CHINESE OPTICS LETTERS

Vol. xx, No. x | March 2024

High-resolution single-photon LiDAR without range-ambiguity using hybrid-mode imaging

Xin-Wei Kong[1,2,3†], Wen-Long Ye[1,2,4†], Wenwen Li[1,2,4], Zheng-Ping Li[1,2,4*], and Feihu Xu[1,2,4*]

[1] Hefei National Research Center for Physical Sciences at the Microscale and School of Physical Sciences, University of Science and Technology of China, Hefei 230026, China

[2] Shanghai Research Center for Quantum Science and CAS Center for Excellence in Quantum Information and Quantum Physics, University of Science and Technology of China, Shanghai 201315, China

[3] School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai 264209, China

[4] Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China

*Corresponding author: lizhp@ustc.edu.cn; feihuxu@ustc.edu.cn

Received Month X, XXXX | Accepted Month X, XXXX | Posted Online Month X, XXXX

We proposed a hybrid imaging scheme to estimate a high-resolution absolute depth map from low photon counts. It leverages measurements of photon arrival times from a single-photon LiDAR and an intensity image from a conventional high-resolution camera. Using a tailored fusion algorithm, we jointly processed the raw measurements from both sensors and output a high-resolution absolute depth map. We scaled up the resolution by a factor of 10, achieving 1300 × 2611 pixels and extending ∼4.7 times unambiguous range. These results demonstrated the superior capability of long-range high-resolution 3D imaging without range ambiguity.

Keywords: single-photon LiDAR, single-photon imaging, hybrid-mode imaging.
DOI: xxxxxxx/COLxxxxxxx.

1. Introduction

Single-photon light detection and ranging (LiDAR) presents high sensitivity and high temporal precision, which has been widely applied in fields such as topographic mapping[1–3], remote sensing[4], target identification[5,6] and underwater imaging[7]. To meet the application demands, long-range and high-resolution single-photon three dimensional (3D) imaging has emerged as a significant trend in the development of the single-photon LiDAR techniques[8,9]. However, it remains challenging to directly achieve rapid and accurate 3D imaging over a wide field-of-view (FoV) and a large depth-of-view (DoV).

Array-based single-photon LiDAR can be used to achieve high-resolution 3D imaging[10]. However, it needs a high-power laser to flood illuminate the scene. Besides, currently available detector arrays have limited size or show a poor time-tagging performance[11]. Therefore, the widely used single-photon LiDAR is typically based on raster scanning[12,13]. But, high-density scanning inevitably leads to a longer imaging time. To mitigate this issue, data fusion techniques have been proposed to merge visible or infrared high-resolution images with single-photon LiDAR data to improve imaging resolution[14–16].

Generally, single-photon LiDAR employs a time-correlated single-photon counting (TCSPC) technique. However, when the target is far away, the photon time of flight (ToF) that extend laser emission periods will be folded, resulting in range ambiguity[17] which leads to difficulties in imaging large DoV. Several approaches have been proposed to mitigate the range ambiguity. Pseudo-random pattern matching scheme[18–21] can identify the exact flight time by correlation between the transmitted and received patterns. Meanwhile, the multi-repetition-rates scheme has also been demonstrated to increase the maximum unambiguous distance beyond 100 kilometers[22] and achieve large DoV imaging[23]. Nonetheless, there is still lacking a comprehensive solution to achieve wide FoV and large DoV simultaneously.

Here, we proposed and demonstrated a fusion method that simultaneously tackled the range-ambiguity and low-resolution bottleneck of single-photon LiDAR. We integrated a multi-repetition-rates single-photon LiDAR and a high-resolution intensity camera on hardware. On software, we developed a tailored fusion algorithm for recovering absolute distance and enhancing the image resolution in the scenarios of low photon counts. We experimentally validated the ability to reconstruct high-resolution absolute depth images. We scaled up the image resolution by a factor of 10 by achieving 1300 × 2611 pixels, and extended ∼4.7 times unambiguous range. Consequently, our method holistically achieved long-range, high-resolution

3D imaging of expansive scenes with high depth accuracy over wide FoV and large DoV.

