APL Photonics

RESEARCH ARTICLE | DECEMBER 02 2024

Real-time scan-free non-line-of-sight imaging

Wenjun Zhang ; Enlai Guo; Shuo Zhu ; Chenyang Huang ; Lijia Chen; Lingfeng Liu; Lianfa Bai; Edmund Y. Lam ; Jing Han

APL Photonics 9, 126101 (2024)
https://doi.org/10.1063/5.0235687

APL Photonics ARTICLE pubs.aip.org/aip/app

Real-time scan-free non-line-of-sight imaging

Cite as: APL Photon. 9, 126101 (2024); doi: 10.1063/5.0235687
Submitted: 29 August 2024 • Accepted: 10 November 2024 •
Published Online: 2 December 2024

View Online Export Citation CrossMark

Wenjun Zhang,[1] Enlai Guo,[1] Shuo Zhu,[2] Chenyang Huang,[1] Lijia Chen,[1] Lingfeng Liu,[1] Lianfa Bai,[1] Edmund Y. Lam,[2] and Jing Han[1,a)]

AFFILIATIONS

[1] Jiangsu Key Laboratory of Spectral Imaging and Intelligent Sense, Nanjing University of Science and Technology, Nanjing 210094, China
[2] Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong, China

[a)] Author to whom correspondence should be addressed: eohj@njust.edu.cn

ABSTRACT

Real-time non-line-of-sight imaging is crucial for practical applications. Among existing methods, transient methods present the best visual reconstruction ability. However, most transient methods require a long acquisition time, thus failing to deal with real-time imaging tasks. Here, we provide a dual optical coupling model to describe the spatiotemporal propagation of photons in free space, then propose an efficient non-confocal transformation algorithm and establish a non-confocal time-to-space boundary migration model. Based on these, a scan-free boundary migration method is proposed. The data acquisition speed of the method can reach 151 fps, which is ~7 times faster than the current fastest data acquisition method, while the overall imaging speed can also reach 19 fps. The background stability brought by fast scan-free acquisition makes the method suitable for dynamic scenes. In addition, the high robustness of the model to noise makes the method have the capability of non-line-of-sight imaging in outdoor environments during the daytime. To further enhance the practicality of this method in real-world scenarios, we exploit the statistical prior and propose a plug-in-and-play super-resolution method to extract higher spatial resolution signals, reducing the detector array requirement from 32 × 32 to 8 × 8 without compromising imaging quality, thus reducing the device expense of detectors.

I. INTRODUCTION

Non-line-of-sight imaging (NLOS) obtains the shape and position of objects outside the line of sight by capturing and analyzing light scattered from multiple surfaces. In recent years, NLOS imaging technology has garnered significant attention in areas such as medical imaging, robot vision, transportation, security monitoring, and fundamental science.[1–6] Various methods have been proposed to achieve NLOS imaging, including speckle correlation,[7–12] wavefront shaping,[13–15] thermal imaging,[16] acoustic imaging,[17,18] occlusion-based imaging,[19–21] and transient imaging.[2,5,22–26] Among these techniques, transient imaging stands out for its ability to generate high-fidelity three-dimensional reconstructions by leveraging rich temporal information acquired from time-resolved sensors. Transient NLOS imaging methods acquire time-resolved images of light propagation by capturing the temporal pulse response of the scene. The time distribution and statistics of the time-of-flight (ToF) histogram returned by the signal photons are related to the length of the travel path in the detection scene. By utilizing computational imaging methods, it is possible to perceive and image objects in various environments based on this path information, thereby retrieving the three-dimensional position information of hidden objects. Most transient methods have high imaging resolution but require dense scanning on the relay wall, and each scanning point needs to be sampled multiple times. As a result, the overall process requires a considerable amount of data acquisition time. However, prolonged exposure times are not always suitable for applications such as autonomous driving, which impedes the practical implementation of NLOS imaging.

To achieve fast imaging in NLOS scenarios, various methods have been proposed. Some methods aim to reduce the number of scanning points to decrease data acquisition time. By using a sparse scanning method based on the correlation between measurement data, the data acquisition time can be reduced by more than 20% while maintaining the exact resolution.[27] By changing the scanning pattern to a circular scan mode, the data acquisition time can be reduced to ~1.6% of the traditional acquisition time, but this approach only allows for the reconstruction of the approximate structure of the object.[28] By introducing virtual confocal signals, a confocal complemented signal-object collaborative regularization algorithm is designed, which can achieve high-quality reconstructions with coarse rather than dense measurements, and the acquisition time can be reduced significantly.[26] By introducing compressed sensing methods and incorporating non-negative constraints into the algorithm SPIRAL-TAP,[29] the number of sampling points can be reduced to 5×5.[30] In addition, it is possible to reconstruct the object from under-sampled data by introducing regularization terms.[31,32] Despite their ability to reduce acquisition time, these iterative methods often suffer from time-consuming reconstruction processes, making it challenging to achieve real-time imaging. To balance acquisition time and reconstruction time, methods based on deep learning have been proposed to learn operators that can recover high spatial resolution signals[33] or learn reconstruction results from under-sampled data directly.[34] Leveraging in-depth physics information is also one of the efficient methods with limited sampling data.[35] These methods can reduce the number of sampling points to a maximum of 4 × 4, enabling near real-time imaging. However, deep learning methods have limited generalization capabilities and may also have certain constraints in practical usage. In addition to reducing the number of scanning points, another approach is to decrease the acquisition time for each individual scanning point. A NLOS method based on time-sequential first photon has been proposed, which enables reconstruction based on sparse photons, thus reducing data acquisition time.[36] However, due to the large variance in histogram signals, there is a certain decrease in reconstruction quality, and the algorithm requires multiple iterations, resulting in a longer processing time compared to direct imaging methods, making real-time imaging difficult to achieve.

Some methods achieve fast measurements by using single-photon avalanche diode (SPAD) arrays or incorporating modulation techniques. The use of spatial multiplexing detection in NLOS imaging enables higher data acquisition efficiency.[37] Compared to point-by-point scanning, the designed orthogonal patterns can increase the data acquisition speed by five times with only a small array of 16 × 16 detection points. However, the overall data acquisition time is still relatively long, and the reconstruction process also takes ~3 s. By remapping the data collected from two specially designed 16 × 1 SPAD array detectors and combining it with fast reconstruction algorithms, it is possible to achieve a data acquisition rate of 5 fps and low-latency reconstruction.[38] The NLOS data acquisition rate can be implemented at 20 fps using a 32 × 32 pixel SPAD array and an optimization-based computational method.[39] Nevertheless, due to the iterative optimization required in the reconstruction process, real-time imaging remains challenging to achieve.

In addition to real-time reconstruction, reducing device costs is an urgent need for the practical implementation of NLOS systems. Using SPAD arrays instead of laser scanning offers advantages in terms of data acquisition speed. However, the imaging resolution of this method is limited by the number of pixels in the SPAD array. To achieve satisfactory imaging results, at least a 32 × 32 pixel SPAD array is required. If further improvements in imaging performance are desired, the demand for a higher number of arrays increases, which imposes stricter manufacturing requirements and overall costs for SPAD devices. Therefore, the important issue to address for the practical application of the system is reducing the number of array SPAD pixels while ensuring imaging quality.

In this paper, a scan-free boundary migration (SCBM) method is proposed, which significantly reduces the acquisition time from tens of minutes to 6.6 ms and achieves real-time object reconstruction. A time-to-space boundary migration model based on non-confocal scan-free scenes is established, and an efficient non-confocal transformation algorithm is proposed, which is the basis of the proposed method. In addition, the proposed SCBM method can image moving objects, and we demonstrate the imaging capability in outdoor daylight environments, which holds significant value across various applications. Furthermore, a self-similarity super-resolution (S^3R) algorithm is proposed to restore high spatial resolution signals, which allows for a reduction in the pixel count of the array detector to 8 × 8 without compromising imaging quality, thereby greatly reducing the device expense of the detector. Our method has been experimentally validated for its imaging capability on objects with different shapes and materials in different environments.

II. METHOD

A. Non-confocal time-to-space boundary migration model

The non-confocal time-to-space boundary migration model and SCBM method proposed in this paper are based on the imaging system shown in Fig. 1. The light emitted by the pulsed laser first illuminates a point on the relay wall, which then diffusely reflects and illuminates the entire hidden object scene. The array of SPAD records the ToF information of photons that diffuse from the object surface to the detection field of view (FoV). Each pixel in the SPAD array represents an independent ToF detection unit, focusing through a lens to detect transient waves containing information about the hidden objects in different small regions within the FoV. The complete three-dimensional measurement is produced with just one exposure of the non-confocal system.

Because light transport is linear in space and time-invariant, the light transport through the scene can be characterized by a response function H.[24,40,41] In the non-confocal NLOS system, the propagation process of photons can be described as follows:[42]

$$Y = H(O) + W, \tag{1}$$

where $Y \in \mathbb{R}^{M\times M\times T}$ is the three-dimensional space–time measurement result received by the SPAD array, which has $M \times M$ pixel size and T time bins. $O \in \mathbb{R}^{X\times Y\times Z}$ is the object, and W represents the additive random noise. Due to the non-confocal configuration of the system, the incident point and detection point of the laser on the relay wall are different. The propagation of photons is modeled as a dual-path coupling process. As shown in Fig. 1, the incident light diffusely reflects off the relay wall and illuminates the entire hidden object scene. Subsequently, the light is reflected by the hidden object and propagated to the relay wall; both of the two processes are free-space propagation. The imaging model assumes no mutual reflection or occlusion between parts of the hidden scene. The first free propagation field is denoted by $\phi_1(x,y,z,t)$, where (x,y,z) denotes the three-dimensional space and t denotes the time. The function $g(x,y,z)$ describes the direction and intensity of photon propagation upon reaching a point on the hidden object, which

