Non-line-of-sight 3D imaging with a single-pixel camera

G. Musarra, ^{1,*} A. Lyons, ¹ E. Conca, ² Y. Altmann, ³ F. Villa, ² F. Zappa, ² M. J. Padgett, ¹ and D. Faccio ^{1,†}

¹ School of Physics & Astronomy, University of Glasgow, Glasgow G12 8QQ, UK.

² Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milano, Italy

³ School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH144AS, UK.

(Dated: March 13, 2019)

Real time, high resolution 3D reconstruction of scenes hidden from the direct field of view is a challenging field of research with applications in real-life situations related e.g. to surveillance, self-driving cars and rescue missions. Most current techniques recover the 3D structure of a non-line-of-sight (NLOS) static scene by detecting the return signal from the hidden object on a scattering observation area. Here, we demonstrate the full colour retrieval of the 3D shape of a hidden scene by coupling back-projection imaging algorithms with the high-resolution time-of-flight information provided by a single-pixel camera. By using a high efficiency Single-Photon Avalanche Diode (SPAD) detector, this technique provides the advantage of imaging with no mechanical scanning parts, with acquisition times down to sub-seconds.

The identification of scenes hidden from the direct lineof-sight, as happens for objects hidden behind an occluder or wall, is a challenging imaging task with applications in defence, surveillance and self-driving vehicles [1]. Non-line-of-sight (NLOS) imaging has been demonstrated by using radar systems [2], wavefront shaping [3] and speckle correlation [1, 4, 5] and recently even with passive cameras capturing light originating from behind a wall using an ordinary digital camera [6]. Most approaches in this field have demonstrated how to identify the hidden scene by collecting the light scattered back by hidden objects with a system similar to light detection and ranging (LIDAR) by using the time-of-flight information of the back scattered signal [7–11]. This technique typically involves a pulsed laser beam pointed on a scattering surface, producing a spherical wave propagating into the hidden scene. When the spherical wave hits the hidden object, the light is then scattered back towards the scattering surface. Collecting the third bounce echo scattered from the hidden object allows the detection and identification of hidden scene by advanced 3D reconstruction algorithms [12]. Past results have demonstrated how this technique can be used for tracking a moving hidden object even over large distances [13, 14] and for the retrieval of the 3D shape of a static hidden object by using back-projection imaging algorithms [12, 15] or ellipsoid mode decomposition for multiple hidden objects [16]. Alternative methods aimed at simplifying or increasing the speed of LIDAR-like NLOS imaging rely on 2D continuous illumination [17], deep learning [18] and confocal illumination/collection [19]. The multiple-bounce backscattered signal is typically very weak so these techniques require high temporal resolution and high speed singlephoton cameras with high detection efficiency.

However, one of the main limitations of current NLOS imaging systems is the finite temporal resolution of the detectors, which in turn determines the spatial resolution of the retrieval and thus the ability to reconstruct satisfactorily the 3D structure of the hidden scene. Fur-

thermore the complete 3D retrieval of the hidden scene often requires prohibitive time resources with the 3D imaging of a moving object, although progress is being made to reduce the acquisition and reconstruction times [6, 19], with the goal of reaching second or even sub-second times. Additionally, the spatial resolution of the retrieval can be improved by using iterative backprojection algorithms that however, typically require longer computational times [20, 21].

Another rapidly evolving imaging technique is based on so-called single-pixel cameras [22]. Standard 2D single-pixel imaging systems recover images by projecting an array of patterns onto the scene and detecting only the total reflected or transmitted light intensity, for which a single pixel is sufficient. The computational image reconstruction can then be achieved by computing a weighted sum of all the illumination patterns, where the weight of each pattern is given by the corresponding measured intensity level. This technique can also be extended for use e.g. in microscopy [23], imaging through scattering media [24] and terahertz imaging [25] or full 3D LIDAR [26–28].

Although this technique requires many consecutive measurements, the consequent long acquisition time required for high spatial resolution images can be significantly reduced by applying compressive sensing [25, 29, 30]. More importantly, the technique provides more flexibility in choosing the optimal (single pixel) detector for the imaging challenge being addressed. Of relevance to this work, this implies that one can choose a single pixel detector with enhanced temporal response and build upon the typically better temporal resolutions that are available in single pixel format (timing resolutions down to picoseconds), when compared to camera/SPAD array technologies.

