



Scan-free time-of-flight-based three-dimensional imaging through a scattering layer

RUJIA DENG,^{1,2} XIN JIN,^{1,*}  DONGYU DU,^{1,2} AND ZHIHENG LI^{1,2}

¹*Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China*

²*Tsinghua Innovation Center in Zhuhai, Zhuhai 519080, China*

*jin.xin@sz.tsinghua.edu.cn

Abstract: Reconstructing an object's three-dimensional shape behind a scattering layer with a single exposure is of great significance in real-life applications. However, due to the little information captured by a single exposure while strongly perturbed by the scattering layer and encoded by free-space propagation, existing methods cannot achieve scan-free three-dimensional reconstruction through the scattering layer in macroscopic scenarios using a short acquisition time of seconds. In this paper, we proposed a scan-free time-of-flight-based three-dimensional reconstruction method based on explicitly modeling and inverting the time-of-flight-based scattering light propagation in a non-confocal imaging system. The non-confocal time-of-flight-based scattering imaging model is developed to map the three-dimensional object shape information to the time-resolved measurements, by encoding the three-dimensional object shape into the free-space propagation result and then convolving with the scattering blur kernel derived from the diffusion equation. To solve the inverse problem, a three-dimensional shape reconstruction algorithm consisting of the deconvolution and diffractive wave propagation is developed to invert the effects caused by the scattering diffusion and the free-space propagation, which reshapes the temporal and spatial distribution of scattered signal photons and recovers the object shape information. Experiments on a real scattering imaging system are conducted to demonstrate the effectiveness of the proposed method. The single exposure used in the experiment only takes 3.5 s, which is more than 200 times faster than confocal scanning methods. Experimental results show that the proposed method outperforms existing methods in terms of three-dimensional reconstruction accuracy and imaging limit subjectively and objectively. Even though the signal photons captured by a single exposure are too highly scattered and attenuated to present any valid information in time gating, the proposed method can reconstruct three-dimensional objects located behind the scattering layer of 9.6 transport mean free paths (TMFPs), corresponding to the round-trip scattering length of 19.2 TMFPs.

© 2023 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

The ability of three-dimensional (3D) imaging through a scattering layer with only a single exposure shows great significance in a wide range of real-life applications, especially when the object is in motion, such as autonomous navigation, rescue operations, and non-destructive testing through packaging. However, 3D imaging through a scattering layer suffers from the scattering attenuation and the coupling of scattered and ballistic photons when diffusing through the scattering layer, as well as the encoding of the object shape information by free-space propagation. With the little information captured by a single exposure while strongly perturbed by the scattering layer and encoded by free-space propagation, 3D reconstruction through a scattering layer in the macroscopic environment remains a huge challenge.

Existing methods can be classified into two categories according to the input data type: intensity-based methods and time-of-flight-based methods. Using the two-dimensional (2D)

intensity image captured by a single shot, speckle correlation methods [1–4] based on the optical memory effect [5,6] explore interference of scattered light to recover a 2D image within a very small memory effect range. Further, some techniques including point spread function manipulation [7], coherence gating [8] and light field estimation [9] can retrieve the object depth information. However, complex experimental configurations and heavy computation workload limit their application in real life. Thus, high-dimensional transient data with time-of-flight information is introduced to alleviate the ill-posedness of the inverse 3D reconstruction problem, leading to many time-of-flight-based methods proposed for 3D scattering reconstruction. Among time-of-flight-based methods, time-gating methods [10–14] isolate the ballistic photons that travel on a direct path without scattering. Although time-gating methods allow for fast data acquisition using detectors such as the single-photon avalanche diode (SPAD) array, they lose effectiveness in 3D shape reconstruction through the multiple-scattering media as the detecting ballistic photons rapidly decay to zero with the strong scattering and the free-space propagation distance. Besides, time-of-flight-based methods based on diffuse optical tomography [15,16] describe the propagation of scattered photons using the time-dependent diffusion equation. Lyons *et al.* [15] estimate the shape and position of the hidden object inside a diffusive material by building a regularized least-square optimization model. However, due to the model assumption and computation approximation, it is only limited to 2D object recovery with a low reconstruction quality. Lindell *et al.* [16] achieve confocal 3D imaging by modeling the confocal scattering light propagation. While it suffers from the confocal scanning scheme to collect data, which requires a long integration time (dozens of minutes) and a stationary scene. Further, non-line-of-sight (NLoS) time-of-flight-based imaging methods [17–24] based on a time-of-flight-based imaging system with a pulsed laser and an ultrafast detector can also be applied to realize 3D imaging through a thin diffuser. Especially, some non-confocal methods [19–24] hold the ability of 3D reconstruction in dynamic NLoS scenarios. Nam *et al.* [21] utilize a specifically designed fast-gated SPAD array for further light efficiency with a fast reconstruction method using the phasor field framework [19,20] to realize NLoS imaging at 5 frames per second. However, all the NLoS methods cannot provide an effective 3D reconstruction through the scattering layer since they only consider the three-bounce reflection without accurately modeling the scattering effect using the specific scattering properties.

Previous time-of-flight-based methods [16–18] use the scanning confocal configuration with a long acquisition time to capture more effective photons and thus produce a good reconstruction, which can recover static scenes but fail for real-time imaging tasks in dynamic scenes. With the development of technology, SPAD arrays used in the scan-free non-confocal setup have been introduced to achieve fast data acquisition [12–15,20–22]. However, the measurements captured in a short acquisition time tend to have low signal strength and low signal-to-noise ratio (SNR), which limit the imaging capability, especially in multiple-scattering scenarios. Overall, existing methods suffers from the tradeoff between detection efficiency and reconstruction quality. The future development trend and goal are to provide both high-speed and high-quality 3D scattering reconstruction methods.

So, in this paper, with a non-confocal imaging system including a 32×32 SPAD array, a 3D reconstruction method based on explicitly modeling and inverting the scan-free non-confocal time-of-flight-based scattering light propagation is proposed to recover the 3D object structure behind a scattering layer. Unlike existing confocal models [16–18] that utilize specific optical path constraints, we completely describe the non-confocal scattering light transport with time. In this case, a non-confocal time-of-flight-based scattering imaging model is developed to map the 3D object shape information to the time-resolved measurements by encoding the 3D object shape into the free-space propagation result and then convolving with the scattering blur kernel derived from the diffusion equation. To invert the effects caused by the scattering diffusion and the free-space propagation, a 3D shape reconstruction algorithm is designed to

apply the deconvolution algorithm to the 3D time-resolved measurements and then implements the diffractive wave propagation, reshaping the temporal and spatial distribution of scattered signal photons and recovering the object shape information. We conduct experiments on a real non-confocal scattering imaging system to demonstrate the effectiveness of the proposed method in 3D scattering reconstruction, with a single exposure time of only 3.5s, 200 times faster than the confocal scanning method [16] taking dozens of minutes to scan. Experimental results show that the proposed method outperforms existing methods in the subjective quality and the objective evaluation of peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). In this system setup, the proposed method can reconstruct 3D objects located behind the scattering layer of 9.6 TMFPs, corresponding to the round-trip scattering length of 19.2 TMFPs.

The remainder of this paper is organized as follows. The proposed method is described in detail in Section 2. Section 3 provides experimental results and discussions. Section 4 concludes the paper.

2. Method

The proposed method is based on the non-confocal scattering imaging system, as illustrated in Fig. 1(a). An ultrafast pulsed laser illuminates a point on the front surface of the scattering layer. At the same time, the ultrafast detector acquires the spatial-temporal response of the scattered photons that diffuse through the scattering layer, propagate through free space to a hidden object and then are reflected, and diffuse back again through the scattering layer, as shown in Fig. 1(b). Based on the space and time resolved measurements, the proposed model is developed by describing the time-of-flight-based scattering light propagation, while the reconstruction method is proposed by explicitly inverting the model to recover the 3D shape of the hidden object behind the scattering layer.

