1. DETAILS OF THE EXPERIMENTAL SETUP

The 355 nm pump laser used in this work is VisUV-355-HP (Picoquant GmbH) with 27 mW average output power at a repetition frequency of 20 MHz. The non-linear crystal cut for type-I SPDC is a β-Barium Borate crystal of size $5 \times 5 \times 1$ mm with a half opening angle of 3 degrees (Newlight Photonics). A 710 nm bandpass filter is placed after the BBO crystal to filter out the spurious pump light. The 780 nm classical laser (PiL079XSM, ALS GmbH) running at a repetition rate of 20 MHz with 40 ps (FWHM) pulse width is coupled with a fiber and then connected to a collimator, the output power of which can be finely tuned. A diffuser (ED1-C50-MD, Thorlabs) is used to create even illumination of the classical light over the object. The 355 nm pump laser always operates as a master to trigger only the camera in the asynchronous measurement, and both camera and the 780 nm laser in the synchronous experiment. 

The camera used in our study is the SwissSPAD2 with microlens on chip. It is composed of 512 × 512 pixels with a pitch of 16.38 µm, a native fill factor of 10.5% and photon detection probability (PDP) of approximately 25% at 700 nm. The camera runs in time-gating mode by scanning gate windows ($\sim 15$ ns wide) continuously with a time step of 18 ps. The starting gate position is tuned to be prior to the first synchronous light reflection by adding an appropriate initial offset to the laser trigger. There is no initial calibration approach implemented in this work to compensate the time-arrival skew from electrical signal (cable length, trigger circuit) and optical signal (fiber) as we only focus on the relative range of different objects. At each gate window, the number of $N$ 8-bit frames are transferred to the computer by USB3 connection and photon correlations are processed on a GPU before the gate shifts to the next successive position to run the same operation. Each 8-bit frame is accumulated by 255 successive 1-bit measurements with 350 ns exposure time. The overall acquisition speed is 370 fps (with 10.2 µs readout time for each bit) and the post-processing time is less than 1 ms per 8-bit frame when running on a GPU. The central $150 \times 150$ pixels are selected for all experiments to minimize the skew influence due to the electrical propagation across the SPAD array. The gate control signal injected to the middle of the pixel array resulted a symmetrical time propagation to the right and left pixels. To remove the hot pixels, we define a threshold at 200 dark counts and set all pixel values above this threshold in each frame to 0, as described in [1]. Then the hot pixels are smoothed by its neighboring pixels [2] for all intensity images shown in the work. The crosstalk effects are also removed by setting the correlation values from direct neighbour pixels to 0. The brightness of the photon pairs source is calculated as $(2.3/4.225%)/(350\text{ns} \times 255) = 610k$ photon pairs per second. These values are estimated from experimental acquisitions performed after by replacing the object by a mirror. In this sense, 2.3 pairs of correlated photons is the maximum number that the camera can detect per 8-bit frame.

2. SPATIAL CORRELATION IMAGE CALCULATION

The photon coincidence processing model used in this study is detailed in [3], in which the spatial joint probability distribution (JPD) $\Gamma(r_i,r_j)$ of entangled photon pairs is measured by multiplying values measured at pixel $i$ and $j$ (in position $r_i$ and $r_j$) between two successive frames:

$$\Gamma(r_i,r_j) = \frac{1}{N} \sum_{\ell=1}^{N} [I_\ell(r_i)I_\ell(r_j) - I_\ell(r_i)I_{\ell-1}(r_j)],$$

where $N$ is the acquired number of frames at each gate position. $I_\ell(r_i)$ and $I_\ell(r_j)$ represent the photon-count value at any pixel $i$ and $j$ (in position $r_i$ and $r_j$) of the $\ell$th frame ($\ell \in [1;N]$). The genuine coincidences that only originates from correlations between entangled photon pairs are
obtained by removing the accidental coincidences resulting from dark counts, after-pulsing, hot pixels, crosstalk, detection of multiple photon pairs and stray light. Both the (i) spatially-averaged correlation and the (ii) spatially-resolved correlation image used in our study can be extracted from the JPD:

(i) The spatially-averaged correlation image (noted $\Gamma_+$) is calculated from JPD $\Gamma$ using the formula:

$$\Gamma_+(r_1 + r_2) = \sum_r \Gamma(r_1 + r_2 - r, r).$$ (S2)

This represents the average number of photon coincidences detected between all pairs of pixels $r_1$ and $r_2$ separated by a distance $r_1 + r_2$.

(ii) The spatially-resolved correlation image is defined as the anti-diagonal component of the JPD $\Gamma(r, -r)$. It represents the number of photon coincidences detected between symmetric pair of pixels.

