An improved algorithm of NLOS imaging based on Bayesian statistics

LUZHE HUANG, XIAOBIN WANG, YIFAN YUAN, SONGYUN GU AND YONGHANG SHEN*

State Key Laboratory of Modern Optical Instrumentation, College of Optical Science and Engineering, Zhejiang University, Hangzhou, 310027, China
physyh@zju.edu.cn

Abstract: The recovery of obscured objects is an important goal in imaging and has been approached by exploiting, e.g., coherence properties, ballistic photons and penetrating wavelengths. In this article, a robust reconstruction non-line-of-sight (NLOS) algorithm was proposed based on the Bayesian statistics using the temporal, spatial and intensity information of each signal. Compared with conventional back-projection methods, this algorithm is able to handle random errors in data, and to image occluded objects with higher quality. Adjustable compensation mechanism in our method is effective in dealing with the diversity of objects in reflective characteristics. Our algorithm was demonstrated to be efficient with both simulated data and experiment data. Additionally, advantages over existing methods and further improvement of the proposed algorithm are discussed.

1. Introduction

Imaging objects through occlusion has been a problem of fundamental significance for a long time. A wide spectrum of methods have been developed, e.g., imaging through scattering medium by optical coherence tomography (OCT) [1], speckle auto-correlations [2, 3] and two-photon "ghost" imaging [4]. Non-line-of-sight (NLOS) imaging technology is to "see" a scene out of sight, for example, objects around a corner or hidden by some shelter. At first, object detection and tracking in NLOS situations have been demonstrated [5–7]. Furthermore, photon time-of-flight (ToF) and corresponding algorithms, which is common in light detection and ranging (LIDAR) systems to recover 3D information of objects through direct reflections [8–11], has also been applied to image beyond the direct line of sight. On this specific problem, many algorithms have been proposed to improve both efficiency and accuracy of imaging [12–20].

One of the major NLOS algorithms is back-projection (BP), and many derivative methods are modified based on BP. Using an ultrafast pulse laser, MIT media lab proposed back-projection algorithm and recovered three-dimensional information of occluded objects based on ToF imaging in 2012 [13, 14]. Their results demonstrated good spatial resolution at cm level. In 2015, single-photon avalanche diodes were utilized to reconstruct objects hidden from view [18]. Velten et al. improved the BP algorithm and performed fast back-projection reconstruction on GPU by computing the intersection of space-time manifolds [19], but their work did not adopt too many changes towards the experiment setups and the rationale compared with BP. In 2018, O'Toole et al. used a co-design of a confocal scanning technique and a computationally efficient inverse method that facilitate fast, high-quality reconstructions of hidden objects [20]. This modality makes use of fast Fourier transform (FFT) to reduce the time consumption of reconstruction algorithm considerably. Though it requires specially designed experiment setups, their method exhibits almost real time reconstruction with good resolution, indicating its potential applications in many realms such as auto-driving.

Nevertheless, there are still some challenges remaining in the field of NLOS imaging. As for the xconventional BP approach, its obvious drawback is the insufficient physical modeling of

NLOS process, e.g., the overlook of randomness in acquired data caused by noise in detectors, the quantum nature of photons, etc. And the lack of modeling reflective characteristics of the wall and objects. Consequently, BP cannot solve the inverse problem accurately. On the other hand, confocal NLOS (CNLOS) method overcomes some problems and consumes less time for reconstruction. However, in spite of the high efficiency of CNLOS, its advantages in time efficiency and reconstruction quality are tightly restricted to confocal systems. And the absence of reflective characteristics in the CNLOS algorithm would affect its performance with diffusal objects to some extent.

In this article, we propose a novel NLOS reconstruction algorithm from a Bayesian statistical perspective. In Section 2, we explicate the theoretical fundament of the proposed method. On the basis of a probabilistic model, the proposed algorithm takes both reflection diversity of objects and inevitable random errors in data into consideration. By introducing a logarithmic-probability loss function, the reconstruction problem can be solved with the same time complexity with BP. Then, in Section 3, the proposed algorithm was performed on both simulation and experiment data. We compare the proposed method with BP to demonstrate its improvement in reconstruction accuracy and quality. Finally, we discussed benefits and drawbacks of the proposed method when comparing with existing methods.