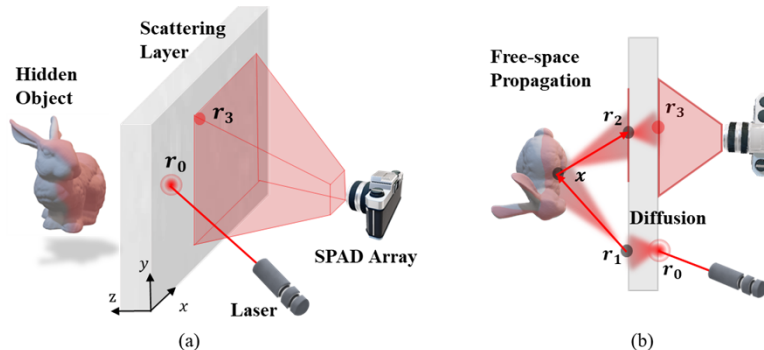


Fig. 1. Schematic diagram for scan-free time-of-flight-based three-dimensional imaging through a scattering layer. (a) The non-confocal scattering imaging system. (b) The complete non-confocal scattering light propagation process.

2.1. Non-confocal time-of-flight-based scattering imaging model

As presented in Fig. 1(b), the light is emitted by the laser to the point r_0 on the front surface of the scattering layer, diffuses through the scattering layer to the back surface of the scattering layer, using r_1 as an example point, reaches the hidden object surface x and reflects back in free space to the back surface of the scattering layer, taking r_2 as an example point, diffuses through the scattering layer, and is finally captured by the SPAD array on the detection area of the front surface of the scattering layer, r_3 as an instance. It is worth mentioning that the confocal scattering imaging model [16] only considers the confocal optical path where the scanning

illuminated point and the detector point are the same, i.e., $\mathbf{r}_0 = \mathbf{r}_3$. While our model utilizes a non-confocal configuration with a single illuminated point \mathbf{r}_0 and a detected area \mathbf{r}_3 , different from \mathbf{r}_0 . Besides, in our model, \mathbf{r}_2 can be different from \mathbf{r}_1 , which means that the free-space propagation of light is non-confocal if taking \mathbf{r}_1 as the emitter and \mathbf{r}_2 as the receiver.

To build our non-confocal model, we first consider the scattering diffusion paths of light traveling through the scattering layer, that is, $\mathbf{r}_0 \rightarrow \mathbf{r}_1$ and $\mathbf{r}_2 \rightarrow \mathbf{r}_3$. Since the scattering layer locates in a physically limited space, the analytical solution of the time-dependent diffusion equation for the slab geometry can be used to represent the intensity $\phi(\mathbf{r}_0, \mathbf{r}_1, t)$ received by the diffused light from the point \mathbf{r}_0 to \mathbf{r}_1 at time t . Specifically,

$$\phi(\mathbf{r}_0, \mathbf{r}_1, t) = \frac{1}{2(4\pi Dc)^{\frac{3}{2}} t^{\frac{5}{2}}} \exp\left(-\mu_a ct - \frac{(\mathbf{r}_{1x} - \mathbf{r}_{0x})^2 + (\mathbf{r}_{1y} - \mathbf{r}_{0y})^2}{4Dct}\right) \cdot \sum_{i=-\infty}^{\infty} \left[(z_d - z_{+,i}) \exp\left(-\frac{(z_d - z_{+,i})^2}{4Dct}\right) - (z_d - z_{-,i}) \exp\left(-\frac{(z_d - z_{-,i})^2}{4Dct}\right) \right], \quad (1)$$

where c , t , z_d correspond to the speed of light within the scattering layer, time, and thickness of the scattering layer, respectively. The position of the point illuminated by the laser is given by $\mathbf{r}_0 = (r_{0x}, r_{0y}, 0)$ while the sampled position on the back surface of the scattering layer is provided by $\mathbf{r}_1 \in \Omega_{z_d} = \{(r_{1x}, r_{1y}, r_{1z}) \in \mathbb{R}^3 \mid r_{1z} = z_d\}$. The diffusion coefficient is $D = 1/3(\mu_a + \mu'_s)$, which is calculated by the absorption coefficient μ_a and the reduced scattering coefficient μ'_s . The positions of $z_{+,i}$ and $z_{-,i}$ can be obtained by the extrapolated boundary condition of the diffusion equation for the slab geometry. We usually truncate the infinite series in Eq. (1) to 7 (i.e., $i = 0, \pm 1, \pm 2, \pm 3$) empirically to reduce the error to a negligible level. This approximate solution of Eq. (1) is also used by the confocal method [16] to solve the equation and model the diffusion.

Similarly, we can produce the analytical expression of $\phi(\mathbf{r}_2, \mathbf{r}_3, t)$ corresponding to the back diffusion propagation where the sampled position on the back surface of the scattering layer is given by $\mathbf{r}_2 \in \Omega_{z_d} = \{(r_{2x}, r_{2y}, r_{2z}) \in \mathbb{R}^3 \mid r_{2z} = z_d\}$ and the imaged position on the detected area is denoted by $\mathbf{r}_3 \in \Omega_{z_d} = \{(r_{3x}, r_{3y}, r_{3z}) \in \mathbb{R}^3 \mid r_{3z} = 0\}$.

For the propagation path $\mathbf{r}_1 \rightarrow \mathbf{x} \rightarrow \mathbf{r}_2$ of light traveling via the hidden object in free space, the free-space propagation operator $H(\mathbf{r}_1, \mathbf{r}_2, t)$ can be expressed by the impulse response function of the hidden object as:

$$H(\mathbf{r}_1, \mathbf{r}_2, t) = \int_{\Psi} f(\mathbf{x}, \mathbf{r}_1) f(\mathbf{x}, \mathbf{r}_2) \delta(ct - \|\mathbf{x} - \mathbf{r}_1\| - \|\mathbf{x} - \mathbf{r}_2\|) d\mathbf{x}, \quad (2)$$

where the function $f(\cdot)$ represents the contribution by the bidirectional reflectance distribution function (BRDF), albedo and distance attenuation from the point $\mathbf{r}_{1/2}$ on the back surface of the scattering layer to the point \mathbf{x} on the hidden object. The Dirac delta function $\delta(\cdot)$ relates the propagation distance with travel time. Finally, an integral is performed with respect to \mathbf{x} over the 3D domain Ψ of the hidden object space behind the scattering layer.

Consisting of light diffusion through the scattering layer from the illuminated point \mathbf{r}_0 to \mathbf{r}_1 , free-space propagation from \mathbf{r}_1 to \mathbf{r}_2 reflected by the hidden object \mathbf{x} , and diffusion back through the scattering layer from \mathbf{r}_2 to \mathbf{r}_3 , the complete non-confocal measurement model can be derived as

$$\tau(\mathbf{r}_0, \mathbf{r}_3, t) = \int_{\Omega_{z_d}} \int_0^{\infty} \left[\int_{\Omega_{z_d}} \int_0^{\infty} \phi(\mathbf{r}_0, \mathbf{r}_1, t'' - t') \cdot H(\mathbf{r}_1, \mathbf{r}_2, t') dt' dr_1 \right] \cdot \phi(\mathbf{r}_2, \mathbf{r}_3, t - t'') dt'' dr_2, \quad (3)$$

which can be further written as a convolution model:

$$\tau(\mathbf{r}_0, \mathbf{r}_3, t) = \phi(\mathbf{r}_2, \mathbf{r}_3, t) * \phi(\mathbf{r}_0, \mathbf{r}_1, t) * H(\mathbf{r}_1, \mathbf{r}_2, t) := \phi * H, \quad (4)$$

where $*$ represents the convolution operator, τ indicates the non-confocal spatial-temporal measurements, ϕ refers to the convolution kernel of light diffusion through the scattering layer and back, and H is the free-space propagation operator. Therefore, the non-confocal model is developed by convolving the analytical solution of the diffusion equation as the blur kernel with the free-space propagation operator, which completely describes the non-confocal scattering light transport with time.