In this work we demonstrate full 3D retrieval of hidden scenes by using a time-resolved single-pixel camera. The choice of a single pixel camera approach allows us to use optimised (sub 30 ps impulse response function) single

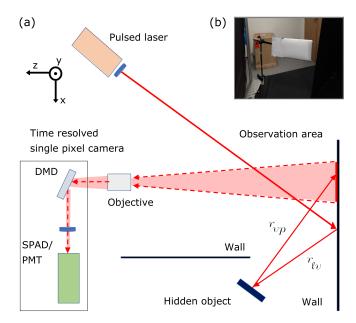


FIG. 1: Schematics of the experimental set-up for non line of sight imaging. a) A pulsed laser scatters on a scattering surface producing a spherical wave in all the surrounding area. The DMD is imaging a 50x50 cm² area of the scattering surface and the time resolved single-pixel camera collects the signal back-scattered from the target in the hidden scene by temporal histograms. We collect the signal scattered back by the hidden target by sequentially collecting light from each of the 20x20 pixels either by raster scan masks or by Hadamard masks. b) Picture of the non cooperative single object scene used for ultra-fast NLOS imaging.

photon detectors in combination with a digital mirror device (DMD) so as to remove the need for any scanning components whilst building upon the 20 kHz refresh rate of the DMD and high single photon sensitivity to reduce acquisition times with good reconstruction fidelity. By employing a white-light laser, we extend the technique to achieve the full RGB colour retrieval of non-cooperative, hidden objects and by choosing high efficiency single photon avalanche diodes (SPADs) we also achieve sub-second acquisition times.

Figure 1(a) shows the experimental set-up: the single-pixel camera is composed of a camera lens objective (8 mm focal length, f/3.5) that images a $50x50 \text{ cm}^2$ portion of a scattering wall onto a DMD (placed 1.16 m from the wall). The DMD then projects only selected portions (masks) of the image onto a single-pixel single-photon detector through a microscope objective. The single-photon detector therefore acts as a bucket detector and has 30 ps impulse response time and an area of $57x57 \mu\text{m}^2$. The DMD masks have 20x20 pixels corresponding to $2.6x2.6 \text{ cm}^2$ pixel areas on the scattering wall. The single-photon data is recorded in time-correlated-single-photon-counting (TCSPC) mode triggered by the illumination laser, as a histogram of photon arrival times with 4096 time bins of 6.1 ps duration each. The laser is di-

rected on the scattering wall 10 cm to the right of the field of view of the single-pixel camera, producing a first scattered spherical wave. Part of the spherical wave hits the hidden object which in turns scatters back into the field of view where it is captured by the time-resolved single-pixel camera. We then explored various imaging scenarios, testing different objects and DMD mask choices.

For the first scenario we investigated a hidden scene of two cooperative objects, i.e. a highly reflective, tin-foil 7.62 cm diameter cylinder and a 2.54 cm diameter mirror placed at different positions and distances from the wall [see Fig. 2(b)]. We used a 120 fs pulsed laser 808 nm wavelength with a repetition rate of 80 MHz and an average power of 800 mW. We collect the third bounce echo with a simple raster scan acquisition on the DMD, one 2.6x2.6 cm² imaging pixel at a time, and collect the reflected light onto a photomultiplier tube (PMT, hybrid photo detector HPM-100-07, Becker&Hickl, 4% efficiency at 808 nm) with optimised temporal response (the measured total impulse response time was 27 ps FWHM). The acquisition time is set by the amount of the back scattered signal detected by the sensor. In this case the acquisition time is 10 seconds per mask (i.e. pixel) for a total acquisition time of 66 minutes. Figure 2(a) shows one time frame of the collected third bounce echos of the two hidden objects.

We proceed with the retrieval of the 3D shape of the hidden objects by applying the back-projection imaging algorithm first introduced by Velten et al. [12, 20]. Although faster retrieval methods are available (see e.g. [15]), our primary focus here is on the hardware rather than on the retrieval software. Regarding the time-offlight evaluation, we consider the moment the laser hits the wall as the zero time reference. We therefore divide the 3D space in 10^6 voxels and we calculate the likelihood of the target to be localized on each voxel by using the time of flight t_p the light takes to cover the distance $r_{\ell v} + r_{vp}$ where $r_{\ell v}$ is the distance between the laser and the voxel and r_{vp} is the distance between the voxel and the pixel. The relation $ct_p = r_{\ell v} + r_{vp}$ indicates that all the possible contributions to a given pixel lie on the surface of an ellipsoid whose foci are the laser spot and the pixel position. The ensemble of the 400 temporal histograms overlap at the scattering object position and thus encodes the 3D geometry of the hidden objects. We assign the likelihood of a voxel by summing the intensity of the pixels that could have received any contribution from the voxel. Following Ref. [12], the final 3D shape of the object is improved by applying a Laplacian filter followed by a threshold selection on the data along the \hat{z} direction of the voxel grid.

