Simultaneously Tracking and Imaging a Moving Object under Photon Crisis

Shuai Sun (孙帅),[1,2] Hong-Kang Hu (胡宏康)●,[1,2] Yao-Kun Xu (徐耀坤),[2,3] Yue-Gang Li (李月刚),[1,2] Hui-Zu Lin (林惠祖),[1,2] and Wei-Tao Liu (刘伟涛)●[1,2,*]

[1]*Department of Physics, College of Liberal Arts and Sciences, National University of Defense Technology, Changsha 410073, China*

[2]*Interdisciplinary Center of Quantum Information, National University of Defense Technology, Changsha, Hunan 410073, China*

[3]*College of Information and Communication, National University of Defense Technology, Xi'an, Shaanxi 710106, China*

In optical sensing, it is a fundamental principle that retrieving a certain quantity of spatial information requires a commensurate number of photons to be detected. However, on the issue of imaging moving objects under low photon flux, when the detectable photon flow is insufficient to recover the information flow formed by both motion and structural images of the object, photon crisis is induced and a more photon-efficient sensing method is urgently expected. Here we implement a simultaneous tracking and imaging method. The motion information and spatial information of a moving object are encoded by sequentially projecting the laterally shifting patterns and recording the results via a photomultiplier. Then both the trajectory and the image of the moving object are reconstructed gradually along with its evolution. As a photon-efficient encoding-decoding method is achieved, our method realizes capturing moving object with angular velocity larger than 45 mrad/s, and works well with the average number of detected photons even down to about 2.4 photons/speckle under single-laser-pulse illumination.

I. INTRODUCTION

In optical sensing, the image information is measured by detecting photon flux, yielding a fundamental and ultimate principle that the more spatial information that needs to be recovered, the more photons are required to be detected [1]. Subject to that, under low photon flux, a hard bound is physically imposed on the speed of photoelectric imaging processes, because a long integration time to collect sufficient photons is indispensable to record or computationally recover the commensurate spatial-resolving information [2,3].

On the issue of imaging a moving object, both motion information and structural image information vary with time and form an increasing information flow. When the photon flow arriving at the detection plane per unit time is insufficient to recover the information flow increased per unit time, a longer integrating detection time will not help to form any image of the moving object, but will only cause aliasing of the blurring information. That is, the detectable photon flux cannot afford the information flux to be recovered and photon crisis breaks out. In addition, as the object moves faster, the increase in the amount of information flow per unit time is larger, causing the

gap between the photon flux and the information flux to grow larger and the photon crisis to appear to be more serious. This fundamental, yet practical difficulty can be hardly alleviated by solely using a photoelectric detector with high detection efficiency. A more essential solution is to develop a photon-efficient optical sensing method, in which not only are the motion and the spatial information of the moving object encoded compressively, but also the information transmitted by the limited number of photons is extracted more efficiently.

Conventionally, a moving object can be captured by sequential images acquired by pixelated sensors at a different transient moment [4–7]. This optical sensing method requires the detectable photon flux to be commensurate with the spatial information of the object within each shutter duration of the sensor, which is required to be shorter subject to the object moving faster, to avoid motion blur. In some scenarios, the detected photon flux within a single shutter duration is rather low due to limited illumination power [8,9], or the low reflectivity of the moving object [10,11], and so on. To obtain a recognizable signal, array sensors with not only a high-speed response but also high sensitivity are concurrently necessary, which is still difficult or expensive to achieve. Furthermore, when the number of photons arriving at the sensor within single shutter duration is insufficient to form any image of the

_____

*wtliu@nudt.edu.cn

object [1], no matter how high detection efficiency or how low dark counts the multi-pixeled sensor features, such conventional methods will fail because multi-exposure will cause motion blur while the weak signal from a single exposure cannot provide a sufficiently high signal-to-noise ratio (SNR) for imaging or even only tracking. Capturing a moving object under photon crisis is essential to improve the operating range or reduce the required radiation dose in optical sensing.

Ghost imaging (GI), as well as single-pixel imaging (SPI), is an optical sensing method capable of reconstructing an image of the object with low photon flux [12–15]. However, sequential projection and the large number of required samplings make it slow [16–25], thus strictly limiting its performance for imaging moving objects. Despite many recent advances in improving the imaging speed [26–42], the spatial information reconstruction of existing GI and SPI is not photon efficient in the case of tracking and imaging a moving object, because the image is still expected to be reconstructed within the time interval that the objects can be treated as "frozen," which also requires detectable photon flux to be commensurate with the spatial image information. Thus, achieving tracking and imaging simultaneously under low photon flux remains an open problem.

