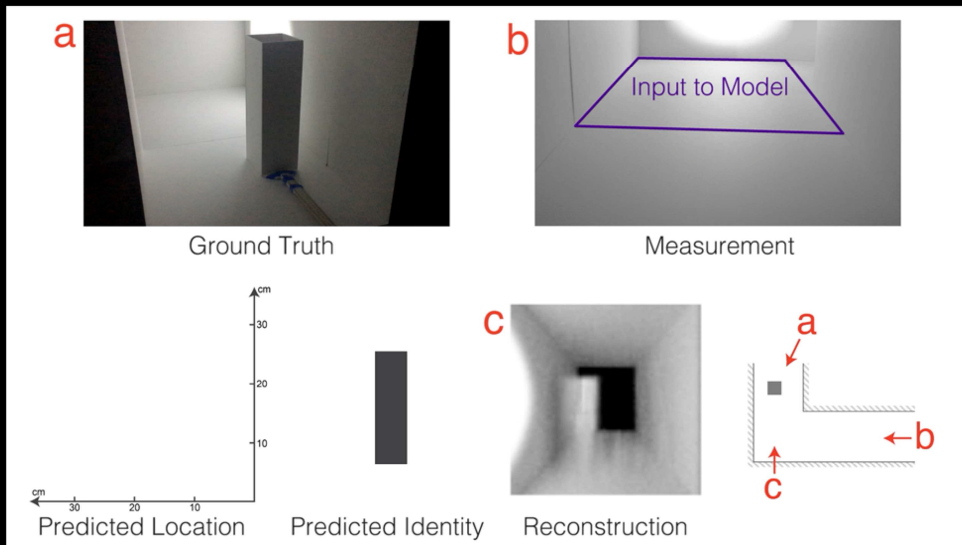
Time-averaged: Data Driven NLOS



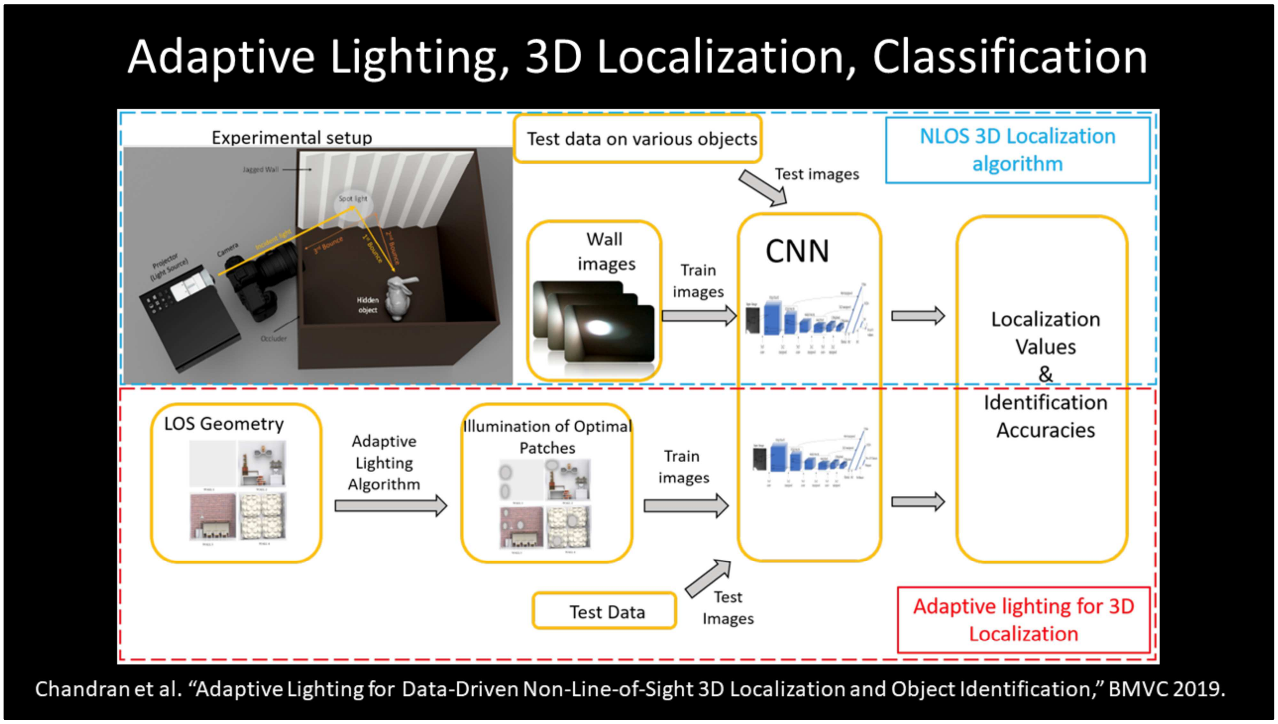
Tancik, Swedish, Satat, Raskar. "Data-Driven Non-Line-of-Sight Imaging with a Traditional Camera," COSI 2018.
Tancik, Satat, Raskar. *Flash Photography for Data-Driven Hidden Scene Recovery*. Arxiv, 2018.

