Single-photon three-dimensional imaging at up to 10 kilometers range

Citation for published version:
Pawlikowska, AM, Halimi, A, Lamb, RA & Buller, GS 2017, 'Single-photon three-dimensional imaging at up to 10 kilometers range', Optics Express, vol. 25, no. 10, pp. 11919-11931.
https://doi.org/10.1364/OE.25.011919

Digital Object Identifier (DOI):
10.1364/OE.25.011919

Link:
Link to publication record in Heriot-Watt Research Portal

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Optics Express

General rights
Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Single-photon three-dimensional imaging at up to 10 kilometers range

AGATA M. PAWLIKOWSKA,1,2,* ABDERRAHIM HALIMI,2 ROBERT A. LAMB,1 AND GERALD S. BULLER2

1Leonardo MW Ltd, Crewe Road North, Edinburgh, EH5 2XS, UK
2Institute of Photonics and Quantum Sciences, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, UK
*agata.pawlikowska@leonardocompany.com

Abstract: Depth and intensity profiling of targets at a range of up to 10 km is demonstrated using time-of-flight time-correlated single-photon counting technique. The system comprised a pulsed laser source at 1550 nm wavelength, a monostatic scanning transceiver and a single-element InGaAs/InP single-photon avalanche diode (SPAD) detector. High-resolution three-dimensional images of various targets acquired over ranges between 800 metres and 10.5 km demonstrate long-range depth and intensity profiling, feature extraction and the potential for target recognition. Using a total variation restoration optimization algorithm, the acquisition time necessary for each pixel could be reduced by at least a factor of ten compared to a pixel-wise image processing approach. Kilometer range depth profiles are reconstructed with average signal returns of less than one photon per pixel.

© 2017 Optical Society of America

OCIS codes: (110.6880) Three-dimensional image acquisition; (100.6890) Three-dimensional image processing; (030.5260) Photon counting; (010.3640) Lidar; (150.5670) Range finding; (040.1345) Avalanche photodiodes (APDs).

References and links

1. Committee on Developments in Detector Technologies, National Research Council, Seeing Photons: Progress and Limits of Visible and Infrared Sensor Arrays (The National Academy Press, 2010).
2. A. McCarthy, R. J. Collins, N. J. Krichel, V. Fernández, A. M. Wallace, and G. S. Buller, “Long-range time-of-flight scanning sensor based on high-speed time-correlated single-photon counting,” Appl. Opt. 48(32), 6241–6251 (2009).
3. M. A. Albota, B. F. Aull, D. G. Fouche, R. M. Heinrichs, D. G. Kocher, R. M. Marino, J. Mooney, N. R. Newbury, M. E. O’Brien, B. E. Player, B. C. Willard, and J. J. Zayhowski, “Three-dimensional imaging laser radars with Geiger-mode avalanche photodiode arrays,” Linc. Lab. J. 13(2), 351–370 (2002).
4. G. S. Buller and R. J. Collins, “Single-photon generation and detection,” Meas. Sci. Technol. 21(1), 012002 (2010).
5. A. McCarthy, X. Ren, A. Della Frera, N. R. Gemmell, N. J. Krichel, C. Scarcella, A. Ruggeri, A. Tosi, and G. S. Buller, “Kilometer-range depth imaging at 1,550 nm wavelength using an InGaAs/InP single-photon avalanche diode detector,” Opt. Express 21(19), 22098–22113 (2013).
6. W. Becker, Advanced Time-Correlated Single-Photon Counting Techniques (Springer, 2005).
7. S. Pellegrini, G. S. Buller, J. M. Smith, A. M. Wallace, and S. Cova, “Laser-based distance measurement using picosecond resolution time-correlated single-photon counting,” Meas. Sci. Technol. 11(6), 712–716 (2000).
8. J. J. Degnan, “Photon-counting multikilohertz microlaser altimeters for airborne and spaceborne topographic measurements,” J. Geodyn. 34(3–4), 503–549 (2002).
9. I. I. Kim, B. McArthur, and E. Korevaar, “Comparison of laser beam propagation at 785 nm and 1550 nm in fog and haze for wireless optical communications,” Proc. SPIE 4214, 26–37 (2001).
10. Safety of laser products, BSI Standards Publication, BS/EN/60825–1 (2014).
11. M. Henriksson, H. Larsson, C. Gronwall, and G. Toft, “Continuously scanning time-correlated single-photon counting single-pixel 3-D lidar,” Opt. Eng. 56(3), 031204 (2016).
12. M. Entwistle, M. A. Itzler, J. Chen, M. Owens, K. Patel, X. Jiang, K. Slomkowski, and S. Rangwala, “Geiger-mode APD camera system for single-photon 3-D LADAR imaging,” Proc. SPIE 8375, 83750D (2012).