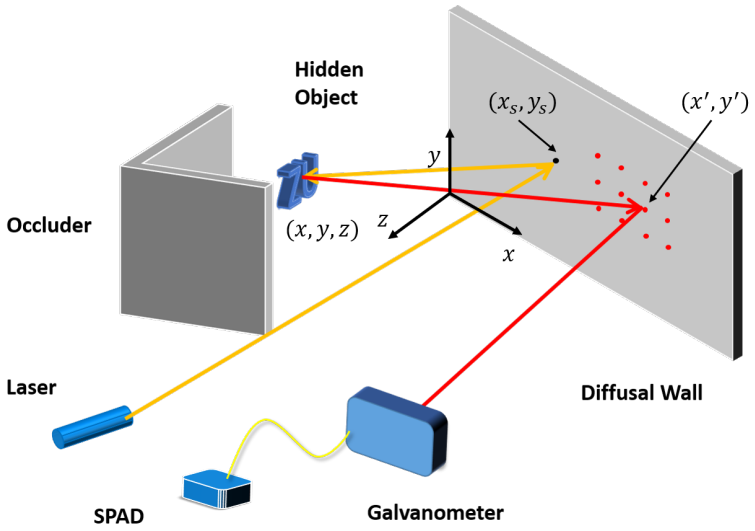

Fig. 1. **The schematic diagram of experiment setups of a B-scan process.** (x, y, z) are coordinates of the object's voxel, (x_s, y_s) refers to the beam spot shone on the wall, and (x', y') are coordinates of scanning point on the wall. (x', y') and (x_s, y_s) would change between scannings.

2. Probability-based Model and Reconstruction

In this section, a probability-based model describing the NLOS process is introduced and the reconstruction algorithm is presented. At first, several basic assumptions of our modality should be mentioned. Basically, our probability-based model requires the following assumptions: no inter-object reflections in the object space, no object occluded by other hidden objects. For most occasions, we generally suppose the wall and the objects have typical Lambertian surface, i.e., isotropic luminance in all directions, to simplify calculations. This model is applicable to other reflectors like retroreflective objects as well, which will be demonstrated in Section 3.

2.1. Bayesian formula and posterior probability

Bayesian statistics are widely used in parameter inference, and machine learning methods [21]. For example, naive Bayesian classifier demonstrates high accuracy in bacterial rRNA sequences assignments [22]. We introduce the concept of prior and posterior probability into this problem, in order to deal with stochastic errors and noises due to the sampling process and the constraints of instruments.

First of all, we should view the three bounces process from a probabilistic perspective. As in the experiment setup illustrated in Fig. 1, a photon is diffusively emitted from the laser spot s at (x_s, y_s) on the wall at first, scattered by the voxel v_i with coordinate (x, y, z), and finally detected at (x', y') on the wall. The corresponding probability $p(x', y'|v_i)$ that this photon is detected flying along the concerning path is proportional to the illuminance distribution and can be easily expressed given various assumption of reflection properties of the objects and wall, e.g., the same albedo and Lambertian reflection for each voxel. In that case, we use $p_l(x', y'|v_i)$ to denote the probability for simplicity and it is expressed as

$$p_l(x', y'|v_i) = \frac{z^2}{(x-x')^2 + (y-y')^2 + z^2} \cdot \frac{z^2}{(x-x_s)^2 + (y-y_s)^2 + z^2}. \tag{1}$$

As mentioned above, owing to inevitable factors including the noise in detectors and the quantum nature of photon, there are always random errors in the measured ToF compared with the ideal ToF. As a matter of fact, these errors prevent object reconstruction with expected accuracy and lead to blurring of reconstructed image. The accuracy and blur issue are two major drawbacks of conventional NLOS algorithms. Here, a normally distributed random variable is introduced to characterize the random error. Such assumption allows enough flexibility and meanwhile is amenable in both analysis and interpretation. Because all photons are independent with each other, the probability of ToF error $p(h(t,n)|h(t_0,n))$ is given by

$$p_h(h(t,n)|h(t_0,n)) = (2\pi\sigma^2)^{-\frac{n}{2}} \mathrm{e}^{-\frac{n}{2\sigma^2}(t-t_0)^2}, \tag{2}$$

where t_0, n respectively refer to the ideal flight time of the reflection process and the number of photons received. $h(t,n)$ denotes a NLOS signal with flight time t and n photons, and σ is the standard variance of the random variable. The ideal ToF t_0 is given as

$$t_0 = \frac{1}{c}(\sqrt{(x-x')^2 + (y-y')^2 + z^2} + \sqrt{(x-x_s)^2 + (y-y_s)^2 + z^2}), \tag{3}$$

where c is the light speed. Eq. (2) models the NLOS signal as a Gaussian random variable with standard variance σ around the ideal ToF.