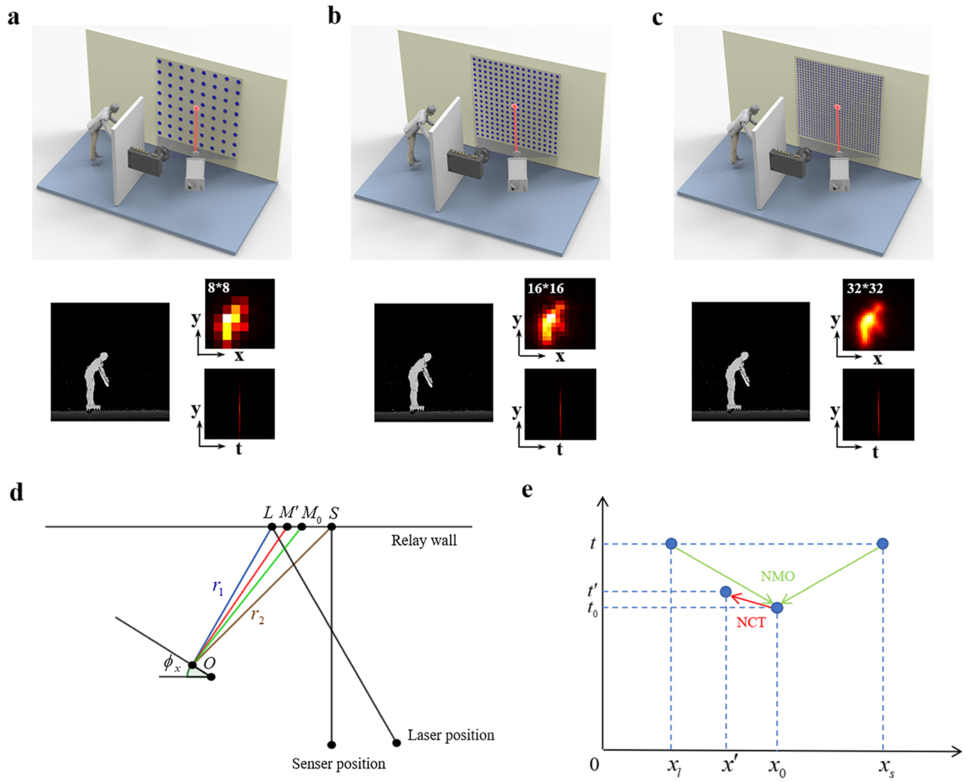

FIG. 1. Non-confocal scan-free imaging system model. (a) Performing numerical simulations to validate the performance of the scan-free NLOS imaging method. We simulate a non-confocal NLOS imaging scenario based on an 8 × 8 SPAD array detector, with a hidden scene being a two-dimensional target in the bottom left corner of the image. The simulated data have a spatial resolution of 8 × 8, and the reconstruction result in the bottom right corner of the image is obtained using the proposed SCBM method. (b) Using the S^3R algorithm to perform twofold super-resolution on the simulated data in (a), the equivalent NLOS imaging scene model diagram and the reconstructed object result are obtained. (c) Performing four-fold super-resolution to the detector data in (a), the equivalent scene model diagram and the reconstructed result are obtained. (d) Corrected optical path diagram. The black line passing through point O represents the tangent to the surface of the hidden object at that point. The green color represents the optical path after NMO correction, where point M_0 is the midpoint between points L and S. The red color represents the optical path after NCT correction. (e) NMO correction diagram and NCT correction diagram.

includes the albedo, surface normal, bidirectional scattering distribution function, and so on. Assuming that the surface structure of the hidden object does not add an additional optical path, i.e., the depth difference of the object is ignored, then we can ignore the hidden object's contribution to the spatial variability in the depth dimension of the field. We repeat the $g(x,y,z)$ in the fourth dimension T times, and $g_T(x,y,z)$ is constructed. The second free propagation field is denoted as $\phi_2(x,y,z,t)$. The response function H can be expressed as

$$H = [\phi_1(x,y,z,t) \times g_T(x,y,z)] * \phi_2(x,y,z,t), \tag{2}$$

where $*$ denotes the convolution.

Directly calculating H and further inverting the imaging model for the reconstruction of hidden objects is a complex task. In order to tackle this issue, a time-to-space boundary migration model based on non-confocal scan-free scenes is established, referring to the $f-k$ method.[2] For simplicity, assume that the relay wall is a plane at $z=0$. Consider the NLOS imaging scenario in which the pulsed laser illuminates a point $L(x_l, y_l, z=0)$ on the relay wall, and a detection point of the SPAD array on the relay wall is denoted as $S(x_s, y_s, z=0)$. The measurements Y can be expressed as the spatial boundary of the field $\phi_2(x,y,z,t)$ after calibrating the distance among the laser, the SPAD, and the front surface of the relay wall, i.e., $Y = \phi_2(x,y,z=0,t)$.

The incident light scatters to the surface of the hidden object. We have assumed that the surface structure of the hidden object does not add an additional optical path. In this instance, the origin of the second free propagation field, i.e., $\phi_2(x,y,z,t=0)$, can be approximated as a steady state field in the time domain, which is defined as the temporal boundary of the field $\phi_2(x,y,z,t)$. Therefore, the temporal boundary $\phi_2(x,y,z,t=0)$ can be denoted as the time integral of the field $\phi_1(x,y,z,t)$ changed by the object[42]

$$\phi_2(x,y,z,t=0) = \int \phi_1(x,y,z,t) \times g_T(x,y,z)dt. \tag{3}$$

Since $\phi_2(x,y,z,t=0)$ represents the main characteristics of the hidden object, it is possible to approximate the object as the temporal boundary of $\phi_2(x,y,z,t)$, i.e., $O \approx \phi_2(x,y,z,t=0)$.

In addition, there are some errors in approximating the object as the temporal boundary of $\phi_2(x, y, z, t)$. In order to tackle this issue, we first need to convert the non-confocal scene into the confocal scene. In confocal NLOS scenes, following O'Toole *et al.*, we interpret the hidden scene as a 3D volume where each point emits a spherical wave at $t = 0$.[23] The first free propagation field is denoted as $\phi_1'(x, y, z, t)$, and the second free propagation field is denoted as $\phi_2'(x, y, z, t)$. Then, the virtual illumination of the object can be seen as the temporal origin of the second free propagation field, i.e., $O = \phi_2'(x, y, z, \quad t = 0)$.[2] Therefore, we can estimate the hidden object O from the measurements in confocal NLOS scenes, which is a boundary value problem that requires us to migrate the field $\phi_2'(x, y, z, t)$ from the spatial boundary (i.e., $z = 0$) to the temporal boundary (i.e., $t = 0$),

$$\phi_2'(x, y, z = 0, t) \Rightarrow \phi_2'(x, y, z, t = 0). \tag{4}$$

To correct non-confocal data to confocal data, the normal moveout (NMO) method is proposed.[2] However, the NMO method requires that the normal of the hidden object be perpendicular to the relay wall, which is not generally applicable in most object scenarios.[43] As shown in Figs. 1(d) and 1(e), taking inspiration from dip moveout correction in seismic imaging, the non-confocal transformation (NCT) correction method is proposed, which allows for the tilted normal of the objects relative to the intermediate surface. By incorporating prior knowledge of the laser point and detection point locations, the accurate correction of data can be achieved.

Considering that the surface normal of the object will not always be perpendicular to the relay wall, let ϕ_x be the tilt angle of a voxel point of the hidden object relative to the relay wall plane; $n_x = \sin\phi_x$ and $n_z = -\cos\phi_x$ are the corresponding normal vectors. The tilt angle of the object is added as a parameter to the correction process, and the relationship between the corrected data Y' and the data Y can be expressed as follows:[44]

$$Y'(t', x') \equiv Y(t, x), \tag{5}$$

where

$$t' = \frac{t_0}{\sqrt{1 + \left(\frac{h}{t_0}\frac{2n_x}{v}\right)^2}}, \tag{6}$$

$$x' = x_0 - \frac{h^2}{\sqrt{1 + \left(\frac{h}{t_0}\frac{2n_x}{v}\right)^2} t_0}\frac{2n_x}{v}, \tag{7}$$

$$t_0 = \sqrt{t^2 - \frac{1}{v^2}\left(|x_s - x_l|^2 + |y_s - y_l|^2\right)}, \tag{8}$$

where h is half of the offset between S point and L point, x' and t' are the confocal position and photon flight time after angle correction, respectively, $x_0 = \frac{x_l + x_s}{2}$, and $v = c$.

In practical systems, the tilt angle of the hidden object relative to the relay wall is unknown, making it impossible to directly calculate the corrected time. We calculate the tilt angle equivalently according to the geometric relationship, apply a stretch to the data along the time domain, use a phase shift in the Fourier domain, and then obtain a closed-form solution to the NCT correction.[45] By incorporating prior knowledge of the laser point and detection point locations, the SPAD-detected data Y can be corrected to confocal scene data Y' accurately. Please refer to the supplementary material for a detailed derivation of the NCT method.

The wave equation is a linear second-order partial differential equation describing the propagation of mechanical waves or light waves in free space. Therefore, the propagation of $\phi_2'(x, y, z, t)$ in space and time is determined by the wave equation

$$\nabla^2\phi_2'(x, y, z, t) - \frac{1}{v^2}\frac{\partial^2\phi_2'(x, y, z, t)}{\partial t^2} = 0, \tag{9}$$

where the Laplace operator $\nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is defined in the space dimension. To match the converted confocal scene, the scalar v is set to half the speed of light in the air.

In order to realize the boundary migration, it is necessary to solve the analytical solution $\phi_2'(x, y, z, t)$ of the wave equation, which can be solved as follows:[2]

$$\phi_2'(x, y, z, t) = \iiint \Phi_2'(k_x, k_y, k_z) e^{2\pi i(k_x x + k_y y + k_z z - f t)} dk_x dk_y dk_z, \tag{10}$$

where $\Phi_2'(k_x, k_y, k_z, t)$ is the spectrum of $\phi_2'(x, y, z, t)$ on the spatial dimensions (x, y, z). $\boldsymbol{k} = 2\pi \cdot (k_x, k_y, k_z)$ is a wave vector representing the propagation direction of a single plane wave, which is related to the wavelength of light λ, $|\boldsymbol{k}| = \frac{2\pi}{\lambda}$. The wave vector $\boldsymbol{k}$ and the frequency f can be related by the dispersion relation $f = v\sqrt{k_x^2 + k_y^2 + k_z^2}$.

When $t = 0$, the function $\phi_2'(x, y, z, t)$ can be expressed as

$$\phi_2'(x, y, z, t = 0) = \iiint \Phi_2'(k_x, k_y, k_z) e^{2\pi i(k_x x + k_y y + k_z z)} dk_x dk_y dk_z. \tag{11}$$

Note that $\phi_2'(x, y, z, t)$ and $\Phi_2'(k_x, k_y, k_z, t)$ are related through a Fourier transform when $t = 0$. With the dispersion relation and its Jacobian, Eq. (10) can be rewritten as

$$\phi_2'(x, y, z, t) = \iiint \overline{\Phi_2'}(k_x, k_y, f) e^{2\pi i(k_x x + k_y y + k_z z - f t)} dk_x dk_y df, \tag{12}$$

where

$$\Phi_2'(k_x, k_y, k_z) = \frac{v|k_z|}{\sqrt{k_x^2 + k_y^2 + k_z^2}} \overline{\Phi_2'}\left(k_x, k_y, v\sqrt{k_x^2 + k_y^2 + k_z^2}\right). \tag{13}$$

When $z = 0$, the function $\phi_2'(x, y, z, t)$ can be expressed as

$$\phi_2'(x, y, z = 0, t) = \iiint \overline{\Phi_2'}(k_x, k_y, f) e^{2\pi i(k_x x + k_y y - f t)} dk_x dk_y df. \tag{14}$$

Note that $\phi_2'(x, y, z, t)$ and $\overline{\Phi_2'}(k_x, k_y, f)$ are related through a Fourier transform when $z = 0$. Therefore, the correspondence between object $\phi_2'(x, y, z, \quad t = 0)$ and measurements $\phi_2'(x, y, z = 0, t)$ is established via interpolation transformation within the frequency domain. The reconstruction of the object is the reverse procedure of the aforementioned derivation, and the detailed process can be found in the supplementary material.