https://doi.org/10.1364/OE.25.011919

Received 6 Feb 2017; revised 30 Mar 2017; accepted 1 Apr 2017; published 12 May 2017
1. Introduction

In recent years there has been increasing interest in the development of single-photon counting lidar for long-range three-dimensional imaging for a number of remote sensing applications. One reason for this is the recent availability of Geiger-mode (Gm) arrays which provide full frame data acquisition with single-photon sensitivity and picosecond resolution. The technology has also found applications in airborne surveillance where long-range target identification through turbulence presents an engineering challenge [1]. Specifically, strong attention is being focused on developing a wide range of applications such as wide field-of-view (FoV) airborne surveillance and long-range target recognition and identification sensors. Although each application will have specific requirements influencing the design and the choice of components, it is clear that there is a need for systems which can provide three-dimensional, high-resolution imaging over long ranges, with night-time imaging capability. The use of lower power laser sources means that single-photon detection will have greater level of covertness and is less likely to exceed eye-safety thresholds. Applications such as airborne surveillance using active imaging impose limits on system weight, size and volume and necessitate low-power laser sources and highly sensitive optical detection [2,3]. Single-photon lidar is a candidate technology that has the potential to meet these challenging requirements.

Time-correlated single-photon counting (TCSPC) is a statistical sampling technique which records the time-of-arrival of photons against the time of emission of the associated laser pulse [4–6]. In contrast to analogue optical detection, the timing resolution of single-photon detection is not limited by the duration or rise time of the voltage pulse but is determined by the variance of the rise time of the detector, or timing jitter [7]. Single-photon detection can therefore provide timing error up to an order of magnitude better than that possible with an analogue optical detector, leading to significantly improved depth resolution [4,8]. In addition, the high sensitivity of single-photon detectors allows lower power laser sources to be used and can permit time-of-flight data to be measured from significantly longer ranges [9]. The possibility of using lower power sources means that single-photon lidar systems can be smaller, lighter and consume less power which is desirable for integration onto airborne platforms.

This paper describes the experimental results and data analysis approaches for depth imaging at kilometer range distances acquired with a single-photon lidar system. The system used a scanning transceiver operating in a mono-static configuration using a 1550 nm
wavelength laser and a single-element, single-photon avalanche diode (SPAD) detector. Compared to shorter wavelengths appropriate for Si SPAD detectors, the benefits of operating at the 1550 nm wavelength include reduced solar background, lower atmospheric loss and increased eye safety thresholds which permit higher optical power levels [10–12].

Implementation of the TCSPC time-of-flight lidar operating at a wavelength of 1550 nm has been reported in several application scenarios. For example, Ren et al. [13] demonstrated laser ranging at a range of 32 meters using a 1 GHz sine-wave InGaAs/InP SPAD with an average laser power of 5 µW. Sub-centimeter resolution ranging at a standoff distance of 330 m was demonstrated using superconducting niobium nitride nanowire single-photon detectors [14]. Similarly, centimeter resolution depth imaging was demonstrated with low signature objects at ranges of up to one kilometer using an InGaAs/InP SPAD at average laser power levels of approximately 1 mW [5,15]. A 32 × 32 InGaAs/InP Gm-array described in [16] was used to demonstrate three-dimensional imaging over ranges of up to 9 km, but required a high power laser with 0.4 W average optical power [17].

This paper presents single-photon time-of-flight depth and intensity imaging acquired over ranges of up to 10.5 km with eye-safe laser powers, with an average optical power of only 10 mW. The measurements were made using an efficient laser scanning approach which, in conjunction with advanced image processing algorithms, can be used to rapidly reconstruct depth profiles with signal levels of less than one photon per pixel. We present an in-depth study and profile analysis of different types of low signature targets, including buildings, an electricity pylon and terrain.

