



# Compressed sensing for active non-line-of-sight imaging

JUN-TIAN YE,<sup>1,2,3</sup> XIN HUANG,<sup>1,2,3</sup> ZHENG-PING LI,<sup>1,2,3</sup> AND FEIHU XU<sup>1,2,3,\*</sup>

<sup>1</sup>Hefei National Laboratory for Physical Sciences at Microscale and Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China

<sup>2</sup>Shanghai Branch, CAS Center for Excellence in Quantum Information and Quantum Physics, University of Science and Technology of China, Shanghai 201315, China

<sup>3</sup>Shanghai Research Center for Quantum Sciences, Shanghai 201315, China

\*e-mail: feihuxu@ustc.edu.cn

**Abstract:** Non-line-of-sight (NLOS) imaging techniques have the ability to look around corners, which is of growing interest for diverse applications. We explore compressed sensing in active NLOS imaging and show that compressed sensing can greatly reduce the required number of scanning points without the compromise of the imaging quality. Particularly, we perform the analysis for both confocal NLOS imaging and active occlusion-based periscopy. In experiment, we demonstrate confocal NLOS imaging with only  $5 \times 5$  scanning points for reconstructing a three-dimensional hidden image which has  $64 \times 64$  spatial resolution. The results show that compressed sensing can reduce the scanning points and the total capture time, while keeping the imaging quality. This will be desirable for high speed NLOS applications.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

## 1. Introduction

Non-line-of-sight (NLOS) imaging refers to reconstructing the image of scene hidden by occluder from both the source and the detector. Using the reflection from the surface of the intermediate wall, it is possible to obtain the unknown information about the hidden object [1,2]. NLOS imaging technology has a wide range of applications in medicine, manufacturing, transportation, public safety, and basic science. Recently, NLOS imaging has received significant attention with solutions that operate on a wide variety of different principles, such as speckle correlations [3–5], wavefront-shaping [6], thermal imaging [7] and transient techniques [8–11].

In particular, the transient techniques have been the most popular research field which has resulted in several important approaches, such as back projection [9,12], light-cone transformation (LCT) [13], Fermat paths [14], temporal focusing [15], phasor-field [16], f-k migration [17] and so forth [18,19]. Transient-based NLOS imaging normally uses pulsed light sources and single-photon counting sensors to measure the time-of-flight information encoded in the multiple scattered light paths, where the hidden scene can be reconstructed using some demultiplexing approaches [9,13,16,20]. Moreover, the occlusion-based imaging technique has been proposed in recent years [21,22], which uses occlusion to undo the mixing of diffuse light. This technique obviates the need for collecting time-of-flight measurements, and it is possible with a single intensity photograph which can be captured with an ordinary digital camera [23].

Existing NLOS techniques generally require dense raster-scanning across the visible wall, which is not favorable for real-time acquisition. Our interest is to realize NLOS imaging with much fewer scanning points, thus reducing the acquisition time. So how many scanning points are capable for NLOS reconstruction without the compromise of image resolution? For scanning line-of-sight imaging [24,25], one needs to raster scan a  $N \times N$  grid to image a scene with  $N \times N$  resolution. However, in transient-based NLOS case, the measured data for each scanning point is related to *all* voxels in the hidden space [26–28]. Hence, it is expected that through the fine

measurements in temporal dimension, with proper reconstruction algorithms, the recovery of high-resolution hidden images is possible with a small number of scanning points in NLOS imaging. Particularly, the compressed sensing techniques [29–32] integrate the compressing and sensing simultaneously through exploiting diverse sparse features of the scene, which can significantly reduce the required number of samples without the compromise of the imaging quality.

Some related works have been done recently. Reference [33] has studied the behavior of the reconstruct when randomly discarding measurement points; Ref. [34] has shown that a circular scan of transients is sufficient to reconstruct a NLOS image. Moreover, for exploiting diverse sparse features of the scene, deep learning has demonstrated a powerful set of techniques to enforce the learned priors in NLOS imaging [35].

In this work, we investigate compressed sensing approaches in NLOS imaging. We show that the transient confocal NLOS measurements with only  $5 \times 5$  scanning points are sufficient for reconstructing 3D hidden images that have  $64 \times 64$  spatial resolution. For occlusion-based NLOS imaging, we can achieve a compression rate of 6.25% while maintaining the image quality. Reference [13] has proved in simulations that longer exposure times for fewer scanning samples would be a preferred trade off to reduce the acquisition time. Here, we evaluate the image quality under the different number of scanning points and photon counts. Through both simulation and experiment, we show that using compressed sensing, the scan points can be significantly reduced without the compromise of image quality. Consequently, the total acquisition time can be greatly reduced.

## 2. Imaging scenario

Most existing NLOS imaging techniques use a generic setting, where a focused scanning source illuminates a visible wall, and a detector captures the reflected photons on the same wall. We first review the general NLOS imaging setup. As shown in Fig. 1(a), there are a total of three diffuse reflections in the NLOS imaging. First, the light emitted by the laser passes through the scanning galvanometer and hits the illumination point  $l$ . Next, the light diffuses at the illumination point as the first reflection and propagates to the target  $x$ . Then, the photons reflect on the target and propagate to the detection point  $p$ . Finally, the photons diffuse at the detection point and enter the single-pixel detector. Mathematically, NLOS imaging can be formulated as the forward model shown in Fig. 1(b).

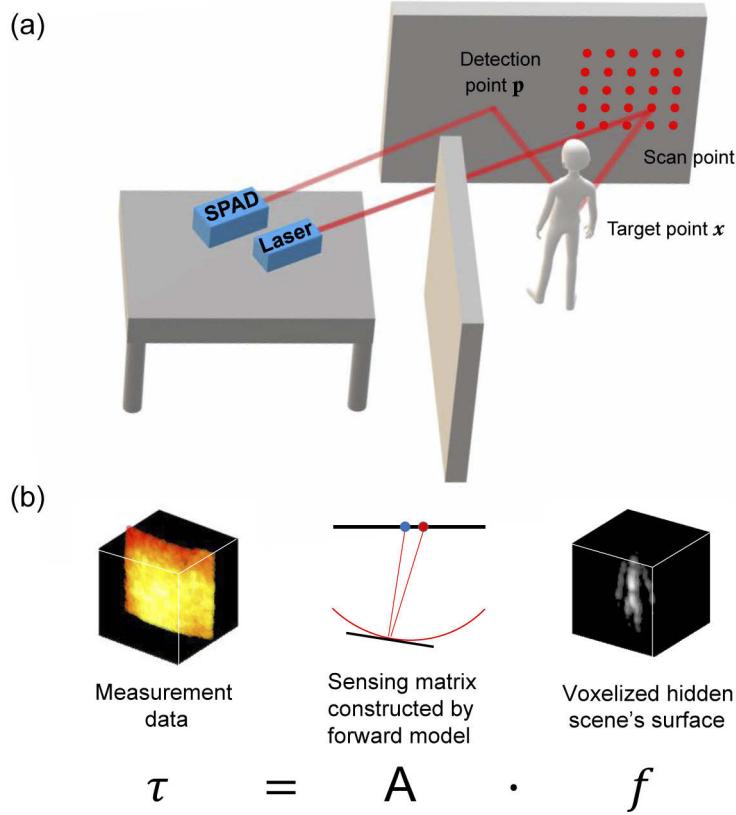
In this work, we perform the analysis for both confocal NLOS method [13] and occlusion-based method [23]. The sensing matrices are constructed according to the light transport model of both NLOS methods. The detailed forward models are described in Appendices.