B. Super-resolution algorithm

In scanning-based NLOS reconstruction, it is necessary to perform scanning on an N × N grid to image the scene with N × N resolution.[46,47] The required number of sampling points is a few times the number of non-zero elements in the object. When the sample count is sufficient, reducing the number of samples has little impact on the reconstruction performance.[48] However, when the number of sampling points is low, e.g., 32 × 32, further reducing the number of sampling can negatively affect the imaging quality.[30] In scan-free non-confocal systems, the pixel count of the array detector determines the number of samples. The commonly used SPAD array pixel numbers are 32 × 32[39] or 64 × 64[49] at present, but such detector arrays are relatively expensive. Some SPAD arrays have pixel counts of only 4 × 4,[50] 8 × 4,[51] or 16 × 1.[52] A lower sampling quantity will affect the resolution of acquired signals and consequently impact the reconstruction quality. As shown in Fig. 1(a), we simulate a non-confocal NLOS imaging scenario based on an 8 × 8 SPAD array detector.[4] It can be observed that the structural information of the object is difficult to recover under low sampling rates. To address this issue, the S^{3}R method is proposed to enhance the imaging resolution while reducing the requirement for pixel count in the SPAD array, ensuring reconstruction accuracy.

The methods of reconstructing hidden objects from low-resolution signals can be expressed as

$$I(x, y, z) = F_l(Y_l'), \tag{15}$$

where $I(x, y, z)$ is the recovered results of the hidden objects, while F_l denotes the imaging algorithms designed for low-resolution signals Y_l'. From a mathematical perspective, NLOS imaging falls within the domain of inverse problems, and the objective is to reconstruct the surface of the hidden target using measured signals. Due to the inherent structure of the physical model[53] and significant measurement noise,[54] NLOS inverse problems exhibit ill-posed. When the number of measurement points is limited, the lack of data leads to an insufficient rank in the measurement matrix, making the reconstruction task more challenging.[32] To address this issue, several methods have been proposed that utilize regularization constraints and iterative optimization for object recovery.[30–32] By scanning only a few sampling points, the total collection time can be reduced. However, this also leads to a decrease in the reconstruction quality.[33]

The methods of reconstructing hidden objects from high-resolution signals can be expressed as

$$I(x, y, z) = F_h(Y_h'), \tag{16}$$

where F_h represents the imaging algorithms designed for high-resolution signals Y_h'. As mentioned above, a lot of fast reconstruction work has been proposed.[2,23,55,56]

To achieve high-quality reconstruction of objects based on a small array detector, the paper proposes a super-resolution algorithm for three-dimensional signals: S^{3}R. This algorithm can map low-resolution signals to high-resolution signals without any external training images based on the self-similarity of signals,

$$Y_h' = T(Y_l'). \tag{17}$$

The specific super-resolution scheme can be divided into the following steps: (1) Dividing the low-resolution signal Y_l' into T low-resolution images based on their temporal distribution and down-sampling the pth image Y_{lp}' $(1 \le p \le T)$ to obtain Y_{lpD}'. (2) For each patch $P(\boldsymbol{t}_i)$ centered at position $\boldsymbol{t}_i = (t_i^x, t_i^y)^T$ in the image Y_{lp}', the PatchMatch algorithm[57,58] is used to estimate the transformation matrix G_i. Considering possible geometric transformations of $P(\boldsymbol{t}_i)$, the matrix G_i maps $P(\boldsymbol{t}_i)$ to its most matching patch $Q_D(\boldsymbol{t}_i, \theta_i)$ in the down-sampled image Y_{lpD}'. (3) Extracting $Q(\boldsymbol{t}_i, \theta_i)$ from the image Y_{lp}'. $Q(\boldsymbol{t}_i, \theta_i)$ is the high-resolution version of the patch $Q_D(\boldsymbol{t}_i, \theta_i)$. (4) Based on $Q(\boldsymbol{t}_i, \theta_i)$ and the inverse of the transformation matrix G_i, the super-resolved result $P_h(\boldsymbol{t}_i)$ in the image Y_{hp}' can be obtained, which is the high-resolution version of the patch $P(\boldsymbol{t}_i)$. (5) Repeating steps 2–4 for all patches to estimate the super-resolved image Y_{hp}'. (6) Repeat steps 2–5 and get all the estimated Y_{hp}', then integrate them into a high-resolution three-dimensional matrix Y_h'.

In the second step, the PatchMatch algorithm is used for nearest-neighbor field estimation,[57,58] and the objective function for the nearest-neighbor field estimation problem takes the form as

$$\min_{\boldsymbol{\theta}_i} \sum_{i \in \Omega} \boldsymbol{E}_{app}(\boldsymbol{t}_i, \boldsymbol{\theta}_i) + \boldsymbol{E}_{scale}(\boldsymbol{t}_i, \boldsymbol{\theta}_i), \tag{18}$$

where θ_i is the unknown parameter set used to construct the transformation matrix G_i that needs to be estimated. $\boldsymbol{E}_{app}$ is used to measure the similarity between the patch $P(\boldsymbol{t}_i)$ and $Q_D(\boldsymbol{t}_i, \theta_i)$, where $Q_D(\boldsymbol{t}_i, \theta_i)$ can be obtained by using the matrix G_i with parameter θ_i. The Gaussian-weighted sum-of-squared distance in the RGB space is used as the metric,

$$\boldsymbol{E}_{app}(\boldsymbol{t}_i, \boldsymbol{\theta}_i) = \| W_i(P(\boldsymbol{t}_i) - Q_D(\boldsymbol{t}_i, \boldsymbol{\theta}_i)) \|_2^2, \tag{19}$$

where the matrix W_i is the Gaussian weights with $\sigma^2 = 3$.

Due to allowing continuous geometric transformations, the nearest-neighbor field often converges to the trivial solution. To avoid such trivial solutions, one approach is to introduce the scale cost $\boldsymbol{E}_{scale}$,

$$\boldsymbol{E}_{scale}(\boldsymbol{t}_i, \boldsymbol{\theta}_i) = \lambda_{scale} \max\left(0, SRF - Scale(G_i)\right), \tag{20}$$

where SRF indicates the desired super-resolution factor, e.g., 2×, 4×, or 8×, and the function $Scale(\cdot)$ denotes the scale estimation of a projective transformation matrix,

$$Scale(G_i) = \sqrt{det\left(\begin{bmatrix} G_{1,1} - G_{1,3}G_{3,1} & G_{1,2} - G_{1,3}G_{3,2} \\ G_{2,1} - G_{2,3}G_{3,1} & G_{3,1} - G_{2,3}G_{3,2} \end{bmatrix}\right)}, \tag{21}$$

where $G_{u,w}$ indicates the value of the uth row and wth column in the transformation matrix G_i with $G_{3,3}$ normalized to one.[59]

The overall flow of the scan-free NLOS imaging method is illustrated in Figs. 2(a)–2(d). Since the time-correlated single-photon counting (TCSPC) of different pixels of the array detector PF32 used operates independently, a preprocessing step is required to synchronize the time bins of each detection point. Then, based on the non-confocal time-to-space boundary migration model, the NCT algorithm is employed to convert non-confocal data into confocal data. Subsequently, a super-resolution algorithm is applied to the low-resolution signals Y_l' to recover high-resolution signals Y_h'. Finally, object recovery can be performed through the boundary migration.

The S^{3}R algorithm can be used to obtain a high-resolution three-dimensional matrix Y_h', leading to improved reconstruction quality. By comparing the signal images in Figs. 2(g) and 2(h), it

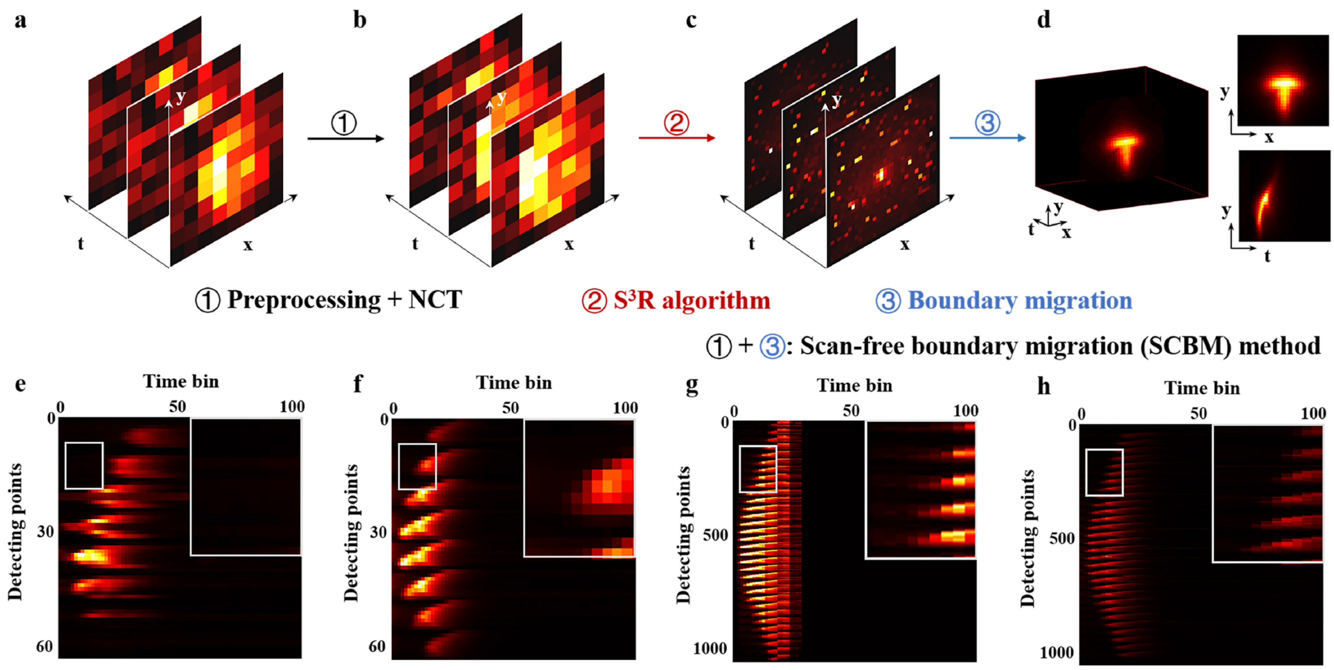

FIG. 2. Schematic diagram of the proposed scan-free NLOS imaging method. (a) Initial data acquired using an 8 × 8 pixels detector. (b) Corrected data obtained by preprocessing and non-confocal correction of the data in (a). (c) High-resolution data obtained by applying 4× super-resolution to the data in (b). (d) Object reconstruction result based on the data in (c). (e)–(g) Two-dimensional display of the data in (a)–(c). (h) Two-dimensional display of the data acquired using a 32 × 32 pixels detector.

can be observed that the S^3R algorithm accurately recovers the ToF information of the object.

