Confocal Non-line-of-sight Imaging

Matthew O'Toole
Stanford University

David B. Lindell
Stanford University

Gordon Wetzstein
Stanford University

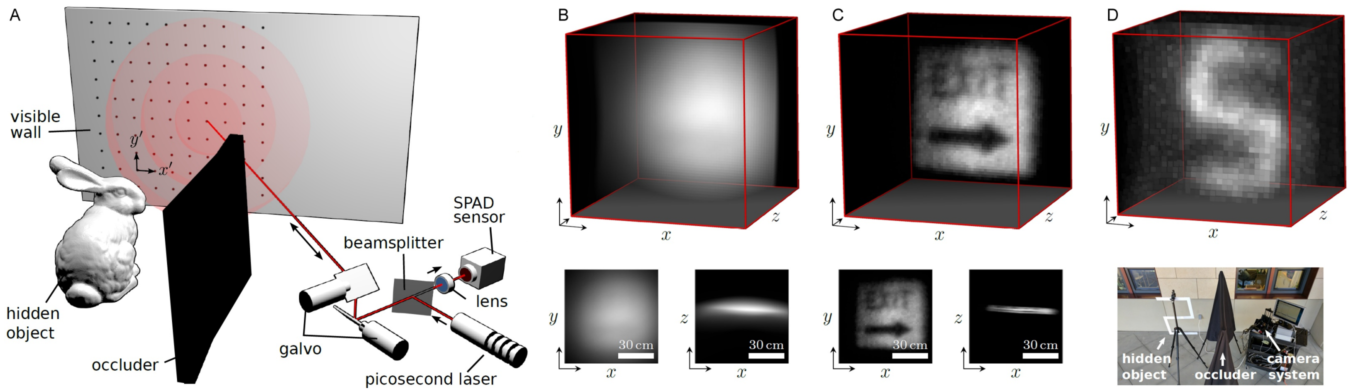

Figure 1: (A) A confocal non-line-of-sight (NLOS) imaging system [O'Toole et al. 2018] comprises a single-photon detector, a pulsed laser, and a galvo mirror to acquire time-resolved measurements of a visible surface. These measurements also contain indirect reflections of hidden objects that are used to recover 3D shape. Experimental results of confocal non-line-of-sight imaging: (B) result of using the backprojection method and (C) result of using the proposed Light Cone Transform reconstruction procedure. (D) Confocal NLOS can also reconstruct the shape and albedo of objects under indirect sunlight.

ABSTRACT

Non-line-of-sight (NLOS) imaging aims at recovering the shape of objects hidden outside the direct line of sight of a camera. In this work, we report on a new approach for acquiring time-resolved measurements that are suitable for NLOS imaging. The system uses a confocalized single-photon detector and pulsed laser. As opposed to previously-proposed NLOS imaging systems, our setup is very similar to LIDAR systems used for autonomous vehicles and it facilitates a closed-form solution of the associated inverse problem, which we derive in this work. This algorithm, dubbed the Light Cone Transform, is three orders of magnitude faster and more memory efficient than existing methods. We demonstrate experimental results for indoor and outdoor scenes captured and reconstructed with the proposed confocal NLOS imaging system.

CCS CONCEPTS

• **Computing methodologies** → **Computational photography**; **3D imaging**;

KEYWORDS

computational photography, single-photon sensors, time-of-flight imaging, non-line-of-sight imaging

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
SIGGRAPH '18 Talks, August 12-16, 2018, Vancouver, BC, Canada
© 2018 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-5820-0/18/08…$15.00
https://doi.org/10.1145/3214745.3214795

ACM Reference format:
Matthew O'Toole, David B. Lindell, and Gordon Wetzstein. 2018. Confocal Non-line-of-sight Imaging. In *Proceedings of SIGGRAPH '18 Talks, Vancouver, BC, Canada, August 12-16, 2018,* 2 pages.
https://doi.org/10.1145/3214745.3214795

1 INTRODUCTION

Imaging objects that are outside the direct line of sight of a camera is a challenging, but important, problem. Emerging non-line-of-sight (NLOS) imaging technology could enable new capabilities for a wide range of applications, including in robotic vision, remote sensing, medical imaging, defense, and autonomous vehicles. A NLOS imaging system typically uses a pulsed laser and a time-resolved detector to measure the time of flight of the light scattered back from the scene to the detector. Whereas light detection and ranging (LIDAR) systems use such measurements to recover the shape of visible objects from direct reflections, NLOS imaging aims at reconstructing the shape and albedo of hidden objects from multiply scattered light [Buttafava et al. 2015; Gupta et al. 2012; Heide et al. 2014; Kirmani et al. 2009; Velten et al. 2012; Wu et al. 2012].

