An improved algorithm of NLOS imaging based on Bayesian statistics

LUZHE HUANG, XIAOBIN WANG, YIFAN YUAN, 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 objects obscured by scattering is an important goal in imaging and has been approached by exploiting, for example, coherence properties, ballistic photons or penetrating wavelengths. Here, we propose a robust reconstruction algorithm based on the Bayesian statistics using the information of each sampling point on the wall, including both the illuminance and the flight time. Compared with traditional backprojection methods, this algorithm handles random errors involved in data well, and images occluded objects with high resolution. Adjustable compensation mechanism in our method is effective in handling the reflection diversity and the illumination attenuation caused by different objects, which will improve the quality of the image. Our method proves to be efficient with both simulated data and experiment data.

1. Introduction

Imaging objects through occlusion has been a problem of fundamental significance for a long time. For example, image through scattering medium by optical coherence tomography (OCT) [1], look through turbid medium and around corner via speckle auto-correlations [2, 3] and observe by two-photon "ghost" imaging [4]. Non-line-of-sight (NLOS) imaging technology is to "see" a scene out of sight, for example around a corner or hidden by some shelter. NLOS imaging reconstructs the shape and albedo of hidden objects from multiple scattered light. NLOS imaging has been demonstrated using both radio and visible wavelengths. At radio wavelengths, systems have been developed to create low resolution images through walls, around corners using specular reflections, and to detect motion around a corner. These systems typically require large apertures, especially when the imaging system is far from the scene to be imaged. Methods of implementing NLOS imaging at visible wavelengths include using a coded controllable light source, such as a projector, to illuminate hidden objects or using specular reflections in a window pane. Moreover, photon time-of-flight, which is typically used for ranging in imaging LIDAR or gated viewing systems, can also be applied to multiply reflected light to image beyond the direct line of sight.

2. Related Works

NLOS has been widely studied in various situations. Many methods of locating and tracking objects in NLOS scene have been proposed. In 1998, H. Koorapaty et al. applied radio signals geolocation to non-line-of-sight scenes by mitigating NLOS errors in location process [5]. J. Klein et al. proposed a novel approach in 2016 to track NLOS object in real time through 2D intensity images and simple experiment setup [6]. In 2015, by utilizing a single photon avalanche diode (SPAD), Gariepy et al. made it possible to detect and track moving objects in real time in NLOS situation [7]. Later in 2017, S. Chan et al. demonstrated a long-distance tracking system [8]. Furthermore, benefiting from high-temporal-resolution detectors, 3D reconstructions can be done by scanning a pulsed laser and using corresponding algorithms, which is common in

light detection and ranging (LIDAR) systems to recover 3D information of objects through direct reflections [9–12]. In 2009, mirror reflective photons were used to image around corners using infrared light based on time-gated detectors [13]. Using ultrafast pulse laser, MIT media lab proposed back-projection algorithm and recovered three-dimensional information of occluded objects based on time-of-flight imaging in 2012 [14, 15]. Their results demonstrated good spatial resolution at cm level. In 2015, the application of single-photon avalanche diode made it possible to reconstruct objects hidden from view [16]. V. Arellano et al. improved the conventional back-projection algorithm and performed fast back-projection reconstruction on GPU by computing the intersection of space-time manifolds in 2017 [17]. 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 [18]. Though it requires specially designed experiment setups, their method exhibits almost real time reconstruction with excellent resolution, indicating its potential applications in many realms such as auto-driving.
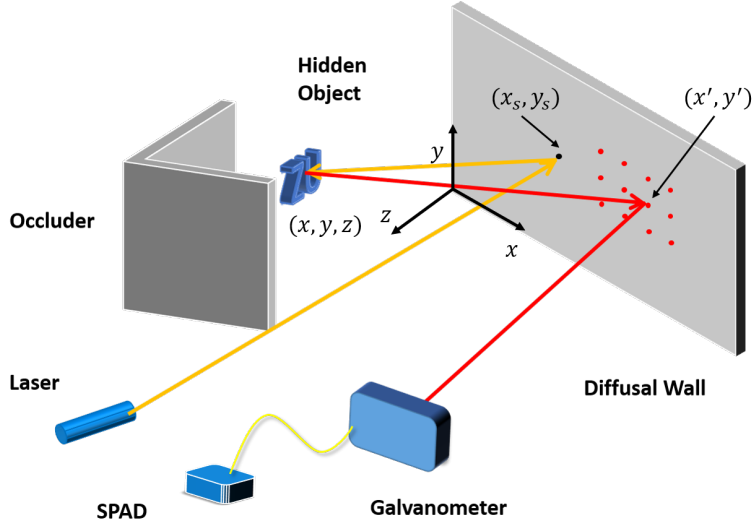