The continuous diffusion operator ϕ and the free-space propagation operator H are implemented with discrete matrix operations as Φ and \mathbf{H} in practice, respectively. The discrete matrix representation of the non-confocal time-of-flight-based scattering imaging model is:

$$\tau = \Phi \mathbf{H} \rho = \Phi \mathbf{y}, \quad (5)$$

where τ denotes the non-confocal 3D time-resolved measurements, ρ represents the hidden object albedo in the 3D voxelized space behind the scattering layer, \mathbf{y} refers to the hidden scene impulse response $\mathbf{H}\rho$ that only encodes the free-space propagation of light without involving the effect of the diffusion operator Φ .

With the above model that maps the 3D object shape information ρ to the non-confocal 3D time-resolved measurements τ , the inverting method can be designed to achieve 3D reconstruction.

2.2. Three-dimensional shape reconstruction algorithm

The pipeline of the proposed three-dimensional shape reconstruction method is shown in Fig. 2. In order to invert the above model, we apply the convolution theorem, the Wiener deconvolution filter and the non-confocal free-space propagation solver \mathbf{H}^{-1} based on phasor-field diffractive wave propagation to derive the following closed-form solution,

$$\hat{\rho} = \mathbf{H}^{-1} \mathbf{F}^{-1} \left[\frac{1}{\Phi^* |\Phi^*|^2 + \frac{1}{\alpha}} \right] \mathbf{F} \tau := \mathbf{H}^{-1} \mathbf{y}, \quad (6)$$

where $\hat{\rho}$ is the reconstructed 3D shape of the hidden object, \mathbf{F} represents the discrete 3D Fourier transform matrix, Φ^* is the diagonal matrix corresponding to the Fourier coefficients of the 3D convolution kernel, and α denotes the frequency-dependent SNR of the Wiener deconvolution filter.

First, the Wiener deconvolution filter using the proposed model's convolution kernel is applied to the 3D time-resolved measurements in Fig. 2(a) to obtain the deconvolved time-resolved result \mathbf{y} in Fig. 2(b). Then, the non-confocal free-space propagation solver \mathbf{H}^{-1} is utilized to image the 3D shape of the hidden object in Fig. 2(c), by introducing a virtual wave field termed phasor field and modeling the diffractive wave propagation of light through the hidden object in free space. This imaging technique [19] used in NLoS imaging can formulate the NLoS light propagation as a virtual line-of-sight wave propagation by converting the backward 3D reconstruction to forward diffractive imaging. According to Fourier optics, the received complex phasor field $P(\mathbf{r}_1, \mathbf{r}_2, t)$ on the virtual aperture \mathbf{r}_2 can be expressed as the convolution of the complex phasor field $P(\mathbf{r}_1, t)$ on the virtual incident light source \mathbf{r}_1 and the hidden scene impulse response function $H(\mathbf{r}_1, \mathbf{r}_2, t)$ at time t :

$$P(\mathbf{r}_1, \mathbf{r}_2, t) = P(\mathbf{r}_1, t) * H(\mathbf{r}_1, \mathbf{r}_2, t) = \left(e^{j\omega t} e^{-\frac{(t-t_0)^2}{2\sigma^2}} \right) * H(\mathbf{r}_1, \mathbf{r}_2, t), \quad (7)$$

where the Gaussian pulse wave with frequency ω and pulse width σ at time t is used as the phasor field $P(\mathbf{r}_1, t)$ on the virtual incident light source, t_0 is the initial emitting time from the virtual incident light source \mathbf{r}_1 . In practical implementation, the descattering result \mathbf{y} in Eq. (6) generated by the deconvolution algorithm is used as the discrete expression of $H(\mathbf{r}_1, \mathbf{r}_2, t)$ in Eq. (7), which is exactly the hidden scene impulse response $\mathbf{H}\rho$ in Eq. (5) in free space without involving the diffusion effect of the scattering layer. So, the physical meaning of Eq. (7) is that,

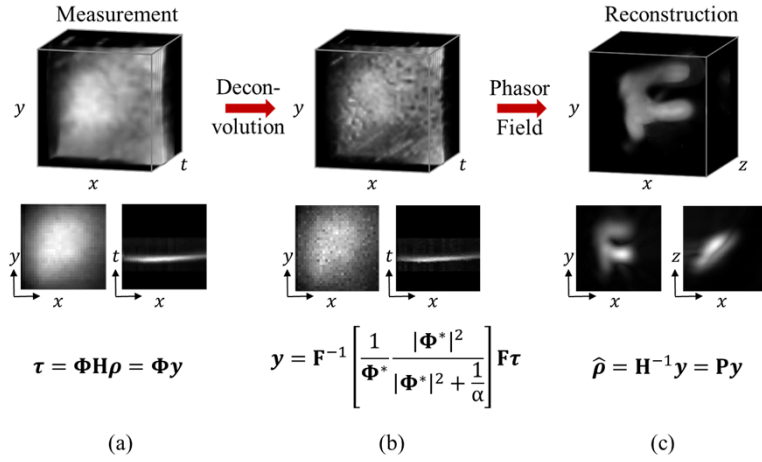


Fig. 2. The pipeline of the proposed three-dimensional shape reconstruction method. (a) The non-confocal 3D time-resolved measurements. (b) The descattering time-resolved result after the deconvolution algorithm. (c) The reconstructed 3D structure after the phasor field.

input with the descattering image y , $P(\mathbf{r}_1, \mathbf{r}_2, t)$ of Eq. (7) provides a forward diffractive imaging result, which is also the complex phasor field form of the 3D shape structure ρ .

Then, we can transform the complex phasor field $P(\mathbf{r}_1, \mathbf{r}_2, t)$ into the intensity image for recovering the 3D shape of the hidden object. Specifically, for the v -th voxel located at \mathbf{x}_v , the corresponding reconstructed albedo ρ_v can be described as

$$\rho_v = \left| \int_{\Omega_{z_d}} \int_{\Omega_{x_d}} P \left(\mathbf{r}_1, \mathbf{r}_2, -\frac{|\mathbf{x}_v - \mathbf{r}_1|}{c} - \frac{|\mathbf{x}_v - \mathbf{r}_2|}{c} \right) d\mathbf{r}_1 d\mathbf{r}_2 \right|^2, \quad (8)$$

which can be discretized and solved by the back-projection method.

Therefore, the 3D albedo $\hat{\rho}$ of the voxelized object space in Fig. 2(c) can be reconstructed by combining the forward diffractive imaging step of Eq. (7) and the back-projection reconstruction step of Eq. (8). In other words, using the descattering result y as input, we apply the diffractive projection operator \mathbf{P} , the discrete form of Eq. (7) and Eq. (8), as the non-confocal free-space propagation solver \mathbf{H}^{-1} to generate the 3D reconstruction $\hat{\rho}$ as

$$\hat{\rho} = \mathbf{H}^{-1} y = \mathbf{P} y. \quad (9)$$

Notably, the most costly step of the proposed method is the application of the phasor field using back-projection in Eq. (9), such that the total computational complexity of the proposed method is $O(32 \times 32 \times N^3)$ where 32×32 comes from the pixel resolution of the SPAD array in our experiments, N is the maximum number of voxels across all dimensions in the reconstructed voxelized space.