Figures 2(c) and (d) show the results on the x-y plane and on the x-z plane, respectively. The dotted line in each figure indicates the actual positions of the targets. Our results show that this technique provides an accurate 3D shape recovering of the hidden (cooperative) targets,

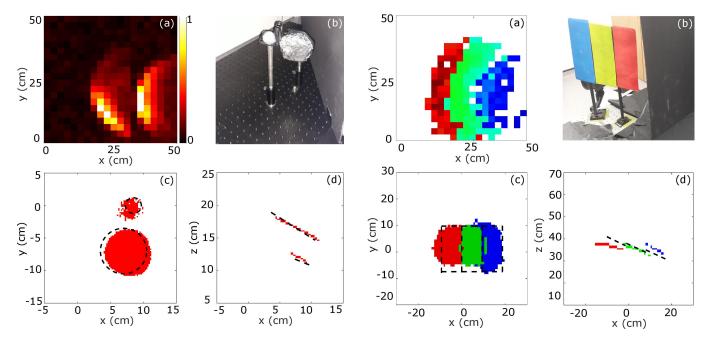


FIG. 2: Cooperative two objects scenario for NLOS single-pixel camera based on a PMT detector. a) Return signal produced by the two hidden objects and retrieved by raster scan imaging onto the DMD and collecting the signal on the single-pixel detector at the time frame $t=7.1~\rm ns.~c-d)$ 3D shape retrieval of the hidden scene in the x-y plane (c) and in the x-z plane (d) after applying backprojection imaging algorithm. The dotted line represents the actual position of the targets. To facilitate the visualization the reflectivity is normalized to the local maxima value. The objects to be recovered are two round targets of 2.54 cm and 7.62 cm of diameter at a varying depth and tilted. The voxels dimension is $0.2 \times 0.2 \times 1~\rm cm^3$ for an area to investigate of $20 \times 20 \times 100~\rm cm^3$. The threshold used for the retrieval is 0.89.

although with relatively long acquisition times.

We then investigate a hidden scene of non-cooperative objects with the same setup (Figure 3). In this case, the scene to be recovered is a red-green-blue (RGB) coloured object, placed outside the direct line-of-sight [Fig. 3(b)], where each coloured region has a rectangular shape of 20x9 cm². In order to retrieve color information we use a supercontinuum laser (SuperK EXTREME/FIANIUM, NKT Photonics, repetition rate 67 MHz, pulse duration ~ 10 ps, average power 100 mW in the 450 nm - 700 nm spectral range). For the RGB retrieval we run a separate measurement for each of the three RGB colours, using a corresponding band-pass spectral filter centred at of 490, 550 and 610 nm (40 nm bandwidth) after the laser source with roughly 20 mW average power for each color. As above, the optimal acquisition time, due to the relatively low laser power, was found to be 10 seconds per mask for an overall acquisition time of 66 minutes. Figure 3(a) shows the three return signals scattered back by the three RGB coloured targets at a time frame of 8 ns. Figures 3(c) and (d) show the reflectivity on the

FIG. 3: Non cooperative objects RGB coloured scenario for NLOS single-pixel camera based on a PMT detector. The red colour corresponds to the recovered target by using the red spectral filter and so on. a) Return signal produced by the hidden object by raster scan acquisition. b) Picture of the RGB coloured target. c,d) RGB retrieval of the 3D shape of the hidden objects in the x-y plane (c) and in the x-z plane (d) after applying backprojection imaging algorithm. The red colour corresponds to the recovered target by using the red spectral filter and so on. The dotted black line represents the actual position of the target. We discretise the space in voxels of 1.4x1.4x1 cm³ for an overall area of 140x140x100 cm³.

x-y plane and on the x-z plane respectively: as can be seen, the retrieved 3D scene corresponds very closely to the ground truth (dashed lines).