Although much progress on advancing NLOS imaging has been made throughout the last few years, this has remained an impractical idea for several reasons. First, the memory and processing requirements of NLOS reconstruction algorithms are prohibitive. Second, the extremely weak signal of multiply scattered light is extremely challenging to detect. In this work, we report on a confocal scanning procedure [O'Toole et al. 2018] that provides a means to address these key challenges (see Fig. 1, left). Confocal scanning facilitates the derivation of a novel closed-form solution to the NLOS reconstruction problem, which requires orders of magnitude less computation and memory than previous reconstruction methods and recovers hidden objects at unprecedented image resolutions. Confocal scanning also uniquely benefits from a sizeable increase in signal and range when imaging retroreflective objects.

2 THE LIGHT CONE TRANSFORM (LCT)

Confocal NLOS (C-NLOS) measurements consist of a 2D set of temporal histograms, acquired by scanning points x', y' on a planar wall at position $z' = 0$. This 3D volume of measurements, or transient image [O'Toole et al. 2017], τ, is given by

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

where c is the speed of light. Every measurement sample $\tau(x', y', t)$ captures the photon flux at point x', y' and time t relative to an incident pulse scattered by the same point at time $t = 0$. Here, the function ρ is the sought-after albedo of the hidden scene at each point in the three-dimensional half-space Ω satisfying $z > 0$.

A remarkable and unique property of Equation (1) is the fact that a change of variables in the integral by $z = \sqrt{u}$, $\frac{dz}{du} = \frac{1}{2\sqrt{u}}$, and $v = \left(\frac{tc}{2}\right)^2$ results in the following formula

$$\underbrace{v^{\frac{3}{2}}\,\tau\left(x', y', \frac{2}{c}\sqrt{v}\right)}_{\mathcal{R}_t\{\tau\}(x',y',v)} = \iiint_\Omega \underbrace{\frac{1}{2\sqrt{u}}\,\rho\left(x, y, \sqrt{u}\right)}_{\mathcal{R}_z\{\rho\}(x,y,u)} \quad (2)$$
$$\underbrace{\delta\left((x'-x)^2 + (y'-y)^2 + u - v\right)}_{h(x'-x,\,y'-y,\,v-u)}\,dx\,dy\,du.$$

This can be expressed as a straightforward 3D convolution where $\mathcal{R}_t\{\tau\} = h * \mathcal{R}_z\{\rho\}$. Here, the function h is a shift-invariant 3D convolution kernel, the transform $\mathcal{R}_z\{\cdot\}$ nonuniformly resamples and attenuates the elements of volume ρ along the z-axis, and the transform $\mathcal{R}_t\{\cdot\}$ nonuniformly resamples and attenuates the measurements τ along the time axis. The inverses of both transforms $\mathcal{R}_z\{\cdot\}$ and $\mathcal{R}_t\{\cdot\}$ also have closed-form expressions. We denote this formulation the *light cone transform* (LCT).

The image formation model can be discretized as $\mathbf{R}_t\boldsymbol{\tau} = \mathbf{H}\mathbf{R}_z\boldsymbol{\rho}$, and inverted using a closed-form solution that is computed with a Wiener filter

$$\boldsymbol{\rho}_* = \mathbf{R}_z^{-1}\mathbf{F}^{-1}\left[\frac{1}{\widehat{\mathbf{H}}} \cdot \frac{|\widehat{\mathbf{H}}|^2}{|\widehat{\mathbf{H}}|^2 + \frac{1}{\mathrm{SNR}}}\right]\mathbf{F}\mathbf{R}_t\boldsymbol{\tau}, \quad (3)$$

where $\mathbf{F}$ is the 3D discrete Fourier transform, $\boldsymbol{\rho}_*$ is the estimated volume of hidden surface albedos, $\widehat{\mathbf{H}}$ is a diagonal matrix containing the Fourier coefficients of the 3D convolution kernel, and SNR represents the frequency-dependent signal-to-noise ratio of the measurements.