2. Approach

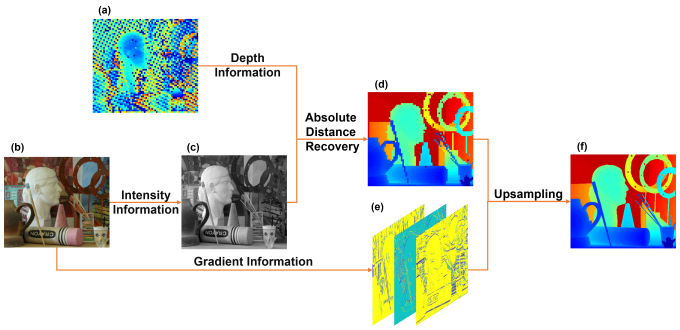

Fig. 1. Schematic diagram of the algorithm. (a) Single-photon LiDAR data acquired by laser source with multiple repetition rates. (b) Image captured by camera. (c) Intensity image of (b). (d) Absolute distance depth image. (e) Horizontal, vertical, and diagonal gradient images from the camera image. (f) High-resolution depth image without range-ambiguity.

In single-photon imaging, the system illuminates the target's p-th pixel with a periodic laser pulse $s(t)$ and then measures the photons scattered back. By recording the time interval t between the arrival of the echo signal and the most recent pulse emission, the depth Z_p and reflectivity α_p of the target's p-th pixel can be estimated. However, when the target is far away, the photon ToF that extend laser emission periods T will be folded, resulting in a Poisson-process rate function as follows:

$$\lambda_p(t) = \eta\alpha_p \sum_{n_p} s(t + n_p T - 2Z_p/c) + B, \quad t \in [0, T), \tag{1}$$

where η is detector's photon-detection efficiency, B represents the average rates of background-light plus dark-count detections, and c is the speed of light. The parameter $\{n_p T\}$ represents the photon ToF being folded.

After N pulsed-illumination trials, the likelihood function for the set of time interval $\{t_p^l\}_{l=1}^{k_p}$ is:

$$P\left(\{t_p^l\}_{l=1}^{k_p}; Z_p, \alpha_p\right) = e^{-\Lambda} \prod_{l=1}^{k_p} N\lambda_p\left(t_p^l\right), \tag{2}$$

where $\Lambda = \int_{\tau=0}^{\tau=T} N\lambda_p(\tau)d\tau$, and k_p is the total number of photons detected at the p-th pixel. Generally, the target distance can be estimated by applying maximum likelihood estimation (MLE):

$$Z_p^{MLE} = \arg\max_{Z_p} \sum_{l=1}^{k_p} log\left(N\left[\eta\alpha_p \sum_{n_p} s\left(t_p^l + n_p T - 2Z_p/c\right) + B\right]\right). \tag{3}$$

Because the maximum likelihood estimator is a periodic function of Z_p, Eq.3 has multiple optimal solutions, which prevents a straightforward calculation of the actual distance to the target and occurs range aliasing.

To overcome this range ambiguity, we use a data acquisition scheme where adjacent pixels are detected through different laser pulse repetition periods, and a data fusion method exploiting images captured by camera. The data acquisition scheme has been extensively detailed in previous paper[23]. Here, we focus on the use of high-resolution images for absolute distance reconstruction and upsampling of single-photon LiDAR data. The schematic of the algorithm is illustrated in Fig.1, and the algorithm can be divided into two steps.

2.1. Resolving range ambiguity guided by intensity image

Upon acquiring the measurements via the multi-repetition-rates scheme, the integration of data from adjacent pixels within the neighborhood Ω through cluster algorithms[20] enables the determination of the absolute distance:

$$\hat{Z}_p = \arg\max_{Z_p} \sum_{q\in\Omega} \omega_{q,p} \sum_{l=1}^{k_q} log\left(N\left[\eta\alpha_p \sum_{n_p} s\left(t_p^l + n_p \cdot T - 2Z_q/c\right) + B\right]\right), \tag{4}$$

