

THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Detection and tracking of moving objects hidden from view

Citation for published version:
Gariepy, G, Tonolini, F, Henderson, R, Leach, J & Faccio, D 2016, 'Detection and tracking of moving objects hidden from view', *Nature Photonics*, vol. 10, pp. 23-26. https://doi.org/10.1038/nphoton.2015.234

Digital Object Identifier (DOI):
10.1038/nphoton.2015.234

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Nature Photonics

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer 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
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Download date: 26. Feb. 2026

Detection and tracking of moving objects hidden from view

Genevieve Gariepy,[1*] Francesco Tonolini,[1] Robert Henderson,[2]
Jonathan Leach[1] and Daniele Faccio[1*]

[1]Institute of Photonics and Quantum Sciences, Heriot-Watt University,
David Brewster Building, Edinburgh, EH14 4AS, UK
[2]Institute for Micro and Nano Systems, University of Edinburgh
Alexander Crum Brown Road, Edinburgh EH9 3FF, UK

*To whom correspondence should be addressed;
E-mail: genevieve2.gariepy@gmail.com; d.faccio@hw.ac.uk.

The ability to detect motion and track a moving object hidden around a corner or behind a wall provides a crucial advantage when physically going around the obstacle is impossible or dangerous. Previous methods have demonstrated that is possible to reconstruct the shape of an object hidden from view. However, these methods do not enable the tracking of movement in real-time. We demonstrate a compact non-line-of-sight laser ranging technology that relies upon the ability to send light around an obstacle using a scattering floor and to detect the return signal from a hidden object with only a few seconds acquisition time. By detecting this signal with a single-photon avalanche diode (SPAD) camera, we follow the movement of an object located a meter away from the camera with centimetre precision. We discuss the possibility of applying this technology to a variety of real-life situations in the

near future.

Recent years have seen remarkable advances in the field of image processing and data acquisition, allowing for a range of novel applications [1–8]. An exciting new avenue is using optical imaging techniques to observe and track objects that are both in movement and hidden from the direct line-of-sight. The ability to detect motion and track a moving object hidden from view would provide a crucial advantage when physically going around the obstacle is impossible or dangerous, for example to detect a person moving behind a wall or a car approaching from behind a blind corner.

Techniques for imaging *static* objects that are hidden from view have been recently demonstrated relying on, for example, radar technology [9, 10], variations of laser illuminated detection and ranging (LIDAR) [3, 5, 11, 12], or speckle-based imaging. The latter approach was first developed for imaging through opaque barriers [13–15], and also allows for imaging around corners [16, 17]. The work of Velten *et al.* [5] and, more recently, Buttafava *et al.* [8] sets out to establish the 3D shape of a static hidden object by collecting the return scattered light with a streak camera or single-photon avalanche diode, respectively. While remarkable 3D reconstruction of objects are achieved with these techniques, Buttafava *et al.* point out that the requirement for scanning and subsequent long acquisition times mean that their technique is currently unsuitable for imaging moving objects.

Notwithstanding these ingenious imaging systems, locating the position of a hidden object in motion and monitoring its movement in real time remains to date a major challenge. We set out to solve the tracking problem and develop a technique based on both hardware and software implementations that are specifically designed for this

purpose. Our solution is based on a LIDAR-like approach where a single-photon avalanche diode (SPAD) camera [7, 18–22] is used to image light that is backscattered from beyond the direct line-of-sight (see Methods for camera details). The high temporal resolution of the camera relies on the fact that each individual pixel is operated in time-correlated single-photon counting (TCPSC) mode [21, 22], and measures the arrival times of single photons with a 45.5 ps time bin. The high sensitivity of the camera allows extremely short acquisition times, which in turn allows one to locate hidden objects on time scales sufficiently short to be able to track their movement. In the following, we first show that we can locate the position of an object hidden behind a wall with centimeter precision, without the need for pre-acquiring a background in the absence of the object. We then show that real-time acquisition is possible for an object moving at a few centimeters per second.