2. Experimental setup

The experimental setup was arranged in a mono-static configuration and incorporated a Peltier-cooled, single-element InGaAs/InP SPAD detector. The SPAD detector module was manufactured by Micron Photon Devices [18] and had a 25 µm active area diameter and timing jitter of 100 ps. The detector used gated quenching where the device is held above avalanche breakdown for a short period (typically 10 ns) around the expected time of arrival of the return photons. The depth profiling system used an erbium-doped fiber laser operating at a wavelength of 1550 nm, which generated 800 ps duration pulses at a repetition rate of 125 kHz. The maximum average laser power used in experiments was 10 mW to comply with laser eye-safety precautions, and was designated as Class 1 as specified by the PD IEC TR 60825-14(2004) standard [19]. This average optical power was equivalent to 0.8 nJ per pulse. Photon returns from a target were collected by the receiver and directed to the SPAD detector as shown in the system layout of Fig. 1. The signal was acquired from the target via a pair of galvanometer mirrors which scan the optical field-of-regard (FoR) directing the outgoing optical pulses in the x and y planes and then directing the target return photons in the return channel.

The laser beam was transmitted from an optical fiber which was coupled into a collimator and a beam expander and then injected into the main system via a small aperture in the annular mirror. The optical components of the collimator and the beam expander were set up to provide the required laser beam divergence for the beam channelled along the direction parallel and displaced from the optical axis of the telescope. The telescope collected a return signal from a target which was then subsequently collimated by an eyepiece. The set of relay lenses projected the image of the exit pupil from the mirror y to the mirror x while the annular mirror reflected it onto the lens L(b) which created a focal spot on the active area of the SPAD detector. In order to suppress out-of-band solar background photons, spectral filters (8 nm bandpass) were placed between the lens L(b) and the SPAD detector. The time interval analyzer used in the setup was a GuideTech GT658PCI, which had two independent input channels and an external “arm” input. The “arm” input enabled the time interval analyzer to start a block of measurements after an external transistor-transistor logic (TTL) signal was supplied. The device provided a count rate of up to 12 Meps per channel and a minimum time-bin width of 80 ps with the maximum number of 1 million time tags for both channels.
This optical system was specifically designed for long-range imaging, and contains a number of key differences from the previous scanning systems described in [2,5,7]. One of the main differences is that the telescope aperture of the new system was nearly three times larger than that described in previous systems [2,5,7], in order to increase the photon collection ability of the system at the increased target range. In addition, the laser power used was significantly greater that the 1mW average powers used in these previous scanning systems.

A commercially sourced digital delay and pulse generator were used to synchronize and gate the SPAD detector and clock the laser pulse at a frequency of 125 kHz. The output power of the laser was adjusted via a current driver controlled by a PC. The laser trigger signal was used to initiate the scanner field-programmable gate array (FPGA), sending TTL pulses to the servo drivers of the scanning mirrors set by control software. The event timing analyzer was used to continuously time-stamp two events: laser trigger pulses and photon events recorded by the detector. The FPGA was set up to trigger scanning mirrors after clocking a certain number of laser trigger pulses and released a TTL pulse after a scanning mirror moved to a new position. This pulse initiated an arm input of the GuideTech which started acquiring time tags. During the time-tag acquisition the GuideTech was synchronized to an internal 50 MHz time base. After each scan point, the time counting process was re-set to a common starting reference at the time-interval analyzer. The delay generator was used to adjust the delay between the SPAD gate trigger pulse with respect to the laser/scanner trigger pulse such that, depending on the range at which the target was located in respect to the optical system, photon returns from the target were collected approximately in the center of the SPAD timing gate. The time tags were collected with 80 ps temporal resolution which was set by the intrinsic limitations of the time interval analyzer.