III. RESULTS AND ANALYSIS

A. Validation experiment of the scan-free NLOS method

For the non-confocal scan-free system adopted in this paper, we utilize a SPAD array to capture measurements, significantly reducing the data acquisition time. However, compared with a single-pixel system, SPAD array detectors have more substantial noise due to the crosstalk between pixels, resulting in a lower signal-to-noise ratio (SNR) in the acquired data, further exacerbating the difficulty of non-confocal reconstruction.[39] We compare the imaging performance of different fast reconstruction algorithms, as shown in Fig. 3. The spatial resolution of the collected data is 32 × 32, with a temporal resolution of 1024, and each temporal bin spans 55 ps. The acquisition time for each object data is 1 s. Considering that the test scenario is non-confocal while the light-cone transform (LCT) method[23] is designed for confocal scenarios, the proposed NCT method is used to perform a non-confocal transformation on the data when using the LCT algorithm. The reconstruction results from the proposed SCBM method show that it has the lowest noise and high-fidelity restoration effects for different objects. The phasor field method[24] exhibits better restoration of object structures but poor robustness to noise. The LCT method performs poorly in restoring object structures, with noise levels falling between the virtual wave and our method. The boundary migration model demonstrates strong robustness to noise without the need for *a priori* SNR, making it highly practical for various scenarios.[2] The experimental results demonstrate that the proposed SCBM method exhibits resistance to noise, possesses robust object recovery capability in low SNR conditions, and has significant advantages in scan-free scenarios.

The proposed scan-free NLOS imaging method based on a small array detector is quantitatively evaluated through the following experiments. The hidden object in this experiment is a diffusive target "T." The initial data acquisition has a spatial resolution of 32 × 32. In order to explore the impact of the pixel quantity in the array detector on the imaging results, the FoV area of the detection region on the relay wall is controlled to remain constant. The spatial dimension of the collected data is down-sampled, and the detection pixels are selected at equal intervals from the horizontal and vertical directions of the detector, resulting in data with spatial resolutions of 16 × 16 and 8 × 8, respectively. The proposed SCBM method is utilized for object recovery, and the reconstruction results of data with different spatial resolutions are shown in Figs. 4(b1)–4(d1). It can be observed that when there are 32 × 32 detection points, the three-dimensional shape of the hidden object can be largely recovered, although there may be some limitations in retrieving fine details. However, when reducing the number of detection points to 16 × 16, the reconstruction quality noticeably decreases. Furthermore, with only 8 × 8 detection points, the object shape becomes difficult to discern.

Using the S^3R algorithm to perform super-resolution by factors of 2, 4, and 8 on the data with different spatial resolutions and

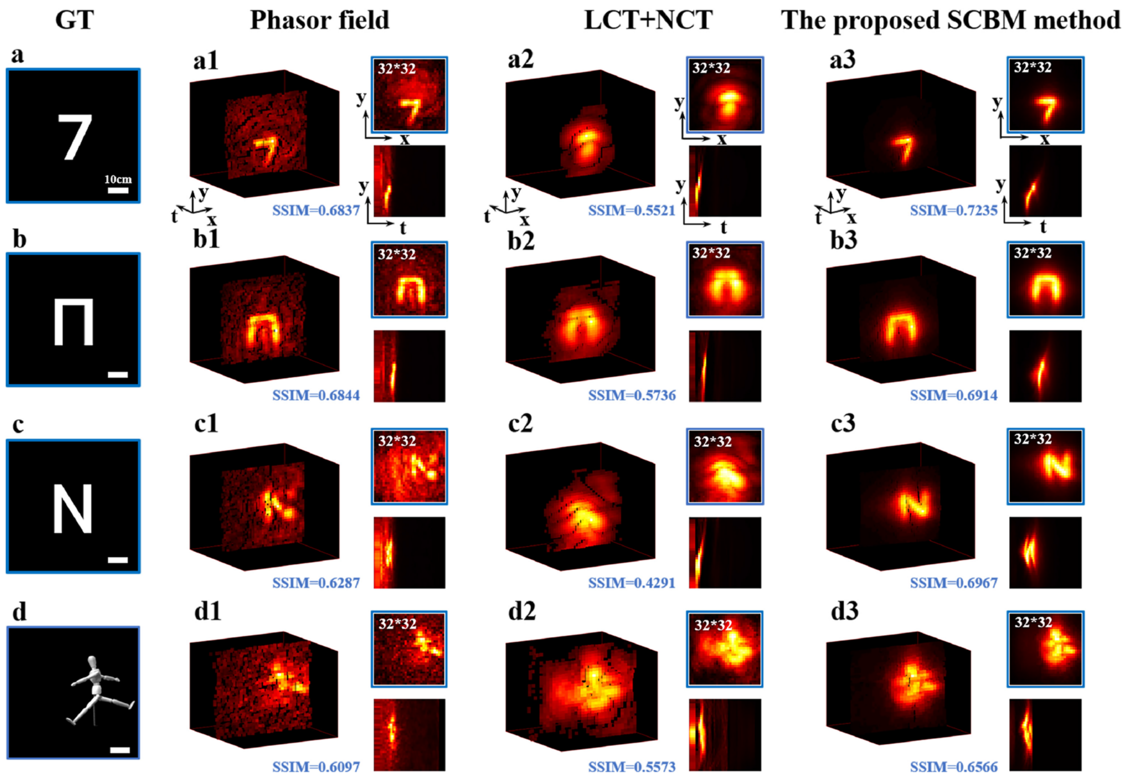

FIG. 3. Comparative results of three fast reconstruction algorithms for different objects. (a)–(d) Hidden diffusive objects with different shapes. (a1)–(d1) Reconstruction results with the phasor field method. (a2)–(d2) Reconstruction results with the LCT method with NCT correction. (a3)–(d3) Reconstruction results with the proposed SCBM method.

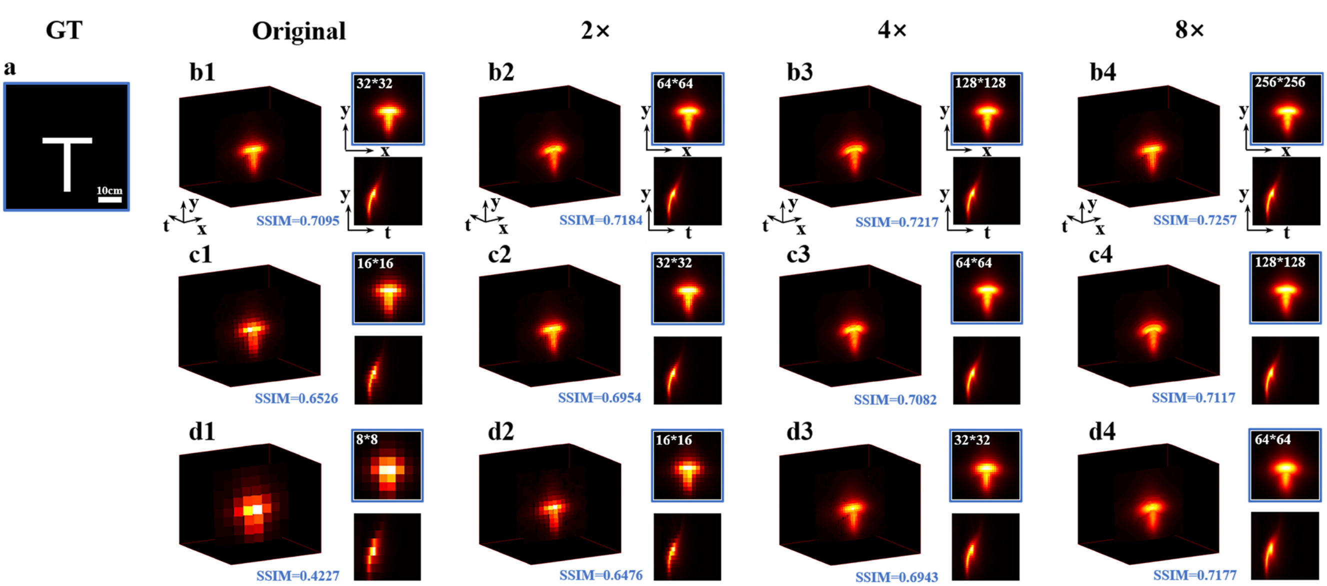

FIG. 4. Comparative results of super-resolution reconstruction for different spatial resolution data. (a) Hidden diffusive target "T." (b1)–(d1) Reconstruction results for data with spatial resolutions of 32×32, 16×16, and 8×8, respectively. The structural similarity index measure between ground truth and the frontal view of the reconstructed object is indicated in blue font below the corresponding figure. (b2)–(b4), (c2)–(c4), and (d2)–(d4) Super-resolution by factors of 2, 4, and 8 applied to the data in (b1)–(d1), followed by reconstruction to obtain object recovery results.

then reconstructing the super-resolved data. It is observed that when there are a sufficient number of detection points (32 × 32), the S^3R algorithm improves the recovery of fine details in the reconstructed results as the super-resolution factor increases. If the number of detection points decreases to 16 × 16, the use of the super-resolution method significantly enhances the object reconstruction capability, with a structural similarity index measure (SSIM) reaching 0.7 between the reconstructed object and ground truth (GT). When the number of detection points drops from 32 × 32 to 8 × 8, the SCBM method struggles to recover the object shape. However, with the S^3R algorithm, there is no significant deterioration in reconstruction quality. This indicates that our method enables accurate object reconstruction using detectors with smaller arrays, thereby greatly reducing the fabrication difficulty and cost of the detector.

