Research Article | Vol. 33, No. 14 / 14 Jul 2025 / *Optics Express* 30783

Optics EXPRESS

# Sub-pixel resolving modulation for non-line-of-sight imaging

WENJUN ZHANG,[1,2] SHUO ZHU,[3] LIJIA CHEN,[1,2] LINGFENG LIU,[1,2] LIANFA BAI,[1,2] EDMUND Y. LAM,[3] ENLAI GUO,[1,2,4] AND JING HAN[1,2,5]

[1] *State Key Laboratory of Extreme Environment Optoelectronic Dynamic Testing Technology and Instrument, Nanjing University of Science and Technology, Nanjing 210094, China*
[2] *Jiangsu Key Laboratory of Visual Sensing and Intelligent Perception, Nanjing University of Science and Technology, Nanjing 210094, China*
[3] *Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong SAR, China*
[4] *njustgel@njust.edu.cn*
[5] *eohj@njust.edu.cn*

**Abstract:** Non-line-of-sight (NLOS) imaging is a rapidly developing technology with significant potential applications. Achieving fast and high-resolution imaging is crucial in practical scenarios. In existing NLOS imaging methods, most approaches rely on time-consuming scanning to detect scenes, which makes it challenging to achieve rapid acquisition. Some methods propose replacing the scanning process with array detectors; however, these approaches are limited by the number of pixels in the array detector and time jitter, making it difficult to achieve high-resolution object reconstruction. This paper introduces digital micromirror device (DMD) modulation into the scan-free NLOS system and proposes a high-resolution imaging method called pixel shift modulation non-line-of-sight (PSM-NLOS) imaging. By adopting a pixel shift modulation acquisition mechanism, we establish a fine scanning acquisition system based on DMD modulation, enabling rapid sub-pixel sampling. The light field evolution characteristic in non-confocal scenarios is investigated, and the effects of illumination and detection regions on time jitter are incorporated into the light field modeling. Considering issues such as ambient light and pixel crosstalk in the array detectors, we achieve high-resolution imaging of hidden objects through model inversion and frequency domain filtering. This method breaks the resolution limit of NLOS systems, enhancing the spatial resolution from 7 cm to 1 cm while significantly improving the reconstruction quality of hidden objects. Moreover, the modulation approach is compatible with various techniques such as the scan-free boundary migration method, the virtual waves method, and the light-cone transform method, and can also serve as an independent modulation strategy to enhance imaging resolution, demonstrating considerable reference value.



## 1. Introduction

Non-line-of-sight (NLOS) imaging enables the visualization of hidden scenes outside the field of view (FOV) of the detector by leveraging indirect reflections of light scattered from the hidden scene onto other surfaces. NLOS imaging technology garners significant attention due to its wide-ranging applications, including autonomous driving, medical imaging, remote sensing, robotic vision, cave exploration, and public safety [1–7]. In recent years, various solutions based on different principles have been proposed for NLOS imaging, including wavefront shaping [8], ordinary camera-based imaging [9–11], thermal imaging [12], acoustic imaging [13], frequency-modulated-continuous-wave techniques [14], and transient imaging based on photon time-of-flight (TOF) [3,15–19]. Among these imaging methods, transient imaging is one of the most widely used in NLOS imaging. Despite the advancements in transient imaging

#569102 https://doi.org/10.1364/OE.569102
Journal © 2025 Received 27 May 2025; revised 30 Jun 2025; accepted 2 Jul 2025; published 11 Jul 2025

technology, achieving both high-resolution and rapid imaging remains challenging due to several limitations. First, NLOS imaging typically requires at least three-bounce reflections, leading to significant spherical wave attenuation. The signal strength of the received light drops by several orders of magnitude, necessitating a relatively long detection duration to collect sufficient signal photons, which limits imaging speed and quality [20]. Second, the commonly used galvanometer scanning technique struggles to balance scanning speed and resolution, rapid imaging often requires reducing the number of sampling points, thereby degrading image resolution [21–27]. Third, if array detectors are used instead of traditional scanning methods, although the acquisition speed is greatly improved, the imaging resolution is limited by the number of detector pixels [28]. Additionally, compared with the scanning system, considerations must be made for the influence of detection FOV on the temporal and spatial broadening of the signal, making high-resolution imaging difficult. Therefore, existing methods often address only single aspects of the problem, failing to balance imaging speed and resolution, which imposes constraints on the practical application of NLOS imaging.

Transient NLOS imaging methods based on photon TOF are limited in spatial resolution by the total time resolution of the system, which restricts resolution to the centimeter scale [14,29,30]. To overcome this limitation, modeling the three-dimensional spatial-temporal kernel matrix and incorporating it into the forward evolution model can mitigate the effect of broadening caused by long distances, thus enhancing resolution [3]. Using backscattered photons from the line-of-sight scenario as a prior for deconvolving the time histogram, the Richardson-Lucy deconvolution method can reduce the spatial resolution to one-third of the original [31]. Deconvolving the instrument response function of the system overcomes the bottleneck in NLOS imaging resolution caused by the system [29]. Designing an up-conversion single-photon detector with a high time resolution of 1.4 ps can enhance the lateral resolution to 2 mm [30]. While these methods effectively improve resolution, they require dense scanning on the relay wall to gather spatial information, resulting in lengthy data acquisition times and difficulty in achieving rapid imaging.

Most NLOS methods rely on galvanometer scanning for extended periods to achieve indirect scene information acquisition, which imposes certain limitations on the development of NLOS technology. If the scanning mode is not changed, reducing the number of scan points to decrease acquisition time may affects imaging resolution and quality [21–27]. Reducing the collection time per detection point requires optimization algorithms for imaging, leading to longer reconstruction times [32]. By utilizing spectral filtering and motion compensation algorithms, the number of scan points can be reduced to quickly acquire low-resolution images for target localization. Subsequently, higher-resolution images can be reconstructed by combining multiple frames of low-resolution image subsets [20]. Although this method incurs some delay, it can achieve rapid imaging through software optimization.

Using digital micromirror device (DMD) modulation [33,34] or array detectors [28,35,36] can partially or completely replace the galvanometer for spatial scanning, further enhancing acquisition speed. Combining a single-photon detector with a DMD reduces acquisition time to 0.8 s, the used DMD has a 22 kHz refresh rate and high single-photon sensitivity [33]. Structured sampling based on single-pixel detectors and DMD can increase acquisition speed by five times [34]. Using array detectors to partially or completely replace scanning offers strong advantages in dynamic scene detection speed. With two custom 16×1 array detectors, sparse scanning with a galvanometer achieves a maximum acquisition speed of 5 fps [35]. To further enhance acquisition speed, the scan-free boundary migration (SCBM) method is presented, which uses a 32×32 detector array instead of laser scanning, achieving a maximum of 151 Hz [28,36]. However, for scan-free NLOS imaging, semiconductor manufacturing processes limit the pixel count of single-photon avalanche diode array detectors [37–39]. To ensure the detection area, lens is needed to focus the detection surface of the array detector on the relay wall, which will expand the detection FOV corresponding to a single pixel of the detector. This affects the temporal and

Optics EXPRESS

spatial information detected, leading to some blurring in the reconstructed object results and impacting imaging resolution [3,40]. Limited by the number of detector pixels and high time jitter in existing scan-free systems, achieving high-resolution imaging is challenging. Using multiple array detectors significantly increases system size and cost. Mechanical movement methods are restricted by the time required for pointing mechanisms, which makes it difficult to ensure rapid imaging [41].