## 3. Compressed sensing

Compressed sensing technology has been proved to be practical and robust under various imaging architectures, such as remote sensing [36,37], infrared imaging [38], and magnetic resonance (MR) imaging [39]. Here we introduce compressed sensing to NLOS imaging. The theory of compressed sensing tells us that the unknown signal can be fully recovered with a small number of samples if two conditions satisfied [32]: the first is that the signal has a sparse representation in some proper basis, and the second is that the sensing matrix is incoherent with the representation basis.

NLOS imaging well meets the conditions of compressed sensing. Firstly, only the surface of the scenes are being sensed (see Fig. 1(b)), which is sparse in the whole 3D volume. Moreover, the NLOS sensing matrix is incoherent (see forward model in Appendices). As a result of compressed sensing [40], the ill condition inverse problem can be solved by a convex program:

$$\min \|f\|_{\ell_1} \text{ subject to } \|Af - \tau\|_{\ell_2} \leq \epsilon, \quad (1)$$



**Fig. 1. three-bounce light trajectory of NLOS imaging.** (a) A focused scanning source illuminates the visible wall at  $l$  and is diffusively reflected (first bounce) toward the hidden scene. Then the light is reflected (second bounce) at target point  $x$  and propagates to the detection point  $p$  (third bounce). Finally, the detector measures the light intensity of the detection point. (b) The voxelized hidden scene  $f$  can be mapped to the measurement result  $\tau$  through a measurement matrix  $A$ .

where  $f$  is the sparse target to be solved, and the  $\ell_1$  norm  $\|\cdot\|_{\ell_1}$  is the sparsity-promoting function used to recover sparse solutions. For reconstruction, the required number of sampling points is only a few times the number of target's non-zero elements, and when the number of samples is sufficient, reducing the sampling number has little influence on reconstruction performance [41]. In this work, we show that for NLOS imaging, the required scan points can be significantly reduced by using the compressed sensing algorithm, and we discuss the influence of Poisson noise on the compression ratio.

#### 4. Algorithm

The compressed sensing algorithm we used for image reconstruction is modified from SPIRAL-TAP [42]. The 3D deconvolution can be solved from the convex optimization program:

$$\hat{f} = \arg \min_{f: f \geq 0} \{-\log P_r(\tau; f) + \lambda \text{pen}(f)\}, \quad (2)$$

where  $f$  represents the voxelized volume of target scenes being sensed (see Fig. 1(b)). The first term is the negative log-likelihood function, which represents the likelihood of the estimated  $\hat{f}$  according to the observed data. We carefully choose the value of penalty parameter  $\lambda$  in

reconstruction (about 0.1-2). A too-small  $\lambda$  will cause over-fitting and noisy results; a too-large  $\lambda$  will cause over-smoothing and blurry results. Since emitted light only interacts with the surface of the hidden scene and the sensed surface is sparse in the whole 3D volume, we use canonical basis to be the sparse basis.

If the noise type in measurements is Gaussian, then the negative log-likelihood function is equivalent to

$$-\log Pr(\tau; f) = \|\tau - \mathbf{DA}f\|_2^2, \quad (3)$$

where  $A$  represents the sensing matrix and the matrix  $D$  represents the downsample operation. This operation is that, if the scanning point is sampled, then the corresponding NLOS measurement data will be retained; otherwise the measurement data will be set to zero.

For Poisson noise, the probability model is:

$$Pr(\tau; f) = \prod_i \frac{(e_i^T \mathbf{DA}f)^{\tau_i}}{\tau_i!} \exp(-e_i^T \mathbf{DA}f), \quad (4)$$

SPIRAL-TAP employs sequential quadratic approximations of the objective function. We get the sub-solution for the next iteration by computing the gradient and Hessian in each iteration. For comparison, we also process the data with various NLOS reconstruction algorithms proposed in recent years.

## 5. Simulation of confocal NLOS imaging

In this section, we simulate the compressed sensing performance in confocal NLOS imaging, and we analyze how scan points can be reduced without compromise of resolution.

### 5.1. Simulation settings

Recent years have witnessed the development of NLOS transient renderer, allowing for real-time runtime [35,43–45]. Here for simplicity, we use the LCT convolution model directly for simulation. The scan pattern used in simulation is a square grid as shown in Fig. 1(a). We use a simple "bowling" scene as hidden objects. The hidden scene was contained within a 1 m × 1 m × 1.2 m volume. Simulated data is processed at a spatial resolution of 64 × 64 and a temporal resolution of 256. Each temporal bin spans 32 ps. The scanning area spans 1 m × 1 m. The full width at half maximum (FWHM) of system jitter is set as 60 ps. We simulate the distance attenuation of light intensity and the Poisson process of photon detection; the average number of photons collected per scan point is approximately 10 K.

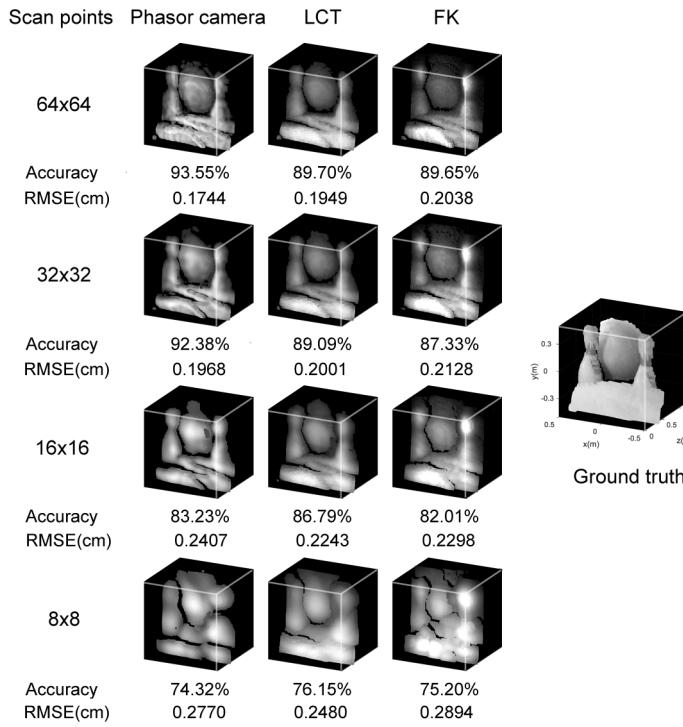
### 5.2. Baseline results

We simulate the NLOS imaging with the number of sampling points set as 64 × 64, 32 × 32, 16 × 16, 8 × 8. We upsample the sparse measurements using nearest neighbor interpolation before applying the algorithms. The reconstruction results are shown in Fig. 2. We use the foreground/background classification accuracy to evaluate the reconstruction performance:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad (5)$$