In order to further validate the generality of the proposed method, the reconstructions of objects with different shapes are conducted. Figure 5 presents the reconstruction results for different objects. The first column shows the reconstruction results when the number of detection points is 32 × 32. The second column shows the reconstruction results when the number of detection points is 8 × 8. By applying twofold and four-fold super-resolution to the data in the second column, the object recovery results shown in the third and fourth columns are obtained, respectively. It can be observed that the SCBM method is capable of effectively restoring the

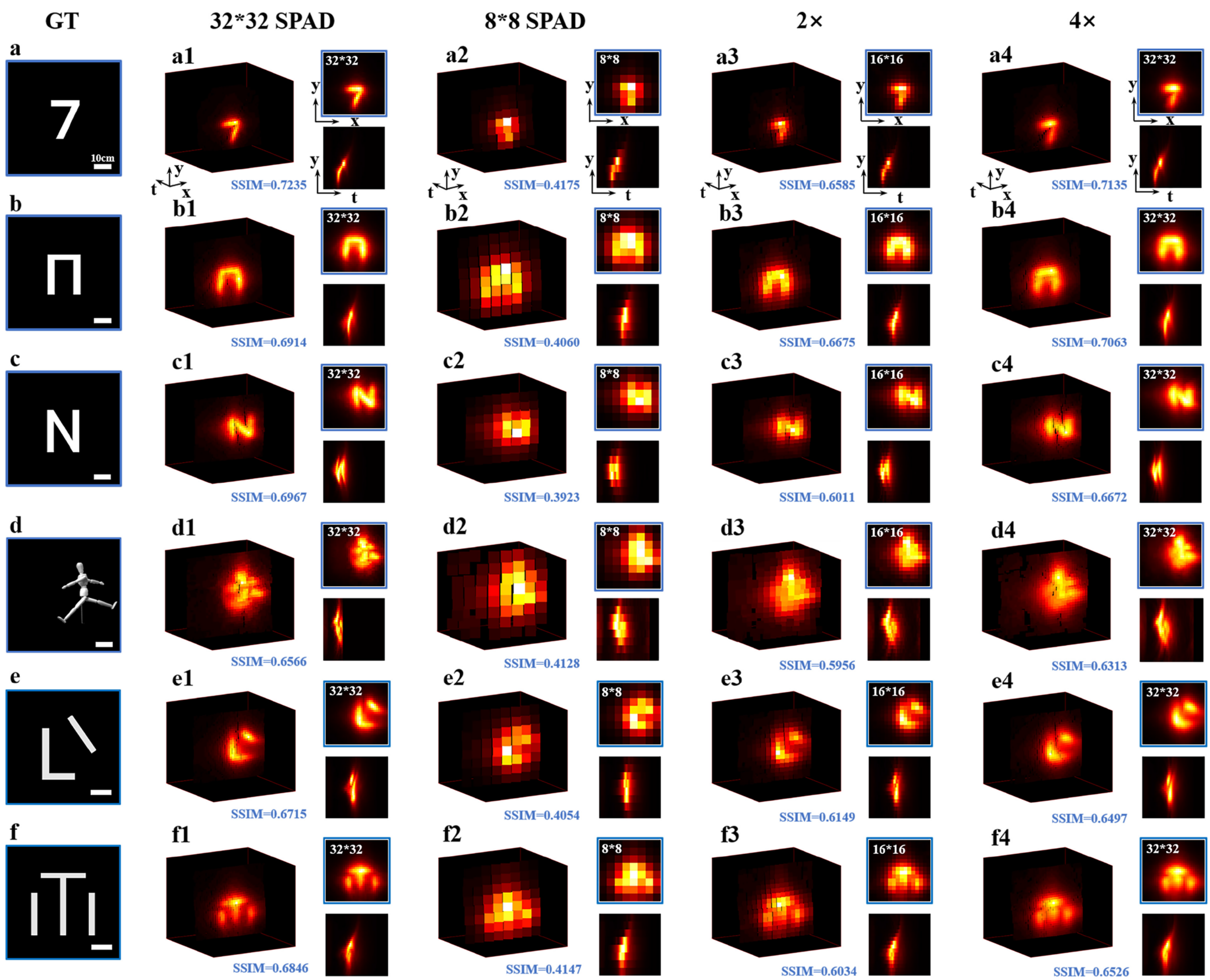

FIG. 5. Comparative results of super-resolution reconstruction for different objects. (a)–(f) Hidden diffusive objects with different shapes. (a1)–(f1) Reconstruction results for data with the spatial resolution of 32 × 32. (a2)–(f2) Reconstruction results for data with the spatial resolution of 8 × 8. (a3)–(f3) Applying twofold super-resolution to the data in (a2)–(f2), followed by reconstruction to obtain object recovery results. (a4)–(f4) Applying four-fold super-resolution to the data in (a2)–(f2), followed by reconstruction to obtain object recovery results. The SSIM values between the front view of the reconstructed results and the GT are listed below the reconstructed results.

structural information of objects with different shapes. Furthermore, by utilizing the S^3R algorithm, it is possible to reconstruct the object structure based on a small number of detection points, achieving imaging results that can essentially match the reconstruction quality of the original data with a spatial resolution of 32 × 32.

The reconstruction results of objects with different materials are shown in the supplementary material. In single-object scenarios where the object material is uniform, there may be variations in the reconstructed details for objects with different materials, but the fundamental shape of the object can still be recovered. In the scenarios of the two objects with different materials, there are differences in the recovery intensities of the two objects in the reconstruction results. By employing the super-resolution reconstruction method, accurate object reconstruction can still be achieved based on a small array detector.

In addition, we test the applicability of the proposed S^3R algorithm to different reconstruction methods, which is shown in the supplementary material. It can be observed that the proposed S^3R method can accurately restore higher resolution data, which can be applied as a plug-in-and-play module and does not restrict the use of reconstruction methods.

B. Dynamic results and the highest achievable frame rate of the imaging system

In order to evaluate whether the proposed scan-free NLOS imaging method can be used for real-time imaging of moving objects, it is necessary to test the overall data acquisition time and object reconstruction time. As shown in Fig. 6, the measurement results of target "T" collected with different exposure times are reconstructed. It can be observed that if the repetition time is not high and is further reduced, the data are more affected by noises, and the histogram obtained by statistics may be difficult to fully reflect the structural information and position of the hidden objects, leading to the decline of the reconstruction quality.[31,38,60] When the repetition time is less than 500, the reconstruction results exhibit significant noise and poor quality. When the number of detection points is reduced from 32 × 32 to 8 × 8, the SCBM method fails to recover the structural information of the object. However, based on the proposed S^3R algorithm in this paper, it is possible to effectively restore the object with a spatial resolution of 32 × 32, achieving recovery results comparable to those of using large-array detectors.

The data acquisition time and reconstruction time for different repetition times in Fig. 6 are listed in Table I, and the frame

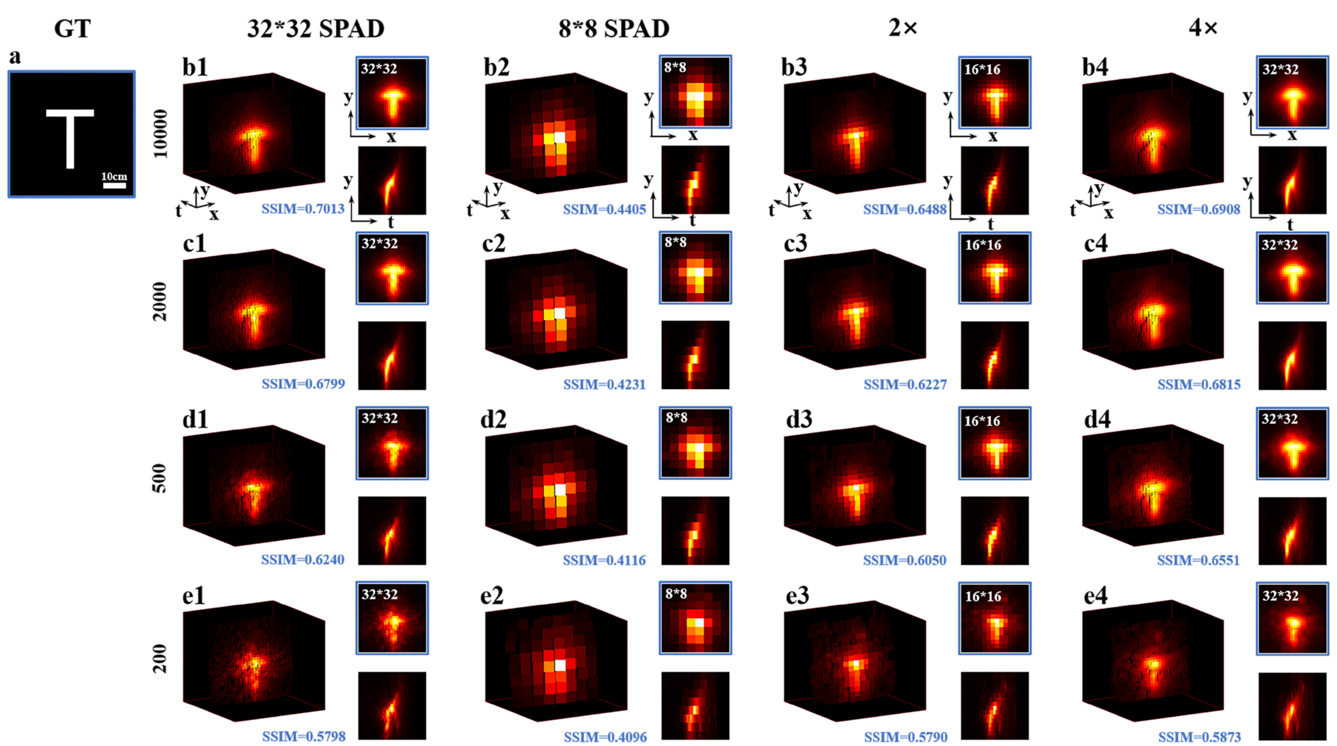

FIG. 6. Comparative results of super-resolution reconstruction for target "T" under different repetition times. (a) Hidden diffusive object. (b1)–(e1) Reconstruction results for data with the spatial resolution of 32 × 32. (b2)–(e2) Reconstruction results for data with the spatial resolution of 8 × 8. (b3)–(e3) Applying twofold super-resolution to the data in (b2)–(e2), followed by reconstruction to obtain object recovery results. (b4)–(e4) Applying four-fold super-resolution to the data in (b2)–(e2), followed by reconstruction to obtain object recovery results. The SSIM values between the front view of the reconstructed results and the GT are listed below the reconstructed results. The repetition times are indicated on the left side of the (b1)–(e1).

TABLE I. Acquisition time and reconstruction time required for data with different repetition times.

Repetition time	Acquisition time (s)	Acquisition frame rate	Reconstruction time (s)	Total time (s)	Imaging frame rate
10 000	0.1023	9.7752	0.0449	0.1472	6.7935
2 000	0.0217	46.0830	0.0450	0.0667	14.9925
500	0.0066	151.5152	0.0459	0.0525	19.0476
200	0.0036	277.7778	0.0447	0.0483	20.7039

rate for data acquisition and the overall imaging frame rate can be calculated. From the reconstruction results in Fig. 6 and the data in Table I, it can be observed that the proposed scan-free NLOS method can detect dynamic scenes at a rate of ~151 fps using an 8 × 8 small detector array. In addition, it can achieve real-time object reconstruction at a rate of 19 fps.