where the weighting factor $\omega_{p,q}$ is used to avoid errors in distance calculation at the edges of objects. Similar with previous paper[23], we leverage the spatial and reflectivity information to evaluate the weighting factor $\omega_{p,q}$ for neighboring pixels. However, due to the reflectivity map of single-photon LiDAR is susceptible to Poisson noise at low photon counts, we use the conventional high-resolution camera images to evaluate reflectivity information of the single-photon LiDAR pixels. Due to the pixel number discrepancy between the conventional camera and single-photon LiDAR, the reflectivity value of the single-photon LiDAR is the weighted average of several conventional camera pixels. A many-to-one pixel mapping scenario arises:

$$I_p = \frac{4}{\sqrt{2\pi}D} \sum_{l=1}^{D} I_p^l e^{-\frac{8*(p-x_p^l)^2}{D^2}}, \tag{5}$$

where $\{x_p^l\}_{l=1}^{D}$ and $\{I_p^l\}_{l=1}^{D}$ correspond to the positions and intensities of the conventional camera images respectively. Therefore, the define of the weighting factor $\omega_{p,q}$ is $\omega_{p,q} = f(|p-q|) \cdot g(|I_p - I_q|)$. Here, f and g are the spatial and reflectivity kernels, respectively, both positively correlated with the Gaussian distribution.

Since the above process of solving $\hat{Z}_p$ requires to integrate the echo signals from the surrounding pixels, this often results in the image becoming overly smoothed, consequently reducing the imaging resolution and affecting the image quality. Here a convex optimization algorithm is employed to further enhance the accuracy of image reconstruction. The folded photon ToF $\{n_p T\}$ for the p-th pixel can be determined as $\hat{n}_p T = \left\lfloor \hat{Z}_p/2c \right\rfloor$. Then, taking advantage of spatial correlations in natural scenes, we select total variation (TV) as the penalization term. Thus, the absolute depth map is derived as follows:

$$Z^{MLE} = \arg\max_{Z} \sum_{p} \sum_{l=1}^{k_p} log\left(N\left[\eta\alpha_p s\left(t_p^l + \hat{n}_p T - 2Z_p/c\right) + B\right]\right) + \beta \cdot penalty(Z). \tag{6}$$

The above equation constitutes a convex optimization problem, and can be solved using a convex optimization algorithms[25] to obtain the final estimated distance value of the target.

2.2. Intensity-image guided upsampling

Furthermore, to improve the resolution of single-photon imaging, we can take advantage of the high resolution offered by the conventional camera images to guide the upsampling of single-photon images. In our framework, Z^H is designated as the high-resolution single-photon depth map we aim to obtain. Correspondingly, the already acquired absolute depth map Z^{MLE}, represents a downsampled mapping of Z^H, and this downsampling satisfies the following relation:

$$Z^{MLE} = f_d\left(Z^H\right) + Z_N, \tag{7}$$

where $f_d(\cdot)$ is the downsampling function that performs pixel-weighted summation using Gaussian weights, and Z_N represents the noise. Assuming the noise follows a Gaussian distribution, its likelihood function can be expressed as follows:

$$L = -\log\left[P\left(Z^H \middle| Z^{MLE}\right)\right] \propto \|Z^{MLE} - f_d\left(Z^H\right)\|_2^2. \tag{8}$$

Thus, by applying MLE, we can obtain the high-resolution single-photon image:

$$\hat{Z}^H = \arg\min_{Z^H}\{L + \beta \cdot penalty(Z^H)\}. \tag{9}$$

Here, we employs a second-order total generalized variation (TGV) regularization as the penalty term to constraint image , which is represented as:

$$penalty\left(Z^H\right) = \alpha_1 \|T^{1/2}(\nabla Z^H - v)\|_1 + \alpha_0 \|\nabla v\|_1, \tag{10}$$

where $T^{1/2}$ is the anisotropic diffusion tensor, v is an auxiliary variable, and the scalars α_1 and α_0 are non-negative weight coefficients. The TGV allows for sharper edge preservation while suppressing noise. Since the problem is convex but non smooth due to TGV regularization term, a primal-dual optimization algorithm is used for solving [14].