In general, adopting a prior such as Laplacian or Gaussian could improve the performance of this algorithm. But here we set prior probability for each voxel in following derivation. Therefore, the posterior probability of voxel v_i can be easily expressed by:

$$\begin{aligned} \pi(v_i|h(t,n;x',y')) = \frac{p(v_i)p(h(t,n;x',y')|v_i)}{\sum_i p(v_i)p(h(t,n;x',y')|v_i)} &\propto p(x',y'|v_i)p_h(h(t,n)|h(t_0,n)) \\ &\propto (\mathrm{e}^{-\frac{1}{2\sigma^2}(t-t_0)^2} \cdot p(x',y'|v_i))^n. \end{aligned} \tag{4}$$

2.2. Log-probability loss function

The formula above is able to model the projection process of a single voxel and is likely to give the true voxel a maximized posterior probability. But in computation, to avoid overflowing, a log-probability loss function is introduced to measure the probability of a certain voxel.

Based on Eq. (4), loss function of each voxel is given by

$$Loss(v_i) = \frac{1}{n_{x'} n_{y'} n_t} \sum_{(x',y',t)} nt^4 \cdot (\log(p(x', y'|v_i)) - \frac{(t - t_0)^2}{2\sigma^2}). \tag{5}$$

Here, the t^4 term compensates the quadratic intensity attenuation when light propagates in free space.

In this way, we are able to calculate a heatmap of the whole object space. In practice, an empirical upper limit is set for the loss of every single signal $h(t, n; x', y')$ because the quadratic loss should be limited locally around its corresponding ground-truth voxel.

$$Loss(v_i) = -\frac{1}{n_{x'} n_{y'} n_t} \sum_{(x',y',t)} nt^4 \cdot \min(\frac{(t - t_0)^2}{2\sigma^2} - \log(p(x', y'|v_i)),\ L), \tag{6}$$

where L stands for the upper limit of loss. According to our test, it is optimal to set the upper limit L at three to five times the value of $\frac{\tau^2}{\sigma^2}$, where τ is the temporal resolution of the system. An inappropriate selection of L may degrade the performance of the proposed algorithm. A schematic diagram is displayed as Fig. 2, and the difference between BP and our method can be clearly seen. In regard of BP, only voxel nearest to the red curve, which is part of the ellipse determined by ToF and two spots, are estimated as contributing voxels. But the ubiquitous random errors in data would inevitably put the accuracy of this estimation in jeopardy. On the contrary, the proposed algorithm assigns a loss value to every voxel, which mitigates the influence of random errors as well as the discretization effect.

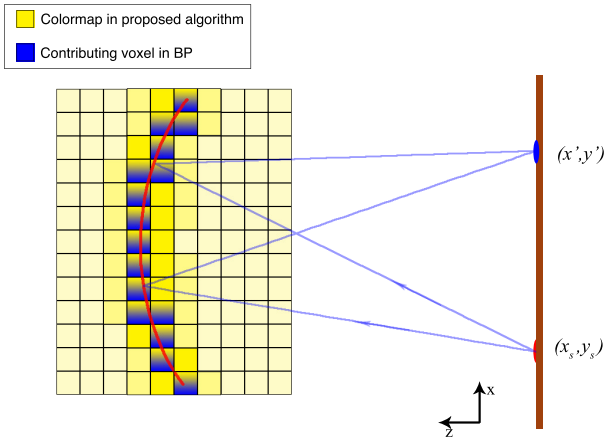

Fig. 2. **The schematic diagram of the reconstruction process.** The red curve is part of an ellipse with foci at the laser spot (x_s, y_s) and the scanning spot (x', y') and focus length equal to the flight length of received photons. Blue voxels represent contributing voxels in BP algorithm. The yellow color indicates the loss in the proposed method for each voxel such that lighter color represents higher loss.

Moreover, by adjusting σ, different weight can be assigned to the influence of reflective properties of the wall and objects. As we decrease σ approxiamtely to 0, this algorithm is basically identical to conventional BP, because only the voxel matches ToF best has a maximal loss value.