Based on the experimental results presented above, it can be concluded that the scan-free method proposed in this paper exhibits

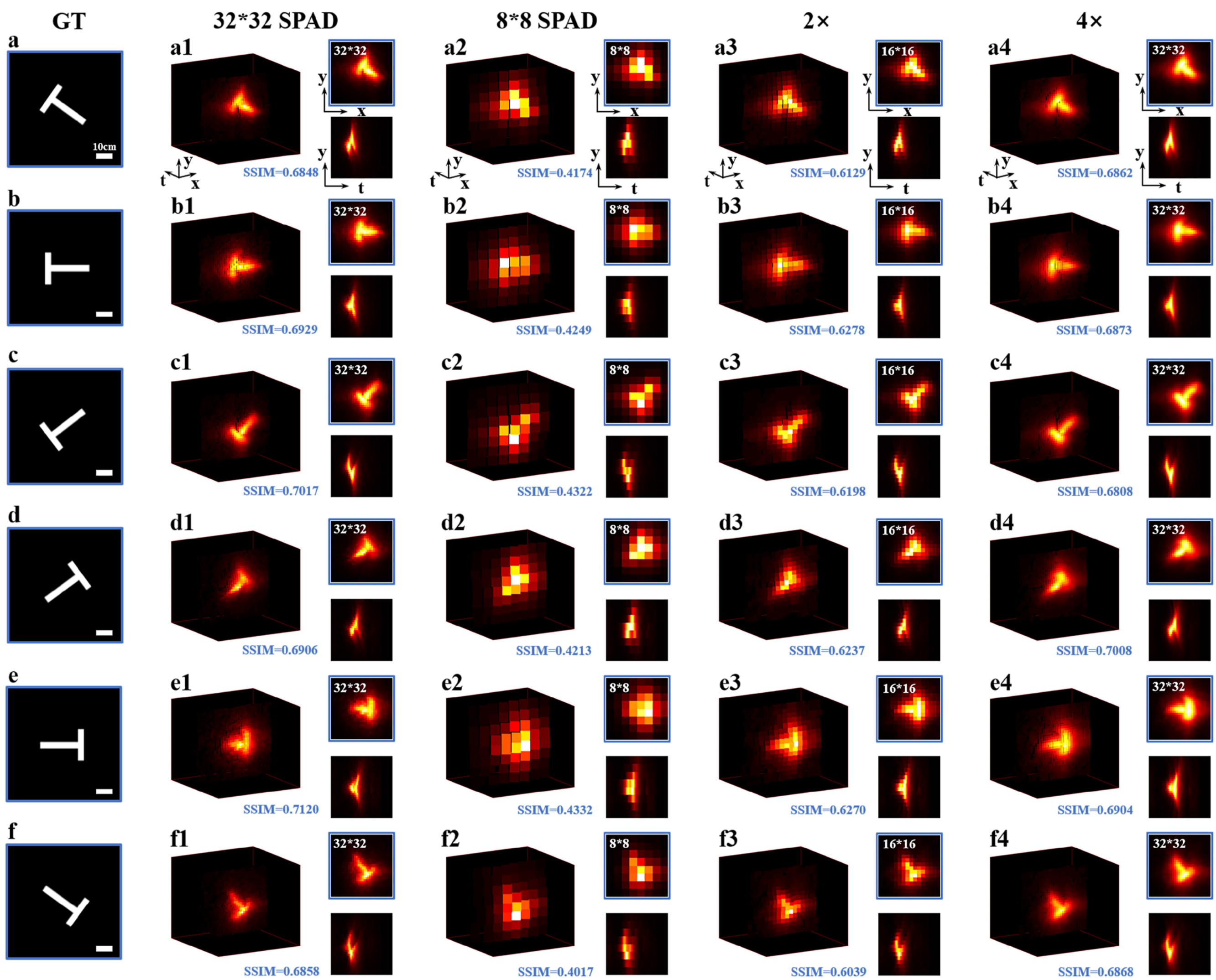

FIG. 7. Comparative results of super-resolution reconstruction for moving objects. (a)–(f) The hidden object with different poses. (a1)–(f1) Reconstruction results for data with the spatial resolution of 32 × 32. (a2)–(f2) Reconstruction results for data with the spatial resolution of 8 × 8. (a3)–(f3) Applying twofold super-resolution to the data in (a2)–(f2), followed by reconstruction to obtain object recovery results. (a4)–(f4) Applying four-fold super-resolution to the data in (a2)–(f2), followed by reconstruction to obtain object recovery results. The SSIM values between the front view of the reconstructed results and the GT are listed below the reconstructed results.

essentially real-time acquisition and imaging capabilities. To test its practical applicability in scenarios involving moving objects, a rotating diffusive "T" is selected as the hidden object. A laser point is illuminated at a fixed position on the relay wall, the object is rotated at a speed of 2.5 r/min, and six different poses of the object are captured during the rotation. The repetition time is 2000, and the reconstruction results of the object are shown in Fig. 7. It can be observed that the proposed method enables real-time imaging of moving objects using a 32 × 32 array detector, and the reconstruction results effectively recover the poses and structural information of the object. For detection arrays with a small number of pixels (such as 8 × 8), the reconstruction can be basically achieved using the S^3R algorithm and SCBM method. Our method can still achieve fast data acquisition, allowing for batch acquisition of dynamic scene data for subsequent imaging. However, the S^3R method currently requires ~50 s for processing, which introduces a certain amount of

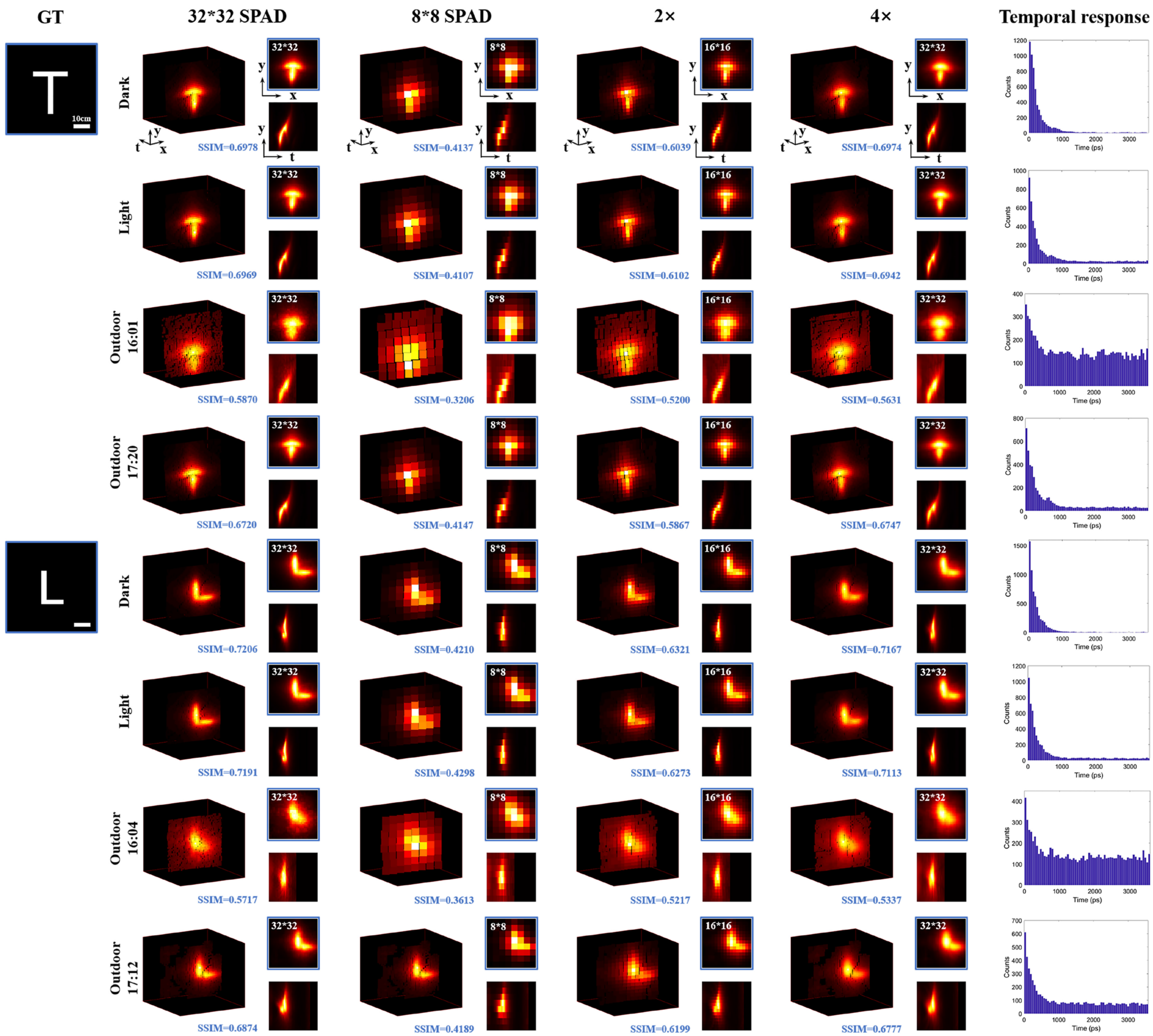

FIG. 8. Comparative results of super-resolution reconstruction for different environments. The SSIM values between the front view of the reconstructed results and the GT are listed below the reconstructed results. The time response of one of the detection points is shown in the last column of Fig. 8, displaying only the signal starting from the second echo.

delay in practical imaging. This will be studied in the subsequent work.

C. Ambient light interference experiment and outdoor experiment

Figure 8 demonstrates the imaging capability of the proposed scan-free NLOS method in the presence of indoor scenes with ambient light interference and outdoor scenes. The object scenes include targets "T" and "L," and their reconstruction ability is tested under four different scenarios: a darkroom, an indoor environment with light, and two different outdoor scenes during daytime. The time response of one of the detection points is shown in the last column of Fig. 8, displaying only the signal starting from the second echo. The measurement for the outdoor daytime scene was conducted on January 24th, 2024, from 4:00 PM to 5:30 PM. The relay wall used is a 1 × 1.5 whiteboard with paper attached to its front surface. From the reconstruction results in Fig. 8, it can be observed that the proposed scan-free NLOS method exhibits resistance to ambient light interference and can be applied in both indoor and outdoor scenarios, effectively recovering the structural information of objects. Even when using an 8 × 8 detector array, the method can still recover the structure of the object. The average SNR of the data collected from the third and seventh rows in Fig. 8 are 0.2862 and 0.2796, respectively, which also represent the SNR boundaries for the imaging capability of the SCBM method.

In addition, we test the object reconstruction capabilities of the three methods in several environments with different levels of noise. As shown in Fig. 9, the target is the letter "L," and we test the object reconstruction abilities in four different scenarios. It can be observed that our SCBM method exhibits high robustness to noise, capable of recovering object information in indoor environments with ambient light or outdoor scenes with partial sunlight, and the details of the target are relatively restored best. The phasor field method is somewhat limited in high environmental noise scenarios and struggles to reconstruct the shape of the object in the outdoor daytime scenario. The LCT method demonstrates slightly better robustness to noise in this type of scenario compared to the phasor field method, and it is able to discern the shape of the object in the outdoor daytime scenario.