3. Simulations

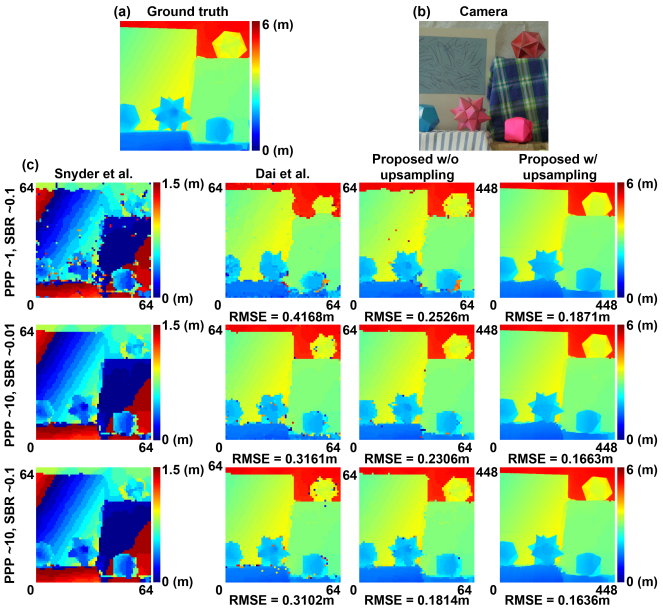

Fig. 2. Simulation results. (a) Ground truth. (b) High-resolution camera image. (c) The simulation results by different methods under various PPP and SBR. From left to right, each column corresponds to PPP $\sim$ 1 with SBR $\sim$ 0.1, PPP $\sim$ 10 with SBR $\sim$ 0.01, and PPP $\sim$ 10 with SBR $\sim$ 0.1, respectively. From top to bottom, each row shows the results reconstructed by Snyder et al., Dai et al., proposed without and with upsampling, respectively.

We conducted simulation experiments using the Middlebury 2007 dataset [26] to validate the effectiveness of our proposed method in reconstructing high-resolution absolute distance images. The resolution of single-photon imaging is set to 64×64 pixels. Considering the depth span of only 6 meters in the simulation scenario, we conducted a downscaled simulation of the imaging system's laser period by a factor of 100. We selected laser periods as 10 ns, 14.3 ns, 15.9 ns, 16.1 ns, and 17.1 ns for the simulation, of which the single period maximum unambiguous range is 2.565 meters. As shown in Fig. 2, we reconstructed the depth map using our method and compared the results with two state-of-the-art methods.

Fig.2(c) demonstrates that conventional algorithm[24] struggle to accurate estimate show the front-to-back position of a target because of range ambiguity. Dai et al.[23] can achieve absolute distance recovery, however, this method leads to the presence of noise in the depth maps. Our proposed method reconstructs absolute distance by combining conventional camera images with single-photon LiDAR, reducing the impact of Poisson noise and thereby achieving higher reconstruction accuracy. Compared with Dai's method, it shows a lower RMSE, which demonstrates superior absolute distance reconstruction capabilities even with low photon counts and low signal-to-background ratio (SBR). Beside, We have used conventional camera images for upsampling, which can enrich target details and remarkably improve image resolution. Compared to the results before upsampling, it has a lower RMSE.

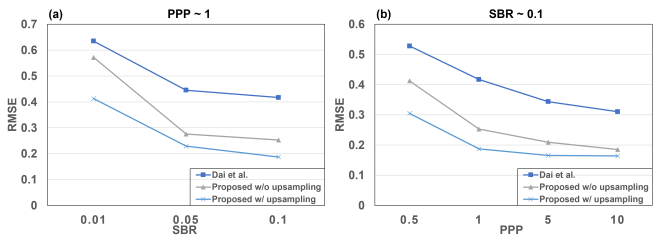

Fig. 3. The RMSE in simulations with different PPP and SBR levels. (a) For PPP $\sim$ 1 with SBR $\sim$ 0.01, 0.05 and 0.1, the RMSE results are calculated by Dai et al., proposed with and without upsampling. (b) For SBR $\sim$ 0.1 with PPP $\sim$ 0.5, 1, 5 and 10, the RMSE results are calculated by Dai et al., proposed without and with upsampling.