The last scenario we investigate is a non-cooperative hidden object (a white paper rectangle of 24x10 cm², see Fig. 1(b)) aimed at the optimising acquisition time. We achieve high speed acquisition by using a high efficiency (70 % peak efficiency at 550 nm) SPAD detector [31] with a measured impulse response function of 30 ps FWHM. The SPAD has a square active area of $57x57 \mu m^2$ and we use a 75 mm focal length lens and a long working distance objective (magnification factor 50) to focus the light after the DMD on to the sensor. The high sensitivity of the detector allows a shorter acquisition time of 1 millisecond per mask, which in this case where chosen as the first 400 Hadamard patterns with the goal of increasing the amount of collected light for each mask (50% of the pixels are always projected onto the detector for each mask). For each Hadamard pattern, one binary mask and its negative are used and combined, leading to a total of 800 patterns. This allowed the total acquisition time to be reduced to only 0.8 s. In this case we used the same supercontinuum laser of the previous sce-

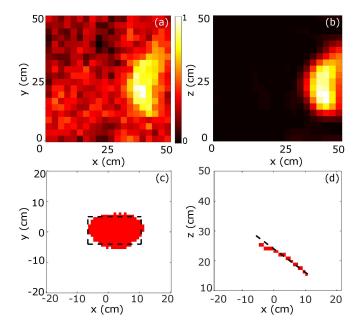


FIG. 4: Ultra-fast NLOS single-pixel camera results. a) Return signal produced by the hidden object by Hadamard acquisition at a time frame of 5.3 ns. b) Return signal obtained by pre-processing filter. c,d) Retrieval of the 3D shape of the hidden target in the x-y plane (c) and in the x-z plane applying backprojection imaging algorithm. The dotted black line represents the actual position of the target. The object to be recovered is a rectangular target of $24x10~\rm{cm}^2$. The voxel dimension is $1.4x1.4x1~\rm{cm}^3$ for an overall of $140x140x100~\rm{cm}^3$. The threshold used for the retrieval is 0.80.

nario with a power of 550 mW at 550 nm. As shown in Fig. 4(a) the collected third bounce echo of the signal is affected by low signal-to-noise ratio due to the short acquisition time. We therefore first apply a denoising algorithm similar to that used in [32] thus returning the signal shown in Fig. 4(b). Precisely, a cost function is defined, accounting for the forward model (including observation noise assumed to be Poisson distributed) relating the image sequence (or video) reaching the DMD and the set of temporal sequences recorded for each Hadamard pattern. The cost function also includes two penalty terms to promote temporal and spatial smoothness after denoising. This is enforced by using a spatial total variation (TV) regularization, as well as a low-pass constraint on the Fourier transform of the temporal intensity profile of each pixel. This cost function is convex and the denoising step, i.e. the cost function minimization, is performed using an alternating direction method of multipliers (ADMM) algorithm [33], as in [34, 35]. Figures 4(c) and (d) show the retrieved reflectivity of the hidden objects on the x-y plane and on the x-z plane respectively. Our results therefore show that this technique provides an accurate 3D shape of a hidden target even with subseconds acquisition times with an average number of only 1.2 photons/pixel in each time frame (with a maximum

peak photon number of ~ 10 photons/pixel).

In conclusion, the high efficiency and the high temporal resolution of single-pixel single-photon detectors allow us to accurately recover the 3D shape of hidden objects even with low resolution masks of 20x20 pixels and no mechanical scanning parts.

The main limitation in the spatial resolution of the retrieval is due to the pixel size on the scattering wall. In our case, a pixel size of $2.6 \times 2.6 \text{ cm}^2$ limits the temporal resolution to 60 ps due to the blurring of the pulse wavefront as it crosses the 3.6 cm pixel collection area (taking the diagonal of the square pixel). This effectively corresponds to an uncertainty in the arrival time of the return pulse and, in analogy with standard LIDAR, will translate into an uncertainty of the object depth location that is 1/2 this value, i.e. 1.8 cm. This can be overcome by decreasing the pixel size however, at the cost of longer acquisition times due to the larger number of Hadamard patterns (or pixels to scan).