As the transmitted beam propagates through the common transmit-return channel there are back-reflections from optical components in the outgoing path. This unwanted back-reflected signal, if not properly attenuated, can cause a significant contribution to the detector count-rate, possibly saturating the detector and inhibiting all measurements of the target return. In order to deal with the back-reflection issue, the electrical detector gating approach employed in these measurements was similar to that previously used by McCarthy et al [5]. The electrical gate was activated around the expected photon return time but is de-activated as the initial optical pulse propagated outward through the optical system, thus removing the possibility of detecting unwanted back-reflections. In addition to this, the detector was also subjected to a “hold-off time” where the detector was de-activated for a pre-determined duration after each detected event, 20 µs in this case. This hold-off time was necessary as
InGaAs/InP SPAD detectors suffer from the deleterious effects of afterpulsing, where carriers are trapped during the avalanche process and later released causing further events which increase the background count level. By introducing a hold-off time which allows trap states to empty prior to the detector being re-activated, the likelihood of afterpulsing can be significantly reduced.

The system was sited in a laboratory which provided access to a mixed urban and rural environment with a range of up to 10.5 km and varying altitude along the line of sight. Measurements were taken in varying weather conditions and different times of the day. The SPAD detector was set to a temperature of 265 K and operated at 3 V excess bias which resulted in a single-photon detection efficiency of ~30%. The detector was gated at the laser repetition frequency of 125 kHz, and the gate width applied to the detector varied between 100 ns and 500 ns which corresponded to an overall image depth of 15 m and 75 m respectively. All the results presented in this paper were acquired with an average laser power of 10 mW. The full-width half-maximum (FWHM) of the instrumental response function (IRF) was measured to be 0.85 ns which is equivalent to 12.7 cm depth. The overall IRF shape was a convolution of the laser pulse width, timing response of the electronics and the SPAD detector jitter.

The telescope used in previous experiments used only refractive optical components. This configuration allowed the optical axis of the transmitted and received beams to directly coincide and propagate along the optic axis of the optical components, simplifying the optical alignment. The aperture of the telescope used in this system was 21 cm, necessitating the use of a reflective telescope due to its reduced mass. However, the complexity of optical alignment significantly increased as the transmitted beam must propagate off-axis through the system due to the telescope’s central obstruction. One of the challenges of this experiment was to develop rigorous alignment procedures in order to ensure that the off-axis transmitted beam would maintain its alignment over the entire FoR at up to 10 km range.

3. Depth profile retrieval

During a depth measurement, the detected photon returns were time-tagged by the GT658PCI and transferred to the control computer where software generated a timing histogram of photon returns for each scanned pixel. The histograms, produced with 80 ps wide bins, were analyzed using least-squares curve fitting algorithm to locate the position of the signal peak in the histogram. Least-squares curve fitting minimizes the square of the error between the experimental data points and the values of the fitting function. In this case the residual of the $n$-th pixel, $res^2$, expressed by [20]

$$\text{res}^2 = \sum_{t=1}^{T} (x_{n,t} - y_{n,t})^2$$

is minimized, where $x_{n,t}$ is the $t$-th element of the $n$-th pixel of the fitting function, $y_{n,t}$ is the $t$-th element of the $n$-th pixel of the data set, i.e. histogram and $T$ is the total number of sampling points in the histogram [20]. A quadratic polynomial is fitted to data set allowing the peak position to be identified with a precision determined by a user-defined number of consecutive data points used in the fitting process [20]. For each peak, the least squares fit was tested against a user defined threshold (typically above the background level), thus, peaks lower than the threshold were ignored.

The algorithm calculated the time-of-flight, $t_f$, corresponding to the identified peak, from which target range, $R$, was determined. There are multiple pulses in transit and therefore establishing the absolute range was not possible without prior information about the target range. In this case the expected range information was known in advance which was then used to estimate the number of gates prior to the gate in which a return from a target is expected. The absolute range was calculated for data cross-correlated with the instrumental response from
\[ R = \frac{(mp + t_f)c}{2} \]  

where \( m \) is the number of a detector gates occurring between the laser pulse being emitted and the return from the target, \( p \) is the gate period which is equal to 8 µs, \( t_f \) is the time-of-flight to the target corresponding to the peak found using the technique described above, and \( c \) is the speed of light in the transmitting medium. In future system designs, we will use a combination of different repetition frequencies or a more sophisticated approach based on pseudorandom pulse trains to provide unambiguous depth measurements [21].