By comparing our method and Dai et al.'s method on root mean square error (RMSE) under same conditions, we find that reconstructions relying purely on LiDAR data, especially in low PPP and low SBR scenarios, tend to have some noisy pixels. By using the upsampling guidance, our algorithm performs well. As shown in Fig.3, our method outperforms Dai's method in terms of RMSE. The trend of our results initially decreases and then stabilizes as SBR/PPP increases, demonstrating that our results achieve the best accuracy.

4. Experiment

4.1. Experimental setup

The schematic of our long-range, high-resolution single-photon imaging system is shown in the Fig.4. We use a digital full-frame camera with a pixel resolution set to 7008×4672. The

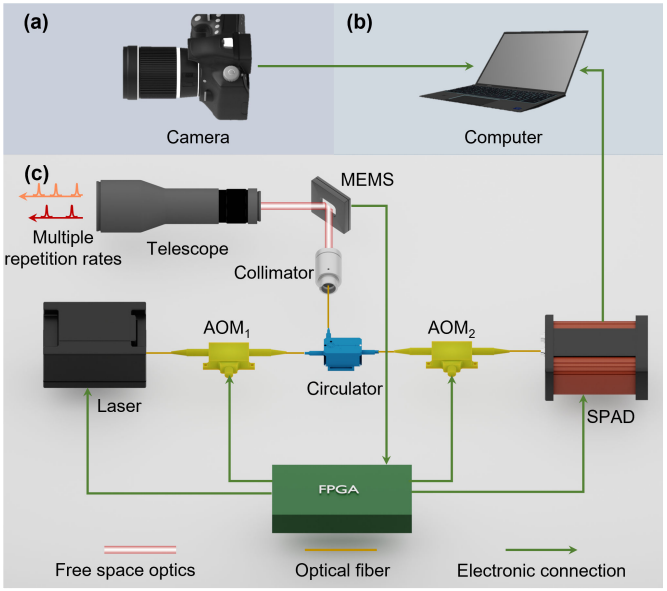

Fig. 4. The layout of the system. (a) Conventional high-resolution camera. (b) Single-photon LiDAR. (c) Data processing system.

focal length of the objective lens of the camera is 400 mm. A raster scanning single-photon LiDAR using laser source with multiple repetition rates provides raw depth data. The scanning interval is set to be 100 μrad. The single-photon LiDAR uses a coaxial design, allowing for highly efficient detection over a wide detection distances. To eliminate the local noise in this coaxial system, we set a temporal separation of laser emission and detection and employ two acousto-optic modulators (AOMs) for noise suppression. The system employs a 1550 nm fiber pulsed laser, and the period is adjustable through an external trigger, which is typically set between 1 and 2 μs. The maximum emission laser power of the system is 250 mW. The system includes a home-made InGaAs/InP single-photon avalanche diode (SPAD) detector with a detection efficiency of 30% and a dark count rate of 1.2 kcps. The system uses a home-made field programmable gate array (FPGA) board for precise timing control. Moreover, we use the pixel signals output from the micro-electromechanical system (MEMS) mirror to discern different pixel information and implement a scanning method where each pixel is illuminated by a specific frequency, with different frequencies employed for adjacent pixels.

4.2. Experimental result

As shown in Fig.5(a), we imaged residential buildings located 0.4 to 1.6 kilometers away. The experiment was conducted under five different laser pulse periods (1 μs, 1.43 μs, 1.59 μs, 1.61 μs, 1.71 μs), with a per-pixel acquisition time of 330 μs. We collected a single-photon image of 128×250 pixels, and the average PPP was ∼4.07. Guided by intensity information from the camera, we obtained absolute depth estimation shown in Fig.5(d). Furthermore, using the extracted contour information of the same image, we successfully generated a depth map