To achieve fast and high-resolution NLOS imaging, this paper combines DMD with array detectors and proposes a high-resolution NLOS imaging method called PSM-NLOS. First, based on the existing array detection, the fine scanning is performed by combining DMD modulation. The movement of detection region can reach sub-pixel precision, which can break through the inherent resolution limitations of the devices and enable rapid sub-pixel sampling. Then, a non-confocal system model is established, which considers the effects of the DMD, laser and other devices and detector FOV on imaging, the temporal and spatial broadening of the signal is modeled as a three-dimensional blur kernel. Finally, high-resolution imaging of hidden scenes is achieved through model inversion and frequency domain filtering. Our method can reduce the impact of signal broadening on imaging, requiring only a minimum of four modulations to achieve scene acquisition, improving the lateral resolution to 1 cm , and enabling high-resolution imaging of hidden objects. It can also achieve imaging under the scene with ambient light interference. The proposed method is the first to introduce DMD modulation into scan-free NLOS systems. Considering the high refresh rate and high single photon sensitivity of DMD, its introduction does not significantly affect system acquisition speed but adds a new modulation capability, hoping to provide a new approach for NLOS imaging.

## 2. Methods

For the traditional scan-free NLOS methods, the number of detector pixels is equivalent to the number of samples, the imaging resolution is still limited by the detector array [28]. Furthermore, the system has a high time jitter. Therefore, it is difficult to achieve high-resolution imaging. To overcome this issue, we propose solutions from both hardware and software perspectives.

In terms of hardware, we achieve sub-pixel sampling based on the DMD spatial modulation. The NLOS imaging model and method based on pixel shift modulation proposed in this paper are based on the system shown in Fig. 1. The imaging system consists of a laser, an array detector, a DMD, a total internal reflection prism, lenses, and a narrowband filter. The laser (BLAZER-GR10) uses a high-power pulsed laser at the 532 nm wavelength band with a maximum power of 8 W, a pulse width of 8 ps, and a repetition frequency of 20 MHz. The object scene hidden from the direct line of sight, and the actual power used by the laser is 1 W. After being incident on the relay wall, the laser scatters, and part of the photons are reflected into the hidden scene, some of them reach the relay wall after scattering again in the hidden scene. This study uses an array detector (Photon Force PF32-USBC-MLA-500k, pixel count: 32×32) combined with DMD modulation for spatial sampling to obtain information about different detection areas on the relay wall. As shown in the lower left corner of Fig. 1, a 50 mm focal length lens is used to image the detection area of the array detector onto a 320×320 pixels range at the center of the DMD chip (pixel count: 1024×768, pixel pitch: 13.68$\mu$m), with one pixel of the detector corresponding to 10×10 pixels of the DMD chip. A total internal reflection prism is placed in front of the DMD to change the direction of the modulated light path. In front of the total internal reflection prism is a 4 mm focal length lens that focuses the detection FOV onto a 120 cm×120 cm area on the relay wall. Spatial encoding modulation can be performed based on the DMD, and the array detector collects TOF data carrying the information of the hidden scene after modulation. The refresh rate of the DMD is as high as 22 kHz, meeting the system's fast detection requirements. A narrowband filter (Thorlabs FLH532-4) is added in front of the detector to remove some environmental interference light. The standard transistor-transistor logic signal output by the laser can serve as the trigger for the array detector. The time-related single-photon counting of different pixels in the array detector operate independently, and the detector has a micro-lens array to improve duty cycle and collection efficiency.

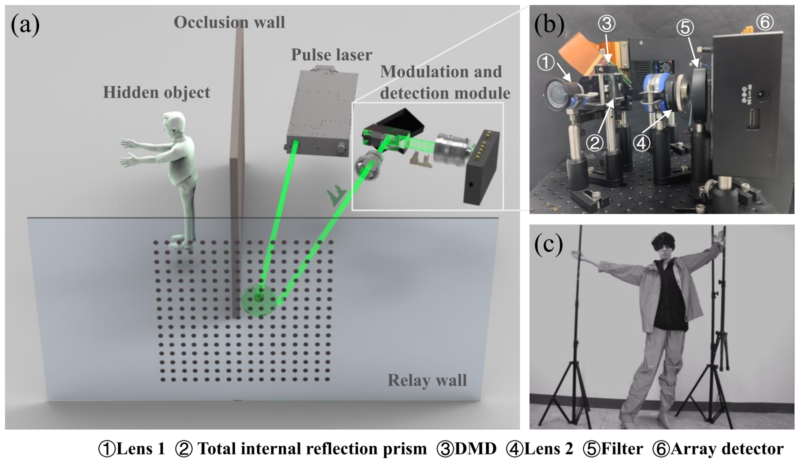


**Fig. 1.** NLOS imaging system model based on DMD modulation. (a) The optical setup of the NLOS imaging system. (b) Experimental setup of the modulation and detection module. (c) Hidden object scene.

It is assumed that there are no occlusions in the hidden scene, and both the relay wall and the hidden objects are Lambertian, ignoring mutual reflections between hidden objects and reflections beyond three times. The system adopts a non-confocal structure, with the laser incident point and the detection point on the relay wall located at different positions. As shown in Fig. 1(a), the modulation and detection module focus the detection plane of the array detector onto the relay wall for measurement. The illumination point of the light is fixed, and each pixel of the detector corresponds to a different pixel area on the DMD, which corresponds to different detection areas on the relay wall by modulation. For a detection pixel on the array detector, the modulation process of the DMD can be expressed as:

$$Y = H\tau_{DMD}, \tag{1}$$

where $H$ is the mask corresponding to the detection point on the DMD, which is an $N \times T$ matrix, $N$ represents the number of modulations and $T$ represents the mask pattern of a 10×10 pixel area on the DMD corresponding to the detection pixel on the array detector. $\tau_{DMD}$ is the spatiotemporal measurement result received by the 10×10 pixel area on the DMD, which is a $T \times S$ matrix, with $S$ representing the number of repeated acquisitions. $Y$ is the three-dimensional spatiotemporal measurement result received at a single detection pixel after spatial modulation by the DMD, and it is an $N \times S$ matrix.

In the system shown in Fig. 1, a single pixel of the array detector is projected onto a 10×10 pixels area on the DMD through a lens with a focal length of 50 mm. A 7×7 pixels region is selected as the detection area for a single pixel, and the detection area undergoes changes across the entire 10×10 region. The array detector has 32×32 pixels, the time-correlated single-photon counting for each pixel operates independently, such changes need to be synchronized to each pixel, ultimately forming a mask on the DMD, each change corresponds to a different mask. By expanding the receiving optical path and ignoring the influence of components like the total internal reflection prism and DMD on the optical path direction, the simplified optical path of the DMD modulation process shown in Fig. 2(a) can be obtained. In the receiving optical path, the signal carrying object scene information first reaches the DMD. After being modulated by different masks on the DMD, the signal undergoes different changes. After each modulation, the array detector detects a set of low-resolution data (32×32× $S$), where $S$ is the number of repeated acquisitions. Figure 2(b) shows the detection region on the relay wall under different modulations. It can be observed that by implementing spatial modulation via the DMD, the movement of the detection area can be accurately controlled to sub-pixel precision. The high refresh rate of the DMD also ensures that this method is applicable to NLOS systems with the fastest detection speeds exceeding 150 fps. By collecting a series of low-spatial-resolution data corresponding to different masks, higher-spatial-resolution data can be calculated through computation. The sub-pixel displacements between these low-resolution data provide frequency aliasing constraints, which are the essence of super-resolution reconstruction [42–45].