3. Experimental results and discussions

3.1. Experimental setup

To demonstrate the effectiveness of the proposed method, the experimental setup in Fig. 3(a) and the prototype of the non-confocal time-of-flight-based laser imaging system in Fig. 3(b) are designed to capture the measurements. A high-power pulsed laser (INNO AMT-532-1W1M) is used with a wavelength of 532 nm, 12ps pulse width, 20 MHz repetition rate and 700 mW average optical

power. The emitted laser beam passes through a beam splitter (Thorlabs PBS251) and then illuminates a spot on the front surface of the scattering layer steered by a pair of galvanometer mirrors (Thorlabs GVS012) controlled with a National Instruments data acquisition device (NI-DAQ USB-6343). In such a non-confocal configuration, the returning light reflected by the scenario travels back and is directly focused by a lens with 2.8-12 mm focal length and 1.6 maximal f-number onto the detector. The ultrafast detector is a SPAD array (Photon Force PF32) with a pixel size of 32×32 . Each pixel operates in time-correlated single photon counting (TCSPC) mode with a temporal resolution of 55ps. A delayer unit (Micro Photon Devices PSD-065-A-MOD) is used to shape the synchronization signal output by the laser into a standard Transistor-Transistor Logic (TTL) signal, which is then used as the acquisition trigger signal of the SPAD array. In this setup, the SPAD array captures photons from all pixels simultaneously and generates 3D data made of two spatial dimensions (32×32 pixels) and one temporal dimension (910 time bins). However, due to the crosstalk between pixels of the SPAD array in low-light environments, more than 100 pixels fail to capture effective signal photons instead of background noise photons, so the actual pixel utilization is less than 32×32 .

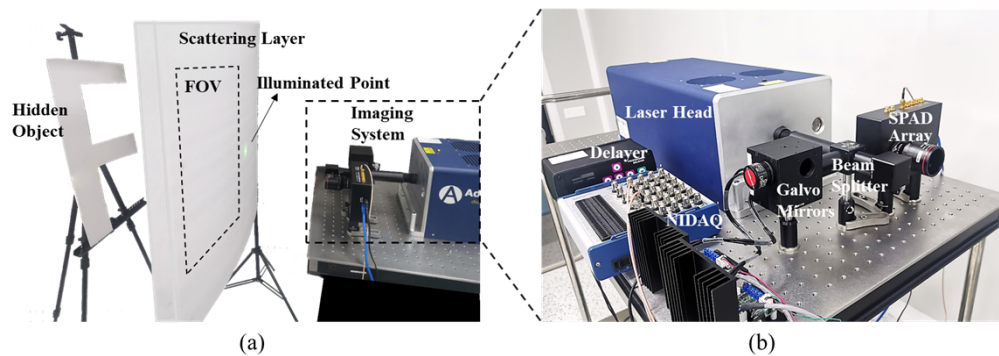


Fig. 3. The experimental setup of the scan-free three-dimensional scattering imaging system. (a) The photograph and layout of the polyethylene foam scattering scenario. (b) The hardware prototype of the non-confocal time-of-flight-based imaging system.

3.2. Implementation details

All time-resolved measurements with a spatial pixel resolution of 32×32 are captured by detecting an $85\text{cm} \times 85\text{cm}$ area with a single exposure of 3.5s. Owing to the depth between the target and the scattering layer, the photons directly reflected by the scattering layer and photons scattered by the hidden scene can be distinguished according to the arrival time, as shown in Fig. 7(a). Inputting the time-resolved measurements, the time-gating method operates by directly removing the first-arriving photon peak reflected by the scattering layer and selecting the rest photons from the hidden scene. As the first step of the proposed method, the descattering method utilizes the Wiener deconvolution filter to remove the scattering effect by deconvolving the time-gated measurements with the diffusion kernel, outputting the descattering result y in Fig. 2(b), that is, the intermediate result of the proposed method. Besides, the phasor field method [19] in NLoS imaging works in a scanning non-confocal configuration by treating the scattering layer as the diffuse wall without the diffusion effect and then modeling the diffractive wave propagation in Eq. (9). Since its principle is independent of the scanning hardware, it can be applied directly to the time-gated measurements in our scan-free non-confocal system. In this case, the time-gated measurement, instead of the descattering result used for the second step of the proposed method, is considered as the input hidden scene impulse response function $H(\mathbf{r}_1, \mathbf{r}_2, t)$ in Eq. (7) to generate the reconstruction of the phasor field in NLoS imaging. While the reconstruction of the

proposed method in Fig. 2(c) can be generated by dividing the reconstructed object space [0 cm, 85 cm] × [0 cm, 85 cm] × [0 cm, 85 cm] into 85 × 85 × 85 voxels and then applying the phasor field method to the descattering result y .

To quantitatively evaluate the reconstruction effect, we use the 2D front views of the reference ground truth and the reconstruction result to calculate the PSNR [25] and SSIM [26], which are the two most common quantitative indicators in scattering imaging and non-line-of-sight imaging research [9,13,16,22]. To make a fair comparison with the 2D reference binary image in Fig. 5(a) and Fig. 6(a) whose values are only 0 and 255, we first compute the gray image of the front view of the reconstruction result and then convert it to a binary image using the threshold of 127.5 before calculating the PSNR and SSIM between them. Specifically,

$$PSNR(x, y) = 10 \log_{10} \left(\frac{M^2}{MSE(x, y)} \right), \quad (10)$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + (k_1M)^2)(2\sigma_{xy} + (k_2M)^2)}{(\mu_x^2 + \mu_y^2 + (k_1M)^2)(\sigma_x^2 + \sigma_y^2 + (k_2M)^2)}, \quad (11)$$

where x and y indicate the binary images of the reference and the reconstruction result. $MSE(x, y)$ calculates the mean square error of x and y , $M = 255$ represents the maximum possible pixel value of the image. μ , σ , σ_{xy} correspond to the mean value, the standard deviation and the covariance of x and y . The constants k_1 and k_2 are empirically set as 0.01 and 0.03.

3.3. Three-dimensional shape reconstructions

The scattering properties of polyethylene foam are estimated using an optimal model [16] to minimize the residual between the temporal measurements and theoretical response. Then the reduced scattering coefficient μ'_s and the absorption coefficient μ_a of the polyethylene foam are estimated as $3.1377 \pm 0.35 \text{ cm}^{-1}$ and $3.3348 \times 10^{-2} \pm 2.6 \times 10^{-3} \text{ cm}^{-1}$. Hence, the length of one transport mean free path (TMFP) is $l^* = 1/\mu'_s = 0.3187 \pm 0.0356 \text{ cm}$, which represents the distance after which a photon's initial propagation direction becomes random. The thickness of polyethylene foam used in this experiment is 2 cm, corresponding to 6.4 TMFPs, so the signal photons are almost scattered into all directions during the propagation. The test Lambertian objects made of white matte papers include single tilted letters 'T', 'K', 'F' in Fig. 4(a) and multi-depth letters 'LL', 'LT', 'LF' in Fig. 5(a). The scenario layout of the target and the polyethylene foam is shown in Fig. 3(a). The objects with different sizes are placed behind the polyethylene foam at different depths, with the front view and top view of the detailed object sizes and depths provided in Fig. 4(a) and Fig. 5(a).