4. Experimental results

4.1 Long range imaging of different types of targets

Different targets were selected for their spatial distribution, reflectivity and structured complexity. The type of objects used to demonstrate versatility of the system included targets with complex spatial and depth features. This included curved as well as multiple angled surfaces, such as a building with an extended depth-profile at a range of ~8.8 km, and sloping terrain at a range of ~10.5 km. In addition, multiple images were tiled in a mosaic to cover an extended field-of-view (FoV). The angular resolution of the system was 28 µrad. We set the inter-pixel spacing to match the measured value for the spatial resolution in all measurements shown in this section regardless of distance. For example, this is equivalent to 28 cm at 10 km. In the next sections, we will show examples of images obtained for a solid target with complex structure, a distributed target and a long-range target.

4.1.1 Solid target with complex structure

A clock tower at 800 meter range shown in Fig. 2(a) was selected and scanned with the lidar to produce a \( 85 \times 85 \) pixel image with an inter-pixel spacing of 2.2 cm. The data was acquired with 170 ms acquisition time per scan point (20 minutes total acquisition time). Depth profiles of the target are shown in Fig. 2(b) and Fig. 2(c). Complex features of the clock tower are shown in Fig. 2(c), such as the shape of the roof, the structure of the ribs, the shape of the gutter, a ventilation opening near the top of the roof and slats in the window shutter. The angle of the slats was estimated to be ~45° with respect to the wall.

![Fig. 2. (a) A two-dimensional visible-band image of the top of the clock tower acquired with a camera lens of f = 200 mm. (b) Depth plot of the top of the clock tower at a range of ~800 m with 85 × 85 scan points and an acquisition time of 170 ms per pixel. (c) Depth plot of the top of the clock tower at a range of ~800 m with 85 × 85 scan points and an acquisition time of 170 ms per pixel showing side view of the target.](image-url)

Figure 3(a) shows a visible-band photograph of a residential building at 8.8 km taken using a camera lens of \( f = 200 \) mm. The object was scanned with 32 × 32 scan points with an inter-pixel spacing of 24.6 cm. Acquisition time per pixel was 0.3 s leading to a total scan time of approximately 5 minutes. Figure 3(b) shows a depth-intensity plot and Fig. 3(c) shows...
depth profile of this building. Although the conventional photographic image is slightly blurred due to the inadequate spatial resolution of the camera lens, the image obtained with the lidar shows clear detail such as the profile of the roof and the size of the window. Reflection from the back wall allows the length of the room to be estimated.

4.1.2 Distributed target (a pylon)

Figure 4(a) shows a close-up, visible-band image of an electricity pylon taken with an $f = 200$ mm camera lens. A $40 \times 80$ scan of this pylon was taken over a distance of $\sim 6.8$ km with an inter-pixel spacing of 19 cm. The data was acquired with 0.23 s acquisition time per scan point, leading to an overall acquisition time of 12 minutes.

Depth-intensity plots of a pylon are shown in Fig. 4(b) and Fig. 4(c). Data was analyzed using the least-squares peak finding algorithm (see section 3) with two different levels of threshold. Figure 4(b) shows the results for a threshold of 5 counts which allows details of the structure to be identified; nevertheless, many data points caused by noise are also present. In Fig. 4(c) the data was analyzed for a threshold of 7 counts which removed the noise at the expense of some detail.
4.1.3 Long range target

A 32 × 32 scan of terrain was taken over a range of ~10.5 km with an inter-pixel spacing of 29 cm, the acquisition time of 0.3 s per scan point and the total scan time of ~5 min. A visible band photograph of the scene taken with an f = 200 mm camera lens is shown in Fig. 5(a). This is a sloping terrain composed of a mix of different materials such as rock and foliage located on a hillside near Edinburgh.

A front view and side view intensity and depth plots of the scene are shown in Fig. 5(b) and in Fig. 5(c) respectively. The depth of terrain stretches across a range of approximately 12 m in the direction of laser beam propagation, its average slope to the line of sight was determined by calculating a linear fit to the data points. A linear equation in the two variables \( x \) and \( y \) is given by \( y = mx + b \) where \( m \) determines the gradient of the line and \( b \) determines the intercept on the \( y \) axis \([22]\). Linear fit to the data points of the top view is given by \( y = 0.77 x + 10501 \). From \( m = \tan \beta \), the approximate angle of the slope in respect to the optical axis of the system is \( \beta = \arctan (0.77) = 38^\circ \).