The proof-of-principle experiments were performed in the laboratory as shown in Fig. 1. The "target" we wish to track is a human form cut in a piece of foam that is 30 cm high, 10 cm wide and 4 cm thick. The target is positioned roughly one meter away from the camera and is hidden from its view by a wall (Fig. 1a,b). As in many real-life situations, there is not always a conveniently placed wall, door or window that can be used as a reflective surface to send and collect light, so we rely only on the presence of the floor, in this case, a piece of white cardboard. The camera is therefore imaging a patch of the floor that is just beyond the edge of the obscuring wall. We then send a train of femtosecond laser pulses on the floor, 15 cm to the left of the field of view of the camera. Light scatters from this point into a spherical wave and propagates behind the wall (see Fig. 1c). Part of the scattered light reaches the hidden target, which then scatters light back into the field of view, approximately as a single spherical wave (see Methods).

In Fig. 1c, we show an example of a single time frame of such a recording (raw acquired data) where the backscattered spherical wave is clearly visible, frozen in time by the camera as it sweeps across the field of view. The target location is then retrieved by utilising the fact that: (i) the time it takes for the light to propagate from the laser to the object and back, similarly to a LIDAR system, gives information about the object's distance and (ii) the curvature and direction with which the spherical wavefront propagates across the camera field-of-view provides information on the target position.

Target position retrieval. The target-position retrieval algorithm therefore relies on both the temporal and spatial information recorded by the SPAD camera. Every pixel i of the 32x32-pixel camera, corresponding to a position $\vec{r}_i = (x_i, y_i)$ in the field of view, records a histogram of photon arrival times (see Fig. 2a). First, we isolate the signal of interest coming from the target alone from the signal coming from unwanted sources in the environment such as the walls and the ceiling. This can be achieved by simply acquiring a background signal in the absence of the target, but is not a practical solution if we are interested in tracking non-cooperative moving targets. Instead, by acquiring data with the target at different positions, we can distinguish the signal that is not changing at each acquisition (generated by the static sources) and the signal that is changing (generated by the target). A median of the temporal histograms for each pixel proves to be a very a good approximation of the background signal [23–25] and allows to effectively isolate the signal generated from the target alone (see Supplementary Information).

Once the target signal is isolated, we proceed similarly to standard time-of-flight measurements and fit a Gaussian function to the temporal histograms, as shown in Fig. 2b [4, 26–28]. For each pixel i, the peak position of the Gaussian fit $\langle t \rangle_i$ is a measure of the

light travel-time from the moment the laser hits the ground, scatters to an object at a point and scatters back to the specific point in the field of view of the camera. There is an ensemble of locations $\vec{r}_o$ that satisfy the condition of equal propagation time, thus forming an ellipse on a plane parallel to the floor defined by the target's height. This ellipse represents a probability distribution for the position of the hidden object: Fig. 2d shows as an example the probabilities calculated from experimental data, correponding to four different pixels indicated in the figure. In order to retrieve the target's position, we then calculate the joint probability density by multiplying the probability densities from all 1024 camera pixels (see Methods for more details).

Results. In a first experiment, we place the target at eight distinct positions and acquire data for three seconds at each position. Using the algorithm detailed above, we retrieve a probability density $P(\vec{r}_o)$ for the eight positions of the target. Figure 3 shows, to scale, the relative positions of the laser illumination spot on the floor, the camera and its field of view, together with the actual positions of the target superimposed on the joint probability distributions in color scale. The method provides an accurate retrieval of the target's position with a precision of approximately ± 1 cm in x and ± 2 cm in y, corresponding to $\sim 20\%$ uncertainty with respect to the target's size in both directions. This precision depends on the target's distance and its position in x and y, due to geometrical arguments (see Supplementary Information). The precision obtained experimentally corresponds to what is expected from simulations. Our results also show that we are able to retrieve target positions when it is not only hidden from view but is actually physically receded behind the end of the wall.

We then placed the target on a moving track directed along the y-direction. Although the target is moving continuously, we record one position every three seconds and we thus retrieve a discrete set of locations, each of which represents the average position of the target during the acquisition time. Figure 4 shows an example of the system tracking a target moving at 2.8 cm/s, where we show a few of the retrieved positions as well as the full evolution of the hidden target's motion, where the maximum probability position is illustrated as a function of time. We see that the motion that we retrieve with real-time acquisition corresponds to the actual target's motion. Data was also recorded at different speeds and for other x positions of the track (see Supplementary Information). The supplementary video clearly shows a difference in the target's speed, highlighting the ability of the camera system to capture with reasonable accuracy both position and speed of a moving target.