The classification combines rates for missing geometry (false negatives/FN), excess (false positives/FP), correct foreground (true positives/TP) and background (true negatives/TN). This criterion is also used in other NLOS literatures [20]. In addition, we use the root-mean-square error(RMSE) to evaluate the depth error for the foreground region.

From Fig. 2 we can see that when the scan points reduce to 32 × 32, the quality of the reconstruction results is not significantly spoiled. That is due to the characteristic of transient-based NLOS imaging—the high resolution sampling in time can compensate for the low resolution



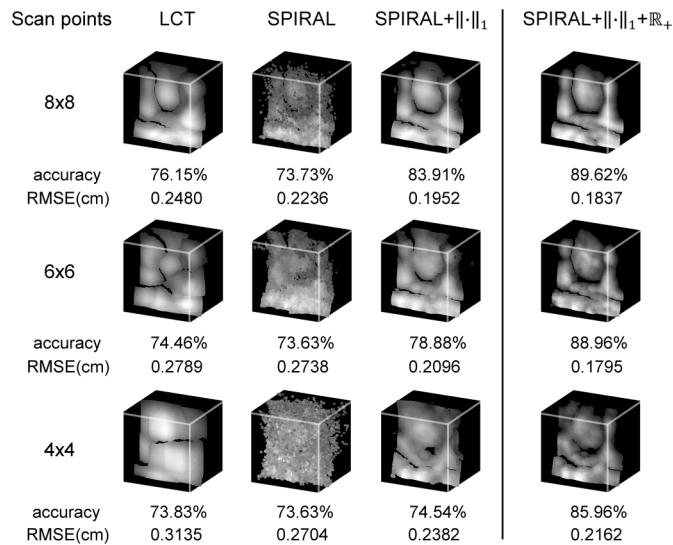
**Fig. 2. Reconstructions with various scan points for confocal NLOS simulation.** The hidden bowling scene is processed at a  $64 \times 64$  spatial resolution and 256 temporal resolution. We fix the number of photons per scan point at about 10 K. The sparse measurements are upsampled using nearest neighbor interpolation before applying the algorithms LCT, FK migration, and phasor camera. The reconstruction quality is clearly degraded when scan points reduce to  $16 \times 16$ .

sampling in space. However, with the scanning point further reducing, the resolution of the result decreases obviously. For NLOS imaging, if we want to reconstruct the hidden scene with  $n_x \times n_y \times n_z$  resolution from the  $n_s \times n_t$  measurement, the sensing matrix is a  $n_s n_t \times n_x n_y n_z$  matrix. This inverse problem can be ill-posed if the number of scan points is not sufficient. So the image quality degrades as the sampling point reduces.

### 5.3. Compressed sensing results

We use the regularized optimization algorithm SPIRAL with  $\ell_1$  norm as regularizer to realize compressed sensing. And in order to speed up the convergence, we delete some elements with small value to 0 after each iteration. We simulate the NLOS imaging with the number of sampling points set as  $8 \times 8, 6 \times 6, 4 \times 4$ . The reconstruction results are shown in Fig. 3. From the results we can see that the image quality is greatly improved due to the effect of the penalty term. This term utilizes the sparse feature of the hidden scene, and finally the scene is recovered with  $8 \times 8$  scan points. Note that the limit of scan points is related to the sparsity of hidden scene. In this case, the bowling scene has about 3000 non-zero elements, while we have 16384 measurements ( $8 \times 8$  scan points with a temporal resolution of 256). Therefore, we can almost reconstruct the scene. As the number of the scan points is further reduced to  $6 \times 6$ , the imaging quality is obviously degraded.

In addition to this, the non-negativity of reflectivity is also useful information for solving the ill-posed inverse problem [46]. We can see from Fig. 3 that the imaging quality is further



**Fig. 3. 3D volume reconstruction results with few scan points.** While the LCT algorithm and the optimizer SPIRAL are failed to recover the bowling scene at  $64 \times 64$  spatial resolution with  $8 \times 8$  scan points, the regularized optimizer SPIRAL+ $\ell_1$  successfully reconstructs the scene. Moreover, by introducing the non-negative constraints, the hidden image can be recovered with  $6 \times 6$  scan points.

improved by introducing the non-negative constraint into algorithm SPIRAL. We expect the reason is that the non-negative constraint is effective for background classification, and thus greatly reduces the number of unknown elements. We also simulate the compressed sensing reconstruction under different bin size condition (see Appendices for more details). The results show that higher time resolution leads to better imaging quality.

The simulation results clearly show the benefit of the compressed sensing techniques under subsampled conditions. Compared with the original  $64 \times 64$  scan points, we reconstruct the hidden scene with only  $6 \times 6$  scan points, which reduces the acquisition time by 114 times.

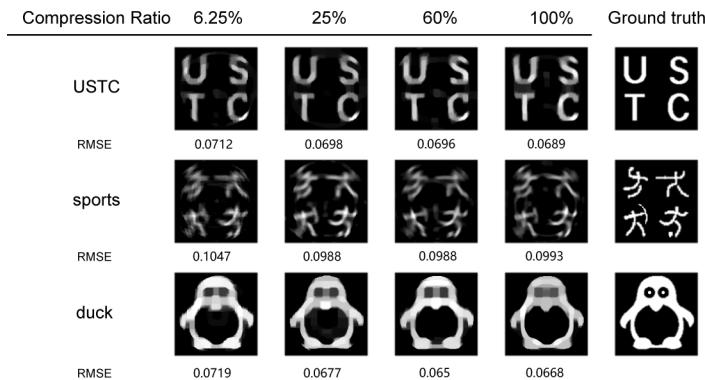
## 6. Simulation of active occlusion-based NLOS imaging

We also performed a numerical simulation to verify compressed sensing reconstruction performance in occlusion-based NLOS imaging systems. We use a 2D grid scan pattern in simulation. The sensing matrix  $A$  is constructed according to the forward model in Appendices. The hidden scenes are three simple patterns and are discretized to  $128 \times 128$  pixels, as Fig. 4 shows.