Overall, the ability to identify a hidden scene by the proposed approach is mainly determined by the time resolution of the detector and by the time required to acquire a significant back-scattered signal. The retrieval of hidden scenes still remains a challenging task due to long acquisition times, low spatial resolution of the retrieval and computational sources, although significant steps have been made recently, see e.g. Refs. [6, 19]. We have shown a NLOS ultra-fast imaging technology which can reliably recover the 3D shape in colour of a scene with high spatial resolution by using single-pixel, single-photon detectors with high temporal resolution. By using a high sensitivity detector, this system is able to retrieve the shape of the hidden target with an improved acquisition time of 0.8 seconds paving the way to real-time 3D shape recovery of hidden objects. The accurate 3D shape recovery of this system could be further improved by fully exploiting the benefits of using single-pixel camera for the acquisition. Indeed a future improvement in this method would be to decrease the acquisition times by using compressive sampling [36]. By combining compressive sensing and constant improvements of detection and computational resources, recovery of the 3D shape of hidden moving objects with high spatial resolution should be possible.

Acknowledgments. The authors acknowledge support from the Royal Academy of Engineering under the Research Fellowship scheme RF201617/16/31 and EPSRC, UK grant EP/M01326X/1.

- * Electronic address: 2376730m@student.gla.ac.uk
- † Electronic address: Daniele. Faccio@glasgow.ac.uk
- D. Faccio and A. Velten. A trillion frames per second: the techniques and applications of light-in-flight photography. Reports on Progress in Physics, 2018.
- [2] 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*, volume 7308, page 73080V. International Society for Optics and Photonics, 2009.
- [3] O. Katz, E. Small, and Y. Silberberg. Looking around corners and through thin turbid layers in real time with scattered incoherent light. *Nature photonics*, 6(8):549– 553, 2012.
- [4] O. Katz, M. Heidmann, P.and Fink, and S. Gigan. Noninvasive single-shot imaging through scattering layers and around corners via speckle correlations. *Nature pho*tonics, 8(10):784, 2014.
- [5] Y. Altmann, S. McLaughlin, M. J. Padgett, V. K Goyal, A.O. Hero, and D. Faccio. Quantum-inspired computational imaging. *Science*, 361(6403):eaat2298, 2018.
- [6] C. Saunders, J. Murray-Bruce, and V. K Goyal. Computational periscopy with an ordinary digital camera. *Nature*, 565(8):472?475, 2019.
- [7] U. Wandinger. Introduction to lidar in lidar-rangeresolved optical remote sensing of the atmosphere. Springer, New York, 2005.
- [8] B. Schwarz. Lidar: Mapping the world in 3d. Nature Photonics, 4(7):429, 2010.
- [9] A. Kirmani, T. Hutchison, J. Davis, and R. Raskar. Looking around the corner using ultrafast transient imaging. *International journal of computer vision*, 95(1):13– 28, 2011.
- [10] M. Buttafava, J. Zeman, A. Tosi, K. Eliceiri, and A. Velten. Non-line-of-sight imaging using a time-gated single photon avalanche diode. *Optics express*, 23(16):20997–21011, 2015.
- [11] A. K. Pediredla, A. Buttafava, M.and Tosi, O. Cossairt, and A. Veeraraghavan. Reconstructing rooms using photon echoes: A plane based model and reconstruction algorithm for looking around the corner. In 2017 IEEE International Conference on Computational Photography (ICCP). IEEE, 2017.
- [12] 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. *Nature communications*, 3:745, 2012.
- [13] G. Gariepy, F. Tonolini, R. Henderson, J. Leach, and D. Faccio. Detection and tracking of moving objects hidden from view. *Nature Photonics*, 10(1):23–26, 2016.
- [14] S. Chan, R. E Warburton, G. Gariepy, J. Leach, and D. Faccio. Non-line-of-sight tracking of people at long range. Optics express, 25(9):10109–10117, 2017.
- [15] Victor Arellano, Diego Gutierrez, and Adrian Jarabo. Fast back-projection for non-line of sight reconstruction. Opt. Express, 25(10):11574–11583, May 2017.
- [16] C. Jin, J. Xie, S. Zhang, Z. Zhang, and Y. Zhao. Reconstruction of multiple non-line-of-sight objects using back projection based on ellipsoid mode decomposition. arXiv preprint arXiv:1805.04518, 2018.
- [17] J. Klein, C. Peters, J. Martin, M. Laurenzis, and M. B. Hullin. Tracking objects outside the line of sight using 2d intensity images. *Scientific reports*, 6:32491, 2016.
- [18] P. Caramazza, A. Boccolini, D. Buschek, M. Hullin, C. F. Higham, R. Henderson, R. Murray-Smith, and D. Faccio. Neural network identification of people hidden from view with a single-pixel, single-photon detector. *Scientific re*ports, 8(1):11945, 2018.
- [19] M. OToole, D.B. Lindell, and G. Wetzstein. Confocal non-line-of-sight imaging based on the light-cone trans-