4.2 Image mosaicking

Image mosaicking refers to the tiling of multiple images into a single composition to display larger target \([23]\). In this way it is possible to build views of scenes which cannot be acquired with a narrow FoV optical system. Using an angular scan step size of 28 µrad - which is equal to the Airy disk diameter produced by the system defining spatial resolution - the maximum number of scan points is 100 × 100, before the effects of vignetting are evident. The object shown in Fig. 6 was imaged with an auxiliary CCD camera in three frames, each containing 100 × 100 scan points collected with an acquisition time of 1.87 ms per scan step. Each scan step corresponded to an inter-pixel spacing of 8.4 cm at the target located at ~3 km.

A riflescope was bore-sighted with the optical axis of the system during the alignment and was used to point the system at the desired part of the target. A test run was performed to establish the relationship between dimensions of the object, the FoR of the system for a 100 ×
100 scan points and the scale of the riflescope cross-hair. A 100 × 100 scan corresponded to the area of the FoR represented by a square shown in Fig. 6. Once this was established the center of the cross-hair of the riflescope was positioned in point A of the scene as illustrated in Fig. 6 and a scan representing part 1 of the image was acquired. Subsequently, the scan was taken with the system positioned with the riflescope centered at the point B, overlapping the previous scan area by approximately 35% (35 pixels). Finally, the system was positioned at the point C, acquiring the scan with about 35% overlap with the scan 2. The overlaps were needed for the alignment of images. A large overlap allows a higher accuracy or image alignment but necessarily requires a longer acquisition time.

The acquired data was analyzed for the three parts of the image. The overlapping parts of the image were then merged and used to produce a depth plot of the building which accounted for 100 × 230 scan points. The total acquisition time required to produce the images was 41 s.

Examples of different viewpoints of the scene are shown in Fig. 7(a) and Fig. 7(b). The total height of the building is nearly 13 m and the width is approximately 9 m.

![Fig. 6. Wide field-of-view (FoV) monochromatic image of the target taken with an auxiliary CCD camera. The scene was divided into three parts which were separately scanned to generate a full image of the object.](image)

![Fig. 7. Depth-plots of a building at 3 km; (a) bottom-right-side view (b) bottom-left-side view. The image is a mosaic of three segments and consists of 100 × 230 scan points. The acquisition time per scan point was 1.8 ms per scan point resulting in 41 s total acquisition time.](image)
The acquisition times for targets reported earlier are too long for a deployed lidar in a number of typical applications. A reduction in the total acquisition time can be achieved by increasing the laser pulse energy or by increasing the repetition rate of the laser source. Increasing the laser power was not desirable due to the laser safety implications. Operating the laser within the class I limit (up to 10 mW at a wavelength of 1550 nm [12]) is an important factor for a deployed lidar because the extended non-ocular hazard distance (ENOHD) is reduced to 0 meters, meaning the deployed system is eye-safe regardless of the position of the observer.

The acquisition time can be reduced by use of image processing algorithms which take into account the spatial correlation in the depth image. In this work, we consider the “Restoration of Depth and Intensity using Total Variation” (RDI-TV) algorithm [24]. Similar techniques have been used to demonstrate intensity and depth profile restoration from sparse single-photon data in underwater imaging [25] and in free-space [26–28].

The RDI-TV algorithm has two main objectives: (i) the restoration of the corrupted depth and intensity images and (ii) the reconstruction of the missing pixels [24]. Indeed, at low acquisition times, a reduced number of photon counts is collected causing many pixels in an image to be empty or less informative. These missing pixels make the depth and intensity estimation of the target impossible without additional information such as the spatial correlation between adjacent pixels. This can be interpreted as an image inpainting problem whose cost function, \( C(d, r) \), is given by [24]

\[
C(d, r) = L(Kd, Kr) + \phi(d, r)
\]  

(3)

where \( d \) (resp. \( r \)) represents an \( N \times 1 \) vector gathering the pixels of the depth (resp. reflectivity) image, \( L \) is the negative log-likelihood function depending on the statistics of the data (Poisson distribution), \( K \) is a \( Q \times N \) binary matrix that contains a single non-zero value on each line to model the loss of some image pixels, \( Q \) is the number of non-empty pixels and \( \phi(d, r) \) is the regularization term that introduces the available prior knowledge about \( d \) and \( r \) [24].