The reconstructions generated by time-gating, descattering, phasor field and the proposed method are provided in Fig. 4 and Fig. 5. It can be seen that the time-gated measurements in Fig. 4(b) and Fig. 5(b) provide only speckle-like patterns without any useful object shape information, as all the target photons are highly scattered by the scattering layer with little ballistic photons existing. The descattering result in Fig. 4(c) and Fig. 5(c) produced by the Wiener deconvolution filter fails to recover the object shape since it only models the diffusion without considering the free-space propagation. The reconstruction by the phasor field in Fig. 4(d) and Fig. 5(d) is also ineffective because it simply regards the scattering layer as the diffuse wall in the NLoS scenario without accurately modeling the diffusion using the specific scattering properties. As the front views and top views of the 3D reconstruction shown in Fig. 4(e) and Fig. 5(e), the proposed method successfully recovers the shape and depth information of all targets, which benefits from completely modeling the non-confocal scattering light propagation, including diffusion and free-space propagation, and accurately solving the inverse problem. Although there is a certain expansion in the depth reconstruction of the top view by the proposed method relative to the reference image, which can also be suppressed by image post-processing such as filtering

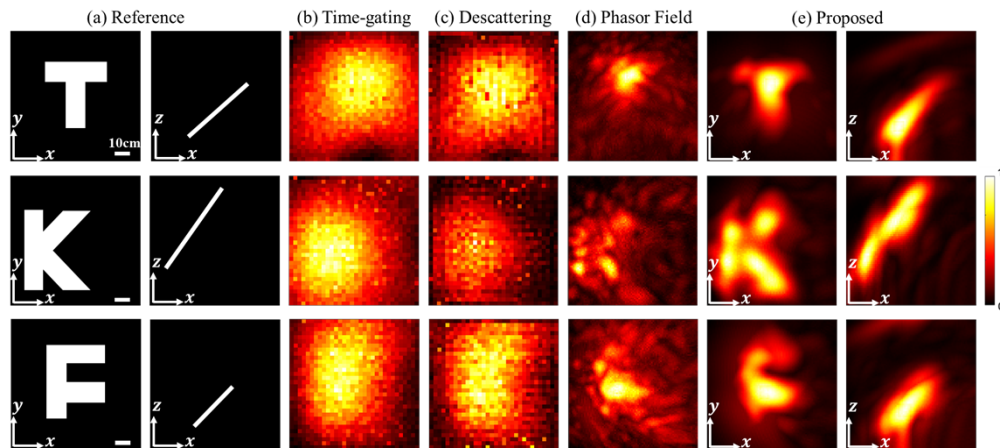


Fig. 4. Comparisons of reconstructing single Lambertian objects.

and brightness correction, the proposed method has the ability to provide the absolute depth information of the object.

It is worth noting that there exist some reconstruction imperfections of the proposed method in Fig. 4(e) and Fig. 5(e), which results from the coupling of data acquisition and the implementation of the proposed method using back projection. At the data acquisition level, as the laser is only incident from one illuminated point, it becomes a spot concentrated in a certain area after being diffused by the scattering layer. In this setup, part of the object (such as the second horizontal line of ‘F’ in Fig. 4(a)) facing the light spot and the FOV of the SPAD array can receive and reflect more light, which contributes more to the measurements on each pixel after passing through the scattering layer. While the intensity contribution from other parts of the object (such as the bottom part of the vertical line of ‘F’ in Fig. 4(a)) is relatively smaller, and it is more difficult to distinguish in the collected data after the diffusion of the scattering layer and thus harder to reconstruct. Further, the proposed method implements the back projection operation $\hat{\rho} = P\mathbf{y}$ on the reconstructed voxel space using the descattering result \mathbf{y} as input, so the reconstruction result $\hat{\rho}$ will be more prominent in the voxels that contribute more to the light intensity of the collected data, sometimes resulting in the uneven brightness of the reconstruction where the second horizontal line of ‘F’ in Fig. 4(e) of the proposed method is very bright while the bottom part of the vertical line of ‘F’ in Fig. 4(e) disappears. For multi-depth objects in Fig. 5, the reconstruction artifacts are more obvious due to the influence of multiple objects with different depths, but the proposed method can still reveal each target while other algorithms only recover part of the objects with low quality.

Table 1 and Table 2 provide the comparison in PSNR and SSIM of different methods where the proposed method can always gain the highest values, which further verifies the effectiveness of the proposed method for shape reconstruction. Therefore, the proposed method outperforms other methods subjectively and objectively in all experiments.

Table 1. Comparison in PSNR(dB) of different methods

Method	T	K	F	LL	LT	LF	Average
Time-gating	6.8535	7.2264	6.9127	5.9346	6.0162	4.7110	6.2757
Descattering	7.7605	7.4729	7.1319	4.1295	7.2348	5.4802	6.5350
Phasor Field	10.2247	7.3570	8.8153	7.1605	8.0537	8.0564	8.2779
Proposed	10.9885	10.7865	10.1041	9.0508	10.0550	10.4283	10.2355

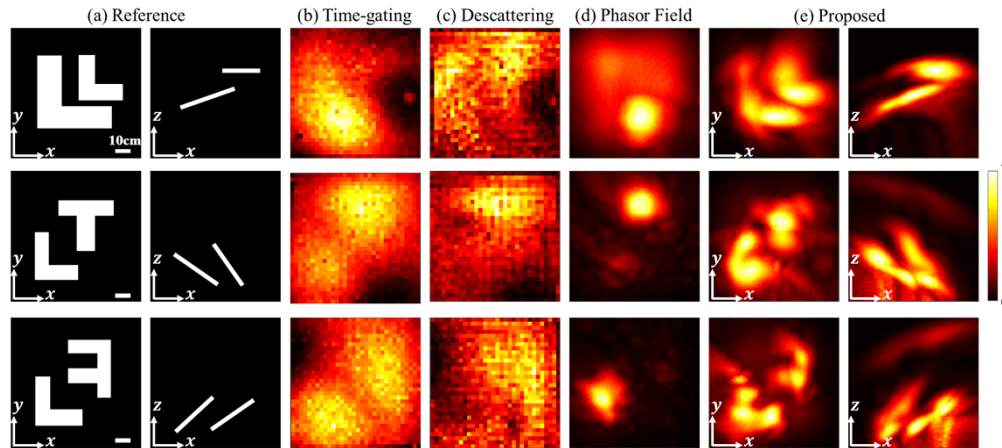


Fig. 5. Comparisons of reconstructing multi-depth Lambertian objects.

Table 2. Comparison in SSIM of different methods

Method	T	K	F	LL	LT	LF	Average
Time-gating	0.7148	0.7072	0.7123	0.6487	0.6376	0.5212	0.6570
Descattering	0.7374	0.7307	0.6914	0.4347	0.7148	0.5418	0.6418
Phasor Field	0.8683	0.7588	0.8132	0.7421	0.7939	0.7891	0.7942
Proposed	0.8832	0.8460	0.8506	0.8008	0.8328	0.8345	0.8413

Table 3 summarizes the exposure time, photon counts per temporal histogram and runtime of the above experiments, which are acquired by running the program on a PC with Intel Core i7-10750 H CPU @2.60 GHz, 16 G RAM and MATLAB R2021a under 64-bit Windows 10 operating system. It can be seen that the average photon count per temporal histogram of all the experiments captured in a single shot of 3.5s is about 1000, which is relatively low compared to the confocal scanning setup with a fast-gated single-pixel SPAD [16]. The numerical cost of the proposed method is close to the phasor field method in the order of seconds, which can be reduced by replacing the back-projection operation in Eq. (8) with the Rayleigh-Sommerfeld diffraction operation [19,20] to realize the diffractive wave propagation in the second step of the proposed method.