3. Experiment Result

In this section, the proposed algorithm was performed on both simulation data and experiment data. Besides, imaging of both Lambertian diffusal and retroreflective objects are presented, in order to verify its adaptability to different characteristics of reflectors.

3.1. Reconstruction based on simulation data

In the simulation, we use a typical three-bounce transient rendering process for the generation of NLOS signals with Poisson noise added to simulate shot noise in the system [23–25]. For data acquisition, the whole process is done by three separate sub-processes. Firstly, the pulsed laser is focused onto a certain pixel on the wall while the galvanometer and the coupled imaging system are focused onto another pixel on the wall. Received light is coupled into a fiber which is connected to a SPAD at the other end. The SPAD will record the received photon numbers and the arrival time during a given time interval. The process above is called A-detect. Then, the galvanometer scans different pixels along a 2D array, which is denoted by B-scan. Moreover, the laser spot is changed for several times after a B-scan, and new scans are repeated. We call it C-frame. This operation could eliminate the influence of unexpected noises and improve the imaging quality.

It is worth noting that even though due to oblique reflection and the fixed angle of view of the system, the acquired illuminance on each pixel is not the actual one. But this effect gets cancelled naturally and received photon numbers do not need angle or distance correction because a larger area on the wall is imaged with respect to further pixels.

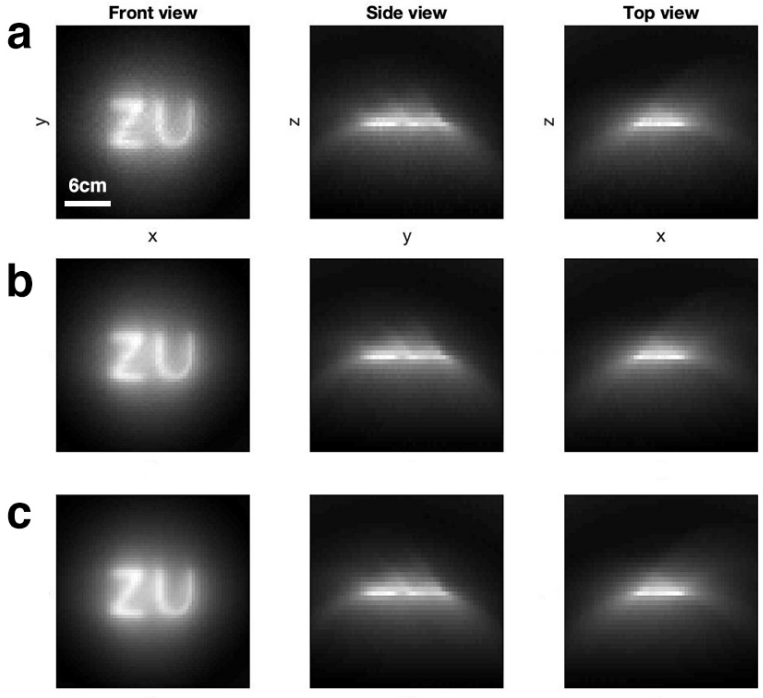

Fig. 3. **Reconstructed object of diffuse ZU letters.** Each row illustrates three views of reconstruction result at various σ. For (a), $\sigma^2 = 0.5$, (b) $\sigma^2 = 1$ and (c) $\sigma^2 = 2$.

Here we select 'Z' and 'U' letters as the objects, and set the parameter of Poisson noise to control the SNR about 25 dB. Each pixel is 8×8 mm^2, and voxel is $4 \times 4 \times 4$ mm^3. 512 bins are

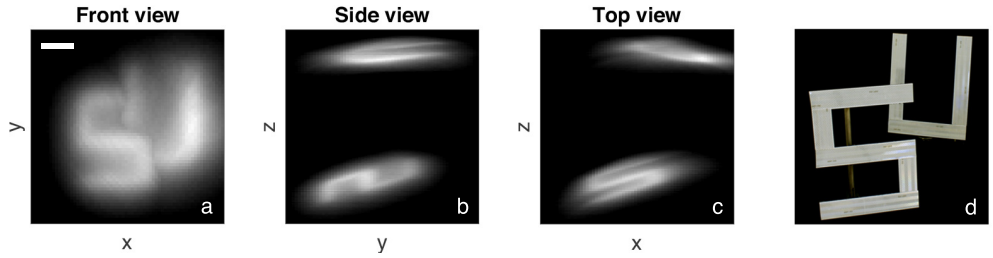

Fig. 4. **Reconstructed object of retroreflective SU letters.** (a), (b), (c) are three views of reconstructed object respectively. (d) is the shot image of real object. In this case, $\sigma^2 = 2$. Scale bar = 10 cm.

captured in each A-detect with a 10 ps temporal resolution at 32×32 sample points on the wall, and a total of $64 \times 64 \times 64$ voxels are gone through in the object space. To improve the image quality, we average the reconstruction results of four C-frames.