The detection range of this technique is ultimately limited by the signal-to-noise ratio (SNR) of the detected signal. We verified from the experimental data that, for a fixed distance between the camera and the field of view, the recorded signal decays as $1/(|r_o - r_l||r_o - r_i|)^2$, where $|r_o - r_l|$ is the distance between the laser spot and the object, and $|r_o - r_i|$ the distance between the object and the field of view. We are currently detecting targets to a range of $|r_0 - r_i| \sim |r_o - r_l| \sim 60$ cm. With the increase of the target's size and possible improvements made to the setup and detection hardware (e.g. increased detector fill factor and optimised wavelength sensitivity), we expect to extend this range to detecting targets at about 10 m distance. A comprehensive analysis of the factors affecting the SNR is presented in the Supplementary Information.

Extending the scope of our current work to include multiple hidden objects is also of considerable interest. We performed preliminary measurements in which we detect two

hidden targets (see the Supplementary Information). The ability of the current technology to track multiple objects is determined by our capacity to distinguish signals from distinct targets. Precise tracking of multiple targets would be enhanced by some relatively straightforward solutions such as increasing the field of view of the system by using large-area arrayed detectors or decreasing the temporal response of the system. Large-format SPAD array cameras with these properties are in development.

SPAD detectors, originally developed as single pixel elements, are gradually becoming widely available as focal plane arrays. The single photon sensitivity and picosecond temporal resolution provide a unique tool for fundamental studies [7] and our results show that they can also enable real-time non-line-of-sight ranging of a moving target. An interesting avenue for future work is to combine these techniques with the simultaneous 3D reconstruction of the target shape. We have shown that we can reliably track the position of a target if it moves by less than its own size during the acquisition time. Our results thus pave the way for tracking hidden objects in real time in a number of real-life scenarios, such as surveillance, rescue missions and implementation in cars for detecting incoming hidden vehicles.

Methods

Laser details: The laser we use in our non-light-of-sight laser ranging system is an 800 nm wavelength femtosecond oscillator that emits pulses of 10 nJ energy and 10 fs duration at a 67 MHz repetition rate (0.67 W average power). A small portion of the laser (8% reflection) is sent to an optical constant fraction discriminator (OCF) that generates a TTL signal then sent to the camera to synchronise the acquisition to the propagation of the laser pulses. We note that the system has been tested with different laser specifications,

e.g. light-in-flight with the same SPAD camera was demonstrated using a portable micro-chip laser, with 4 kHz repetition rate [7].

Camera details: The camera is a 32x32-pixel array of Si CMOS single-photon avalanche diodes (SPAD) that are individually operated in time-correlated single-photon counting (TCSPC) mode: every time a photon is detected by a pixel, the time difference between its arrival and the arrival of the TTL trigger from the OCF is measured and stored in the time histogram. Each histogram has 1024 time pixels with a time-bin of 45.5 ps. The time resolution is limited by the electronic jitter of the system, which is $\sim$110 ps (measured at full-width-half-maximum). This impulse response corresponds to a spatial (depth) resolution of a 1.65 cm, i.e. of the same order of magnitude of our target, allowing us to approximate the back scattering as a single spherical wave originating from the target. A standard Nikon-mount lens is attached to the camera (Samyang, 8 mm focal length, F3.5).

The histograms are recorded over 10,000 laser pulses and the camera is operated at its minimum operating exposure time of 300 μs, so that each acquisition takes 3 seconds. The operating frame rate is only limited by the camera's USB connection to the computer, so that the minimum exposure time is currently 300 μs (frame rate of 3 kHz). The next generation of SPAD camera will be implemented with USB3.0 which will allow to reach higher operating rates, up to the limit of 1 MHz set by the camera's internal functioning. The limit at which data can be acquired will then be set by the amount of signal scattered back to the field of view of the camera, but we expect to be able to record data for one position within less than a second.