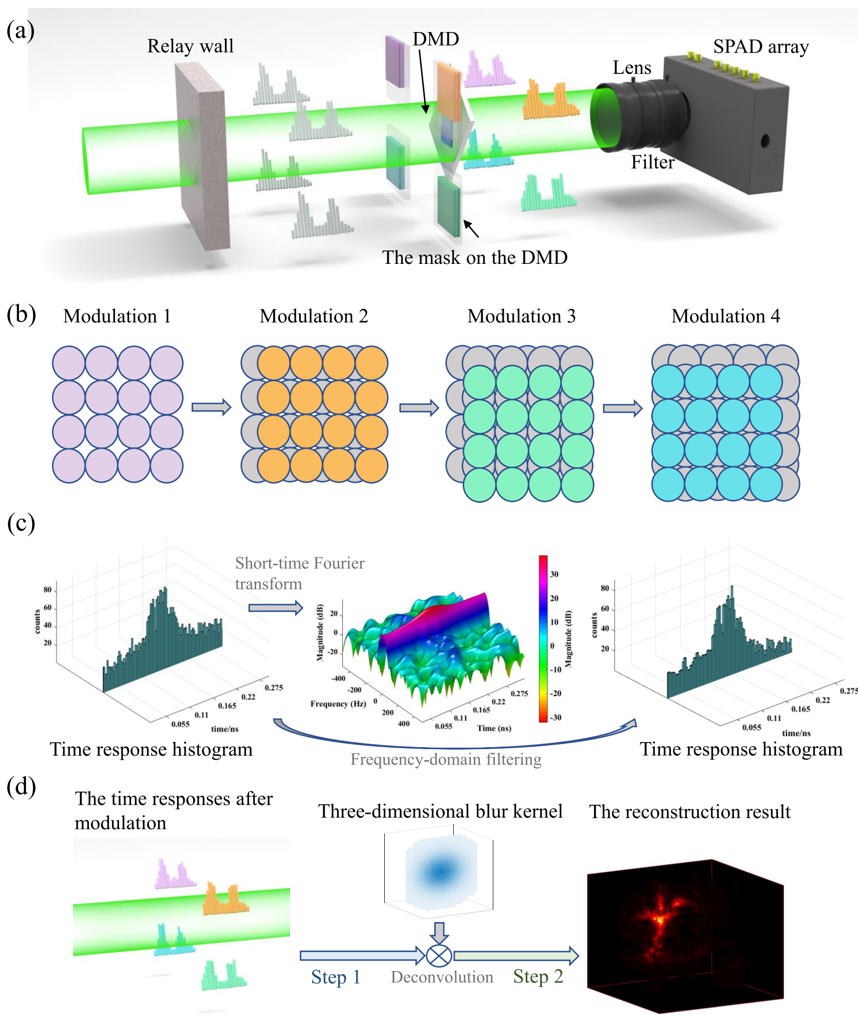


**Fig. 2.** The schematic of the PSM-NLOS method. (a) The simplified optical path of the DMD modulation process. (b) The detection region on the relay wall under different modulations. (c) The influence of Fourier-domain filtering on the time response histogram collected by a pixel of the array detector. (d) The 3D reconstruction algorithm process of the PSM-NLOS method.

For each independent pixel, the imaging of the detector onto the DMD can divide one original detector pixel into 10×10 virtual detector pixels. By designing the DMD mask, signals from only 7×7 virtual detector pixel regions can be collected each time. Each modulation of the DMD can shift the virtual detection area by one or two virtual detector pixels, thus achieving sub-pixel detection for actual detection points and obtaining more detailed information [46]. The combination of multi-frame low-resolution data detected into high-resolution data $Y_h$ can overcome the inherent resolution limitation of the imaging system. Note that the number of DMD pixels corresponding to a single detector pixel, the number of DMD modulations, and the number of virtual detector pixels moved each time can all be designed according to actual needs. The number of corresponding pixels on the DMD should consider both the imaging effect of a single detection and the required number of modulations. The number of modulations should balance imaging quality with acquisition time, generally, good results can be achieved with four modulations. Increasing the number of modulations will increase the acquisition time and computational load. However, if higher resolution is required for the scene, more modulations can be considered. The number of virtual detector pixels moved each time should be designed based on the corresponding pixel count on the DMD and the number of modulations. The DMD-modulated subpixel acquisition method proposed in this paper can be applied to any TOF reconstruction method, such as the SCBM method [28], light-cone transform (LCT) method [16], phasor field method [18] and so on.

In software, considering the impact of time jitter on imaging, we improved the modeling for NLOS scenarios based on array detectors and achieved high-resolution imaging of hidden scenes based on model inversion and frequency-domain filtering. For simplicity, assume that the relay wall is a flat plane located at $z = 0$. The light emitted by the pulsed laser illuminates a point $(x_l, y_l, 0)$ on the relay wall. After diffuse reflection from the relay wall, the entire hidden scene is illuminated. The array detector can record photon TOF information containing structural information of the hidden object.

Let the distance from the laser irradiation point $(x_l, y_l, 0)$ to a point $(x, y, z)$ on the object be denoted as $r_1$, and the distance from a detection point $(x_s, y_s, 0)$ on the DMD to the point $(x, y, z)$ be denoted as $r_2$. The data $\tau_{\mathrm{DMD}}$ detected from the DMD can be expressed as [3,16,36]:

$$\tau_{\mathrm{DMD}}(x_s, y_s, t) = \phi(x_s, y_s, t) * \iiint_{\Omega} \frac{1}{r_1^2 r_2^2} \rho(x, y, z) \delta(r_1 + r_2 - tc) dxdydz, \tag{2}$$

where

$$\phi = \phi_{xy} \cdot \phi_t, \tag{3}$$

$$\phi_{xy} = exp(-\frac{x_s^2}{2\sigma_x^2} - \frac{y_s^2}{2\sigma_y^2}), \tag{4}$$

$$\phi_t = exp(-\frac{t^2}{2\sigma_t^2}). \tag{5}$$

Here, $c$ is the speed of light, and $\rho$ is the albedo of the hidden scene in the three-dimensional half-space $\Omega$. The Dirac delta function $\delta$ simulates the propagation of light from the relay wall to the hidden object and back to the relay wall.$\phi$ represents the three-dimensional blur kernel, $\phi_{xy}$ denotes the spatial intensity distribution of the detection FOV on the relay wall corresponding to a single detection pixel, while $\phi_t$ represents the time jitter of the entire system. Both $\phi_{xy}$ and $\phi_t$ can be expressed as Gaussian functions with standard deviations $\sigma_x$, $\sigma_y$, and $\sigma_t$, respectively. $*$ represents convolution.