Table 3. Comparison of exposure time, photon count and runtime for different objects

Object	Exposure Time (s)	Photon Count	Runtime (s)		
			Descattering	Phasor Field	Proposed
T	3.5	1238	0.0469	2.4312	2.4304
K	3.5	876	0.0587	2.0188	1.7846
F	3.5	885	0.0487	2.4663	2.4209
LL	3.5	871	0.0570	2.3830	2.4601
LT	3.5	1658	0.0593	2.3826	2.3955
LF	3.5	1176	0.0575	2.3766	2.4581

3.4. Imaging limit analysis

In this section, the imaging limit of the proposed method is analyzed using the non-confocal imaging system, compared with other algorithms including time-gating, descattering and phasor field. We first set the Lambertian letter 'F' with the shape and depth information shown in Fig. 4(a) behind the polyethylene foam with different thicknesses ranging from 1 cm to 4 cm, corresponding to the one-way scattering length from 3.2 TMFPs to 12.8 TMFPs, provided in Fig. 6. The measurements are collected in a single shot with the exposure time of 3.5s and input to the reconstruction algorithms. As shown in Fig. 6(a-b), the time-gating and descattering methods are unable to distinguish any features of the object. The reconstruction of the phasor field in Fig. 6(c) recovers the object located behind the 1.5 cm thick foam, corresponding to the imaging limit of 4.8 TMFPs. In contrast, the proposed method in Fig. 6(d) can successfully reconstruct the object shape at the foam thickness of 9.6 TMFPs, corresponding to the round-trip scattering length of 19.2 TMFPs. When the thickness of the foam exceeds 12.8 TMFPs, the reconstruction of the proposed method is less effective. So, the imaging limit of the proposed method in this system setup is 9.6 TMFPs.

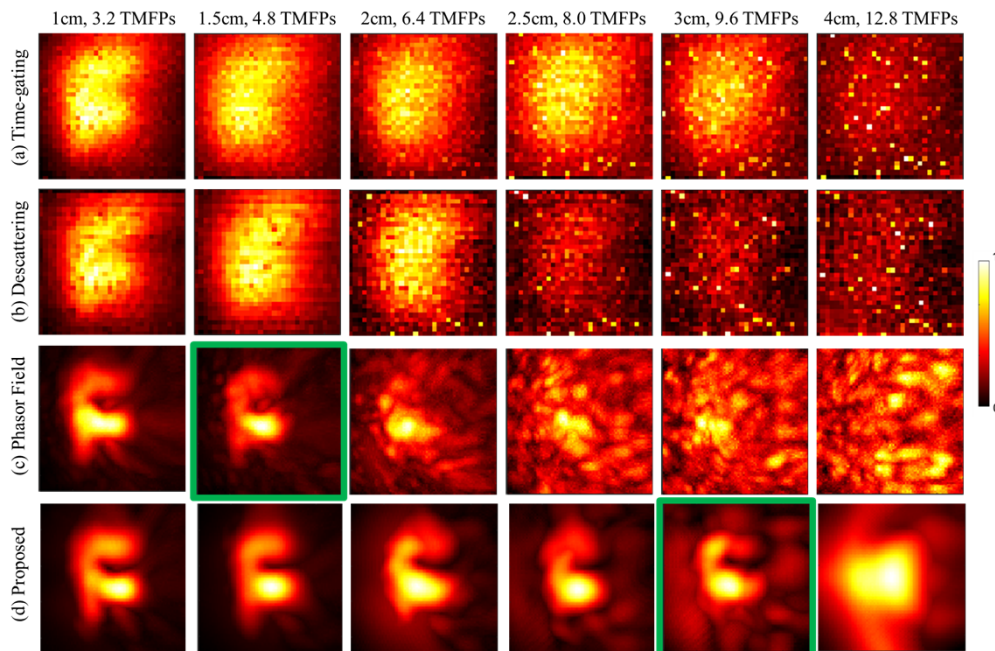


Fig. 6. Comparisons of reconstructing the Lambertian object 'F' behind the scattering layer with different thicknesses.

Theoretically, the validity range of the proposed method is unbounded since the main assumption of the modeling in Eq. ((5)–(6)) is the diffusion approximation from solving the diffusion equation, which is actually more accurate for large TMFP lengths, as demonstrated in several studies [27–29]. Besides, there is no numerical approximation in the proposed method. However, in practical experiments, the reconstruction quality is limited by the signal intensity of measurements, which is highly related to the TMFP lengths of the scattering layer. As the temporal responses of one pixel collected at different TMFP lengths shown in Fig. 7, the signal photons at 6.4 TMFPs in Fig. 7(a) have a certain magnitude and can be clearly distinguished. For 12.8 TMFPs in Fig. 7(b), the number of the signal photons is much reduced and the temporal distribution of the signal photons is wider, making signal photons too overwhelmed by the

background noise photons to provide enough signal information and form a good reconstruction. Besides, it can be clearly seen that the magnitude of the collected signal photons is much lower than the return photons directly from the scattering layer, because the SPAD array used in the experiment has no gating mode, resulting in an insufficient detection of target signal photons. Since the reconstruction quality is highly related to the signal-to-noise ratio level, a better reconstruction and a higher imaging limit can always be achieved by improving the quality of the collected measurements in the physical setup, such as increasing the incident power of the laser, using more sensitive photon detectors with the gated operation or increasing the exposure time.

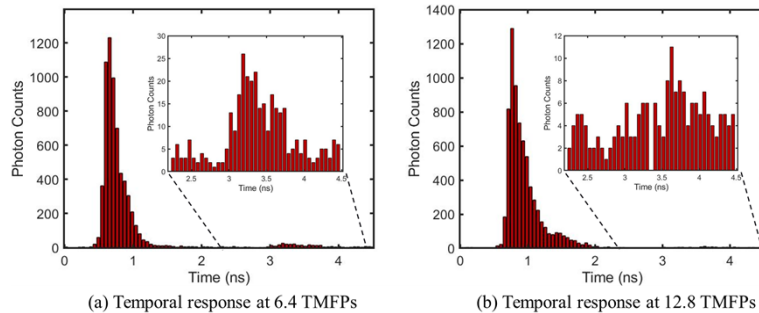


Fig. 7. Comparisons of the temporal response of the same pixel for different foam thicknesses.

To analyze the imaging capability of the proposed method with different exposure times, we conduct the experiments to reconstruct the same Lambertian letter 'F' in Fig. 4(a) behind the 3 cm thick foam of 9.6 TMFPs with a varying exposure time of 1s, 3.5s, 7s, 10s, 15s and 30s. As shown in Fig. 8, the proposed method in Fig. 8(d) can reveal most of the object features until the exposure time is less than 3.5s while other methods in Fig. 8(b-c) perform poorly with no legible object information. As the exposure time increases, the intensity of the object signal photons in Fig. 8(a) increases and the reconstruction effect in Fig. 8(d) improves a bit but not significantly. This is because the scattering strength of 9.6 TMFPs has reached the imaging limit of the proposed method, even if the exposure time is increased beyond 3.5s, the captured signal intensity and SNR are relatively low and thus limit the reconstruction. In the general case of 2cm-thick foam corresponding to 6.4 TMFPs, the proposed method can work well in a single shot with an exposure time of less than 3.5s. To recover a single Lambertian letter 'T' with an approximate length of ~45 cm in the reconstructed volume of 60cm×60cm×50 cm, the scanning confocal diffuse tomography (CDT) method [16] uses a single-pixel fast-gated SPAD with an exposure time of 3.5s at each scan position with a 32×32 scanning grid. Here, the proposed method only requires a single shot with an exposure time of 3.5s, which speeds up more than 200 times than the scanning CDT method [16] with a total scanning acquisition time of dozens of minutes. Thus, the proposed method has the potential to observe the dynamic scenes using a short acquisition time in the order of seconds.