Retrieval details: For a pixel i, the peak position of the Gaussian fit $\langle t \rangle_i$ is used to determine the total photon flight time, with an uncertainty that is taken to be the Gaussian standard deviation σ_{t_i}. The time $\langle t \rangle_i$ is a measure of the light travel-time from the moment the laser hits the ground, scatters to an object at a point $\vec{r}_o = (x_o, y_o)$ and scatters back to the specific point $\vec{r}_i$ in the field of view of the camera. There is a an ensemble of locations $\vec{r}_o$ that satisfy this condition, forming a three-dimensional ellipsoid which collapses to a two-dimensional ellipse on a plane parallel to the floor defined by the target's height, where we restrict our search (see Supplementary Information). This ellipse is defined by $|\vec{r}_o - \vec{r}_l| + |\vec{r}_o - \vec{r}_i| = \langle t \rangle_i \times c$, where $|\vec{r}_o - \vec{r}_l|$ and $|\vec{r}_o - \vec{r}_i|$ are the distances from the laser point $\vec{r}_l$ on the floor to the target and from the target to the point $\vec{r}_i$, respectively, as illustrated in Fig. 2c. This ellipse represents a probability distribution for the position of the hidden object with uncertainty σ_{t_i}:

$$P_i^{\text{ellipse}}(\vec{r}_o) \propto \exp\left[-\frac{(\varepsilon/c - \langle t \rangle_i)^2}{2\sigma_{t_i}^2}\right] \tag{1}$$

where ε is the ellipsoidal coordinate $\varepsilon = |\vec{r}_o - \vec{r}_l| + |\vec{r}_o - \vec{r}_i|$. We therefore calculate the probability distributions $P_i^{\text{ellipse}}(\vec{r}_o)$ for every pixel i of the field of view. Figure 2d shows as an example four of the $P_i^{\text{ellipse}}(\vec{r}_o)$ probabilities calculated from experimental data, correponding to four different pixels indicated in the figure. In order to retrieve the target's position, we calculate the joint probability density by multiplying the probability densities from all 1024 camera pixels:

$$P(\vec{r}_o) = N \prod_{i=1}^{1024} P_i(\vec{r}_o). \tag{2}$$

$P(\vec{r}_o)$ determines the overall probability distribution of the location of the target, and N is a normalisation constant. A complete mathematical development and details about the

form of $P_i(\vec{r}_o)$ are given in Supplementary Information.

Acknowledgments

We acknowledge support from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC GA 306559, the Engineering and Physical Sciences Research Council (EPSRC, UK, Grants EP/M006514/1, EP/M01326X/1, EP/ K03197X/1), and ST Microelectronics, Imaging Division, Edinburgh, for their support in the manufacture of the Megaframe chip. The Megaframe project has been supported by the European Community within the Sixth Framework Programme IST FET Open.

References

[1] Kirmani, A. *et al.* First-photon imaging. *Science* **343**, 58–61 (2014).

[2] Sun, B. *et al.* 3d computational imaging with single-pixel detectors. *Science* **340**, 844–847 (2013).

[3] Wandinger, U. *Introduction to lidar* (Springer, 2005).

[4] Massa, J. S., Wallace, A. M., Buller, G. S., Fancey, S. J. & Walker, A. C. Laser depth measurement based on time-correlated single-photon counting. *Opt. Lett.* **22**, 543–545 (1997).

[5] Velten, A. *et al.* Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nat Commun* **3**, 745 (2012).

[6] Gao, L., Liang, J., Li, C. & Wang, L. V. Single-shot compressed ultrafast photography at one hundred billion frames per second. *Nature* **516**, 7477 (2015).

[7] Gariepy, G. *et al.* Single-photon sensitive light-in-flight imaging. *Nat Commun* **6**, 6021 (2015).

[8] Buttafava, M., Zeman, J., Tosi, A., Eliceiri, K. & Velten, A. Non-line-of-sight imaging using a time-gated single photon avalanche diode. *Opt. Express* **23**, 20997 (2015).

[9] Sume, A. *et al.* Radar detection of moving objects around corners. In *Proc. SPIE*, 73080V–18 (2009).

[10] Chakraborty, B. *et al.* Multipath exploitationwith adaptivewaveform design for tracking in urban terrain. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, 3894–3897 (2010).