The lidar observation \( y_{n,t} \) (i.e., the histogram) represents the number of photon counts within the \( t \)-th bin of the \( n \)-th pixel. According to [29], each photon count, \( y_{n,t} \), is assumed to be drawn from the Poisson distribution \( P(s_{n,t}) \) with an average \( s_{n,t} = d_n r_n \) related to the impulse response of the system, the target depth \( d_n \) and reflectivity \( r_n \). Reference [24] described some justified assumptions that simplify the Poisson-based negative log-likelihood function as follows (see [29] for more details)

\[
L(d, r) = \sum_{n=1}^{N} \sum_{t=1}^{T} \left[ s_{n,t} - y_{n,t} \log(s_{n,t}) \right] + \text{cst}
\]  

(4)

Regarding the regularization term, \( \phi(d, r) \), a total variation (TV) is considered. The latter assumes spatially correlated pixels (using a four neighborhood structure) leading to the cost function given by:

\[
C_{TV}(d, r) = L(Kd, Kr) + \tau_1 TV(d) + \tau_2 TV(r)
\]  

(5)

where \( \tau_1 > 0 \) and \( \tau_2 > 0 \) are regularization parameters and \( TV(s) = \sum_{n=1}^{N} \sqrt{(\Delta^h n \cdot s)^2 + (\Delta^v n \cdot s)^2} \) where \( \Delta^h n \) and \( \Delta^v n \) are the horizontal and vertical first order differences at the \( n \)-th pixel [24]. In order to minimize, \( C_{TV}(d, r) \), an alternating direction method of multipliers (ADMM) algorithm is used [24]. This algorithm was applied to restore the lost depth information due to the reduction of acquisition time of a clock face shown in Fig. 8. The target was scanned over a range of ~800 m with 50 × 50 scan points, inter-pixel spacing of 2.2 cm and 10 mW average laser power.
Fig. 8. Wide field-of-view, visible band image of a clock face on a tower taken with an f = 200 mm camera lens.

Depth plots of the target were generated using the classical approach, which consists of the cross-correlation of the histogram with the instrumental response function (see Fig. 9 a1-g1). These figures are compared to those obtained when considering the RDI-TV algorithm, as shown in Fig. 9 (a2-g2). Note that both algorithms were applied on images with an acquisition times per pixel varying from 5.3 ms to 13.75 µs. The acquisition time is the time over which photons were recorded and does not relate to the mechanical scan time.

Fig. 9. Depth profile measurements of a clock face on a tower over a range of ~800 m acquired with 50 × 50 scan steps analyzed using cross-correlation (a1) – (g1) and RDI-TV (a2) – (g2).

At the 5.3 ms acquisition time per scan point the scatterplots generated with the cross-correlation and the RDI-TV show comparable image quality, i.e. the surface of the clock appears to be smooth, the edges of the detail are sharp and allow the detail of the clock such as the hands and the circular shape of the clock face to be identified. For the cross-correlation approach, as the acquisition time decreases the identification of detail is significantly degraded; at 55 µs only 30% of the pixels have generated a depth measurement and these pixels contain significant errors in depth, preventing target identification. The RDI-TV algorithm, however, restores empty pixels in the data array and reconstructs the depth relationship between the pixels allowing some of the main detail of the target, such as the clock hands, to be recovered at 27.5 µs acquisition time per pixel. The roughly circular shape of the clock face can also be restored from the RDI-TV generated depth map. This demonstrates that the RDI-TV algorithm allows detailed identification of the clock face and
restoration of the missing pixels with the acquisition time of 27.5 µs per pixel. For an image comprising 50 × 50 scan point this yields the total acquisition time of approximately 69 ms. The image quality degrades significantly for an acquisition time of 13.75 µs per pixel for the cross-correlation method where nearly 90% of the image pixels are missing. Nevertheless, the RDI-TV allows more than a half of the image to be restored with some of the clock hands being identified. These results confirm that use of the RDI-TV algorithm has significant potential for the acquisition time reduction in single-photon depth imaging.