Modeling non-confocal scenarios directly is rather complex and computationally expensive. We propose the non-confocal transformation (NCT) correction method [28] to transform the

Optics EXPRESS

data $\tau_{\mathrm{DMD}}$ detected on the DMD into confocal scenario data $\tau'_{\mathrm{DMD}}$, correcting the position of the detection points to $(x_c, y_c, 0)$. The confocal scenario data $\tau'_{\mathrm{DMD}}$ can be expressed as:

$$\tau'_{\mathrm{DMD}}(x_c, y_c, t) = \phi(x_c, y_c, t) * \iiint_\Omega \frac{1}{r_c^4}\rho(x, y, z)\delta(2r_c - tc)dxdydz = \phi(x_c, y_c, t) * I'(x_c, y_c, t), \quad (6)$$

where $r_c = \sqrt{(x_c - x)^2 + (y_c - y)^2 + z^2}$, $I'(x_c, y_c, t) = \iiint_\Omega \frac{1}{r_c^4}\rho(x, y, z)\delta(2r_c - tc)dxdydz.$

Based on the LCT model, a change of variables is performed by letting $u = z^2$ and $v = (\frac{tc}{2})^2$, and Eq. (6) can thus be rewritten as [16]:

$$\begin{aligned}&\tau'_{\mathrm{DMD}}(x_c, y_c, \frac{2\sqrt{v}}{c})\\ &= \tilde{\phi}(x_c, y_c, \frac{2\sqrt{v}}{c}) * [v^{-3/2} \iiint_\Omega \frac{1}{2\sqrt{u}}\rho(x, y, \sqrt{u})\delta((x_c - x)^2 + (y_c - y)^2 + u - v)dxdydu]\\ &= \tilde{\phi}(x_c, y_c, \frac{2\sqrt{v}}{c}) * \tilde{I}'(x_c, y_c, \frac{2\sqrt{v}}{c}),\end{aligned} \quad (7)$$

where $\tilde{\phi} = \tilde{\phi}_{xy} \cdot \tilde{\phi}_t$, $\tilde{\phi}_{xy} = exp(-\frac{x_c^2}{2\sigma_x^2} - \frac{y_c^2}{2\sigma_y^2})$, $\tilde{\phi}_t = exp(-\frac{2v}{c^2\sigma_t^2})$.

Letting $R_t\{\tilde{I}'\} = v^{3/2}\tilde{I}'(x_c, y_c, \frac{2\sqrt{v}}{c})$, $R_z\{\rho\} = \frac{1}{2\sqrt{u}}\rho(x, y, \sqrt{u})$, $h = \delta((x_c - x)^2 + (y_c - y)^2 + u - v)$, and the following formula can be obtained: $R_t\{\tilde{I}'\} = h * R_z\{\rho\}$. Where $h$ is a shift-invariant three-dimensional convolution kernel, $\rho$ represents the albedo of the hidden object, the transform $R_z$ nonuniformly resamples and attenuates the elements of volume $\rho$ along the $z$ axis, and the transform $R_t$ nonuniformly resamples and attenuates $\tilde{I}'$ along the time axis [16].

The discrete free-space propagation operator can be expressed as $\tilde{\boldsymbol{I}}' = \boldsymbol{R}_t^{-1}\boldsymbol{H}\boldsymbol{R}_z\boldsymbol{\rho} = \boldsymbol{A}\boldsymbol{\rho}$, where matrix $\boldsymbol{H}$ is the shift-invariant three-dimensional convolution operation, the vectorized volume of the albedo of the hidden object is represented as $\boldsymbol{\rho}$, and $\boldsymbol{A} = \boldsymbol{R}_t^{-1}\boldsymbol{H}\boldsymbol{R}_z$ is the light transport matrix [16]. The discrete imaging model can be expressed as:

$$\boldsymbol{\tau}'_{\mathrm{DMD}} = \boldsymbol{\Phi}\tilde{\boldsymbol{I}}' = \boldsymbol{\Phi}\boldsymbol{A}\boldsymbol{\rho}, \quad (8)$$

where the matrix $\boldsymbol{\Phi}$ represents the shift-invariant three-dimensional convolution operation (i.e., a convolution with a discretized version of the kernel $\tilde{\phi}$) [16].

By inverting this model and applying the convolution theorem, the Wiener deconvolution filter, and a free-space propagation solver based on LCT, the albedo $\boldsymbol{\rho}$ of the hidden object can be recovered [47].

The overall workflow of the method is shown in Fig. 2. First, based on DMD modulation, sub-pixel acquisition as shown in Fig. 2(b) is implemented. After each modulation, the array detector will detect a set of low-resolution data (32×32× $S$). The multi-frame low-resolution data is combined into high-resolution data to obtain a higher resolution reconstruction result. As shown in Fig. 2(c), due to the introduction of DMD, total internal reflection prism and lens, both detection efficiency and noise are affected. There is also pixel crosstalk in the array detector, and the detected signal is greatly affected by noise. In addition, the applicability of the algorithm under environments with lights on needs to be considered. For this problem, as shown in Fig. 2(c), we use short-time Fourier transform for time-frequency analysis. From the histogram of the time response collected by one pixel of the array detector, it can be found that the detected signals contain substantial noise with a low signal-to-noise ratio. In addition to uniformly distributed noise, there exist sharp noise components, resulting in noise distribution across both low frequency and high frequency. As shown in the time-frequency analysis results in Fig. 2(c), the uniform noise is distributed in the low-frequency range, while the signal frequencies are slightly higher than this noise component. However, significant spectral overlap exists between them, making it difficult to separate noise from signal by simply removing partial low-frequency components. Some sharp noise components are distributed in high frequencies, and this portion of noise can be removed through low-pass filtering to better extract object information. Finally, leveraging the modeling with the blur kernel and the non-confocal imaging model, the Wiener deconvolution filter and the free-space propagation solver based on LCT are utilized for object reconstruction. As shown in Fig. 2(d), after post-processing such as grayscale stretching, the high-resolution three-dimensional imaging results of the object are obtained.

## 3. Experiment

This paper introduces DMD modulation into scan-free NLOS imaging systems for the first time. The system uses an array detector instead of galvanometer scanning to realize rapid data acquisition, and realizes sub-pixel sampling based on DMD spatial modulation. The high precision and refresh rate of DMD modulation allow the system to preserve the acquisition speed advantage of the array detector while capturing more detailed information and overcoming the inherent resolution limitations of the imaging system to obtain high-resolution data. For the system used in this study, factors such as pixel crosstalk among array detectors, the influence of the FOV of individual detection points on the temporal and spatial broadening of the signal, and the impact of introducing components like DMD, total internal reflection prisms, and lenses on system time jitter, detection efficiency, and noise interference must be considered [28,36,48]. We model the temporal and spatial broadening of the signal as a three-dimensional blur kernel and incorporate it into the system model. Object scene recovery is achieved through model inversion and frequency-domain noise filtering.

We demonstrate the reconstruction capability of our method in the non-confocal scenario based on array detectors, and compare the imaging effects of our method with that of the scan-free NLOS imaging method based on an array detector—the SCBM method, NLOS classical method—the phasor field method and the LCT algorithm. We show the enhancement of modulation on imaging resolution and showcase the compatibility of our modulation approach with different methods. Moreover, we test the imaging resolution of our method, its ability to distinguish human poses, its performance under various acquisition times, and its imaging capabilities under environment with light. These results are compared with those of other methods, which also show the advantages of our approach.

### *3.1. Imaging performance verification of the proposed PSM-NLOS method*

In order to verify the actual imaging performance of our method, we present the imaging results of our method and different fast imaging methods, as shown in Fig. 3. The hidden scene consisted of four handmade targets: “L”, “O”, “$\mu$”, and a dual-target “LU”. Since the LCT method applies to confocal scenes, we used the NCT algorithm to correct the data into confocal data before performing the reconstruction. The overall detection area on the relay wall is 120 cm ×120 cm. If only single modulation is performed, the spatial resolution of the collected data is 32×32, and the spatial resolution of the imaging result is also 32×32. By comparing the imaging effects of different methods, it can be observed that the phasor field method recovers object structures with significant missing parts and high background noise. The LCT method is able to basically restore the shape of objects but also has high background noise. The SCBM method has less background noise compare to the other two methods, but it also has issues with structural loss. All three methods mentioned above have problems with blurry object structures and noise interference in imaging. Our method comprehensively considers the effects of signal broadening and noise on imaging, allowing it to basically recover the clear structure of hidden objects and significantly reduce the interference of background noise on imaging, demonstrating strong noise robustness and significant imaging advantages. By comparing the last two columns of reconstruction results, it can be seen that multiple acquisitions and modulations can recover more detailed information of objects, improving the imaging effect.