3. DETAILS ON DATA PROCESSING

The intensity images shown in this work are pixel-wise summed by the acquired number $N$ frames. The value of hot pixels is set to be 0 for correlation calculation and then interpolated with the neighboring pixels to show a better intensity image. The spatially-averaged correlation image shown in the context is cropped from the full $300 \times 300$ data for better visualization purpose. The results shown in Fig. S1 are the full spatially correlation data corresponding to the 4 selected gate positions in Fig. 2. In Fig. S1c and d, the correlation peaks are well visible, while the background fluctuation in Fig. S1d is smaller as no classical light is collected at that gate position [4].

![Fig. S1. Full spatially-averaged correlation images. (a), (b), (c) and (d) The full 300 × 300 spatially-averaged correlation images at 4 gate positions corresponding to those in Fig. 2. The red curves on the walls are the projections of the data across the correlation peak.](image)

The time arrival information of the quantum illuminated object can be located by finding the falling edge of the correlation peak profile. In order to locate the middle point at the falling edge, we fitted the fluctuated correlation peak data by the the error function $\text{erf}(.)$ [5]. From the fitted curves shown in Fig. S2a and b, we obtained the falling time of 0.882 ns for the synchronous case and 0.864 ns for the asynchronous case (from 90% to 10%). The tiny variation (1 gate position) also proves the reliability of the correlation peak profile used for object ranging. As the falling edge is not perfectly sharp, we recovered the intensity images of the classical object and quantum object by subtracting the intensity image at the gate position 45 gates before the middle point from the one 45 gates after. This allows us to achieve a subtracted intensity image with better contrast by avoiding using the images at the falling edge. However, the falling edge profile can be improved by optimizing the electrical gate shape of the camera and decreasing the width of the laser pulse.

4. CORRELATION PEAK SNR ANALYSIS

To analyse the visibility of the correlation peak in the spatially-averaged correlation image, we calculate the single-to-noise ratio (SNR) defined as the correlation peak value divided by the standard deviation of the background noise surrounding it. Fig. S3 shows the SNR of the corresponding result in Fig. 2 and the average intensity plotted as reference. We observe that the SNR increases above 1 when quantum light arrives at the camera 11 ns (classical light already present). In addition, the SNR is further improved when the classical light disappears at 17 ns because the background noise decreases.

In order to further evaluate the visibility of the correlation peak in the spatially-average correlation image, we measure the correlation peak SNR for various classical light over quantum
light intensity ratios ($I_c/I_q$) by keeping the pump laser power constant and tuning the classical laser power. The experiments are performed with the same camera configurations as in Fig. 4. As shown in Fig. S4a and b, the superimposed images are acquired at 13.5 ns gate position where the gate window captures both classical light and quantum light. It is difficult to distinguish the traffic stop sign when $I_c/I_q$ is high (3.48) in the superimposed image, while it is better resolved after subtracting the classical light acquired at gate position of 24.66 ns. From the spatially-averaged correlation image we can see that the background is more noisy when classical light intensity is higher as shown in the profiles across the center. Fig. S4c shows the measured correlation peak SNR values with various $I_c/I_q$ and the fitting curve with the theoretical model described in [4]. All the results are based on the measurements with fixed $N = 3000$ 8-bit frames.

A. Improving the quantum LiDAR acquisition speed

A.1. Reducing the number of frames

In Fig. 2, 5000 frames are acquired in 13.5 s at each gate position to measure a spatially-averaged correlation image and identify the peak with a SNR on the order of 30. As shown in Figure S5, the peak is still well visible (SNR on the order of 6) if the number of acquired frame per gate is reduced to 300, which strongly lower the acquisition time down to 0.81 s per gate.

A.2. Using a correlation-driven algorithm

To further improve the quantum LiDAR speed, we developed a coincidence-driven algorithm inspired by the binary search process used in successive-approximation register (SAR) analog-to-digital converter. To cover the range of the object, the scanning time range should be larger than the laser pulse period. Here, we use 2800 gate positions corresponding to 50.4 ns. As shown in Fig. S6a, we initially scan 3 gate positions with 700 gates (12.6 ns) interval dividing the scanning range to 4 parts equally. As the width of the gate window is 15.066 ns, we can make sure that at least one of the initially scanned gates can capture the reflected quantum light pulse, thus a
Fig. S4. SNR analysis with various classical over quantum light intensity ratios. (a) The intensity images and spatially-averaged correlation images with $I_c/I_q = 0.75$. (b) The intensity images and spatially-averaged correlation images with $I_c/I_q = 3.48$. The insets in the correlation image are the central $9 \times 9$ data and the curves are the vertical profiles across the center. (c) More measured correlation SNR values with different $I_c/I_q$ ratios and the fitted curve with a theoretical model.