In the simulation, the measured photon counts are drawn randomly from the binomial distribution. The number of scan points is set as  $32 \times 32$ ,  $64 \times 64$ ,  $100 \times 100$ ,  $128 \times 128$ . We allow the reconstruction algorithm to run enough time to ensure approximate convergence. The reconstruction is performed with the compression ratio ranging from 6.25% to 100%. The RMSE is averaged over 40 repeated simulations considering the random process. For each reconstruction, the regularization parameter of the algorithm is optimized. Figure 4 shows the reconstruction examples of three scenes with different compression ratios.

It is evident from Fig. 4 that the reconstruction fidelity does not degrade as the number of measurements decrease from  $128 \times 128$  to  $32 \times 32$ , which shows that accurate reconstructions can be obtained with fewer measurements using compressed sensing.

We furthermore investigate the compressed sensing performance under various detected photons per pixel (PPP) (See more details in Appendices). The results show that the reconstruction error



**Fig. 4. Reconstruction results with different compression ratios for occluder-based NLOS simulation.** The three scenes are of  $128 \times 128$  resolution. The number of scan points is set as  $32 \times 32$ ,  $64 \times 64$ ,  $100 \times 100$ ,  $128 \times 128$  corresponding to the compression ratio 6.25%, 25%, 60%, 100%, and the reconstruction fidelity does not degrade as the number of measurements decreases.

will increase when PPP decreases. Moreover, under different photon levels, the reconstruction error does not increase when the compression ratio goes from 100% to 6.25%.

## 7. Experimental demonstration

In order to experimentally validate the performance of compressed sensing in NLOS imaging, we built a confocal NLOS system.

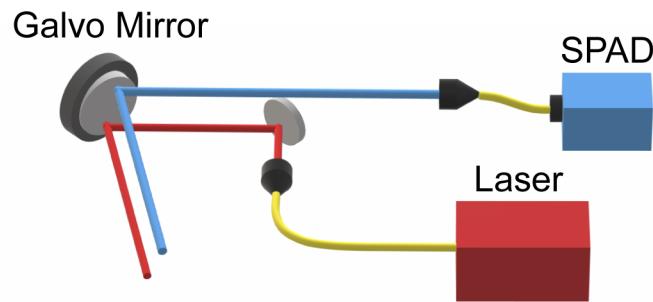
### 7.1. Confocal NLOS imaging setup

Figure 5 shows our experiment NLOS imaging system. A 1560 nm femtosecond laser(MenloSystems C-Fiber) emits pulses at 100 MHz repetition frequency and 20 mW average power. The pulses pass through a collimator, a two-axis raster-scanning galvo mirror, and are sent out to the visible wall. The diffuse photons are collected by the same galvo mirror and coupled to a multimode fiber directed to the detector. The illuminated point and detected point are misaligned in order to reduce the measured intensity of direct reflection light. The detector is a homemade, multimode fiber-coupled free-running InGaAs/InP single-photon avalanche diode (SPAD) detector, which has  $\sim 20\%$  detection efficiency and  $\sim 210$  ps timing jitter. A time-correlated single photon counter or TCSPC (PicoQuant HydraHarp 400) records the photon-detection signals from the SPAD and outputs measurement data to the computer. We raster scan a square grid of points across a  $1.2 \text{ m} \times 1.2 \text{ m}$  area on the visible wall. The NLOS scene is a structure made of white papers, which is placed about 0.7 m away from the visible wall.

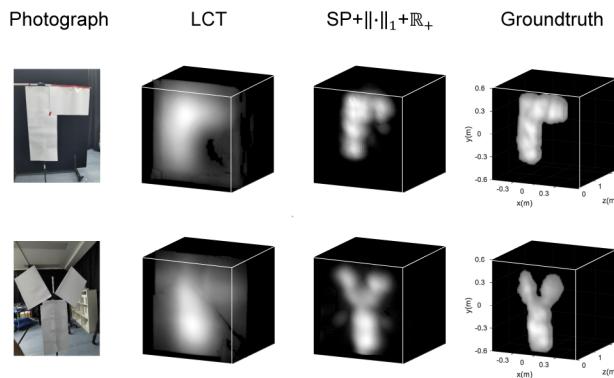
### 7.2. Results

We demonstrate the ability to perform NLOS imaging with only  $5 \times 5$  scan points in Fig. 6. The captured data has 1024 temporal bins with bin size of 16 ps. The background noise is about 10 K per second and the signal photons is about 1 K per second. The signal PPP are about 16 K on average. As we can see in Fig. 6, the compressed sensing algorithm successfully recover the scene while the LCT algorithm failed.

NLOS imaging suffers from Poisson noise. The intensity of Poisson noise depends on the acquired photons number, and thus we always need a certain exposure time to capture enough photons for NLOS imaging. We verify comp reconstruction performance under different photon



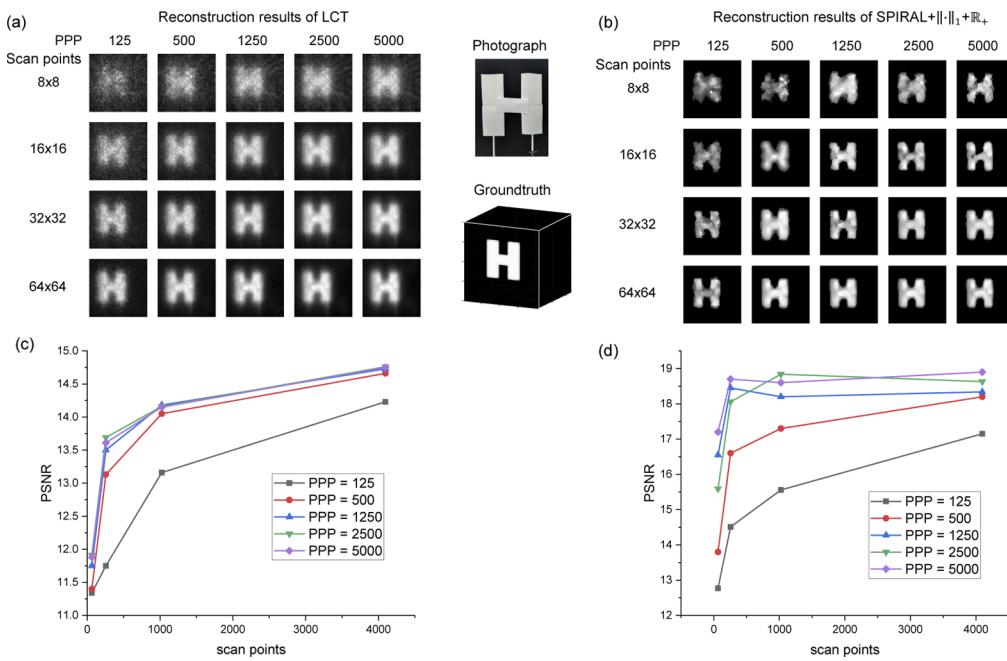
**Fig. 5. Experimental confocal NLOS imaging system.** The pulses emitted by the laser pass through a galvo mirror to scan the scene. Direct and indirect reflected lights are collected by the same galvo mirror and coupled to a multimode fiber, then measured by a single-photon avalanche diode (SPAD). The whole system is operated at 1560 nm with a femtosecond laser and an InGaAs/InP SPAD.