[11] Gupta, O., Willwacher, T., Velten, A., Veeraraghavan, A. & Raskar, R. Reconstruction of hidden 3d shapes using diffuse reflections. *Opt. Express* **20**, 19096–19108 (2012).

[12] Repasi, E. *et al.* Advanced short-wavelength infrared range-gated imaging for ground applications in monostatic and bistatic configurations. *Appl. Opt.* **48**, 5956–5969 (2009).

[13] Vellekoop, I. M. & Mosk, A. P. Universal optimal transmission of light through disordered materials. *Phys. Rev. Lett.* **101**, 120601 (2008).

[14] Mosk, A. P., Lagendijk, A., Lerosey, G. & Fink, M. Controlling waves in space and time for imaging and focusing in complex media. *Nat Photon* **6**, 283–292 (2012).

[15] Bertolotti, J. *et al.* Non-invasive imaging through opaque scattering layers. *Nature* **491**, 232–234 (2012).

[16] Katz, O., Small, E. & Silberberg, Y. Looking around corners and through thin turbid layers in real time with scattered incoherent light. *Nat Photon* **6**, 459553 (2012).

[17] Katz, O., Heidmann, P., Fink, M. & Gigan, S. Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations. *Nature Photonics* **8**, 784–790 (2014).

[18] Richardson, J. *et al.* A 32x32 50ps resolution 10 bit time to digital converter array in 130nm cmos for time correlated imaging. In *Custom Integrated Circuits Conference, 2009. CICC '09. IEEE*, 77–80 (2009).

[19] Richardson, J., Grant, L. & Henderson, R. Low dark count single-photon avalanche diode structure compatible with standard nanometer scale cmos technology. *Photonics Technology Letters, IEEE* **21**, 1020–1022 (2009).

[20] Niclass, C., Rochas, A., Besse, P.-A. & Charbon, E. Design and characterization of a cmos 3-d image sensor based on single photon avalanche diodes. *Solid-State Circuits, IEEE Journal of* **40**, 1847–1854 (2005).

[21] O'Connor, D. & Phillips, D. *Time-Correlated Single Photon Counting* (Academic Press, 1984).

[22] Becker, W. *Advanced time-correlated single photon counting techniques* (Springer, 2005).

[23] Anderson, B. D. & Moore, J. B. *Optimal Filtering* (Prentice-Hall, 1979).

[24] Cucchiara, R., Grana, C., Piccardi, M. & Prati, A. Detecting moving objects, ghosts, and shadows in video streams. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **25**, 1337–1342 (2003).

[25] Cutler, R. & Davis, L. View-based detection and analysis of periodic motion. *Pattern Recognition, International Conference on* **1**, 495 (1998).

[26] Buller, G. S. & Wallace, A. M. Ranging and three-dimensional imaging using time-correlated single-photon counting and point-by-point acquisition. *Selected Topics in Quantum Electronics, IEEE Journal of* **13**, 1006–1015 (2007).

[27] Ho, C. *et al.* Demonstration of literal three-dimensional imaging. *Appl. Opt.* **38**, 1833–1840 (1999).

[28] Albota, M. A. *et al.* Three-dimensional imaging laser radar with a photon-counting avalanche photodiode array and microchip laser. *Appl. Opt.* **41**, 7671–7678 (2002).