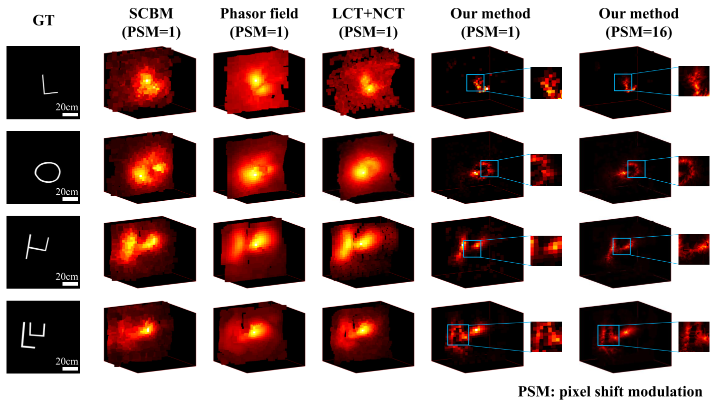


**Fig. 3.** Comparison of imaging performance of different fast imaging methods.

For the closely spaced dual-target "LU", due to the difference in distance between the two targets and the relay wall, the reconstruction results of the two targets obtained by SCBM, phasor field and LCT methods all have intensity differences, and the target with lower intensity is more affected by noise. Additionally, the structure of the recovered two targets is partially adhered, so it is difficult to distinguish the structure of the targets accurately. Under single modulation, our method can recover the basic shape of the targets, but the two targets still exhibit some degree of adhesion. By performing spatial dimension modulation multiple times using DMD, more spatial dimensional information of the objects can be explored, overcoming the inherent resolution limitations of the imaging system to achieve high-resolution imaging. From the last column of the imaging results, it can be observed that multiple modulations could solve the issue of slight adhesion in the reconstructed result and accurately restore the two independent objects.

To investigate the influence of the modulation times on imaging performance, as shown in Fig. 4, the reconstruction results of the LCT method and our proposed method under different modulation times are shown. The spatial resolution of the imaging results is 32×32, 64×64, and 128×128 for 1, 4, and 16 times of collecting modulations, respectively. It can be observed that the more modulation times applied, the more complete the recovery of object structural details. For objects with relatively simple structures or not tightly structured, four modulation times are sufficient to restore the structural information of the objects. For the dual-target shown in the last line of Fig. 4, it can be found that there is slight adhesion between the two targets under single modulation, while the two targets are basically separated after four modulations. After 16 modulations, the two independent targets can be clearly restored. The modulation method can also serve as an independent step and is applicable to other methods such as LCT, which can improve the imaging details and resolution.

### *3.2. Resolution capability verification of the proposed PSM-NLOS method*

To verify the imaging resolution capability and resolution advantage of our method, as well as to validate the applicability and effectiveness of the modulation method for different methods, we investigate the system's resolution limitations and experimentally test the lateral resolutions of SCBM, phasor field, LCT methods, and our proposed method. For the NLOS system based on array detectors, the lateral resolution of the system satisfies [16]:

$$\Delta x \geq \frac{c \cdot \sqrt{w^2 + z^2}}{2w} \cdot \gamma, \tag{9}$$

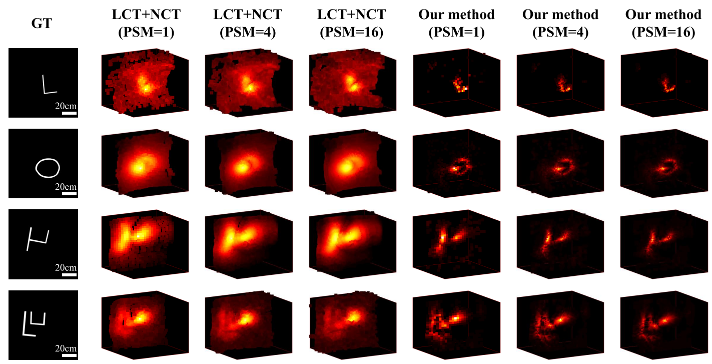


**Fig. 4.** The influence of modulation times on imaging.

where $c$ is the speed of light, the total detection area on the relay wall is $2w \times 2w$, $z$ is the distance between the object and the relay wall, and $\gamma$ is half the full width at half maximum (FWHM) of the system's overall time jitter. Additionally, the number of spatial dimension samples affects the lateral resolution. If the sampling is performed at N×N, the resolution of the reconstructed object scene will also be N×N. However, if the sampling interval exceeds the theoretical resolution, the actual resolution will decrease accordingly [24,36]. For the NLOS system shown in Fig. 1, the total detection area of the relay wall is 120 cm ×120 cm, and the overall time jitter of the system is 330 ps. If the distance between the object and the relay wall is 50 cm, the theoretical lateral resolution of the imaging can be calculated to be 6.44 cm according to Eq. (9). If only single DMD modulation is performed, the sampling interval is 3.125 cm. Our method can weaken the impact of signal temporal and spatial broadening on imaging and improve the actual lateral resolution of the imaging based on multiple DMD modulations. The reconstruction results of the SCBM, phasor field, LCT method, and our method under different DMD modulation times are shown in Fig. 5. The hidden objects are comprised three foam resolution bars aligned along the same axial position, the two leftmost bars maintained a fixed 7 cm spacing. The position of the rightmost resolution bar is adjusted multiple times, achieving lateral distances of 6 cm, 4 cm, 2 cm, and 1 cm relative to the central resolution bar. From the imaging results, it can be found that DMD modulation can significantly improve the imaging resolution and is applicable to all methods. Our method can further reduce the impact of factors such as time jitter and noise on imaging, enabling clearer differentiation of scene information. The imaging resolution can reach as low as 1 cm, which can basically meet the needs of most practical scenarios.

To validate the practicality of the proposed method, we conduct imaging experiments using a mannequin model as the hidden object, capturing various poses. Figure 6 compares the reconstruction results of five different poses using SCBM, phasor field, LCT, and our method. Prior to data acquisition, ground truth (GT) references are obtained by photographing the object with a conventional camera. Columns 2-4 in Fig. 6 display the imaging results of SCBM, phasor field, and LCT under single DMD modulation. Due to signal broadening and noise interference, these methods struggle to accurately resolve the object's pose. Our approach effectively mitigates signal broadening and noise, enabling clear discrimination of different poses. Furthermore, we adopt a sub-pixel acquisition scheme with four DMD modulation times, and the imaging results are shown in the last column of Fig. 6. It can be found that with the increase of modulation times, more details of spatial dimensions can be collected, and the spatial resolution of imaging results is also improved.