Fig. S5. Results with reduced number of frames per gate. The average intensity and the correlation peak over 1500 gate positions are measured corresponding to the scenario in Fig. 2. At each gate position, $N = 300$ frames are acquired with an acquisition time of 0.81 seconds. The fitted curve of the correlation peak falling edge is shown with green line.
Fig. S6. Correlation-driven scanning approach and the result. (a) The correlation-driven scanning principle. (b) Only 8 scanning gate positions (300 frames acquired per gate) are required to locate the range of the object, which costs about 6.56 seconds. (c) The spatially-averaged correlation image at scanning position 2 in (b). The inset is the central $9 \times 9$ data. (d) shows the subtracted intensity image and the arrival time of the reflected quantum light from the object, 25.524 ns.

corresponding higher correlation peak will be obtained. Since the falling edge is not perfectly sharp, a following scanning process is implemented to check if the gate position with higher correlation peak is from the falling edge. To avoid such false locating, we scanned the gate position 45 gates before and 45 gates after the target position from the last scanning. The target range defined for the next scanning is between the last gate position with higher correlation peak and the scanned gate position just after it. When the target range is narrow enough the fitting method can be applied to achieve the falling edge according to the discrete scanned points. An example is depicted in Fig. S6b, where only 6.56 seconds is consumed with 8 scanning points to locate the range of the object. Fig. S6c shows the projected coincidences at scanning point 2, in which the coincidence peak is obvious. Fig. S6d is the subtracted intensity image and the measured relative range is 25.524 ns. Note that the initial offset for measurements in Fig. 2 and Fig. S6 are different resulting different relative time ranges.

By reducing the number of acquired frames to 300 and using a correlation-driven algorithm, we retrieve the quantum-illuminated object and its depth in 7 seconds. In the current implementation, it is not necessary to consider the gating profile variation over different pixels since the object is based on a 2 dimensional mask, which simplifies the processing and makes the proposed algorithm effective. However, the correlation-driven algorithm can be also extended to quantum LiDAR applications with 3 dimensional objects by applying in-pixel successive approximation. The similar approach has been implemented for conventional TCSPC-based LiDAR to reduce the output bandwidth [6].

Fig. S7. Different synchronous scenarios. (a) The response profiles of classical light and quantum light are separated. (b) The response profiles of classical light and quantum light are overlap, while the classical light arrives into the camera first similar to the scenario in Fig. 2. (c) The response profiles of classical light and quantum light are overlap, while the quantum light arrives into the camera first. (d) The two profiles are entirely overlap or the distance is smaller than the depth resolution of the camera, which can be only distinguished with spatially-resolved correlation image as shown in Fig. 3.
B. Analysis of different scenarios

According to the different arrival time of the reflected photons from the classical light and quantum light, different scenarios are shown in Fig. S7. For the first three scenarios shown in Fig. S7a, b and c, the target quantum object can be located by the falling edge of the correlation peak profile and subtracted from the intensity image after the falling edge no matter it is background light (in a and b) or classical light (c). However, if the distance between the two classical and quantum objects are smaller than the depth resolution of the camera, the spatially-resolved correlation image has to be performed for distillation, which takes more time.

REFERENCES

1. H. Defienne, J. Zhao, E. Charbon, and D. Faccio, “Full-field quantum imaging with a single-photon avalanche diode camera,” Phys. Rev. A 103, 042608 (2021).
2. 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, 1–8 (2016).
3. H. Defienne, M. Reichert, and J. W. Fleischer, “General model of photon-pair detection with an image sensor,” Phys. review letters 120, 203604 (2018).
4. H. Defienne, M. Reichert, J. W. Fleischer, and D. Faccio, “Quantum image distillation,” Sci. advances 5, eaax0307 (2019).
5. S. Chan, A. Halimi, F. Zhu, I. Gyongy, R. K. Henderson, R. Bowman, S. McLaughlin, G. S. Buller, and J. Leach, “Long-range depth imaging using a single-photon detector array and non-local data fusion,” Sci. reports 9, 1–10 (2019).
6. B. Kim, S. Park, J.-H. Chun, J. Choi, and S.-J. Kim, “7.2 a 48 × 4013.5 mm depth resolution flash lidar sensor with in-pixel zoom histogramming time-to-digital converter,” in 2021 IEEE International Solid-State Circuits Conference (ISSCC), vol. 64 (IEEE, 2021), pp. 108–110.