Considering the considerable challenges inherent in time-averaged measurements, data-driven methods have been proposed to learn approximate inverse mappings, translating observations directly into reconstructions, localization, or classification of the hidden scene. These methods are trained using simulated data, and have been shown to transfer to real-world experimental data.

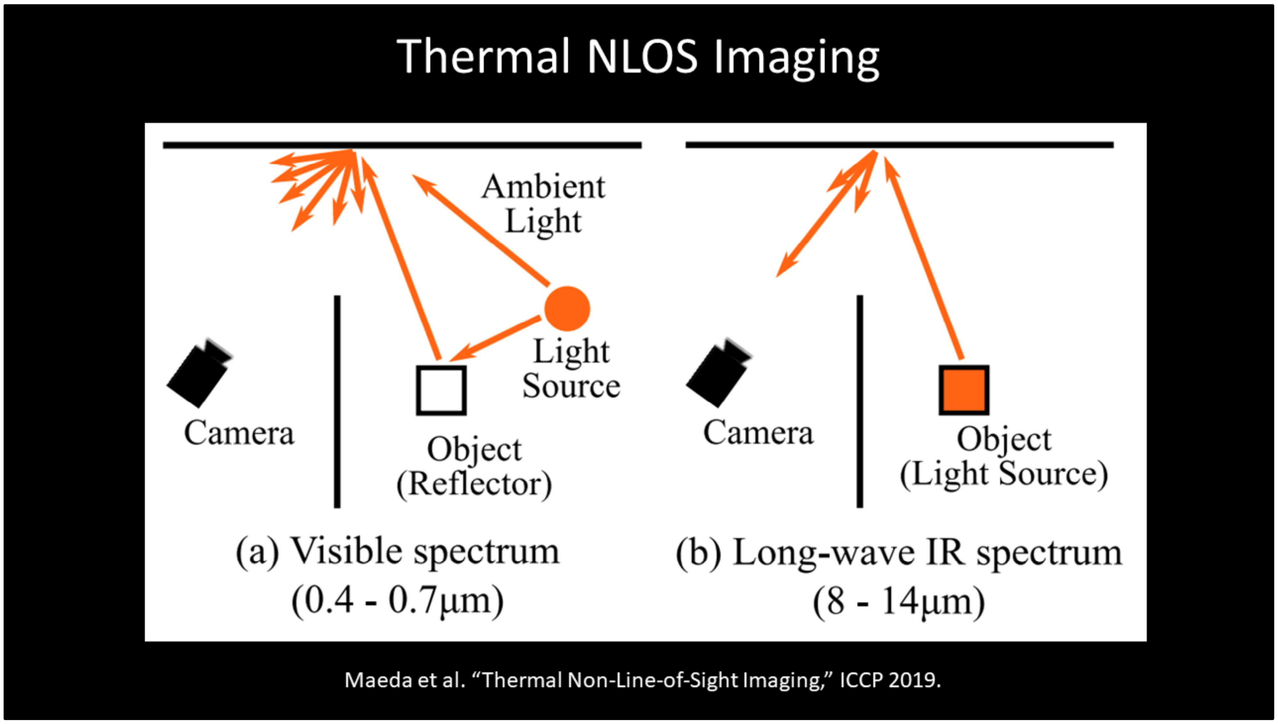
Time-averaged: Data Driven NLOS



Tancik et al. “Flash Photography for Data-Driven Hidden Scene Recovery,” ArXiv 2018.