To verify the robustness of the proposed method for reconstructing objects with different albedo surfaces, we test the reconstruction limit of the retroreflective object 'F' made of microprism reflective films in the same setup with the same size and location in Fig. 4(a) and Fig. 6. Since the retroreflective surface can considerably increase the direct-reflected target signal photons with a much larger difference from the background scattered ones, the reconstruction performance of all algorithms is improved to some extent in Fig. 9 compared to the Lambertian object reconstruction in Fig. 6. The time-gated and descattering result in Fig. 9(a-b) can provide partial discernible object features while still cannot recover the exact shape and size of the object. Except for the reconstruction artifact of the extra horizontal line in the first column of Fig. 9(c) and Fig. 9(d) for

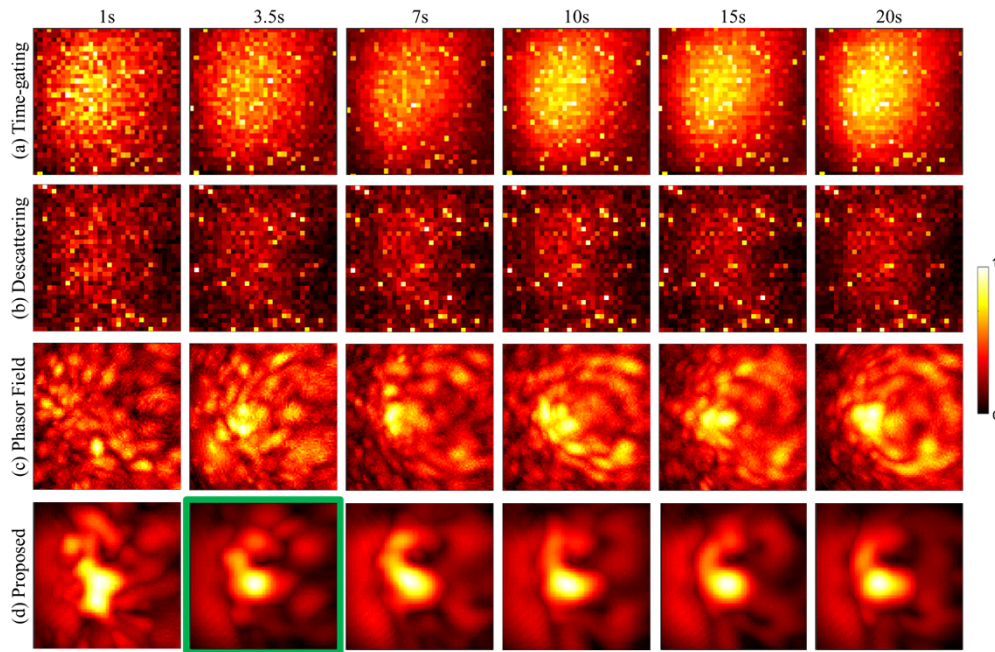


Fig. 8. Comparison of reconstructing the Lambertian object 'F' behind the 3 cm thick foam of 9.6 TMFPs with different exposure times.

3.2 TMFPs, the reconstruction quality of the proposed method in Fig. 9(d) is always better than that of the phasor field method in Fig. 9(d), especially for higher TMFPs where the diffusion of the scattering layer is not taken into account by the phasor field method. Therefore, the imaging limits of the phasor field in Fig. 9(c) remain at 4.8 TMFPs while the proposed method in Fig. 9(d) can accurately reconstruct the object shape at the foam thickness of 9.6 TMFPs. It is found in the two experiments of Fig. 6 and Fig. 9 that the proposed method outperforms other algorithms in the reconstruction quality and imaging limit, which further demonstrates the effectiveness of the proposed method in 3D scattering reconstruction.

Similar to the phasor field method [19], the spatial resolution limit of the proposed method in the non-confocal imaging system is also related to the Rayleigh diffraction limit $\Delta_x = 0.61c\sigma L/d$, where c is the speed of light in vacuum, σ represents the full width at half maximum of the pulse, L is the imaging distance, d denotes the width of the detected field of view (FOV) of the SPAD array on the scattering layer. Therefore, as the imaging distance L increases, or the width of FOV d decreases, the spatial resolution limit decays linearly. The choice of the object size on the order of 10 cm is mainly limited by the spatial resolution of the system $\Delta_s = 1.65$ cm (corresponding to the time resolution of the system 55ps) and the spatial resolution of the proposed method $\Delta_x = 0.61c\sigma L/d$. The proposed method cannot recover the object structure with its size smaller than the minimal spatial resolution, i.e., $\min(\Delta_s, \Delta_x)$. For the object with a bigger size, the imaging distance L is increased, the width of FOV d should be increased to maintain the same spatial resolution and produce a good reconstruction. Besides, the maximum detectable depth in the system can be calculated by the number of time bins and the time resolution of the SPAD array, which corresponds to the round-trip distance of $910 \times 55\text{ps} \times 3 \times 10^8\text{m/s} = 15.015$ m while the depth of field of the proposed method is mainly limited by the reduction in spatial resolution and the detected photon intensity reflected by the object that can be detected by the SPAD array. Therefore, the object cannot be too close to the scattering layer to receive the light of the laser

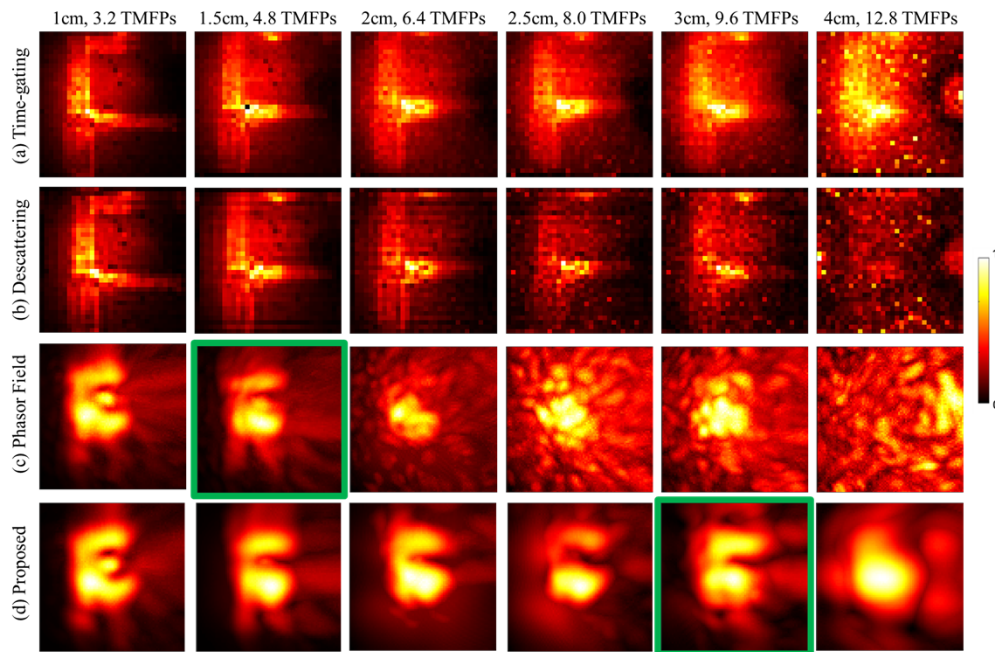


Fig. 9. Comparisons of reconstructing the retroreflective object 'F' behind the scattering layer with different thicknesses.

incident from only one point, nor too far away to reflect enough light intensity to be collected through the scattering layer and form a clearly identifiable reconstruction.