Based on the comparison and analysis of the experiment mentioned above, it can be observed that the SCBM method exhibits

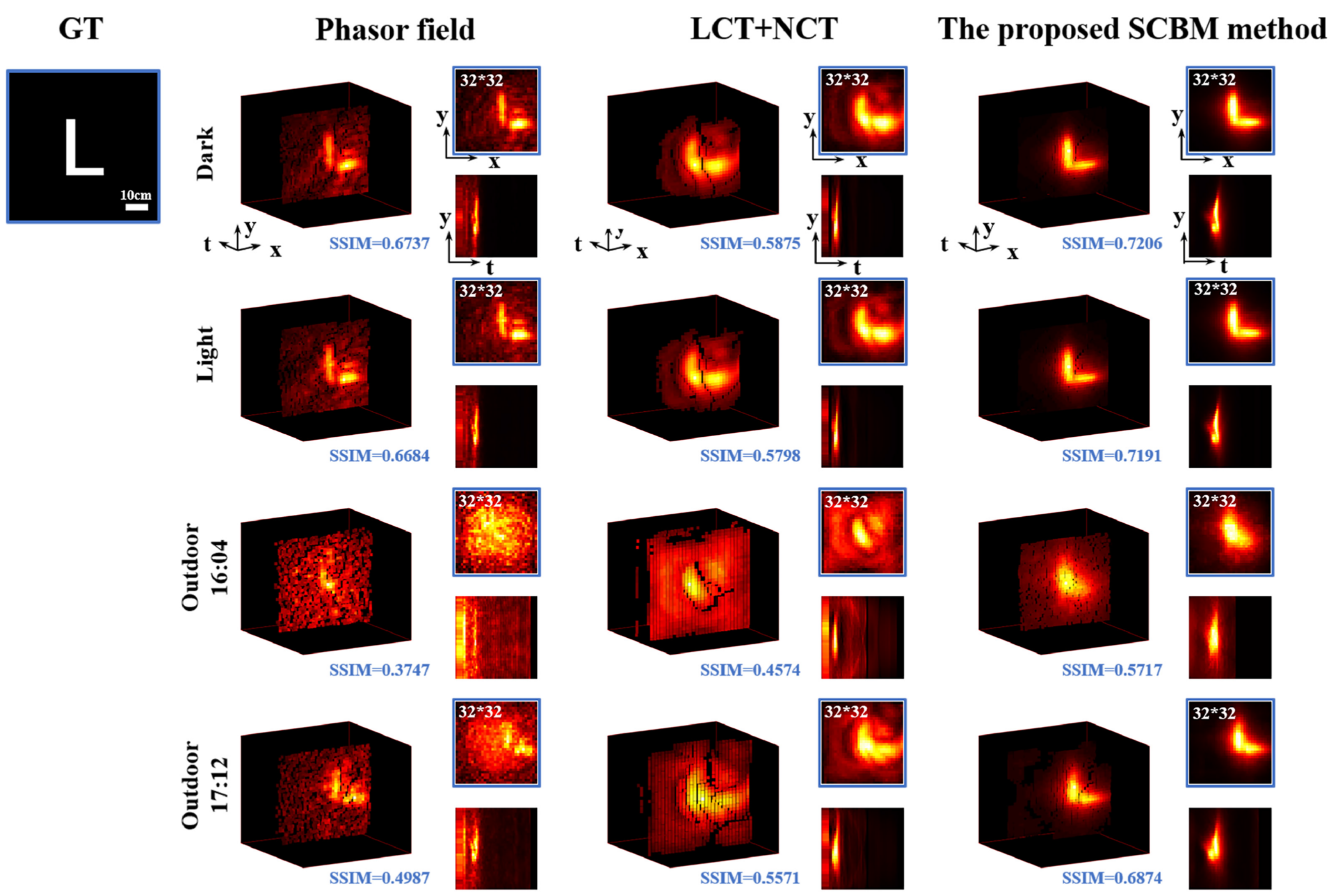

FIG. 9. Comparative results of three fast reconstruction algorithms for different environments. The SSIM values between the front view of the reconstructed results and the GT are listed below the reconstructed results.

higher robustness to noise compared to phasor field and LCT methods. In outdoor scenes during the daytime, our method has certain advantages. If the amount of information detected is sufficient, our method can be applied to environments where the average SNR of the collected data is higher than 0.2796 dB.

IV. DISCUSSION

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

(i) In order to provide fast and high-quality NLOS reconstruction, we use a non-confocal imaging system and propose a SCBM method. First, the propagation process of photons in the non-confocal NLOS imaging system is simulated, modeling the propagation of photons in the non-confocal imaging system as a dual optical coupling process. Subsequently, for non-confocal acquisition scenarios, an efficient non-confocal transformation algorithm NCT is proposed, and based on this, a non-confocal boundary migration model is established to convert the measured values obtained by the array detector into object information in the Fourier domain. Finally, a non-confocal imaging system is built to verify the effectiveness of the proposed method. Experimental results show that the data acquisition time is only 0.0066 s, which is over 10 000 times faster compared to the confocal scenario, making it currently the fastest data acquisition method. The overall imaging acquisition rate reached 19 fps, essentially achieving real-time imaging.

(ii) Considering the cost of devices, a super-resolution method called S^3R is proposed to explore higher spatial resolution signals, reducing the array size requirement for detectors from 32 × 32 to 8 × 8 without compromising image quality. The S^3R method is an independent module that processes only the collected data, which can be applied as a plug-in-and-play module and does not restrict the use of reconstruction methods. In addition, the scan-free system significantly simplifies the sampling process of the system. By further converting the output synchronization signal of the laser into the signal type received by the detector, the system does not need to use a picosecond delay unit. Therefore, with only the laser and array detector as key components, the system is simple in structure and makes it possible to miniaturize the NLOS imaging device. Combining the low-cost detector requirements and real-time imaging performance, the scan-free imaging method proposed in the paper can greatly advance the practical application of NLOS systems.

(iii) The background stability caused by rapid acquisition makes our method suitable for dynamic scenes and can recover the structural information of rotating moving objects with high fidelity. In addition, most ambient noise information is discarded during the interpolation process of the SCBM method. Therefore, this method is not sensitive to noise and is capable of recovering object information in indoor environments with ambient light or outdoor scenes with partial sunlight. However, in outdoor scenes with strong sunlight, the SNR of the acquired data is low, and noise can affect the extraction of effective structural information of the object, ultimately impacting the quality of the reconstruction. To further enhance the practicality of this method, several improvements can be considered. Establishing a noise model based on the actual scenario can be helpful for noise separation, enhancing the SNR, and thereby making the method applicable to all-day scenarios. Considering that actual NLOS scenarios may involve smoke or biological tissue interference, incorporating modeling of different scattering scenarios can achieve object recovery in more general situations.[61,62] In addition, utilizing multi-band lasers can be considered for the recovery of hidden objects with color information; increasing the laser power appropriately could allow for a shorter acquisition time. With these enhancements, the proposed method in this paper can be more practical, providing a feasible solution for the practical implementation of NLOS methods.

V. CONCLUSION

This paper proposes a scan-free NLOS imaging method, which can achieve fast acquisition and imaging of objects based on a small array detector. The data acquisition speed of the method can achieve 151 fps, making it the fastest method in terms of acquisition speed currently. The overall acquisition and imaging time can also reach 19 fps, essentially achieving real-time imaging. This method allows real-time reconstruction of moving objects, has high robustness to noise, and is suitable for indoor lighting conditions and outdoor daytime environments, thus greatly advancing the practical implementation of NLOS systems.

SUPPLEMENTARY MATERIAL

The supplementary material provides the functioning details of the time-to-space boundary migration and the NCT algorithm, the experimental setup of the non-confocal scan-free NLOS imaging system, the analysis of the influence of object materials on the reconstruction, and the applicability of the proposed S3R algorithm to different reconstruction methods. The supplementary material also provides a video of the dynamic reconstruction results.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Nos. 62031018, 62101255, and U23A20283), the Jiangsu Provincial Key Research and Development Program (No. BE2022391), and the Fundamental Research Funds for the Central Universities (No. 30924010812).

The authors thank Xiaorui Tian, Yi Wei, and Jinye Miao for their technical support and experiment discussion.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Wenjun Zhang and Enlai Guo contributed equally to this work.

Wenjun Zhang: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Enlai Guo**: Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal); Writing – review & editing (equal). **Shuo Zhu**: Conceptualization (equal); Formal analysis (equal); Supervision (supporting); Writing – review & editing (equal). **Chenyang Huang**: Data curation (equal); Formal analysis (supporting); Investigation (supporting). **Lijia Chen**: Data curation (supporting); Methodology (supporting). **Lingfeng Liu**: Conceptualization (supporting); Data curation (supporting). **Lianfa Bai**: Funding acquisition (lead); Project administration (equal). **Edmund Y. Lam**: Writing – review & editing (supporting). **Jing Han**: Conceptualization (equal); Funding acquisition (equal); Project administration (equal); Resources (equal); Supervision (equal).

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to state restrictions such as privacy or ethical restrictions.

REFERENCES

[1] S. Bauer, R. Streiter, and G. Wanielik, "Non-line-of-sight mitigation for reliable urban GNSS vehicle localization using a particle filter," in *2015 18th International Conference on Information Fusion (Fusion)* (IEEE, 2015), pp. 1664–1671.

[2] D. B. Lindell, G. Wetzstein, and M. O'Toole, "Wave-based non-line-of-sight imaging using fast *f-k* migration," ACM Trans. Graphics **38**, 1–13 (2019).

[3] N. Scheiner, F. Kraus, F. Wei, B. Phan, F. Mannan, N. Appenrodt, W. Ritter, J. Dickmann, K. Dietmayer, and B. Sick, "Seeing around street corners: Non-line-of-sight detection and tracking in-the-wild using Doppler radar," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2020), pp. 2065–2074.

[4] M. Isogawa, Y. Yuan, M. O'Toole, and K. Kitani, "Optical non-line-of-sight physics-based 3D human pose estimation," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2020), pp. 7011–7020.

[5] C. Wu, J. Liu, X. Huang, Z.-P. Li, C. Yu, J.-T. Ye, J. Zhang, Q. Zhang, X. Dou, V. K. Goyal, F. Xu, and J.-W. Pan, "Non-line-of-sight imaging over 1.43 km," Proc. Natl. Acad. Sci. U. S. A. **118**, e2024468118 (2021).

[6] F. Mu, S. Mo, J. Peng, X. Liu, J. H. Nam, S. Raghavan, A. Velten, and Y. Li, "Physics to the rescue: Deep non-line-of-sight reconstruction for high-speed imaging," IEEE Trans. Pattern Anal. Mach. Intell. (published online, 2022).

[7] 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**, 784–790 (2014).

[8] 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**, 63–71 (2020).