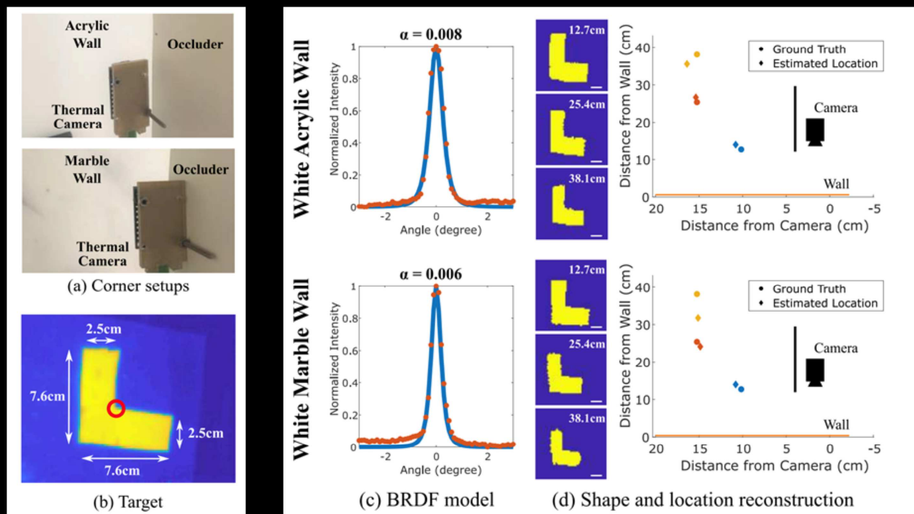

Adaptive Lighting, 3D Localization, Classification
Experimental setup
Test data on various objects
NLOS 3D Localization algorithm
Test images
Wall images
Train images
CNN
Localization Values & Identification Accuracies
LOS Geometry
Adaptive Lighting Algorithm
Illumination of Optimal Patches
Train images
Test Data
Test Images
Adaptive lighting for 3D Localization
Chandran et al. "Adaptive Lighting for Data-Driven Non-Line-of-Sight 3D Localization and Object Identification," BMVC 2019.

OTHER METHODS FOR SEEING AROUND CORNERS – THERMAL



Another modality that has been explored is long wave infrared radiation. LWIR has a much longer wavelength, greatly reducing scattering, and making the resulting scattering much more specular for many common materials used to construct relay-walls. In this configuration, a hidden object (e.g. a human) emits LWIR, and a camera is able to observe the scene convolved with the relay wall BRDF. By fitting measurements to a physics-based microfacet BRDF model, both the direction and distance of the hidden emitter can be recovered.

Thermal NLOS Imaging

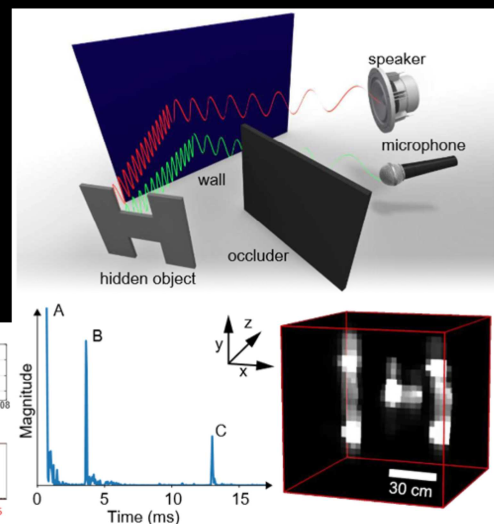


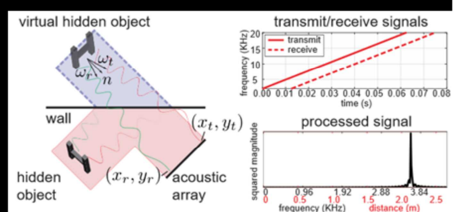
Maeda et al. “Thermal Non-Line-of-Sight Imaging,” ICCP 2019.