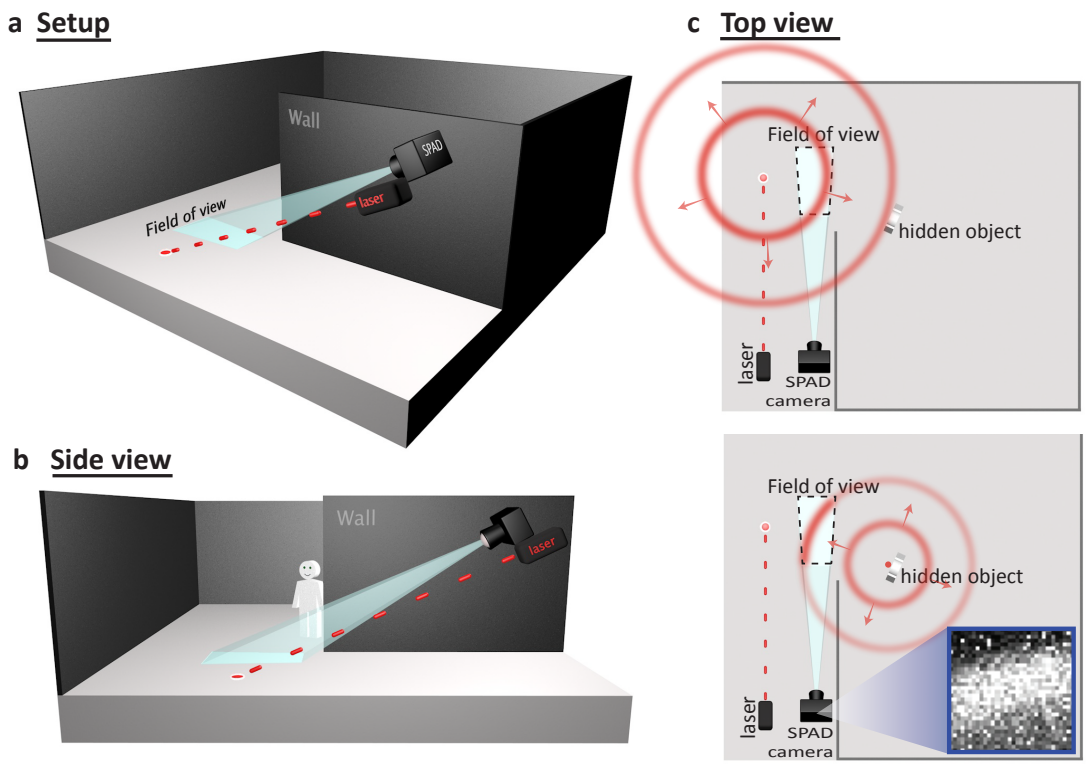


Figure 1: **Looking around a corner.** Our setup re-creates, at a $\sim$ 5x reduced scale, a situation where a person is hidden from view by a wall or an obstacle. a) The camera is positioned on the side of the wall and is looking down at the floor: it cannot see what is behind the wall but its field of view is placed beyond the end of the obstacle. b) A side view shows that the target is hidden behind the wall. In order to see the hidden target around the corner, laser pulses are sent to the floor. c) The light then scatters off the floor and propagates as a spherical wave behind the obstacle, reaching the hidden object. This light is then in turn scattered back into the field of view of the camera. The SPAD camera records both spatial and temporal information on the propagating spherical light wave as it passes through the field of view, creating an elliptical pattern where it intercepts the floor. An example of the spatially resolved raw data, as recorded by the camera for a fixed time frame as the ellipse passes in the field of view is shown in the inset.