**Fig. 6. Experimental results with  $5 \times 5$  scan points.** The hidden scene is a structure made of white papers processed at  $64 \times 64$  resolution. The groundtruth is generated with long exposure time and  $64 \times 64$  scan points. While the LCT algorithm fails to recover the outline of the scene, the compressed sensing algorithm successfully reconstructs the image of hidden scenes with only  $5 \times 5$  scan points.

counts levels. We change the PPP from 125 to 5000, The results are shown in Fig. 7. We can see that as the acquired number of photons decreases, the reconstruction quality of LCT and SPIRAL both degrade. On the other hand, if we reduce the scan points, for results of LCT, the Peak Signal to Noise Ratio (PSNR) decreases; while for the case of compressed sensing, PSNR remains almost unchanged as long as there are sufficient photon counts.

If we want to reduce the acquisition time of NLOS imaging, would it be better to reduce the exposure time with fixed number of scan points, or to subsample the wall while keeping the exposure time per scan fixed? As we can see from Fig. 7, for compressed sensing, our results suggest that it is beneficial to reduce the number of samples rather than the exposure time per scan point. This conclusion is also verified and proved by some other works [13,47]. An intuitive explanation is as following. If we keep the total exposure time constant while reducing the number of scan points, each scan point will have longer acquisition time and thus more signal photon counts, which helps the reconstruction.

**Fig. 7. Experimental results with different number of scan points and photon counts.**

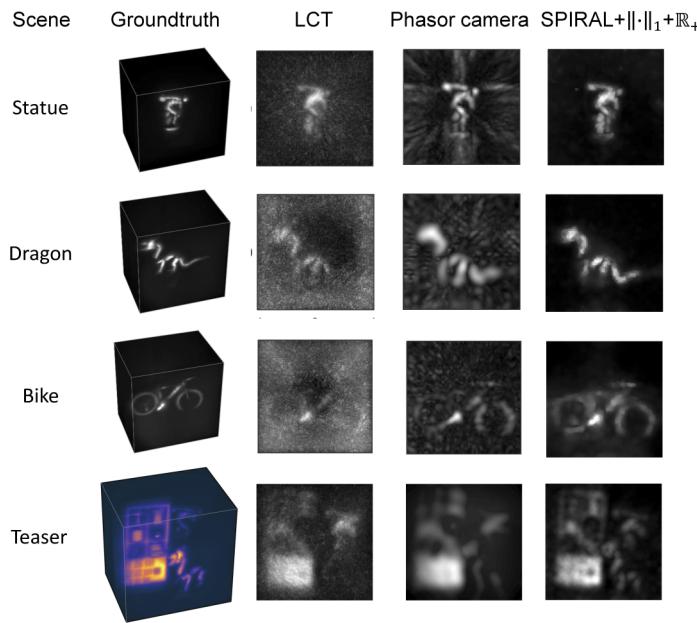
The hidden scene is a letter "H" processed at  $64 \times 64$  resolution, we capture confocal NLOS measurement data with varying scan points  $8 \times 8, 16 \times 16, 32 \times 32, 64 \times 64$  and PPP 125, 500, 1250, 2500, 5000. (a), Reconstruction results using LCT algorithm. (b), Reconstruction results using compressed sensing algorithm SPIRAL +  $\|\cdot\|_1 + \mathbb{R}_+$ . (c), PSNR of the results using the LCT algorithm, the PSNR decrease as the scan points reduce. (d), PSNR of the results using the compressed sensing algorithm. While PPP is sufficient, PSNR remains almost unchanged as the scan points reduce from  $64 \times 64$  to  $16 \times 16$ . Compared to shorter exposure times with more samples, longer exposure times with fewer samples yields higher PSNR.

## 8. Results for public data

We verified our compressed sensing method using public available NLOS data. The public confocal NLOS data is from the work [17], and we evenly extracted  $16 \times 16$  ( $32 \times 32$  for the bike data) measurement data from  $128 \times 128$  scan points measurements. The result is shown in Fig. 8. We can see from the result that, when LCT and phasor camera algorithms give blurry results, the compressed sensing algorithm yields better reconstruction quality and the objects can still be recognized. And as the scan points reduce to  $16 \times 16$ , the acquisition time can be 64 times faster.

The running time of our algorithm for the sparse "dragon" scene ( $128 \times 128 \times 512$  resolution, 10 iterations) is 192 seconds, while for the complex "teaser" scene ( $128 \times 128 \times 512$  resolution, 40 iterations), the running time is 710 seconds, which is tested on Intel Core i7-10875H @ 2.3GHz CPU. The reconstruction resolution and the sparsity of hidden scenes will greatly affect the running time.

We also apply compressed sensing method to public occluder-based NLOS data set [22] (see Appendices for more details). The results show that the reconstruction fidelity does not degrade as the number of measurements decreases from 10000 to 625, which means using the compression sensing method, the NLOS imaging acquisition can be carried out at a much faster speed than before.



**Fig. 8. Reconstruction results with confocal NLOS public data.** We evenly extracted  $16 \times 16$  ( $32 \times 32$  for the bike data) measurement data from  $128 \times 128$  scan points public dataset as subsampled measurements. The data is processed using LCT, phasor camera, and SPIRAL algorithms, the compressed sensing algorithm shows better reconstruction results.

## 9. Conclusion

We focus on the problem of reconstructing the hidden scenes from fewer measurements. We investigate the compressed sensing technique in NLOS imaging. By simulation, we show that compressed sensing can handle the ill-posed inverse problem in reconstruction, and reduce the requirement for scan points. We also show the benefit to introduce non-negative constrain in the algorithm. By experiment, We demonstrate that confocal measurements with only  $5 \times 5$  sampling points are sufficient for reconstructing 3D hidden scenes with  $64 \times 64$  spatial resolution. Our results suggest that, compared to densely scanning across the wall, longer exposure times with fewer samples yields better imaging quality and thus would be a preferred tradeoff when trying to minimize acquisition time. Using compressed sensing, we can reduce the scan points without the obvious compromise of imaging quality, and the acquisition time of NLOS imaging can be greatly reduced. In future, by combining compressive sensing and line or circular scanning [34], the imaging speed can be greatly enhanced for fast real-time capture of NLOS scenes.