Fig. 1. **The schematic diagram of experiment setups of a B-scan process.**
(x, y, z) are coordinates of the object's voxel, (x_s, y_s) are coordinates of beam spot shone on the wall, and similarly, (x', y') are coordinates of scanning point on the wall. (x', y') varies with A-scans, and (x_s, y_s) changes among different C-frames.

3. Probability-based Model and Reconstruction

Conventional NLOS methods such as back-projection (BP) require assumptions about the light path as well as the reflective characteristics of objects and wall. These assumptions are generally held in experiments and real scenes. Similarly, our probability-based method makes the following assumptions: no inter-object reflections in the object space, no object occluded by other hidden objects and with normally distributed errors. Moreover, we generally suppose the wall and the objects have typical Lambertian surface, which have isotropic luminance, to simplify calculations.

3.1. Bayesian formula and posterior probability

Bayesian statistics are widely used in parameter inference, and conventional machine learning methods [19]. For example, naive Bayesian classifier demonstrates high accuracy in bacterial rRNA sequences assignments [20]. 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.

Firstly, as in the experiment setup illustrated in Fig. 1, we denote the prior probability of each voxel $v_i = (x, y, z)$ as $p(x, y, z)$, which are usually set equal in that we lack prior inference about hidden objects. Secondly, we calculate the probability distribution of a photon emitted from v_i and detected at (x', y'), and we use $w(x', y'|v_i)$ to represent that event. The corresponding probability $p(w(x', y'|v_i))$ is proportional to the illuminance distribution and can be easily expressed given various assumption of reflection characteristics of the objects and wall, e.g. the same albedo and Lambertian reflection for each voxel. In that case, $p(w(x', y'|v_i))$ can be simplified as:

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

where $w = (x_s, y_s)$ denotes the laser spot coordinates on the wall, which serves as a point light source in the following process. Moreover, we use $p_{lambert}(x', y'|v_i)$ to denote the probability distribution in the case of Lambertian reflection.

Nevertheless, because of quantum effect, photon detection is a Poisson process which results in random errors of time of flight (ToF) compared with ideal value. Let alone the temporal resolution limit of detector, e.g. single photon avalanche diode (SPAD). These errors prevent object reconstruction with expected accuracy and lead to significant blur of reconstructed image. The accuracy and blur issue are two major drawbacks with respect to ordinary backprojection algorithm. In constrast, the model proposed here tackles with errors well. Here we introduce a normal distribution to depict the random error. Such assumption allows enough flexibility and meanwhile is amenable in both analysis and interpretation. In that all photons are independent with each other. The probability of ToF error $p(h(t, n)|h(t_{ideal}, n))$ is given by

$$p(h(t, n)|h(t_{ideal}, n)) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{n}{2\sigma^2}(t - t_{ideal})^2} \tag{2}$$

where t_{ideal} denotes the ideal flight time of a photon from w to v_i and then to (x', y'), n is the number of photons received.

Therefore, the posterior probability of voxel v_i is expressed by:

$$\pi(v_i|h(t, n; x', y')) = C * (e^{-\frac{1}{2\sigma^2}(t - t_{ideal})^2} \cdot p_{lambert}(x', y'|v_i))^n \tag{3}$$

where C is a constant making probability standardized, and is often neglected in the following formula and algorithm.

3.2. Log-probability loss function

It is obvious the formula above models the projection process of a single voxel well and is likely to give the true point a maximized posterior probability. However, what should be emphasized is, generally, more than one signal are received at a single pixel on the wall. Moreover, it is another complex problem to judge which signal belongs to which voxel. Thus, to solve this problem and make our method work, a log-probability loss function is introduced to measure the probability of a certain voxel.

Loss function of each voxel is given by

$$Loss(v_i) = \frac{1}{n_{x'}n_{y'}n_t} \sum_{(x', y', t)} nt^4 * max(-\frac{(t - t_{ideal})^2}{2\sigma^2} + log(p_{lambert})), L) \tag{4}$$

where σ is a coefficient controlling the variance of the error, and L a constant restricting the range of loss function.