OTHER METHODS FOR SEEING AROUND CORNERS – ACOUSTIC

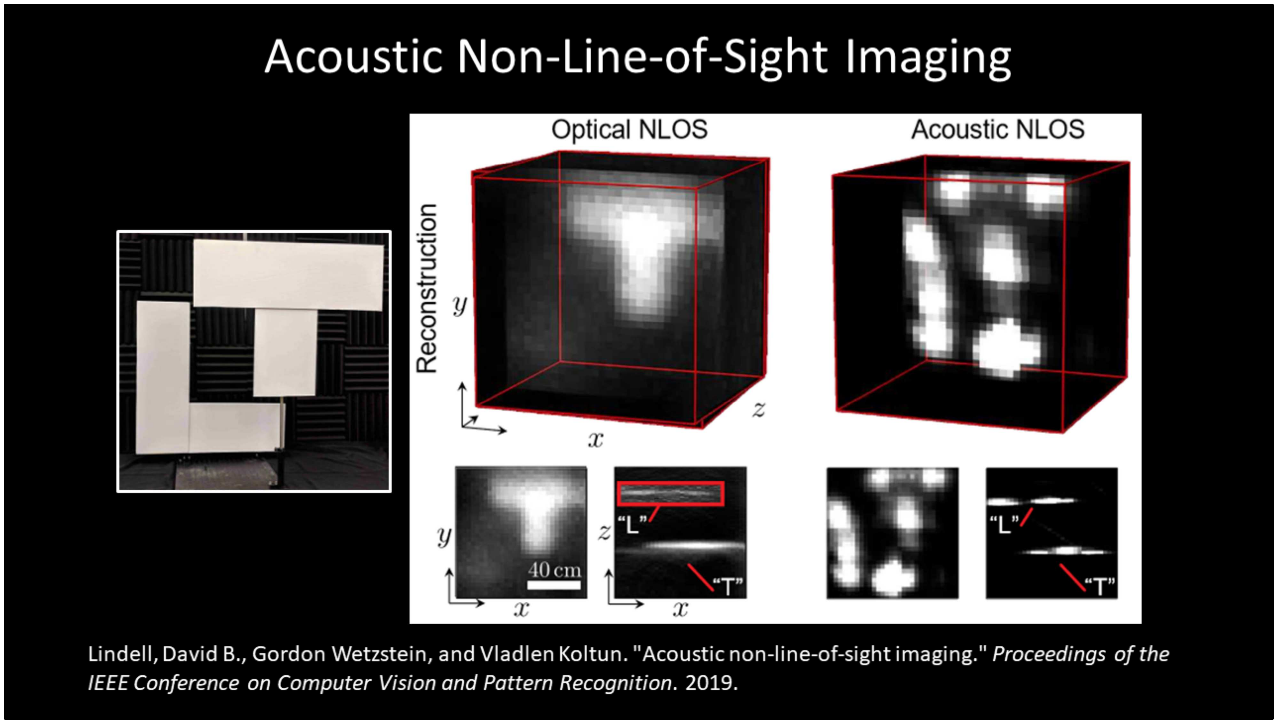
Acoustic Non-Line-of-Sight Imaging

- "Seeing" around corners with sound waves
- Emitter and microphone arrays, chirp signals
- Walls are "mirrors" for sound waves





Lindell, David B., Gordon Wetzstein, and Vladlen Koltun. "Acoustic non-line-of-sight imaging." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 2019.

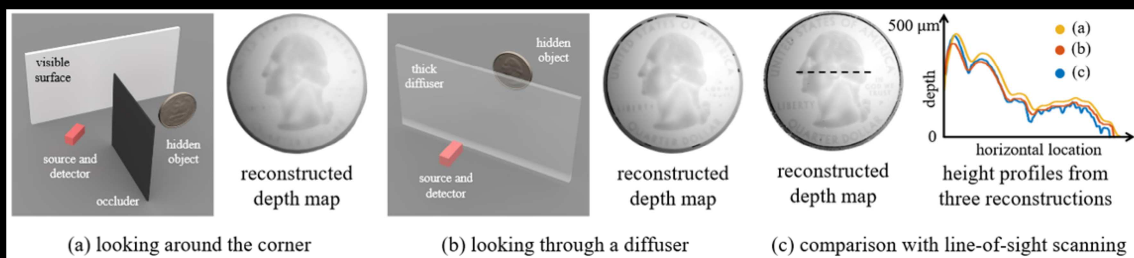


One disadvantage: sound waves are susceptible to wind and other air movement, therefore probably not feasible for outside applications

OTHER METHODS FOR SEEING AROUND CORNERS – INTERFEROMETRIC

Interferometric Non-Line-of-Sight Imaging

- Data acquisition with an interferometer allows for NLOS depth resolution down to 10 µm



Xin, Shumian, et al. "A theory of fermat paths for non-line-of-sight shape reconstruction." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2019

Gkioulekas, Ioannis, et al. "Micron-scale light transport decomposition using interferometry." *ACM Transactions on Graphics (ToG)* 34.4 (2015): 1-14.

Also highly recommended reading: Willomitzer, Florian, et al. "Synthetic Wavelength Holography: An Extension of Gabor's Holographic Principle to Imaging with Scattered Wavefronts." *arXiv preprint arXiv:1912.11438* (2019).