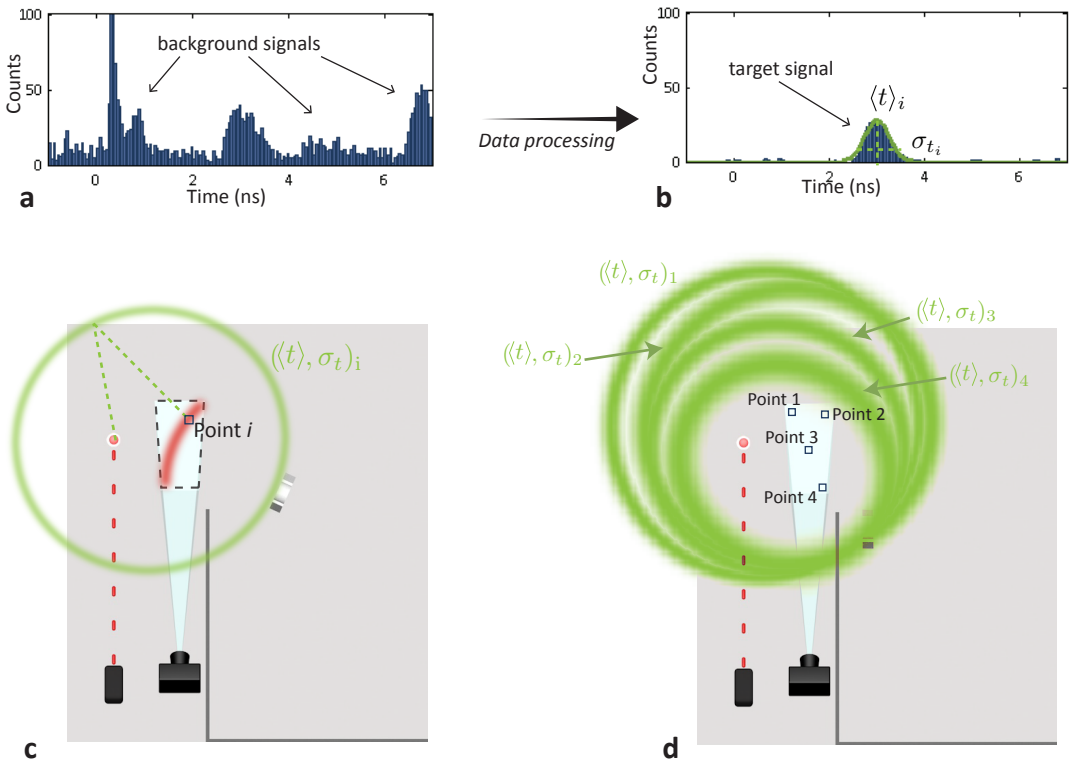


Figure 2: **Retrieving a hidden object's position.** a) A histogram of photons arrival times is recorded for every pixel (here for pixel i as indicated in c). This experimental histogram contains signals both from the target and unwanted background sources. b) Background subtraction and data processing allows to isolate the signal from the target and fit a Gaussian to its peak, centered at $\langle t \rangle_i$ with a standard deviation of σ_{t_i}. c) The time of arrival $\langle t \rangle_i$ is used to trace an ellipse of possible positions of the target which would lead to a signal at this time. d) Ellipses calculated from different pixel (experimental data) give slightly displaced probability distribution that intercepts at a given point. The area where the ellipses overlap indicates the region of highest probability for the target location. Multiplying these probability distributions (also with all other similar distributions from all 1024 pixels of the camera) provides an estimate of the target location.