[9] S. Zhu, E. Guo, J. Gu, L. Bai, and J. Han, "Imaging through unknown scattering media based on physics-informed learning," Photonics Res. **9**, B210–B219 (2021).

[10] W. Zhang, S. Zhu, L. Liu, L. Bai, J. Han, and E. Guo, "High-throughput imaging through dynamic scattering media based on speckle de-blurring," Opt. Express **31**, 36503–36520 (2023).

[11] S. Zhu, C. Wang, J. Huang, P. Zhang, J. Han, and E. Y. Lam, "Neuromorphic encryption: Combining speckle correlography and event data for enhanced security," Adv. Photonics Nexus **3**, 056002 (2024).

[12] K. Wang, L. Song, C. Wang, Z. Ren, G. Zhao, J. Dou, J. Di, G. Barbastathis, R. Zhou, J. Zhao, and E. Y. Lam, "On the use of deep learning for phase recovery," Light: Sci. Appl. **13**, 4 (2024).

[13] 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**, 549–553 (2012).

[14] A. P. Mosk, A. Lagendijk, G. Lerosey, and M. Fink, "Controlling waves in space and time for imaging and focusing in complex media," Nat. Photonics **6**, 283–292 (2012).

[15] R. Cao, F. de Goumoens, B. Blochet, J. Xu, and C. Yang, "High-resolution non-line-of-sight imaging employing active focusing," Nat. Photonics **16**, 462–468 (2022).

[16] 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.

[17] I. Dokmanić, R. Parhizkar, A. Walther, Y. M. Lu, and M. Vetterli, "Acoustic echoes reveal room shape," Proc. Natl. Acad. Sci. U. S. A. **110**, 12186–12191 (2013).

[18] D. B. Lindell, G. Wetzstein, and V. Koltun, "Acoustic non-line-of-sight imaging," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2019), pp. 6773–6782.

[19] C. Thrampoulidis, G. Shulkind, F. Xu, W. T. Freeman, J. H. Shapiro, A. Torralba, F. N. C. Wong, and G. W. Wornell, "Exploiting occlusion in non-line-of-sight active imaging," IEEE Trans. Comput. Imaging **4**, 419–431 (2018).

[20] C. Saunders, J. Murray-Bruce, and V. K. Goyal, "Computational periscopy with an ordinary digital camera," Nature **565**, 472–475 (2019).

[21] F. Xu, G. Shulkind, C. Thrampoulidis, J. H. Shapiro, A. Torralba, F. N. C. Wong, and G. W. Wornell, "Revealing hidden scenes by photon-efficient occlusion-based opportunistic active imaging," Opt. Express **26**, 9945–9962 (2018).

[22] 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**, 745 (2012).

[23] M. O'Toole, D. B. Lindell, and G. Wetzstein, "Confocal non-line-of-sight imaging based on the light-cone transform," Nature **555**, 338–341 (2018).

[24] X. Liu, I. Guillén, M. La Manna, J. H. Nam, S. A. Reza, T. Huu Le, A. Jarabo, D. Gutierrez, and A. Velten, "Non-line-of-sight imaging using phasor-field virtual wave optics," Nature **572**, 620–623 (2019).

[25] 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 *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2019), pp. 6793–6802.

[26] X. Liu, J. Wang, L. Xiao, Z. Shi, X. Fu, and L. Qiu, "Non-line-of-sight imaging with arbitrary illumination and detection pattern," Nat. Commun. **14**, 3230 (2023).

[27] Z. Zhao, Q. Zhang, X. Li, Y. Guo, M. Pu, F. Zhang, H. Guo, Z. Wang, Y. Fan, M. Xu, and X. Luo, "High-resolution non-line-of-sight imaging based on liquid crystal planar optical elements," Nanophotonics **13**, 2161–2172 (2024).

[28] M. Isogawa, D. Chan, Y. Yuan, K. Kitani, and M. O'Toole, "Efficient non-line-of-sight imaging from transient sinograms," in *Computer Vision—ACCV 2020* (Springer, 2020), pp. 193–208.

[29] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "This is spiral-tap: Sparse Poisson intensity reconstruction algorithms—theory and practice," IEEE Trans. Image Process. **21**, 1084–1096 (2012).

[30] J.-T. Ye, X. Huang, Z.-P. Li, and F. Xu, "Compressed sensing for active non-line-of-sight imaging," Opt. Express **29**, 1749–1763 (2021).

[31] R. Ding, J. Ye, Q. Gao, F. Xu, and Y. Duan, "Curvature regularization for non-line-of-sight imaging from under-sampled data," IEEE Trans. Pattern Anal. Mach. Intell **46**, 8474–8485 (2024).

[32] X. Liu, J. Wang, L. Xiao, X. Fu, L. Qiu, and Z. Shi, "Few-shot non-line-of-sight imaging with signal-surface collaborative regularization," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2023), pp. 13303–13312.

[33] J. Wang, X. Liu, L. Xiao, Z. Shi, L. Qiu, and X. Fu, "Non-line-of-sight imaging with signal superresolution network," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2023), pp. 17420–17429.

[34]S. Zhu, Y. M. Sua, T. Bu, and Y.-P. Huang, "Compressive non-line-of-sight imaging with deep learning," Phys. Rev. Appl. **19**, 034090 (2023).
[35]S. Zhu, E. Guo, W. Zhang, L. Bai, H. Liu, and J. Han, "Deep speckle reassignment: Towards bootstrapped imaging in complex scattering states with limited speckle grains," Opt. Express **31**, 19588–19603 (2023).
[36]Z. Li, X. Liu, J. Wang, Z. Shi, L. Qiu, and X. Fu, "Fast non-line-of-sight imaging based on first photon event stamping," Opt. Lett. **47**, 1928–1931 (2022).
[37]W. Yang, C. Zhang, W. Jiang, Z. Zhang, and B. Sun, "None-line-of-sight imaging enhanced with spatial multiplexing," Opt. Express **30**, 5855–5867 (2022).
[38]J. H. Nam, E. Brandt, S. Bauer, X. Liu, M. Renna, A. Tosi, E. Sifakis, and A. Velten, "Low-latency time-of-flight non-line-of-sight imaging at 5 frames per second," Nat. Commun. **12**, 6526 (2021).
[39]C. Pei, A. Zhang, Y. Deng, F. Xu, J. Wu, D. U.-L. Li, H. Qiao, L. Fang, and Q. Dai, "Dynamic non-line-of-sight imaging system based on the optimization of point spread functions," Opt. Express **29**, 32349–32364 (2021).
[40]P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. P. A. Lensch, "Dual photography," ACM Trans. Graphics **24**, 745–755 (2005).
[41]M. O'Toole, F. Heide, L. Xiao, M. B. Hullin, W. Heidrich, and K. N. Kutulakos, "Temporal frequency probing for 5d transient analysis of global light transport," ACM Trans. Graphics **33**, 1–11 (2014).
[42]D. Du, X. Jin, and R. Deng, "Non-confocal 3d reconstruction in volumetric scattering scenario," IEEE Trans. Comput. Imaging **9**, 732–744 (2023).
[43]Ö. Yilmaz, *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data* (Society of Exploration Geophysicists, 2001).
[44]J. L. Black, K. L. Schleicher, and L. Zhang, "True-amplitude imaging and dip moveout," Geophysics **58**, 47–66 (1993).
[45]B. Zhou, I. M. Mason, and S. A. Greenhalgh, "An accurate formulation of log-stretch dip moveout in the frequency-wavenumber domain," Geophysics **61**, 815–821 (1996).
[46]A. Kirmani, D. Venkatraman, D. Shin, A. Colaço, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal, "First-photon imaging," Science **343**, 58–61 (2014).
[47]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**, 1532–1540 (2020).
[48]X. Jiang, G. Raskutti, and R. Willett, "Minimax optimal rates for Poisson inverse problems with physical constraints," IEEE Trans. Inf. Theory **61**, 4458–4474 (2015).
[49]P.-Y. Jiang, Z.-P. Li, W.-L. Ye, Y. Hong, C. Dai, X. Huang, S.-Q. Xi, J. Lu, D.-J. Cui, Y. Cao *et al.*, "Long range 3D imaging through atmospheric obscurants using array-based single-photon LiDAR," Opt. Express **31**, 16054–16066 (2023).
[50]M. A. Albota, B. F. Aull, D. G. Fouche, R. M. Heinrichs, D. G. Kocher, R. M. Marino, J. G. Mooney, N. R. Newbury, M. E. O'Brien, B. E. Player *et al.*, "Three-dimensional imaging laser radars with Geiger-mode avalanche photodiode arrays," Lincoln Lab. J. **13**, 351–370 (2002).
[51]A. Rochas, M. Gosch, A. Serov, P.-A. Besse, R. S. Popovic, T. Lasser, and R. Rigler, "First fully integrated 2-D array of single-photon detectors in standard cmos technology," IEEE Photonics Technol. Lett. **15**, 963–965 (2003).
[52]M. Renna, J. H. Nam, M. Buttafava, F. Villa, A. Velten, and A. Tosi, "Fast-gated 16 × 1 SPAD array for non-line-of-sight imaging applications," Instruments **4**, 14 (2020).
[53]X. Liu, S. Bauer, and A. Velten, "Analysis of feature visibility in non-line-of-sight measurements," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2019), pp. 10132–10140.
[54]X. Feng and L. Gao, "Improving non-line-of-sight image reconstruction with weighting factors," Opt. Lett. **45**, 3921–3924 (2020).
[55]X. Liu, S. Bauer, and A. Velten, "Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems," Nat. Commun. **11**, 1645 (2020).
[56]S. I. Young, D. B. Lindell, B. Girod, D. Taubman, and G. Wetzstein, "Non-line-of-sight surface reconstruction using the directional light-cone transform," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2020), pp. 1404–1413.
[57]C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman, "Patchmatch: A randomized correspondence algorithm for structural image editing," ACM Trans. Graphics **28**, 1–11 (2009).
[58]J.-B. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2015), pp. 5197–5206.
[59]O. Chum and J. Matas, "Planar affine rectification from change of scale," in *Computer Vision—ACCV 2010* (Springer, 2011), pp. 347–360.
[60]Z. Wang, X. Li, M. Pu, L. Chen, F. Zhang, Q. Zhang, Z. Zhao, L. Yang, Y. Guo, and X. Luo, "Vectorial-optics-enabled multi-view non-line-of-sight imaging with high signal-to-noise ratio," Laser Photonics Rev. **18**, 2300909 (2024).
[61]D. B. Lindell and G. Wetzstein, "Three-dimensional imaging through scattering media based on confocal diffuse tomography," Nat. Commun. **11**, 4517 (2020).
[62]D. Du, X. Jin, R. Deng, J. Kang, H. Cao, Y. Fan, Z. Li, H. Wang, X. Ji, and J. Song, "A boundary migration model for imaging within volumetric scattering media," Nat. Commun. **13**, 3234 (2022).