- form. Nature, 555(7696):338, 2018.
- [20] M. La Manna, F. Kine, E. Breitbach, J. Jackson, T. Sultan, and A. Velten. Error backprojection algorithms for non-line-of-sight imaging. *IEEE transactions on pattern analysis and machine intelligence*, 2018.
- [21] A. C. Kak. Algorithms for reconstruction with nondiffracting sources. Principles of computerized tomographic imaging, pages 49–112, 2001.
- [22] M. P. Edgar, G. M. Gibson, and M. J. Padgett. Principles and prospects for single-pixel imaging. *Nature photonics*, 13:13–20, 2019.
- [23] N. Radwell, K. J. Mitchell, G. M Gibson, M. P Edgar, R.. Bowman, and M. J Padgett. Single-pixel infrared and visible microscope. *Optica*, 1(5):285–289, 2014.
- [24] E. Tajahuerce, V. Durán, P. Clemente, E. Irles, F. Soldevila, P. Andrés, and J. Lancis. Image transmission through dynamic scattering media by single-pixel photodetection. *Optics Express*, 22(14):16945–16955, 2014.
- [25] Wai Lam Chan, Kriti Charan, Dharmpal Takhar, Kevin F Kelly, Richard G Baraniuk, and Daniel M Mittleman. A single-pixel terahertz imaging system based on compressed sensing. Applied Physics Letters, 93(12):121105, 2008.
- [26] G. M Gibson, B. Sun, M. P Edgar, D. B Phillips, G. T. Hempler, N.and Maker, G. PA Malcolm, and M. J. Padgett. Real-time imaging of methane gas leaks using a single-pixel camera. *Optics express*, 25(4):2998–3005, 2017.
- [27] M. J. Sun, M. P. Edgar, G. M. Gibson, B. Sun, N. Radwell, R. Lamb, and M. J Padgett. Single-pixel threedimensional imaging with time-based depth resolution. *Nature communications*, 7:12010, 2016.
- [28] M. P Edgar, M. J. Sun, G. M. Gibson, G. C. Spalding, D. B. Phillips, and M. J. Padgett. Real-time 3d video utilizing a compressed sensing time-of-flight single-pixel camera. In *Optical Trapping and Optical Micromanip*ulation XIII, volume 9922, page 99221B. International Society for Optics and Photonics, 2016.
- [29] M. F Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE signal process*ing magazine, 25(2):83–91, 2008.
- [30] M. Satat, G. Tancik and R. Raskar. Lensless imaging with compressive ultrafast sensing. *IEEE Transactions* on Computational Imaging, 3(3):398–407, 2017.
- [31] M. Sanzaro, F. Tosi A. Croce G. Gattari, P. Villa, and F. Zappa. Single-photon avalanche diodes in a 0.16 μm bcd technology with sharp timing response and redenhanced sensitivity. *IEEE Journal of Selected Topics in* Quantum Electronics, 24(2):1–9, 2018.
- [32] P. Caramazza, K. Wilson, G. Gariepy, J. Leach, S. McLaughlin, D. Faccio, and Y. Altmann. Enhancing the recovery of a temporal sequence of images using joint deconvolution. *Scientific Reports*, 8, March 2018.
- [33] M. A. T. Figueiredo and J. M. Bioucas-Dias. Restoration of poissonian images using alternating direction optimization. *IEEE Transactions on Image Processing*, 19(12):3133–3145, Dec 2010.
- [34] Y. Altmann, A. Maccarone, A. Halimi, A. McCarthy, G. Buller, and S. McLaughlin. Efficient range estimation and material quantification from multispectral lidar waveforms. In 2016 Sensor Signal Processing for Defence (SSPD), pages 1–5, Sep. 2016.
- [35] R. Tobin, Y. Altmann, X. Ren, A. McCarthy, Robert

A. Lamb, S. McLaughlin, and G. S. Buller. Comparative study of sampling strategies for sparse photon multispectral lidar imaging: Towards mosaic filter arrays. *Journal of Optics*, 19(9), 8 2017.
[36] Q. Chen, S. K. Chamoli, P. Yin, X. Wang, and X. Xu.

Imaging of hidden object using passive mode single pixel imaging with compressive sensing. Laser Physics Letters, $15(12):126201,\ 2018.$