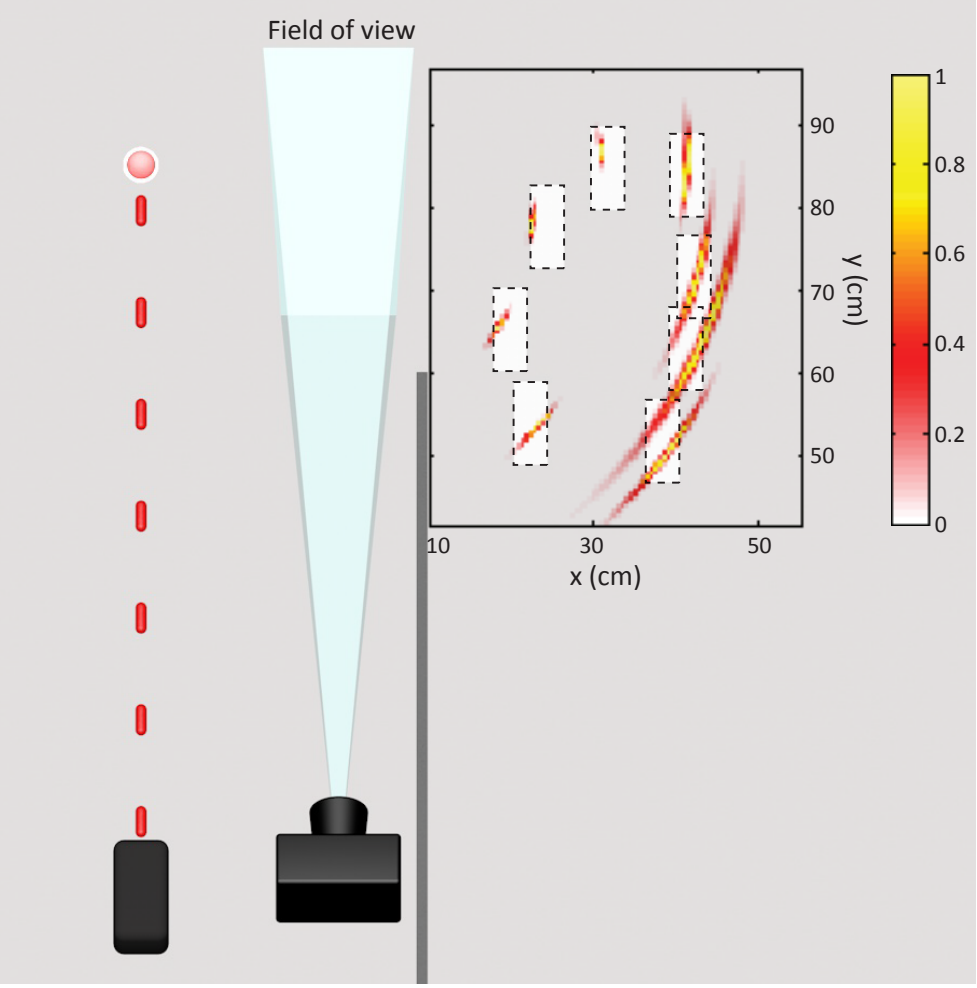


Figure 3: **Experimental results of hidden object's position retrieval.** Experimental layout and results showing the retrieved locations for eight distinct positions of the target, approximately one meter away from the camera (distances indicated in the figure are measured from the camera). The coloured areas in the graph indicate the joint probability distribution for the target location whose actual positions are shown as white rectangles. Each $P(\vec{r}_o)$ peak value is individually normalised to one.

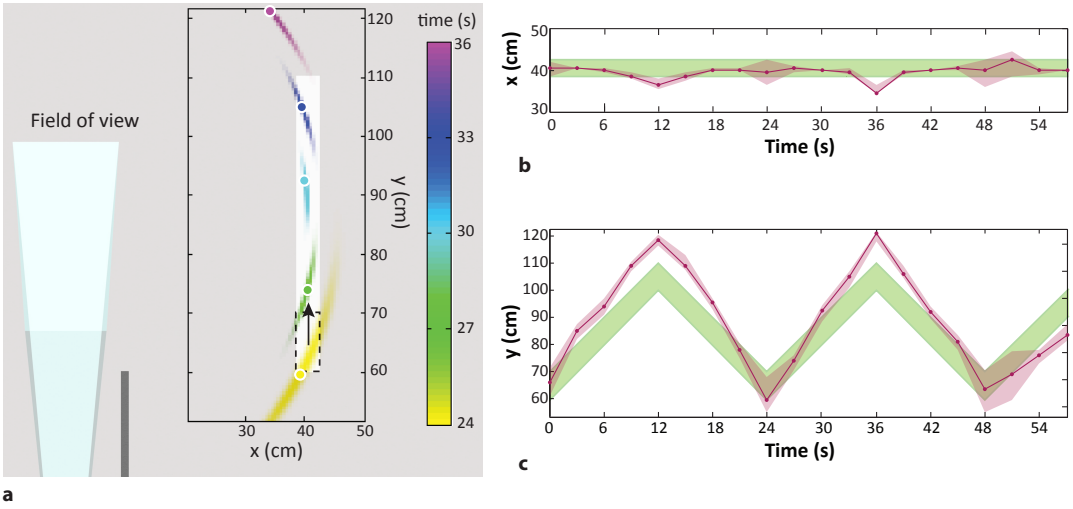


Figure 4: **Non-line-of-sight tracking of a moving target.** Distances in the graph are measured from the camera position. a) The object is moving in a straight line along the y-direction, from bottom to top (as represented by the dashed rectangle and the arrow), at a speed of 2.8 cm/s. The coloured areas represent the retrieved joint probability distributions: the point of highest probability, indicating the estimated target location, is highlighted with a filled circle. The colours correspond to different acquisition "start" times, as indicated in the colorbar: successive measurements are each separated by 3 second intervals, i.e. the data acquisition time as explained in the text. We show the retrieved positions in x and y as a function of time in b) and c), respectively. The dots in (b) and (c) show the points of maximum probability together with the 50% confidence bounds (red shaded area). The green area shows the actual position of the target.