Optics EXPRESS

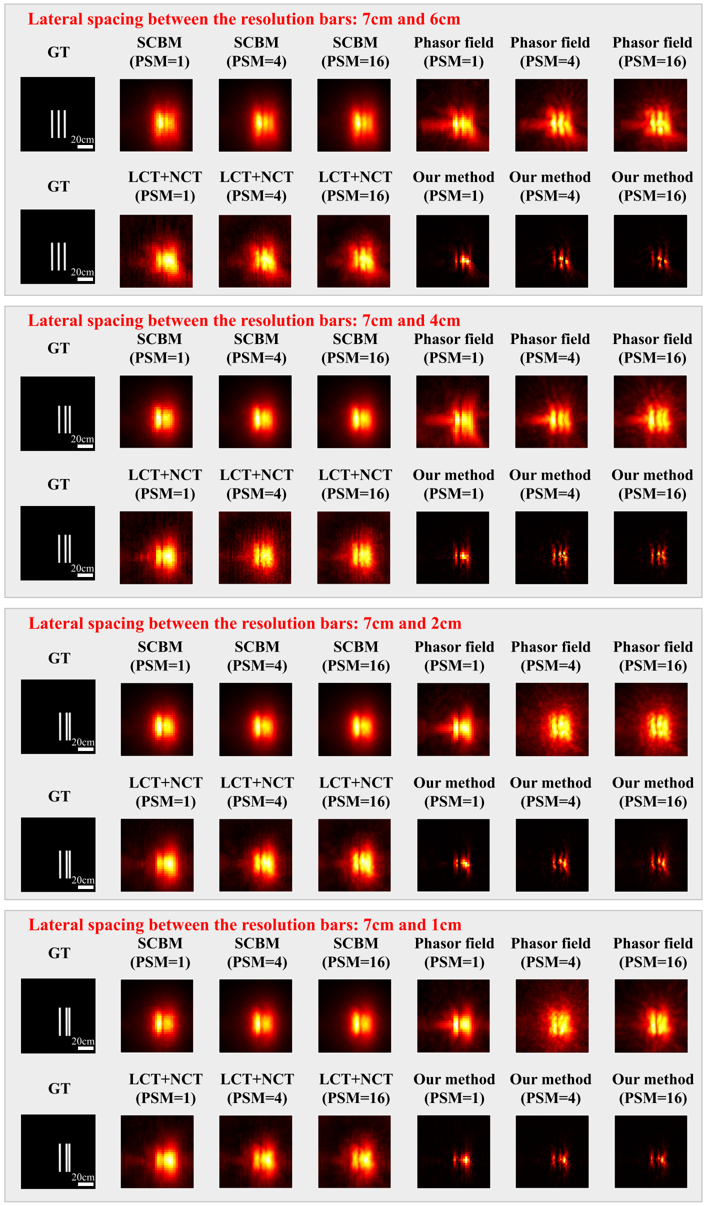


**Fig. 5.** Lateral resolution performance comparison of different methods under varying modulation times.

### 3.3. *Influence of repetition times and modulation times on imaging*

In this paper, an array detector combined with DMD modulation is adopted to replace galvanometer scanning, enabling rapid acquisition of hidden scenes. We evaluate the minimum acquisition time required for targets “L” and “LU” under different modulation times, with results presented in Fig. 7. It can be observed that, for the single target “L”, if single modulation only repeats sampling 1000 times, i.e., the collection time is about 0.0112 s, the basic structure of the object can be restored, though with some loss of fine details. When the repetitions are increased to 2,000, i.e., 0.0217 s, the structural information can be substantially recovered. When performing 4 or

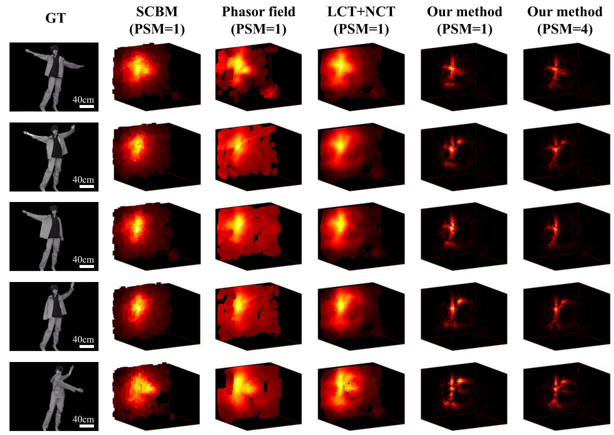


**Fig. 6.** Comparison of mannequin model pose reconstruction performance using different methods.

even 16 modulations, the object structure can still be effectively restored with 1000 repetitions per modulation. The refresh rate of DMD can reach up to 22 kHz, which hardly affects continuous rapid acquisition, facilitating system assembly and synchronized integration. For more complex targets such as "LU", the required acquisition time will increase. In the case of 4 modulations, the object structure can be well restored when the number of repetitions per modulation is 2000.

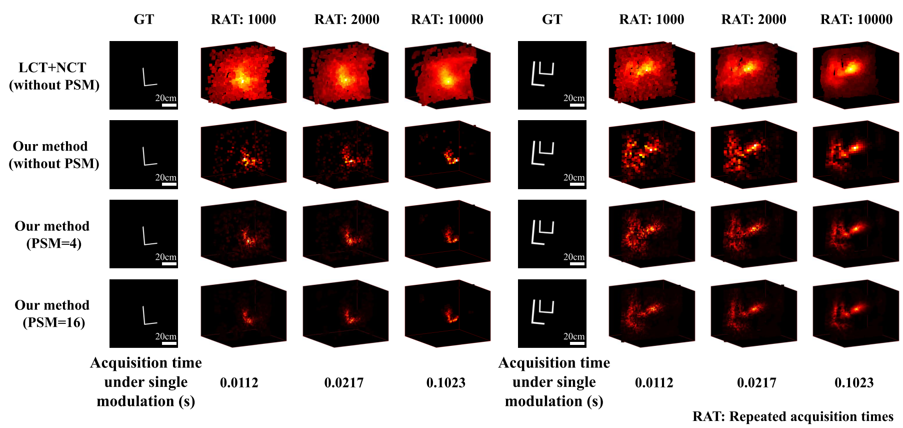


**Fig. 7.** The influence of repetition times and modulation times on imaging.

However, as the number of modulations increases, combining low-spatial-resolution data obtained under different masks to generate higher-spatial-resolution data not only leads to an increase in acquisition time but also significantly increases the computational load for subsequent imaging algorithms, thereby prolonging the overall imaging time. For instance, calculating data with a spatial resolution of 32×32 requires our imaging algorithm to take 0.024 s computation time. If the spatial resolution is increased to 64×64, the required computation time becomes 0.107 s. For a data with a spatial resolution of 128×128, the computation time extends to 0.469 s.

The introduction of DMD modulation can enhance the spatial resolution and imaging details of the image, but it will also increase the time cost to a certain extent. Therefore, the selection of the modulation times also needs to be considered comprehensively. Combined with the experimental results and analysis of Figs. 3–7, four modulations can achieve rapid high-resolution imaging of the object. If the object structure is relatively compact or the spacing between objects is small, and these scenarios require higher imaging resolution, the number of modulations can be appropriately increased to obtain imaging results with higher spatial resolution.

### 3.4. Imaging capabilities of the proposed PSM-NLOS method under the scene with ambient light interference

We also test the effect of turning on the lights on imaging. As shown in Fig. 8, the hidden scene includes single targets "U" and "$\mu$", dual-target "LU", as well as one pose of the mannequin model, with an illuminance of 248 lx under lighted conditions. It can be observed that turning on the lights introduces more environmental noise, thereby increasing the interference to imaging. The filtering performed in the frequency domain in this paper can remove some of the environmental noise, thereby better restoring the information of the target. Therefore, under the scenario with ambient light interference, our method can still basically eliminate the noise influence and achieve high-fidelity restoration of the hidden scene.