## Appendices

### 9.1. Forward model for confocal NLOS imaging

For confocal NLOS imaging [13], the illuminated point  $l$  and detect point  $p$  are co-located where  $l = p$ . The detector then measures the transient response,  $\tau(l, t)$ , which represents the amount of light detected at scan point  $l$  at time  $t$ . Here  $t$  represents the time interval between the first reflection and the third reflection. The laser raster scans a regular 2D grid of points on the wall. The image formation model can be expressed as:

$$\tau(\mathbf{l}, t) = \iiint_{\Omega} \rho(\mathbf{x}) \frac{\delta(2\|\mathbf{l} - \mathbf{x}\| - tc)}{\|\mathbf{l} - \mathbf{x}\|^4} d\mathbf{x}, \quad (6)$$

As discussed by O'Toole et al. [13], with a change of variables by  $v = (tc/2)^2$ , the integration can be expressed as a straightforward 3D convolution

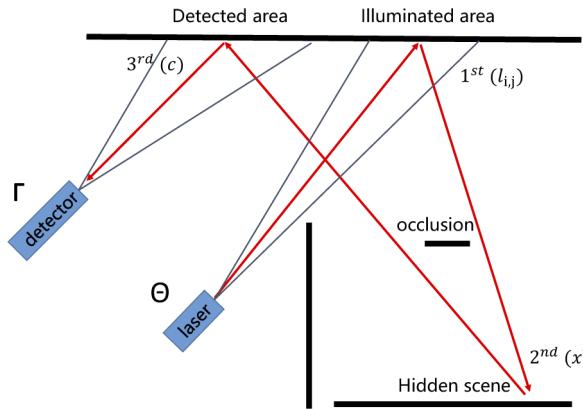
$$\mathcal{R}_t\{\tau(x, y, v)\} = h * \mathcal{R}_z\{\rho(x, y, z^2)\}, \quad (7)$$

where  $h(x, y, v) = \delta(x^2 + y^2 - v)$ , and the equation above is called light cone transform (LCT).

### 9.2. Forward model for active occlusion-based NLOS imaging

Here we introduce the forward model of active occlusion-based NLOS imaging developed in [21,22].

As shown in Fig. 9, there are a total of three diffuse reflections in the occluder-based NLOS imaging light path. During the imaging process, the detector's field of view and the position of the obstacle remains the same. The laser scans a square grid on the visible wall surface, and the occluder blocks part of the light from the first and second diffuse reflections. At each scanning point, the detector measures photon counts without time-of-flight information.



**Fig. 9.** Geometry for occluder-based NLOS imaging.

Discretize the illumination area into  $N \times N$  grid points, and use  $i, j$  to represent the horizontal and vertical indicators of these grid points. Let  $\Theta, l_{ij}, x, c, \Gamma$  correspond to the position vector of the laser, illumination area, hidden target, detection area, and single-pixel detector area. Record the reflectivity of the target area as  $f(x)$ , and the measurement result as  $Y_{ij}$ .

The effect of diffuse reflection can be expressed by the Lambertian scattering function (BRDFs)  $G_{\Theta, l_{ij}, x, c, \Gamma}$ , and the influence of obstacles can be introduced into the visual function  $V(l_{ij}, x, c)$  to describe whether the propagation of light is obscured by obstacles. We can get the relationship between the measurement result  $Y$  and the target reflectance  $f$ :

$$Y_{ij} = \int_S dx \int_C dc \int_D d\Gamma \frac{f(x)V(l_{ij}, x, c)G_{\Theta, l_{ij}, x, c, \Gamma}}{\|l_{ij} - x\|^2 \|x - c\|^2 \|c - \Gamma\|^2}, \quad (8)$$

The squared term in the denominator indicates the square of the light intensity as a function of the propagation distance. For a specific grid point  $(i, j)$  on the illuminated area, define the corresponding row of the system sensing matrix  $A$  as:

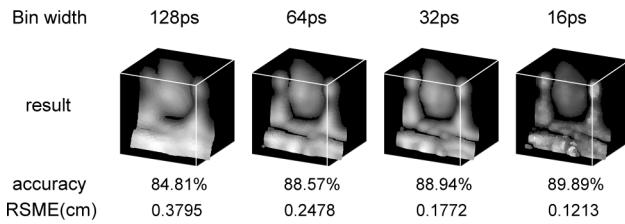
$$A_{ij}(x) = \int_C dc \int_D d\Gamma \frac{V(l_{ij}, x, c)G_{\Theta, l_{ij}, x, c, \Gamma}}{\|l_{ij} - x\|^2 \|x - c\|^2 \|c - \Gamma\|^2}, \quad (9)$$

We can have a discretization representation of the optical forward model by discretizing the target region, the model can be written in matrix form:

$$Y^{ij \times 1} = A^{ij \times kl} f^{kl \times 1}, \quad (10)$$

### 9.3. Compressive sensing for confocal NLOS imaging under different time-bin sizes

In confocal NLOS imaging [13], the measurement data contains temporal information. To evaluate the influence of temporal resolution for compressed sensing, we generate the simulated measurement data with the time-bin size of 128 ps, 64 ps, 32 ps, 16 ps. We fix the scan points to  $6 \times 6$ , and for simplicity, the Poisson noise is not added. The hidden scene is placed within a  $1 \text{ m} \times 1 \text{ m} \times 1.2 \text{ m}$  volume and is discretized at a spatial resolution of  $64 \times 64$ . The reconstruction results are shown in Fig. 10. The reconstruction results indicate that smaller bin size brings better reconstruction results, as expected. Since higher temporal resolution means finer sampling, we have more measurements to solve the ill-posed inverse problem. Thus the sparse hidden scene can be better recovered.



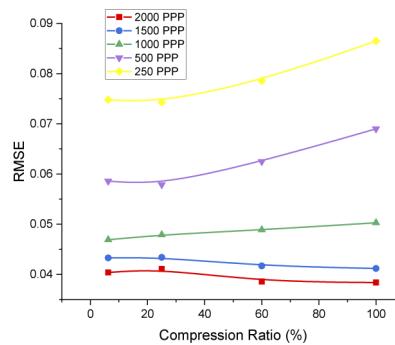
**Fig. 10. Reconstruction results under different bin size condition.** We evaluate the influence of temporal resolution for compressed sensing, the number of scan points is fixed to  $6 \times 6$ . The hidden scene is processed at a spatial resolution of  $64 \times 64$ . For reconstruction results, the accuracy increases while the temporal resolution is smaller, as is expected.