Here we use a trick to replace σ^2 in Eq. (1) with λ before $log(p_{lambert})$. Thus, Eq. (1) can be rewritten as

$$Loss(v_i) = \frac{1}{n_{x'}n_{y'}n_t} \sum_{(x',y',t)} nt^4 * max(-(t - t_{ideal})^2 + \lambda log(p_{lambert})), L) \tag{5}$$

The constraint L is essential because it makes loss function approximately convex when it maintains the quadratic form. To make it easier to understand, one can imagine two signals received at the same pixel. A proper L results in two separated ideal signals otherwise only a illusive signal in the middle position of the two is retrieved without L. And a t^4 term is introduced here to compensate the quadratic attenuation of light propagation in free space.
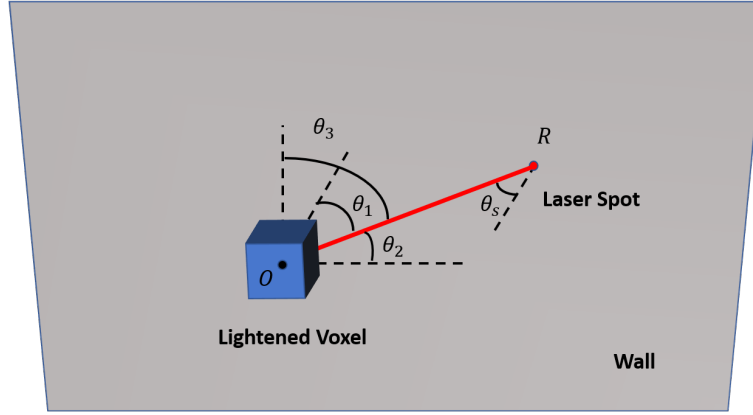

Fig. 2. **The schematic diagram of simulation.** $\theta_s, \theta_{o1}, \theta_{o2}$ and θ_{o3} are defined as shown above. $\theta'_s, \theta'_{o1}, \theta'_{o2}$ and θ'_{o3} are defined similarly with respect to the voxel and another pixel (j,k)

4. Simulation Process

Assume the size of the voxel, the basic component unit of the object, is $v * v * v(mm^3)$ and the size of the area pixel, the basic component unit of the wall, $p * p(mm^2)$. Also assume the object and the wall are ideal Lambertian and the reflection coefficient of the object and the wall a constant. If the luminous flux of the laser spot on the wall is $M(lm)$, then we can get its luminance as

$$L_w = \frac{M}{\pi} \tag{6}$$

In the following we take a voxel as an example to simulate the whole process, which can be segmented as three sub-process. Firstly calculate the illuminance from the wall to the voxel, and then the illuminance from the voxel to the wall. Secondly take influence of spatial-temporal resolution of the sampling process into consideration, on both the illuminance and the flight time of photons. Finally we will discretize the results above and append Poisson noise due to the quantum properties of photons.

Suppose each voxel has three normalized surfaces lightened by the point laser source on the wall. According to the relation between illumination and luminance, the illuminance of these three surfaces is

$$E_{o1} = L_w \frac{\cos \theta_s \cos \theta_{o1}}{|\overrightarrow{R_{so}}|^2}, E_{o2} = L_w \frac{\cos \theta_s \cos \theta_{o2}}{|\overrightarrow{R_{so}}|^2}, E_{o3} = L_w \frac{\cos \theta_s \cos \theta_{o3}}{|\overrightarrow{R_{so}}|^2} \tag{7}$$

Here, $\overrightarrow{R_{so}}$ is the vector from the laser spot on the wall to the center of the voxel, θ_s is the angle between the normal vector of the wall and $\overrightarrow{R_{so}}$. θ_{o1}, θ_{o2}, and θ_{o3} are angles between normal vectors of the lightened voxel surfaces and $-\overrightarrow{R_{so}}$.