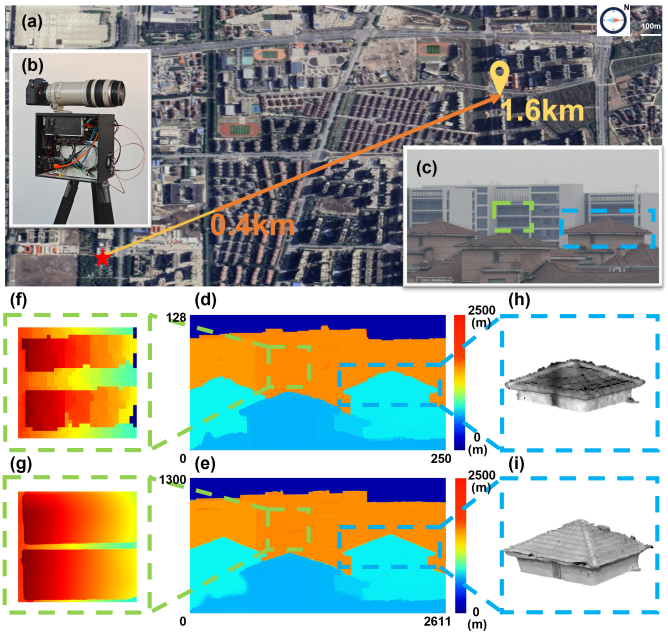

Fig. 5. The experimental results. (a) The target's location on the map. (b) Photograph of our system. (c) High-resolution camera image of target. (d), (e) The results using our proposed method without and with upsampling. (f), (g) Closeup views of the building details in depth reconstructions (area highlighted by green rectangle in (c)). (h), (i) 3D profiles of the eaves details in depth reconstructions (highlighted by blue rectangle in (c)).

with tenfold higher resolution (1300×2611) while maintaining high depth accuracy as illustrated in Fig.5(e). By comparing Fig.5(f) and Fig.5(g), our method displays better detail of the building after upsampling. The comparison between Fig.5(h) and Fig.5(i) shows an superiority on capturing detailed 3D surfaces in complex urban environments. These results prove the robustness and accuracy of our method in practical applications.

5. Conclusion

We proposed and validated a fusion long-range 3D imaging method to overcome the challenge of range ambiguity and low-resolution issue. The outdoor experimental results extended ∼4.7 times unambiguous range and imaged with over 3 megapixels (1300 × 2611), a 10-fold increase in resolution. By providing accurate depth perception and fine spatial awareness, the results may offer enhanced methods for rapid, high resolution, long-range 3D imaging for large-scale scenes. These are essential for target identification and environmental mapping in complex areas.

Acknowledgement

The authors acknowledge insightful discussions with Chen Dai, Wei Li, Xin Huang and Pengyu Jiang. This work was supported by the Key-Area Research and Development Program of Guangdong Province (2020B0303020001), the

National Natural Science Foundation of China (62031024, 12104443), the Innovation Program for Quantum Science and Technology (2021ZD0300300), the Shanghai Municipal Science and Technology Major Project (2019SHZDZX01), the Shanghai Science and Technology Development Foundation (22JC1402900), the Shanghai Academic/Technology Research Leader (21XD1403800), the Shanghai Sailing Program (21YF1452600), the Natural Science Foundation of Shanghai (21ZR1470000).

†These authors contributed equally to this work.