In Fig. 3, reconstruction results with different parameters are exhibited. The result exhibits good reconstruction performance. One can easily recognize the letters, "Z" and "U". Besides, by comparing Fig. 3a and c, we can conclude σ smoothes images, leading to better recontruction quality. The contrast of recovered image can be further enhanced by applying a proper filter.

3.2. Reconstruction based on confocal data

In this section, we conduct our algorithm on confocal data of retroreflective 'SU' letters, which is provided by O'Toole et al. of Stanford University in Ref. [20]. The data is sampled on a 70×70 cm^2 area at 64×64 sample points with 2048 time bins for the temporal histograms, and their experiment apparatuses are stated in Ref. [20]. The temporal resolution of the data is 4 ps, but we downsample it to 16 ps, which means the downsampled signal size is $64 \times 64 \times 512$. And reconstruction is conducted on a $64 \times 64 \times 64$ voxel grid.

It is worthy to note that although their experiment setups are based on the CNLOS method, our algorithm can be adapted to their data. Because of the retroreflective characteristic of objects, we modify compensation item in Eq. (5) from t^4 to t^2, and adjust $p(x', y'|v_i)$ correspondingly. The result exhibits clearly two letters at different distances as shown in Fig. 4.

Application of filters such as Laplacian [13] is promising to improve the reconstruction quality. For the purpose of assessing the performance of the proposed algorithm, comparison between conventional BP [13, 14] and light cone transform (LCT) algorithm is presented in Fig. 5. By comparing Fig. 5a and b, reconstruction of the proposed method without filtering is able to mitigate the blurring and reveal letter "U" in the behind. With respect to LCT algorithm in Fig.

Table 1. **Comparison of algorithms on "SU" data.** The "SU" data is sampled on a 70 cm * 70 cm area, and the temporal resolution is 16 ps. All running time is measured on a quad-core Intel i5 CPU at 2.3 GHz.

Algorithm	Proposed algorithm	Backprojection	LCT
Data size	64*64*512		
Number of voxels	64*64*160	64*64*512	64*64*512
Voxel size (mm)	10*10*4	10*10*2.4	10*10*2.4
Time consumption (s)	~ 700	~ 600	~ 1

5d, which intrinsically involves a 3D Wiener filter, by applying a Laplacian of Gaussian filter along z axis, the proposed method can achieve alike resolution and clarity, but using fewer grid voxels. Besides, due to compensation scheme in our method, objects in the distance are imaged with equal brightness so more details of distant objects can be imaged.

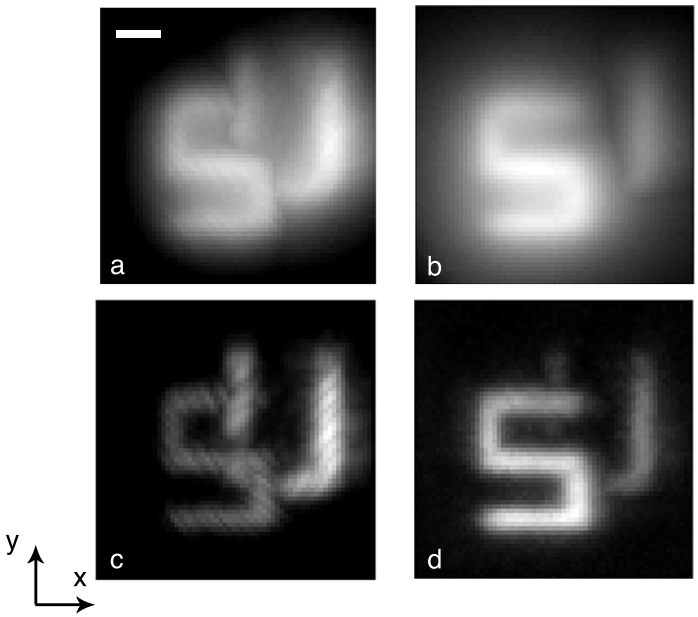

Fig. 5. **Comparison between various algorithms.** Reconstruction result of "SU" letters by (a) the proposed method, (b) BP algorithm, (c) the proposed method with LoG filter and (d) LCT method. Scale bar = 10 cm.