The restoration quality was also evaluated using the reconstruction signal-to-noise ratio (RSNR) for both the cross-correlation and RDI-TV. The RSNR is defined as follows [24]

$$RSNR = 10 \log_{10} \left( \frac{||d_{\text{ref}}||_2^2}{||d_{\text{ref}} - \hat{d}||_2^2} \right)$$

where $d_{\text{ref}}$ is the reference depth defined as the processed depth images for the highest acquisition time, $\hat{d}$ is the restored depth and $||x||_2$ denotes the $l_2$ norm given by $\sum x^2$.

Table 1(d) shows the RSNR values calculated for the cross-correlation and the RDI-TV algorithm for an acquisition times per pixel varying between 5.3 ms – 13.75 µs. As expected, the RSNR reduced as the acquisition time is shortened for both cross-correlation and RDI-TV approaches. For acquisition time less than 1 ms, the RSNR is considerably higher for RDI-TV than for the cross-correlation and clearly demonstrates the benefits of use of this algorithm for accurate image recovery from sparse data sets. In terms of the quality of the raw data, Table 1(c) shows signal-to-background ratio (SBR = $N_S/N_B$), where $N_S$ and $N_B$ denote the number of counts associated with the signal and background, respectively [8]. As expected, this ratio decreases with reduced acquisition time and becomes less precise for acquisition times per pixel lower than 55 µs.

The improved image quality at very low acquisition times possible with RDI-TV comes at an expense of the processing time. Table 1(e) shows the processing time of the image (comprising 50 × 50 pixels) when considering the cross-correlation and the RDI-TV methods. These values were obtained by using MATLAB R2015a on a computer with Intel(R) Core(TM) i7- 4790 CPU@3.60GHz and 32GB RAM. For the cross-correlation method, the total processing time does not depend on the acquisition time per pixel and is less than 0.2 s. In the case of the RDI-TV, the total processing time increases for a decreasing acquisition time per pixel since there are more pixels to reconstruct. The time required to process the image using the RDI-TV technique is ~10 s for 5.3 ms per-pixel acquisition time and
increases four-fold up to approximately 39 s for 13.75 µs pixel acquisition time. Note that the processing times can be improved by optimizing the current code and by using software tools such as “C” instead of MATLAB. Table 1(b) lists the average number of photons per pixel for each acquisition time. This average is calculated by evaluating the number of photons in an individual pixel for a time window of 5.6ns centered on the expected photon return time. Using this approach, the number of photons in each cross-correlation is then summed for all pixels and divided by the number of pixels to provide an average number of photons per pixel. For acquisition times of 100 µs or less (i.e. corresponding to less than 250 ms for the total 50 × 50 frame acquisition time) the average number of photons per pixel is less than 1. This shows that the RDI-TV algorithm provides significant potential for a complex scene reconstruction when operating in the sparse photon regime.

5. Conclusions

We have demonstrated kilometer-range high-resolution three-dimensional imaging using time-of-flight time-correlated single-photon counting. High-resolution three-dimensional images of various types of targets acquired over ranges between 800 meters and 10.5 km show that long-range data acquisition is feasible in a practical system and that the three-dimensional images generated show potential for target recognition and identification. It was also shown that by use of a total variation restoration optimization algorithm the acquisition time necessary for each pixel could be reduced by a factor of ten compared to a pixel-wise image processing approach. The total variation restoration algorithm has shown promising results in reconstructing images using data with much less than one photon per pixel. This could facilitate the transition of the system into a deployed lidar where imaging from fast moving platforms is required. Although not described in this paper, this optical system can be re-configured for operation with a 32 × 32 Geiger mode array. By incorporating an interchangeable lens, the two configurations were designed to provide identical pixel resolution for both the single-element systems and the Gm-array configurations in order to permit a performance comparison to be conducted, which is a subject of our future work. Furthermore, the full potential of the different restoration algorithms [24, 26, 28] would be realized when applied to Geiger-mode (Gm) array data, which is the subject of future work.

Funding

UK Engineering and Physical Sciences Research Council awards: EP/N003446/1, EP/M01326X/1, EP/K015338/1 and EP/M006514/1.

Acknowledgments

Agata M. Pawlikowska acknowledges EPSRC for support via the Engineering Doctorate Centre in Optics and Photonics Technologies.