Then regard the three lightened surfaces of the voxel as a secondary light source, we can similarly get the luminance

$$L_{o1} = \rho_O \frac{E_{o1}}{\pi}, L_{o2} = \rho_O \frac{E_{o2}}{\pi}, L_{o3} = \rho_O \frac{E_{o3}}{\pi} \tag{8}$$

Now the illuminance from the voxel to an area pixel (j, k) of the wall can be calculated.

$$E_w(j,k) = \frac{(L_{o1} \cos \theta'_{o1} + L_{o2} \cos \theta'_{o2} + L_{o3} \cos \theta'_{o3}) \cos \theta'_s}{|\overrightarrow{R_{ow}(j,k)}|^2} \tag{9}$$

where ρ_O is the reflection coefficient of the object, $\overrightarrow{R_{ow}(j,k)}$ the vector from the center of the voxel to the unit area pixel(j, k) on the wall. θ'_s is the angle between the normal vector of the wall and $-\overrightarrow{R_{ow}}$, and $\theta'_{o1}, \theta'_{o2}, \theta'_{o3}$ are defined as the angle between the normal vector of the voxel and $\overrightarrow{R_{ow}}$ respectively. The range of j, k depends on the size of the scanning area.

Consider the fixed relative position between the wall and the detector and assume the reflection coefficient of the wall a constant, $E_w(j,k)$can be regarded as the illuminance effect of a single voxel on the detector after ignoring a constant coefficient. In addition, ignoring a constant flight time from the wall to the detector, the flight time of photons can also be seen as

$$t(j,k) = \frac{|\overrightarrow{R_{so}}| + |\overrightarrow{R_{ow}(j,k)}|}{c} \tag{10}$$

Furthermore we discretize the result, $E_w(j,k)$ and $t(j,k)$, according to the energy of a photon and the minimum resolution time of the detector. Finally, we obtain a $N'_x \times N'_y \times N_t$ spatial-temporal signal matrix, where N'_x, N'_y, N_t denote the number of sample points along each dimension. A Poisson noise is added to the matrix to simulate the quantum properties of the photons.
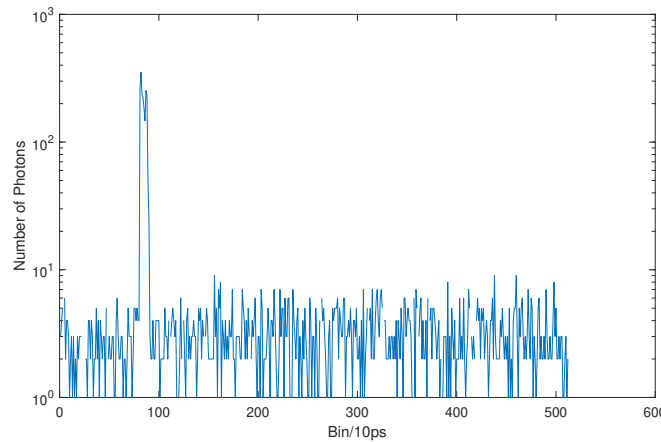

Fig. 3. **Simulated signal with Poisson noise.** The mean value of Poisson noise is about 3.

5. Reconstruction Result

The whole measuring process is as follows: 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 A-detect. Then, the galvanometer scans different pixels along a 2-D array, which is denoted by B-scan. Moreover, unlike conventional NLOS methods using backprojection algorithm, the laser spot is changed for several times after a B-scan, and new scans are repeated. We call it C-frame. This operation eliminates influences of noise and improves imaging quality significantly.

It is worth noting that even though one cannot acquire the illuminance on each pixel in direct methods but by using a detector with fixed view angle, the received photon numbers do not need angle and distance correction because a larger area on the wall is imaged with respect to further pixels.
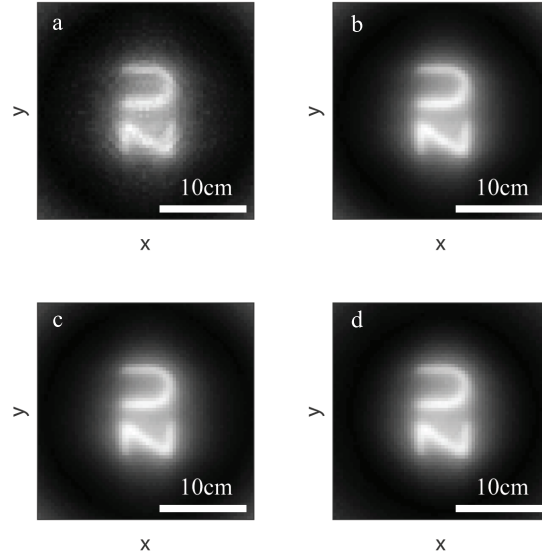

Fig. 4. **Reconstructed object of diffuse ZU letters.**
For (a), $\lambda = 0$ and $L = 1$ to approximate conventional BP algorithm. (b) $\lambda = 0$ and $L = 3$.
(c) $\lambda = 0$ and $L = 5$. (d) $\lambda = 1$ and $L = 5$. Proper L provides better reconstruction quality.