3 ASSESSMENT

Solving the NLOS problem with the LCT involves evaluating Equation (3) in three steps: (i) resampling and attenuating the measurements $\boldsymbol{\tau}$ with the transform $\mathbf{R}_t$, (ii) applying the Wiener filter to the result, and (iii) applying the inverse transform $\mathbf{R}_z^{-1}$ to recover $\boldsymbol{\rho}$. These three steps can be efficiently evaluated in terms of memory and number of operations required. The most expensive step occurs when applying the Wiener filter, which requires $O(N^3 \log N)$ operations for the 3D fast Fourier transforms and which has a memory footprint of $O(N^3)$, where N is the maximum number of elements across all dimensions in space-time. In comparison, existing backprojection-type reconstructions [Buttafava et al. 2015; Gupta et al. 2012; Velten et al. 2012] require $O(N^5)$ operations, and methods based on inversion are significantly more expensive both in their memory and processing requirements [Gupta et al. 2012; Heide et al. 2014; Wu et al. 2012]. The LCT is about three orders of magnitude faster and more memory efficient than existing methods that use backprojection-type [Buttafava et al. 2015; Gupta et al. 2012; Velten et al. 2012] or iterative inverse methods [Gupta et al. 2012; Heide et al. 2014; Wu et al. 2012].

Bringing NLOS imaging outdoors requires detecting indirect light of the hidden object in the presence of strong ambient illumination. To accomplish this, confocal NLOS takes advantage of the high light throughput associated with retroreflective objects. Figure 1 (right) presents an outdoor NLOS experiment under indirect sunlight (approx. 100 lux). The dimensions of the hidden retroreflective object are 0.76×0.51 m with 32×32 sampled locations over a 1×1 m area. The exposure is 0.1 s per sample, with a total exposure time of 1.7 min. MATLAB reconstructs a $32 \times 32 \times 1024$ volume in 0.5 s. We also show a comparison of a conventional backprojection-type reconstruction and the LCT for a retroreflective traffic sign captured indoors. Refer to O'Toole et al. [2018] for additional results.

4 DISCUSSION

We demonstrate that the co-design of a confocal scanning technique and a computation-efficient inverse method facilitates fast, high-quality reconstructions of hidden objects. The proposed techniques enable NLOS imaging with commodity hardware at significantly faster speeds, with lower memory footprint, with less power, over a longer range, under ambient lighting, and at higher resolution than any existing approach.

REFERENCES

Mauro Buttafava, Jessica Zeman, Alberto Tosi, Kevin Eliceiri, and Andreas Velten. 2015. Non-line-of-sight imaging using a time-gated single photon avalanche diode. *Opt. Express* 23, 16 (2015), 20997–21011.

Otkrist Gupta, Thomas Willwacher, Andreas Velten, Ashok Veeraraghavan, and Ramesh Raskar. 2012. Reconstruction of hidden 3D shapes using diffuse reflections. *Opt. Express* 20, 17 (2012), 19096–19108.

Felix Heide, Lei Xiao, Wolfgang Heidrich, and Matthias B. Hullin. 2014. Diffuse mirrors: 3D reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors. *Proc. CVPR* (2014), 3222–3229.

Ahmed Kirmani, Tylor Hutchison, James Davis, and Ramesh Raskar. 2009. Looking around the corner using transient imaging. *Proc. ICCV* (2009), 159–166.

Matthew O'Toole, Felix Heide, David B. Lindell, Kai Zang, Steven Diamond, and Gordon Wetzstein. 2017. Reconstructing transient images from single-photon sensors. *Proc. CVPR* (2017).

Matthew O'Toole, David B. Lindell, and Gordon Wetzstein. 2018. Confocal non-line-of-sight imaging based on the light-cone transform. *Nature* 555 (2018), 338–341.

Andreas Velten, Thomas Willwacher, Otkrist Gupta, Ashok Veeraraghavan, Moungi G. Bawendi, and Ramesh Raskar. 2012. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Communications* 3 (2012).

Di Wu, Gordon Wetzstein, Christopher Barsi, Thomas Willwacher, Matthew O'Toole, Nikhil Naik, Qionghai Dai, Kyros Kutulakos, and Ramesh Raskar. 2012. Frequency analysis of transient light transport with applications in bare sensor imaging. *Proc. ECCV* (2012), 542–555.