### 9.4. Compressed sensing for occlusion-based NLOS imaging under various photon number conditions

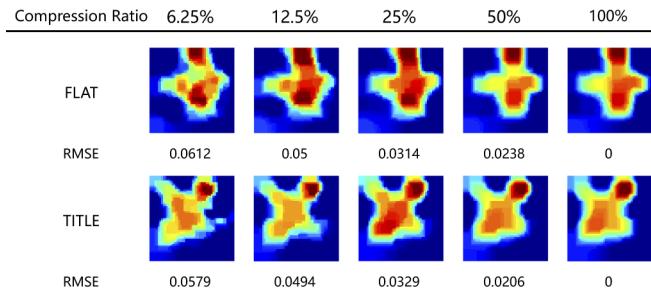
For occluder-based NLOS imaging [23], we investigate compressed sensing performance under various photon number conditions. Since the measurement data  $Y$  contains Poisson noise, we do a simulation to quantify the effect of Poisson noise on compressed sensing performance. By changing the number of emitted pulses from the laser, we generate several sets of measurement data with 250, 500, 1000, 1500, 2000 PPP on average, respectively. For each compression ratio and PPP, the RMSE is computed by averaging over 40 repeated simulations. The simulation results are shown in Fig. 11. And the reconstruction error is calculated under different compression ratios and varying PPP. The results show that the reconstruction error will increase when PPP decreases, but at all photon count levels, the reconstruction error does not increase while the compression ratio goes from 100% to 6.25%.

### 9.5. Results for occlusion-based NLOS public data

We show the application of compressed sensing method using public occluder-based NLOS dataset [22]. The dataset has two measurement data captured with different hidden scenes. Each measurement data has  $100 \times 100$  intensity measurements corresponding to different scan points, and each scan point has about 480 photon counts on average. We evenly extracted  $m = 625, 1250, 2500, 5000$  measurements data from  $100 \times 100$  intensity measurements. Figure 12 shows the recovered images of two hidden scenes with varying compression ratio, and we take the reconstruction result with 100% compression ratio as the groundtruth to compute the RMSE of each recovered image. As we can see from Fig. 12, the reconstruction fidelity does not degrade as the number of measurements decreases from 10000 to 625, which means using the compression sensing method, only 625 scan points are sufficient to reveal the hidden scenes and each of the scan points only has about 480 photon counts on average.

**Fig. 11. Reconstruction error under different compression ratios and varying PPP.**

We evaluate the influence of PPP for compressed sensing, and the RMSE is computed by averaging over 40 repeated simulations. For all PPP conditions, the RMSE does not increase when the compression ratio goes from 100% to 6.25%.

**Fig. 12. Reconstruction results with occluder-based NLOS public data.** For public occluder-based NLOS dataset [22], we evenly extracted  $m = 625, 1250, 2500, 5000$  measurements from  $100 \times 100$  intensity measurements. The data is processed using the compressed sensing algorithm, and we take the reconstruction result with 100% compression ratio as the groundtruth to compute the RMSE of every recovered image. The reconstruction fidelity does not degrade as the number of measurements decreases from 10000 to 625.

**Funding.** National Key Research and Development Program of China (2018YFB0504300); National Natural Science Foundation of China (61771443, 62031024); Shanghai Municipal Science and Technology Major Project 2019SHZDZX01; Special Project for Research and Development in Key areas of Guangdong Province 2020B0303020001; Shanghai Science and Technology Development Foundation 18JC1414700.

**Disclosures.** The authors declare no conflicts of interest.

## References

1. T. Maeda, G. Satat, T. Swedish, L. Sinha, and R. Raskar, “Recent advances in imaging around corners,” arXiv preprint arXiv:1910.05613 (2019).
2. D. Faccio, A. Velten, and G. Wetzstein, “Non-line-of-sight imaging,” *Nat. Rev. Phys.* **2**(6), 318–327 (2020).
3. J. Bertolotti, E. G. Van Putten, C. Blum, A. Lagendijk, W. L. Vos, and A. P. Mosk, “Non-invasive imaging through opaque scattering layers,” *Nature* **491**(7423), 232–234 (2012).
4. O. Katz, P. Heidmann, M. Fink, and S. Gigan, “Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations,” *Nat. Photonics* **8**(10), 784–790 (2014).
5. C. A. Metzler, F. Heide, P. Rangarajan, M. M. Balaji, A. Viswanath, A. Veeraraghavan, and R. G. Baraniuk, “Deep-inverse correlography: towards real-time high-resolution non-line-of-sight imaging,” *Optica* **7**(1), 63–71 (2020).
6. O. Katz, E. Small, and Y. Silberberg, “Looking around corners and through thin turbid layers in real time with scattered incoherent light,” *Nat. Photonics* **6**(8), 549–553 (2012).
7. T. Maeda, Y. Wang, R. Raskar, and A. Kadambi, “Thermal non-line-of-sight imaging,” in *2019 IEEE International Conference on Computational Photography (ICCP)*, (IEEE, 2019), pp. 1–11.