5.1. Reconstruction based on simulation data

The data used in this section is simulated according to processes above. Here we select 'Z' and 'U' letters as the objects, and control the parameter of Poisson noise at a reasonal range to make the SNR about 30-50. Each pixel is 8*8 mm^2, and voxel is 4*4*4 mm^3. A total of 512 bins are captured in each A-detect with a 10 ps temporal resolution. A typical received signal is simulated, as shown in Fig.3. The data is captured on 32*32 sample points in B-scan process. To improve the image quality, we average the reconstruction results of four C-frames.

The result exhibits excellent reconstruction performance. One can easily recognize the letters, 'Z' and 'U', and further enhance the contrast through edge detection or image growth algorithm.

In addition, we test the algorithm with different parameters. We control the parameter L in Eq. (5) to mitigate artificial noises in the reconstruction image. A proper L exhibits excellent reconstruction images with sharp edges as well as continuous inner regions. Fig.4 presents the effect of L. When L approaches 0, Eq. (2) becomes the conventional backprojection (BP) formula with Lambertian prior. Apparently conventional BP results in a worse result under the same voxel number and size. In general, conventional BP algorithm requires smaller and far more voxels, which accounts for its low efficiency. Besides, conventional BP cannot handle random errors of the recorded time and the Lambertian characteristic of reflectors. Hence, our algorithm is more efficient and less costly in memory, which benefits its applications in real-time situation. In our test, it requires about 110s for a single C-frame to reconstruct Fig.4 when running MATLAB R2017b on a Surface Pro 4 (Intel 1.5GHz Dual Core). The running time is significantly shorter than conventional BP methods because our algorithm calculates fewer voxels without reducing reconstruction quality.

The result seems relatively insensitive to λ, while in fact λ reduces noise along z axis, which is shown in Fig.5.
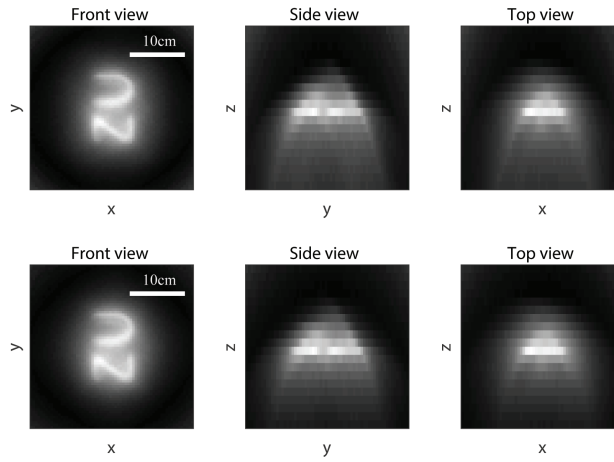

Fig. 5. **Reconstructed object of diffuse ZU letters with different λ.**
Upper row used $\lambda = 0$ and $\lambda = 1$ for lower one. $L = 5$ for both rows. λ controls image blurring along z axis. Three images for each row are three views of reconstructed object respectively.

5.2. Reconstruction based on confocal data

In this section, we conduct our algorithm on confocal data from a retroreflective 'SU' letter, which is provided by Matthew O'Toole et. al of Stanford University in Ref. [18]. The data is sampled on a 70cm*70cm area at 64*64 sample points with 4096 time bins for the temporal histograms. The temporal resolution of the data is 4ps, but we downsample it to 16ps, which means the downsampled signal size is $64 * 64 * 1024$. It is worthy to note that their experiment setups are based on a novel C-NLOS method. But with minor modification, our algorithm can be adapted on their data. Reconstructed objects are limited in a 64*64*64 space with each voxel size $10 * 10 * 10mm^2$. By modifying compensation item from t^4 to t^2, the result exhibits clearly two letters at different distances as shown in Fig.6.