References

1. J. J. Degnan, "Photon-counting multikilohertz microlaser altimeters for airborne and spaceborne topographic measurements" , Journal of Geodynamics **34**, 503 (2002).
2. R. M. Marino and W. R. Davis, "Jigsaw: a foliage-penetrating 3D imaging laser radar system" , Lincoln Laboratory Journal **15**, 23 (2005).
3. T. Markus, T. Neumann, A. Martino, W. Abdalati, K. Brunt, B. Csatho, S. Farrell, H. Fricker, A. Gardner, and D. Harding, "The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): science requirements, concept, and implementation" , Remote Sens. Environ. **190**, 260 (2017).
4. C. L. Glennie, W. E. Carter, R. L. Shrestha, and W. E. Dietrich, "Geodetic imaging with airborne LiDAR: the Earth's surface revealed" , Rep. Prog. Phys. **76**, 086801 (2013).
5. A. B. Gschwendtner and W. E. Keicher, "Development of coherent laser radar at Lincoln Laboratory" , Lincoln Laboratory Journal **12**, 383 (2000).
6. Y. Hong, Y. Li, C. Dai, J.-T. Ye, X. Huang and F. Xu, "Image-free target identification using a single-point single-photon LiDAR" Opt. Express **31**, 30390 (2023)
7. A. Maccarone, K. Drummond, A. McCarthy, U. K. Steinlehner, J. Tachella, D. A. Garcia, A. Pawlikowska, R. A. Lamb, R. K. Henderson, and S. McLaughlin, "Submerged single-photon LiDAR imaging sensor used for real-time 3D scene reconstruction in scattering underwater environments" , Opt. Express **31**, 16690 (2023).
8. M. Laurenzis, F. Christnacher, and D. Monnin, "Long-range three-dimensional active imaging with superresolution depth mapping" , Opt. Lett. **32**, 3146 (2007).
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. Express **21**, 8904 (2013).
10. J. Tachella, Y. Altmann, N. Mellado, A. McCarthy, R. Tobin, G. S. Buller, J.-Y. Tourneret, and S. McLaughlin, "Real-time 3D reconstruction from single-photon lidar data using plug-and-play point cloud denoisers" , Nat. Commun. **10**, 4984 (2019).
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. M. Pawlikowska, A. Halimi, R. A. Lamb, and G. S. Buller, "Single-photon three-dimensional imaging at up to 10 kilometers range", Opt. Express **25**, 11919 (2017).
13. Z.-P. Li, X. Huang, P.-Y. Jiang, Y. Hong, C. Yu, Y. Cao, J. Zhang, F. Xu, and J.-W. Pan, "Super-resolution single-photon imaging at 8.2 kilometers" , Opt. Express **28**, 4076 (2020).
14. D. Ferstl, C. Reinbacher, R. Ranftl, M. Rüther, and H. Bischof, "Image guided depth upsampling using anisotropic total generalized variation" , in Proceedings of the IEEE International Conference on Computer Vision, 993 (2013).
15. D. B. Lindell, M. O'Toole, and G. Wetzstein, "Single-photon 3D imaging with deep sensor fusion" , ACM Trans. Graph. **37**, 1 (2018).
16. I. Gyongy, S. W. Hutchings, A. Halimi, M. Tyler, S. Chan, F. Zhu, S. McLaughlin, R. K. Henderson, and J. Leach, "High-speed 3D sensing via hybrid-mode imaging and guided upsampling" , Optica **7**, 1253 (2020).
17. W. H. Long, D. H. Mooney, and W. A. Skillman, "Pulse doppler radar" , Radar Handbook **2** (1990).
18. Q. Zhang, H. W. Soon, H. Tian, S. Fernando, Y. Ha, and N. G. Chen, "Pseudo-random single photon counting for time-resolved optical measurement" , Optics Express **16**, 13233 (2008).
19. P. A. Hiskett, C. S. Parry, A. McCarthy, and G. S. Buller, "A photon-counting time-of-flight ranging technique developed for the avoidance of range ambiguity at gigahertz clock rates" , Optics Express **16**, 13685 (2008).
20. N. J. Krichel, A. McCarthy, and G. S. Buller, "Resolving range ambiguity in a photon counting depth imager operating at kilometer distances" , Opt. Express **18**, 9192 (2010).
21. Y. Liang, J. Huang, M. Ren, B. Feng, X. Chen, E. Wu, G. Wu, and H. Zeng, "1550-nm time-of-flight ranging system employing laser with multiple repetition rates for reducing the range ambiguity" , Opt. Express **22**, 4662 (2014).
22. Z.-P. Li, J.-T. Ye, X. Huang, P.-Y. Jiang, Y. Cao, Y. Hong, C. Yu, J. Zhang, Q. Zhang, and C.-Z. Peng, "Single-photon imaging over 200 km" , Optica **8**, 344 (2021).
23. C. Dai, W.-L. Ye, C. Yu, X. Huang, Z.-P. Li, and F. Xu, "Long-range photon-efficient 3D imaging without range ambiguity" , Opt. Lett. **48**, 1542 (2023).
24. D. L. Snyder and M. I. Miller, *Random point processes in time and space* (Springer Science & Business Media, 2012).
25. 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 (2011).
26. D. Scharstein and C. Pal, "Learning conditional random fields for stereo" , in 2007 IEEE Conference on Computer Vision and Pattern Recognition, (2007), 1.