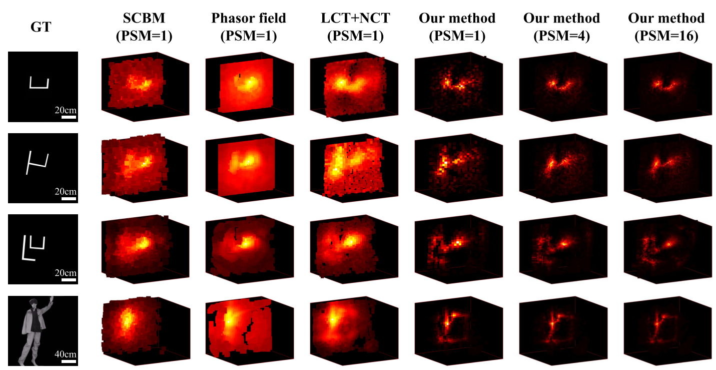


**Fig. 8.** Comparison of the imaging capabilities of different methods under the scene with ambient light interference.

## 4. Discussion

According to the theoretical analysis and experimental demonstration, several discussions are clarified as follows:

(i) Considering the difficulty of rapid data acquisition in scanning-based NLOS methods [16–18] and the limited imaging spatial resolution caused by detector arrays in scan-free approaches [28,36], we propose for the first time the integration of a DMD with an array detector to develop a high-resolution NLOS imaging method based on pixel-offset modulation, called PSM-NLOS. The large pixel array and precision of the DMD enable the system to achieve accurate sub-pixel sampling, while its high refresh rate satisfies the demand for rapid data acquisition in systems based on array detectors, thus allowing our method to achieve fast and high-resolution imaging. Compared to conventional scan-free systems [28], the implementation of DMD modulation requires the introduction of additional devices. In addition to considering pixel crosstalk among array detectors and the influence of the FOV of individual detection points on the temporal and spatial broadening of the signal, the addition impact of introducing components like DMD, total internal reflection prism, and lenses on system time jitter, detection efficiency, and noise interference must be considered. For these challenges, this paper establishes a non-confocal system model. By modeling the broadening as a three-dimensional blur kernel and combining it with spatial sub-pixel sampling, we achieve high-resolution imaging of the hidden scene through model inversion and frequency domain filtering.

(ii) The fine sub-pixel acquisition with fine sub-pixel precision based on DMD modulation can capture more spatial dimension details. Experimental results demonstrate that the modulation method enhances imaging resolution and is applicable to various methods such as SCBM, virtual wave, and LCT. To verify the performance of the proposed PSM-NLOS method, we test its ability to recognize different poses of the mannequin model, its imaging performance under different acquisition times, and its applicability under the scene with ambient light interference. We also demonstrate the imaging advantages of our method compared to SCBM, virtual wave, and LCT in various scenarios. Our method mitigates the impact of signal broadening on imaging, achieving a lateral resolution of 1 cm, which meets the needs of most practical scenarios and enables accurate identification of mannequin model poses. With single modulation, the acquisition time required to restore the structural information of the object is approximately 0.0112 s. With four modulations, image quality can be improved and high-resolution imaging of the object can be achieved. Furthermore, this paper analyzes noise conditions under the scene with ambient light interference and performs frequency domain filtering to partially suppress noise interference, endowing our method with strong noise robustness and improving structural information extraction capability.

(iii) There are various methods to enhance the lateral resolution in imaging, including interpolation [24] and different super-resolution techniques [21,49]. These methods are based on models for data estimation, and the effect depends on the accuracy of the model. However, more refined modeling demands longer computational time [42,43] and often struggles to achieve the required effect and stability for practical detection tasks. In this paper, based on the variation of the DMD mask, the movement of the detection area is precise to sub-pixels, which can obtain more detailed information of the scene and overcome the inherent resolution limitations of the imaging system to obtain high-resolution data. This hardware-based modulation approach enables rapid and accurate high-resolution data acquisition without requiring additional modeling, while maintaining stability and adaptability across diverse scenarios. Our work represents the first integration of DMD modulation into NLOS systems based on array detectors, introducing a novel modulation strategy for this architecture. Future research could explore additional DMD-based modulation schemes [50] or incorporate alternative modulation methods [51,52] into NLOS systems, potentially further enhancing system performance and practical utility. Moreover, recent studies have demonstrated the feasibility of TOF-based imaging through a scattering layer using array detectors [53]. Existing modulation methods, including the proposed DMD modulation method, could be further explored for broader applications.

## 5. Conclusion

To achieve rapid and high-resolution imaging, this paper introduces the DMD modulation for the first time into the scan-free NLOS system, proposing a PSM-NLOS imaging method. The three-dimensional blur kernel derived from signal broadening modeling is incorporated into the non-confocal scene modeling. By leveraging the rapid modulation of the DMD, fine sub-pixel sampling in the spatial dimensions of the scene is achieved. Finally, the object's fine structure is obtained through model inversion. This method achieves an imaging resolution of up to 1 cm and can accurately identify different poses of mannequin models. The time required for a single modulation acquisition is only 0.0112 s, and high-resolution imaging of the object can be achieved with as few as four modulations, which preserves the speed advantage of the system based on the array detector. This method is also applicable in scenes with ambient light interference and

Optics EXPRESS

has certain noise robustness. Additionally, the modulation method is compatible with different methods such as SCBM, virtual waves, and LCT, and can serve as an independent modulation method to enhance imaging resolution.

**Funding.** National Natural Science Foundation of China (62427803, 62031018, U23A20283); Jiangsu Provincial Key Research and Development Program (BE2022391); Fundamental Research Funds for the Central Universities (30924010812).

**Acknowledgments.** The authors thank Yi Wei and Jinye Miao for technical support and experiment discussion.

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

**Data availability.** Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

## References

1. N. Scheiner, F. Kraus, F. Wei, *et al.*, "Seeing around street corners: Non-line-of-sight detection and tracking in-the-wild using doppler radar," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2020), pp. 2065–2074.
2. D. B. Lindell and G. Wetzstein, "Three-dimensional imaging through scattering media based on confocal diffuse tomography," Nat. Commun. **11**(1), 4517 (2020).
3. C. Wu, J. Liu, X. Huang, *et al.*, "Non-line-of-sight imaging over 1.43 km," Proc. Natl. Acad. Sci. **118**(10), e2024468118 (2021).
4. L.-C. Liu, C. Wu, W. Li, *et al.*, "Active optical intensity interferometry," Phys. Rev. Lett. **134**(18), 180201 (2025).
5. M. Isogawa, Y. Yuan, M. O'Toole, *et al.*, "Optical non-line-of-sight physics-based 3d human pose estimation," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2020), pp. 7011–7020.
6. W. Chen, F. Wei, K. N. Kutulakos, *et al.*, "Learned feature embeddings for non-line-of-sight imaging and recognition," ACM Transl. on Graph. **39**(6), 1–18 (2020).
7. S. Chan, R. E. Warburton, G. Gariepy, *et al.*, "Non-line-of-sight tracking of people at long range," Opt. Express **25**(9), 10109–10117 (2017).
8. R. Cao, F. de Goumoens, B. Blochet, *et al.*, "High-resolution non-line-of-sight imaging employing active focusing," Nat. Photonics **16**(6), 462–468 (2022).
9. F. Willomitzer, P. V. Rangarajan, F. Li, *et al.*, "Fast non-line-of-sight imaging with high-resolution and wide field of view using synthetic wavelength holography," Nat. Commun. **12**(1), 6647 (2021).
10. C. A. Metzler, F. Heide, P. Rangarajan, *et al.*, "Deep-inverse correlography: towards real-time high-resolution non-line-of-sight imaging," Optica **7**(1), 63–71 (2020).
11. R. Czajkowski and J. Murray-Bruce, "Two-edge-resolved three-dimensional non-line-of-sight imaging with an ordinary camera," Nat. Commun. **15**(1), 1162 (2024).
12. T. Maeda, Y. Wang, R. Raskar, *et al.*, "Thermal non-line-of-sight imaging," in *International Conference on Computational Photography* (IEEE, 2019), pp. 1–11.
13. D. B. Lindell, G. Wetzstein, and V. Koltun, "Acoustic non-line-of-sight imaging," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2019), pp. 6773–6782.
14. X. Huang, R. Ye, W. Li, *et al.*, "Non-line-of-sight imaging and vibrometry using a comb-calibrated coherent sensor," Phys. Rev. Lett. **132**(23), 233802 (2024).
15. A. Velten, T. Willwacher, O. Gupta, *et al.*, "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," Nat. Commun. **3**(1), 745 (2012).
16. 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).
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. X. Liu, I. Guillén, M. La Manna, *et al.*, "Non-line-of-sight imaging using phasor-field virtual wave optics," Nature **572**(7771), 620–623 (2019).
19. X. Liu, J. Wang, L. Xiao, *et al.*, "Non-line-of-sight imaging with arbitrary illumination and detection pattern," Nat. Commun. **14**(1), 3230 (2023).
20. J. Ye, Y. Sun, W. Li, *et al.*, "Real-time non-line-of-sight computational imaging using spectrum filtering and motion compensation," Nat. Comput. Sci. **4**(12), 920–927 (2024).
21. J. Wang, X. Liu, L. Xiao, *et al.*, "Non-line-of-sight imaging with signal superresolution network," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023), pp. 17420–17429.
22. Z. Zhao, Q. Zhang, X. Li, *et al.*, "High-resolution non-line-of-sight imaging based on liquid crystal planar optical elements," Nanophotonics **13**(12), 2161–2172 (2024).
23. M. Isogawa, D. Chan, Y. Yuan, *et al.*, "Efficient non-line-of-sight imaging from transient sinograms," in *Computer Vision* (2020), pp. 193–208.
24. J. Ye, X. Huang, Z.-P. Li, *et al.*, "Compressed sensing for active non-line-of-sight imaging," Opt. Express **29**(2), 1749–1763 (2021).