4. Conclusions and discussions

In this paper, with a non-confocal time-of-flight-based imaging system, a scan-free time-of-flight-based 3D reconstruction method is proposed to recover the 3D object structure behind the scattering layer. To completely describe the non-confocal scattering light propagation with time, we first build the non-confocal time-of-flight-based scattering imaging model, mapping the 3D object shape information to the time-resolved measurements. Then a 3D reconstruction method consisting of the deconvolution and diffractive wave propagation is developed to invert the effects caused by the scattering diffusion and the free-space propagation, reshaping the scattered measurements to recover the object information. Experiments on a real scattering imaging system show that the proposed method outperforms existing methods in terms of three-dimensional reconstruction accuracy and imaging limit. In this system setup, three-dimensional objects located behind the scattering layer of 9.6 TMFPs, corresponding to the round-trip scattering length of 19.2 TMFPs, can be reconstructed by the proposed method.

Although the proposed method relies on pre-calibration or prior knowledge of the scattering layer including the thickness and the scattering property, the proposed method is also robust to a certain degree of parameter perturbation and can give considerable reconstruction results. To achieve the high-quality reconstruction of completely unknown scattering media, the blind deconvolution algorithm should be explored and used in the first step of the proposed method to better remove the blurring effect of scattering media. Besides, the proposed method uses an exposure time of 3.5s in the experiment, much faster than confocal methods taking dozens of minutes to scan, which provides a potentially promising direction for 3D scattering reconstruction

in dynamic scenes. While the proposed method with the second-order exposure time is still not instantaneous and thus places restrictions on the moving speed of the object to be reconstructed in dynamic scenes. To achieve real-time 3D scattering reconstruction in future work based on the proposed method, the exposure time should be further shortened by improving the experimental setup to detect more effective signal photons with some approaches, such as using a more sensitive gated SPAD array, increasing the incident power of the laser and so on. In addition, data preprocessing operations can also be applied to the low SNR measurements with a very short exposure time to increase the data quality and thus enhance the 3D reconstruction effect.

Funding. Shenzhen Project (JSGG20210802154807022); NSF of Guangdong Province (2023A1515012716); Zhuhai Basic and Applied Basic Research Foundation (ZH22017003210067PWC).

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. J. Bertolotti, E. G. Van Putten, C. Blum, A. Lagendijk, W. L. Vos, and A. P. Mosk, "Non-invasive imaging through opaque scattering layers," *Nature* **491**(7423), 232–234 (2012).
2. O. Katz, E. Small, and Y. Silberberg, "Looking around corners and through thin turbid layers in real time with scattered incoherent light," *Nat. Photonics* **6**(8), 549–553 (2012).
3. O. Katz, P. Heidmann, M. Fink, and S. Gigan, "Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations," *Nat. Photonics* **8**(10), 784–790 (2014).
4. E. Edrei and G. Scarcelli, "Optical imaging through dynamic turbid media using the Fourier-domain shower-curtain effect," *Optica* **3**(1), 71–74 (2016).
5. S. Feng, C. Kane, P. A. Lee, and A. D. J. P. r. l. Stone, "Correlations and fluctuations of coherent wave transmission through disordered media," *Phys. Rev. Lett.* **61**(7), 834–837 (1988).
6. I. Freund, M. Rosenbluh, and S. Feng, "Memory effects in propagation of optical waves through disordered media," *Phys. Rev. Lett.* **61**(20), 2328–2331 (1988).
7. X. Xie, H. Zhuang, H. He, X. Xu, H. Liang, Y. Liu, and J. Zhou, "Extended depth-resolved imaging through a thin scattering medium with PSF manipulation," *Sci. Rep.* **8**(1), 4585 (2018).
8. O. Salhov, G. Weinberg, and O. Katz, "Depth-resolved speckle-correlations imaging through scattering layers via coherence gating," *Opt. Lett.* **43**(22), 5528–5531 (2018).
9. X. Jin, Z. Wang, X. Wang, and Q. Dai, "Depth of field extended scattering imaging by light field estimation," *Opt. Lett.* **43**(20), 4871–4874 (2018).
10. L. Wang, P. Ho, C. Liu, G. Zhang, and R. Alfano, "Ballistic 2-D imaging through scattering walls using an ultrafast optical Kerr gate," *Science* **253**(5021), 769–771 (1991).
11. A. Redo-Sanchez, B. Heshmat, A. Aghasi, S. Naqvi, M. Zhang, J. Romberg, and R. Raskar, "Terahertz time-gated spectral imaging for content extraction through layered structures," *Nat. Commun.* **7**(1), 12665 (2016).
12. A. Maccarone, A. McCarthy, X. Ren, R. E. Warburton, A. M. Wallace, J. Moffat, Y. Petillot, and G. S. Buller, "Underwater depth imaging using time-correlated single-photon counting," *Opt. Express* **23**(26), 33911–33926 (2015).
13. G. Satat, M. Tancik, and R. Raskar, "Towards photography through realistic fog," in *International Conference on Computational Photography* (IEEE, 2018), pp. 1–10.
14. A. Maccarone, F. M. Della Rocca, A. McCarthy, R. Henderson, and G. S. Buller, "Three-dimensional imaging of stationary and moving targets in turbid underwater environments using a single-photon detector array," *Opt. Express* **27**(20), 28437–28456 (2019).
15. A. Lyons, F. Tonolini, A. Boccolini, A. Repetti, R. Henderson, Y. Wiaux, and D. Faccio, "Computational time-of-flight diffuse optical tomography," *Nat. Photonics* **13**(8), 575–579 (2019).
16. D. B. Lindell and G. Wetzstein, "Three-dimensional imaging through scattering media based on confocal diffuse tomography," *Nat. Commun.* **11**(1), 4517 (2020).
17. 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).
18. 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).
19. 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**(7771), 620–623 (2019).
20. X. Liu, S. Bauer, and A. Velten, "Phasor field diffraction based reconstruction for fast non-line-of-sight imaging systems," *Nat. Commun.* **11**(1), 1645 (2020).
21. 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**(1), 6526 (2021).

22. C. Pei, A. Zhang, Y. Deng, F. Xu, J. Wu, U. David, 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**(20), 32349–32364 (2021).
23. C. Wu, J. Liu, X. Huang, Z.-P. Li, C. Yu, J.-T. Ye, J. Zhang, Q. Zhang, X. Dou, and V. K. Goyal, "Non-line-of-sight imaging over 1.43 km," *Proc. Natl. Acad. Sci.* **118**(10), e2024468118 (2021).
24. P. Luesia, M. Crespo, A. Jarabo, and A. Redo-Sanchez, "Non-line-of-sight imaging in the presence of scattering media using phasor fields," *Opt. Lett.* **47**(15), 3796–3799 (2022).
25. O. S. Faragallah, H. El-Hoseny, W. El-Shafai, W. Abd El-Rahman, H. S. El-Sayed, E.-S. M. El-Rabaie, F. E. Abd El-Samie, and G. G. Geweid, "A comprehensive survey analysis for present solutions of medical image fusion and future directions," *IEEE Access* **9**, 11358–11371 (2021).
26. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Process.* **13**(4), 600–612 (2004).
27. K. Yoo, F. Liu, and R. Alfano, "When does the diffusion approximation fail to describe photon transport in random media?" *Phys. Rev. Lett.* **64**(22), 2647–2650 (1990).
28. T. G. Papazoglou, W. Liu, A. Vasilio, R. Grassmel, E. Papagiannakis, and C. Fotakis, "Limitations of diffusion approximation in describing femtosecond laser transillumination of highly scattering media of biological significance," *Appl. Phys. Lett.* **67**(25), 3712–3714 (1995).
29. S. A. Carp, S. A. Prahl, and V. Venugopalan, "Radiative transport in the delta-P 1 approximation: accuracy of fluence rate and optical penetration depth predictions in turbid semi-infinite media," *J. Biomed. Opt.* **9**(3), 632–647 (2004).