Parameters and running time of results are listed in Table 1. All algorithms are programmed in MATLAB and run on a MacBook Pro with Intel i5 quad-core processor at 2.3 GHz (up to 3.8 GHz) and 16 GB memory. Though the running time is slightly longer than BP, theoretically, the proposed method has a time complexity of $O(N^5)$, assuming the hidden volume of resolution $N \times N \times N$ and the acquired data of size $N \times N \times N$. This means the proposed algorithm has equal computational complexity with BP [20] except for a constant difference. Besides, owing to the probability model mitigating the inaccuracy induced by discretization, the voxel grid used for the proposed algorithm is only about one third of that of the other two algorithms, which means less memory demand.

4. Discussion

Basically, in this article, we proposed a novel probability-based NLOS reconstruction algorithm. The proposed Bayesian statistical model enables us to handle various factors in NLOS system, including the stochastic errors in data. It is also versatile to various reflective characteristics of reflectors as well as both confocal and conventional scanning modes. The essence of utilizing Bayesian statistics lies on the clearness of forward NLOS process, and thus the solution of the inverse problem can be easily calculated through Bayes principle.

Compared with existing algorithms, the proposed algorithm has a better physical model depicting the NLOS process. Moreover, through the introduction of the local quadratic loss function, our algorithm is robust to the minor random noise in NLOS data, which is ubiquitous due to lots of factors such as jitter of detectors, quantum nature of photons and the system's

fixed angle of view, etc. Another principle advantage is that our algorithm models all kinds of reflectors directly regardless of whether they are Lambertian or retroreflective reflectors, and is able to reconstruct objects with good quality in both situations.

Though employing a comprehensive probability model, this algorithm has the same time complexity of $O(N^5)$ with BP and does not need complicated optimization. Previous research [20, 26] of O'Toole et al. utilized a probability model to handle the Poisson-distributed noise of the photodetector. However, their method mainly takes the randomness of detectors into account and entails a time-consuming global optimization algorithm. Based on experiment, our proposed algorithm is able to image occluded objects with alike quality with LCT, with about one third memory requirement of that. It is true that LCT algorithm runs faster on confocal data at $O(N^3 \log N)$, but the limit of LCT method such as the complicated confocal system exposing the detector to direct reflections, restricts its applicability and thus cannot be ignored. Further optimization of the current proposed algorithm on confocal setups such as adopting FFT and GPU acceleration, the time consumption of the proposed algorithm may possibly decrease by orders of magnitude.

5. Conclusion

In this article, we derive a probability-based model for solving NLOS problem to handle random error in data acquisition and variable reflective characteristics of reflectors, and propose a new algorithm correspondingly. We also validate its applicability on both Lambertian as well as retroreflective objects. The result demonstrates good image quality than conventional BP methods. Besides, this algorithm is also demonstrated applicable to CNLOS setups. Finally, comparison with existing NLOS algorithms is shown and benefits of the proposed algorithm are discussed.

As a universal method, the proposed algorithm and model can be applied to various NLOS situations. Providing better reconstruction quality and less memory consumption, this work is able to advance the application of NLOS technologies in a variety of fields, such as battlefield surveillance, medical diagnosis, archaeological detection and automatic driving. We hope that our work will spur more research into this promising technology.

Code Availability

The MATLAB code for our proposed algorithm is availbale from the authors on request.

Acknowledgment

This program is sponsored by Zhejiang University through the National Undergraduate Training Programs for Innovation and Entrepreneurship. The authors are grateful to M. O'Toole et al. for sharing of their confocal data.

Disclosures

The authors declare that there are no conflicts of interest related to this article.

References

1. D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gergory, C. A. Puliafito, and J. G. Fujimoto, "Optical coherence tomography," Science **254**, 1178–1181 (1991).
2. 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 (2014).
3. 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).
4. R. S. Bennink, S. J. Bentley, and R. W. Boyd, ""Two-photon" coincidence imaging with a classical source," Phys. Rev. Lett. **89**, 113601 (2002).
5. J. Klein, C. Peters, J. Martín, M. Laurenzis, and M. B. Hullin, "Tracking objects outside the line of sight using 2D intensity images," Sci. Rep. **6**, 32491 (2016).