8. A. Kirmani, T. Hutchison, J. Davis, and R. Raskar, "Looking around the corner using transient imaging," in *Proc. 12th IEEE Int. Conf. Computer Vision*, (IEEE, 2009), pp. 159–166.
9. A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," *Nat. Commun.* **3**(1), 745 (2012).
10. M. Buttafava, J. Zeman, A. Tosi, K. Elceiri, and A. Velten, "Non-line-of-sight imaging using a time-gated single photon avalanche diode," *Opt. Express* **23**(16), 20997–21011 (2015).
11. M. O'Toole, F. Heide, D. B. Lindell, K. Zang, S. Diamond, and G. Wetzstein, "Reconstructing transient images from single-photon sensors," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2017), pp. 1539–1547.
12. V. Arellano and D. Gutierrez and A. Jarabo, "Fast back-projection for non-line of sight reconstruction," *Opt. Express* **25**(10), 11574–11583 (2017).
13. M. O'Toole, D. B. Lindell, and G. Wetzstein, "Confocal non-line-of-sight imaging based on the light-cone transform," *Nature* **555**(7696), 338–341 (2018).
14. S. Xin, S. Nousias, K. N. Kutulakos, A. C. Sankaranarayanan, S. G. Narasimhan, and I. Gkioulekas, "A theory of fermat paths for non-line-of-sight shape reconstruction," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2019), pp. 6800–6809.
15. A. Pediredla, A. Dave, and A. Veeraraghavan, "Snlss: Non-line-of-sight scanning through temporal focusing," in *2019 IEEE International Conference on Computational Photography (ICCP)*, (IEEE, 2019), pp. 1–13.
16. X. Liu, I. Guillén, M. La Manna, J. H. Nam, S. A. Reza, T. H. Le, A. Jarabo, D. Gutierrez, and A. Velten, "Non-line-of-sight imaging using phasor-field virtual wave optics," *Nature* **572**(7771), 620–623 (2019).
17. D. B. Lindell, G. Wetzstein, and M. O'Toole, "Wave-based non-line-of-sight imaging using fast fk migration," *ACM Trans. Graph.* **38**(4), 1–13 (2019).
18. G. Gariepy, F. Tonolini, R. Henderson, J. Leach, and D. Faccio, "Detection and tracking of moving objects hidden from view," *Nat. Photonics* **10**(1), 23–26 (2016).
19. F. Heide, M. O'Toole, K. Zang, D. B. Lindell, S. Diamond, and G. Wetzstein, "Non-line-of-sight imaging with partial occluders and surface normals," *ACM Trans. Graph.* **38**(3), 1–10 (2019).
20. J. G. Chopite, M. B. Hullin, M. Wand, and J. Isenberghausen, "Deep non-line-of-sight reconstruction," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (2020), pp. 960–969.
21. C. Thrampoulidis, G. Shulkind, F. Xu, W. T. Freeman, J. H. Shapiro, A. Torralba, F. N. Wong, and G. W. Wornell, "Exploiting occlusion in non-line-of-sight active imaging," *IEEE Trans. Comput. Imaging* **4**(3), 419–431 (2018).
22. F. Xu, G. Shulkind, C. Thrampoulidis, J. H. Shapiro, A. Torralba, F. N. Wong, and G. W. Wornell, "Revealing hidden scenes by photon-efficient occlusion-based opportunistic active imaging," *Opt. Express* **26**(8), 9945–9962 (2018).
23. C. Saunders, J. Murray-Bruce, and V. K. Goyal, "Computational periscopy with an ordinary digital camera," *Nature* **565**(7740), 472–475 (2019).
24. A. Kirmani, D. Venkatraman, D. Shin, A. Colaço, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal, "First-photon imaging," *Science* **343**(6166), 58–61 (2014).
25. Z.-P. Li, X. Huang, Y. Cao, B. Wang, Y.-H. Li, W. Jin, C. Yu, J. Zhang, Q. Zhang, C.-Z. Peng, F. Xu, and J.-W. Pan, "Single-photon computational 3D imaging at 45 km," *Photonics Res.* **8**(9), 1532–1540 (2020).
26. O. Gupta, T. Willwacher, A. Velten, A. Veeraraghavan, and R. Raskar, "Reconstruction of hidden 3D shapes using diffuse reflections," *Opt. Express* **20**(17), 19096–19108 (2012).
27. M. Laurenzis and A. Velten, "Nonline-of-sight laser gated viewing of scattered photons," *Opt. Eng.* **53**(2), 023102 (2014).
28. F. Heide, L. Xiao, W. Heidrich, and M. B. Hullin, "Diffuse mirrors: 3D reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, (2014), pp. 3222–3229.
29. E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.* **59**(8), 1207–1223 (2006).
30. M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.* **25**(2), 83–91 (2008).
31. J. Romberg, "Imaging via compressive sampling," *IEEE Signal Process. Mag.* **25**(2), 14–20 (2008).
32. R. M. Willett, R. F. Marcia, and J. M. Nichols, "Compressed sensing for practical optical imaging systems: a tutorial," *Opt. Eng.* **50**(7), 072601 (2011).
33. X. Liu and A. Velten, "The role of Wigner Distribution Function in Non-Line-of-Sight Imaging," in *2020 IEEE International Conference on Computational Photography (ICCP)*, (IEEE, 2020), pp. 1–12.
34. M. Isogawa, D. Chan, Y. Yuan, K. Kitani, and M. O'Toole, "Efficient Non-Line-of-Sight Imaging from Transient Sinograms," in *European Conference on Computer Vision*, (Springer, 2020), pp. 193–208.
35. W. Chen, F. Wei, K. N. Kutulakos, S. Rusinkiewicz, and F. Heide, "Learned feature embeddings for non-line-of-sight imaging and recognition," *ACM Trans. Graph.* **39**(6), 1–18 (2020).
36. W.-K. Yu, X.-F. Liu, X.-R. Yao, C. Wang, Y. Zhai, and G.-J. Zhai, "Complementary compressive imaging for the telescopic system," *Sci. Rep.* **4**(1), 5834 (2015).
37. W. Gong, C. Zhao, H. Yu, M. Chen, W. Xu, and S. Han, "Three-dimensional ghost imaging lidar via sparsity constraint," *Sci. Rep.* **6**(1), 26133 (2016).

38. M. P. Edgar, G. M. Gibson, R. W. Bowman, B. Sun, N. Radwell, K. J. Mitchell, S. S. Welsh, and M. J. Padgett, "Simultaneous real-time visible and infrared video with single-pixel detectors," *Sci. Rep.* **5**(1), 10669 (2015).
39. J. P. Haldar, D. Hernando, and Z.-P. Liang, "Compressed-sensing mri with random encoding," *IEEE Trans. Med. Imaging* **30**(4), 893–903 (2011).
40. E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.* **25**(2), 21–30 (2008).
41. X. Jiang, G. Raskutti, and R. Willett, "Minimax optimal rates for poisson inverse problems with physical constraints," *IEEE Trans. Inf. Theory* **61**(8), 4458–4474 (2015).
42. Z. T. Harmany, R. F. Marcia, and R. M. Willett, "This is spiral-tap: Sparse poisson intensity reconstruction algorithms—theory and practice," *IEEE Trans. on Image Process.* **21**(3), 1084–1096 (2012).
43. A. Jarabo, J. Marco, A. Muñoz, R. Buisan, W. Jarosz, and D. Gutierrez, "A framework for transient rendering," *ACM Trans. Graph.* **33**(6), 1–10 (2014).
44. A. Pediredla, A. Veeraraghavan, and I. Gkioulekas, "Ellipsoidal path connections for time-gated rendering," *ACM Trans. Graph.* **38**(4), 1–12 (2019).
45. J. Isenberghausen and M. B. Hullin, "Non-line-of-sight reconstruction using efficient transient rendering," *ACM Trans. Graph.* **39**(1), 1–14 (2020).
46. D. Shin, A. Kirmani, and V. K. Goyal, "Low-rate poisson intensity estimation using multiplexed imaging," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, (IEEE, 2013), pp. 1364–1368.
47. M. Raginsky, R. M. Willett, Z. T. Harmany, and R. F. Marcia, "Compressed sensing performance bounds under poisson noise," *IEEE Trans. Signal Process.* **58**(8), 3990–4002 (2010).