The reconstruction time is relatively longer compared with the confocal method [18] in that their method apply FFT to reconstruction process, which improves complexity from $O(N^2)$ to
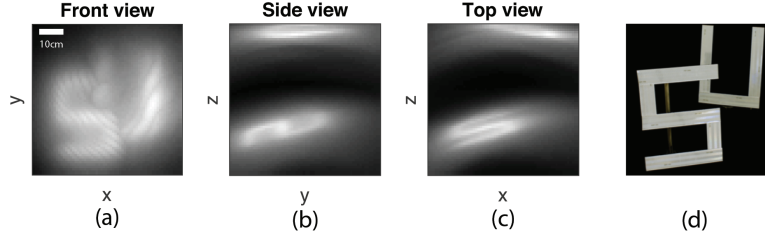
Fig. 6. **Reconstructed object of retroreflective SU letters.**
(a), (b), (c) are three views of the reconstructed object respectively. (d) is the shot image of real object. In this case, $\lambda = 2$ and $L = 8$

$O(NlogN)$. N is the total number of voxels. However, further adaption of our method to the confocal setups is promising to reduce time consumption to several seconds and this will be the subject of our future work.

6. Conclusion

We derived a probability-based model for solving NLOS problem to handle random error in data acquisition and variable reflective characteristics of reflectors, and proposed a new algorithm and the sampling method. We also validate its applicability on both simulated and experiment data. As for experiment data, we test our algorithm on the condition of retroreflective objects. The result demonstrates good image quality and higher efficiency in memory and time consumption than conventional BP methods. Besides, this algorithm can be applied to confocal NLOS setups.

NLOS imaging technology is very important for battlefield surveillance, search and rescue and can be applied into the fields of medical diagnosis or archaeological research. We hope that our work will spur more research into the potential applications.

Acknowledgment

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. Gregory, C.A. Puliafito, "Optical coherence tomography," Science **254**(5035), 1178-81 (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**(10), 784–790 (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**(8), 549–553 (2012).
4. R.S. Bennink, S.J. Bentley and R.W. Boyd, ""Two-photon" coincidence imaging with a classical source," Phys. review letters **89**(11), 113601 (2002).
5. H. Koorapaty, H. Grubeck, M. Cedervall, "Effect of biased measurement errors on accuracy of position location methods," Proc. IEEE Glob. Telecommun. Conf. **3**, 1497-1502 (1998).
6. J. Klein, C. Peters, J. Martín, M. Laurenzis, M.B. Hullin, "Tracking objects outside the line of sight using 2D intensity images," Sci. reports **6**, 32491 (2016).
7. G. Gariepy, F. Tonolini, R. Henderson, J. Leach, D. Faccio, "Detection and tracking of moving objectshidden from view," Nat. Photonics **10**(1), 23-U32 (2015).

8. S. Chan, R.E. Warburton, G. Gariepy, J. Leach, D. Faccio, "Non-line-of-sight tracking of people at long range," Opt. Express **25**(9), 10109-17 (2017).

9. B. Schwarz, "LIDAR: Mapping the world in 3D," Nat. Photonics **4**(7), 429 (2010).

10. 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**(7), 8904-8915(2013)

11. A. Kirmani, D. Venkatraman, D. Shin, A. Colaço, F.N. Wong, J.H. Shapiro and V.K. Goyal, "First-photon imaging" Science 1246775(2013).

12. 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. communications 7, 12046(2016)

13. A. Sume, M. Gustafsson, A. Jänis, S. Nilsson, J. Rahm, A. Örbom, "Radar detection of moving objects around corners," Proc. of SPIE 7308, 73080V–73080V, 2009.

14. A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M.G. Bawendi, R. Raskar, "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," Nat. Commun. **3**(1), 745 (2012).

15. O. Gupta, T. Willwacher, A. Velten, A. Veeraraghavan, R. Raskar, "Reconstruction of hidden 3D shapes using diffuse reflections," Opt. Express **20**(17), 19096-19108 (2012).

16. M. Buttafaca, J. Zeman, A. Tosi, K. Eliceiri, A. Velten, "Non-line-of-sight imaging using a time-gated single photon avalanche diode," Opt. Express **23**(16), 20997-21011 (2015).

17. V. Arellano, D. Gutierrez and A. Jarabo, "Fast back-projection for non-line of sight reconstruction," Opt. Express **25**(10), 11574-11583 (2017).

18. M. O'Toole, D. Lindell and G. Wetzstein, "Confocal non-line-of-sight imaging based on the light-cone transform," Nature **555**(7696), 338-341 (2018).

19. H.X. Gao, "Applied Multivariate Statistical Analysis," (Peking University, 2005).

20. 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. environmental microbiology **73**(16), 5261-5267 (2007).