6. G. Gariepy, F. Tonolini, R. Henderson, J. Leach, and D. Faccio, "Detection and tracking of moving objects hidden from view," Nat. Photonics **10**, 23–26 (2016).
7. S. Chan, R. E. Warburton, G. Gariepy, J. Leach, and D. Faccio, "Non-line-of-sight tracking of people at long range," Opt. Exp. **25**, 10109–10117 (2017).
8. B. Schwarz, "LIDAR: Mapping the world in 3D," Nat. Photonics **4**, 429 (2010).
9. A. McCarthy, N. J. Krichel, N. R. Gemmell, X. Ren, M. G. Tanner, S. N. Dorenbos, V. Zwiller, R. H. Hadfield, and G. S. Buller, "Kilometer-range, high resolution depth imaging via 1560 nm wavelength single-photon detection," Opt. Exp. **21**, 8904–8915 (2013).
10. A. Kirmani, D. Venkatraman, D. Shin, A. Colaço, F. N. Wong, J. H. Shapiro, and V. K. Goyal, "First-photon imaging," Science **343**, 58–61 (2014).
11. D. Shin, F. Xu, D. Venkatraman, R. Lussana, F. Villa, F. Zappa, V. K. Goyal, F. N. Wong, and J. H. Shapiro, "Photon-efficient imaging with a single-photon camera," Nat. Commun. **7**, 12046 (2016).
12. A. Sume, M. Gustafsson, A. Jänis, S. Nilsson, J. Rahm, and A. Örbom, "Radar detection of moving objects around corners," in *Radar Sensor Technology XIII,* vol. 7308 (International Society for Optics and Photonics, 2009), p. 73080V.
13. 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).
14. O. Gupta, T. Willwacher, A. Velten, A. Veeraraghavan, and R. Raskar, "Reconstruction of hidden 3D shapes using diffuse reflections," Opt. Exp. **20**, 19096–19108 (2012).
15. M. Laurenzis and A. Velten, "Nonline-of-sight laser gated viewing of scattered photons," Opt. Eng. **53**, 023102 (2014).
16. F. Heide, L. Xiao, W. Heidrich, and M. B. Hullin, "Diffuse mirrors: 3D reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition,* (2014), pp. 3222–3229.
17. A. Kadambi, H. Zhao, B. Shi, and R. Raskar, "Occluded imaging with time-of-flight sensors," ACM T. Graph. **35**, 15 (2016).
18. M. Buttafava, J. Zeman, A. Tosi, K. Eliceiri, and A. Velten, "Non-line-of-sight imaging using a time-gated single photon avalanche diode," Opt. Exp. **23**, 20997–21011 (2015).
19. V. Arellano, D. Gutierrez, and A. Jarabo, "Fast back-projection for non-line of sight reconstruction," Opt. Exp. **25**, 11574–11583 (2017).
20. M. O'Toole, D. B. Lindell, and G. Wetzstein, "Confocal non-line-of-sight imaging based on the light-cone transform," Nature **555**, 338 (2018).
21. G. Huixuan, *Applied multivariate statistical analysis* (Peking University, 2005).
22. Q. Wang, G. M. Garrity, J. M. Tiedje, and J. R. Cole, "Naive bayesian classifier for rapid assignment of rrna sequences into the new bacterial taxonomy," Appl. Environ. Microbiol. **73**, 5261–5267 (2007).
23. A. Smith, J. Skorupski, and J. Davis, "Transient rendering," Tech. rep. (2008).
24. A. Jarabo, J. Marco, A. Muñoz, R. Buisan, W. Jarosz, and D. Gutierrez, "A framework for transient rendering," ACM T. Graph. **33**, 177 (2014).
25. A. Jarabo, B. Masia, J. Marco, and D. Gutierrez, "Recent advances in transient imaging: A computer graphics and vision perspective," Vis. Informatics **1**, 65–79 (2017).
26. M. O'Toole, F. Heide, D. B. Lindell, K. Zang, S. Diamond, and G. Wetzstein, "Reconstructing transient images from single-photon sensors," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition,* vol. 1 (2017).