25. R. Ding, J. Ye, Q. Gao, *et al.*, "Curvature regularization for non-line-of-sight imaging from under-sampled data," IEEE Trans. Pattern Anal. Mach. Intell. **46**(12), 8474–8485 (2024).
26. X. Liu, J. Wang, L. Xiao, *et al.*, "Few-shot non-line-of-sight imaging with signal-surface collaborative regularization," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023), pp. 13303–13312.
27. S. Zhu, Y. M. Sua, T. Bu, *et al.*, "Compressive non-line-of-sight imaging with deep learning," Phys. Rev. Appl. **19**(3), 034090 (2023).
28. W. Zhang, E. Guo, S. Zhu, *et al.*, "Real-time scan-free non-line-of-sight imaging," APL Photonics **9**(12), 126101 (2024).
29. D. Wang, W. Hao, Y. Tian, *et al.*, "Enhancing the spatial resolution of time-of-flight based non-line-of-sight imaging via instrument response function deconvolution," Opt. Express **32**(7), 12303–12317 (2024).
30. B. Wang, M.-Y. Zheng, J.-J. Han, *et al.*, "Non-line-of-sight imaging with picosecond temporal resolution," Phys. Rev. Lett. **127**(5), 053602 (2021).
31. C. Jin, J. Xie, S. Zhang, *et al.*, "Richardson–lucy deconvolution of time histograms for high-resolution non-line-of-sight imaging based on a back-projection method," Opt. Lett. **43**(23), 5885–5888 (2018).
32. Z. Li, X. Liu, J. Wang, *et al.*, "Fast non-line-of-sight imaging based on first photon event stamping," Opt. Lett. **47**(8), 1928–1931 (2022).
33. G. Musarra, A. Lyons, E. Conca, *et al.*, "Non-line-of-sight three-dimensional imaging with a single-pixel camera," Phys. Rev. Appl. **12**(1), 011002 (2019).
34. W. Yang, C. Zhang, W. Jiang, *et al.*, "None-line-of-sight imaging enhanced with spatial multiplexing," Opt. Express **30**(4), 5855–5867 (2022).
35. J. H. Nam, E. Brandt, S. Bauer, *et al.*, "Low-latency time-of-flight non-line-of-sight imaging at 5 frames per second," Nat. Commun. **12**(1), 6526 (2021).
36. C. Pei, A. Zhang, Y. Deng, *et al.*, "Dynamic non-line-of-sight imaging system based on the optimization of point spread functions," Opt. Express **29**(20), 32349–32364 (2021).
37. R. H. Hadfield, J. Leach, F. Fleming, *et al.*, "Single-photon detection for long-range imaging and sensing," Optica **10**(9), 1124–1141 (2023).
38. J. Tachella, Y. Altmann, N. Mellado, *et al.*, "Real-time 3d reconstruction from single-photon lidar data using plug-and-play point cloud denoisers," Nat. Commun. **10**(1), 4984 (2019).
39. W.-L. Ye, P.-Y. Jiang, Z.-P. Li, *et al.*, "Large-fov 3d imaging of single-photon lidar at up to 12 km," Opt. Lett. **49**(23), 6621–6624 (2024).
40. M. O'Toole, F. Heide, D. B. Lindell, *et al.*, "Reconstructing transient images from single-photon sensors," in *Conference on Computer Vision and Pattern Recognition* (IEEE, 2017), pp. 2289–2297.
41. D. Shin, F. Xu, D. Venkatraman, *et al.*, "Photon-efficient imaging with a single-photon camera," Nat. Commun. **7**(1), 12046 (2016).
42. S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," IEEE Signal Process. Mag. **20**(3), 21–36 (2003).
43. S. Farsiu, D. Robinson, M. Elad, *et al.*, "Advances and challenges in super-resolution," Int. J. Imaging Syst. Tech. **14**(2), 47–57 (2004).
44. B. Sun, M. P. Edgar, R. Bowman, *et al.*, "3d computational imaging with single-pixel detectors," Science **340**(6134), 844–847 (2013).
45. Z.-P. Li, X. Huang, P.-Y. Jiang, *et al.*, "Super-resolution single-photon imaging at 8.2 kilometers," Opt. Express **28**(3), 4076–4087 (2020).
46. S. Khademi, A. Darudi, and Z. Abbasi, "A sub pixel resolution method," arXiv (2010).
47. W. Zhang, S. Zhu, L. Chen, *et al.*, "High-resolution and real-time non-line-of-sight imaging based on spatial correlation," Opt. Lasers Eng. **193**, 109100 (2025).
48. M. jie Sun, M. P. Edgar, D. B. Phillips, *et al.*, "Improving the signal-to-noise ratio of single-pixel imaging using digital microscanning," Opt. Express **24**(10), 10476–10485 (2016).
49. L. Xiao, J. Wang, Y. Wang, *et al.*, "Fast non-line-of-sight imaging with hybrid super-resolution network over 18 m," IEEE Trans. Comput. Imaging **10**, 1439–1448 (2024).
50. S. Jiang, X. Li, Z. Zhang, *et al.*, "Scan efficiency of structured illumination in iterative single pixel imaging," Opt. Express **27**(16), 22499–22507 (2019).
51. J. Wu, F. Yang, and L. Cao, "Resolution enhancement of long-range imaging with sparse apertures," Opt. Lasers Eng. **155**, 107068 (2022).
52. A. Kadambi, R. Whyte, A. Bhandari, *et al.*, "Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles," ACM Trans. Graph. **32**(6), 1–10 (2013).
53. R. Deng, X. Jin, D. Du, *et al.*, "Scan-free time-of-flight-based three-dimensional imaging through a scattering layer," Opt. Express **31**(14), 23662–23677 (2023).