In this paper, we demonstrate a simultaneous tracking and imaging scheme. Laterally shifting patterns are sequentially projected onto a translationally moving object, thus both its motion information and spatial information are encoded into the bucket signals. The velocity vector of the moving object is decoupled and calculated from temporal correlation of the bucket signals independently. Then, taking the motion information into account, both the image and trajectory of the moving object are recovered simultaneously along with its evolution. As the photon flux required in this method is just commensurate with the velocity information of the object, which is far less than the information of a clear spatial-resolving image [43–47], our scheme is valid when the angular speed of the moving object is larger than 45 mrad/s and works well with an average number of detected photons as low as about 2.4 photons/speckle under single-laser-pulse illumination. Under the same condition, no image can be recorded by a common CCD, thus we state that simultaneous tracking and imaging of a moving object is achieved via our scheme under photon crisis. Although a single translational object is demonstrated, this scheme provides an initial inspiration to develop technology to capture moving objects via an optical method under low photon flux.

II. THEORETICAL ANALYSIS

To avoid motion blur, the signal acquisition process in photoelectric imaging is required to be finished within the duration that the object can be treated as "frozen," which

also means neither the motion information nor the spatial image information are increased within that time interval. Accordingly, to reconstruct the image of a moving object from N samplings via an SPI system, the angular speed of the object $\omega(t)$ with respect to the imaging system is quantitatively limited by

$$\int_0^{N/f} \omega(t)\, dt \leq \theta, \tag{1}$$

with θ and f being the angular resolution and the sampling repetition frequency of the system, respectively. Considering that several thousands of samplings is barely sufficient to reconstruct an image, the displacement of the object during every individual sampling should be much smaller than the resolution, which strictly limits the speed of the moving object that can be captured. In addition, it also demands that the image of the object can be formed within a short time, which raises a requirement on the detectable photon flux which thus will fail in weak illumination cases.

In SPI system, sequential optical patterns are illuminated onto the object and the reflected light is collected by a single-pixel detector. In this way, the spatial information of the object is encoded into a single temporal waveform, which is called a bucket signal. In particular, when the object is moving at a constant velocity $\mathbf{v}_0$ under the illumination of a random pattern laterally shifting with $\mathbf{v}_i$ (Fig. 1), the bucket signals is described as

$$B(t) = \int T(\mathbf{r}_o) I(\mathbf{r}_o - \Delta\mathbf{v}t)\, d\mathbf{r}_o, \tag{2}$$
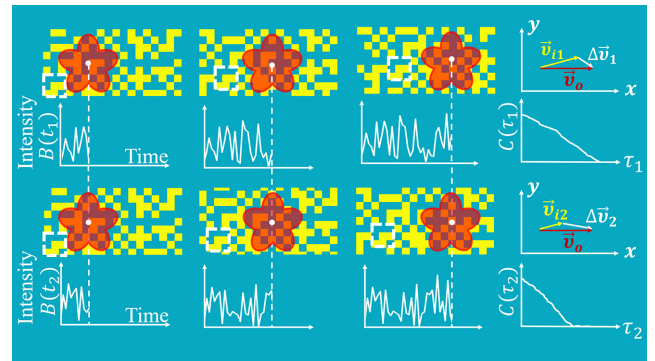

FIG. 1. Schematic of the velocity calculation process. The object is moving under the laterally shifting pattern (indicated by the white dotted box). Along with the motion of the object, the bucket signal $B(t)$ is recorded. The temporal correlation of the bucket signal $C(\tau)$ is directly related to the relative velocity $\Delta\mathbf{v}$ between the object and the shifting pattern, as shown by the right most column. Thus $\Delta\mathbf{v}$ can be calculated from the HWHM of $C(\tau)$. Then the velocity of object $\mathbf{v}_0$ can also be obtained according to the velocity of illumination pattern $\mathbf{v}_i$ and the calculated $\Delta\mathbf{v}$.

where $T(\mathbf{r}_o)$ is the reflectivity of the object, $I(\mathbf{r}_o)$ is the intensity distribution of the illumination pattern, and $\Delta\mathbf{v} = \mathbf{v}_0 - \mathbf{v}_i$. Suppose the structure of the shifting pattern remains the same, the degree of the correlation of the pattern projected within the object area will decrease with the relative displacement increasing (Fig. 1). In addition, the decorrelation time depends on the relative velocity. Consequently, the velocity of the object can be calculated by analyzing the temporal correlation of the bucket signals, which is expressed as

$$C(\tau) = \langle B(t)B(t+\tau)\rangle_{t_m} - \langle B(t)\rangle_{t_m}\langle B(t+\tau)\rangle_{t_m}, \quad (3)$$

where t_m is defined as velocity measurement interval, within which the structure of the shifting pattern keeps unchanged. Here $C(\tau)$ will decorrelate when the relative displacement is nearly equal to the size of the object in the motion direction, which is a constant for a certain translational object. If the velocity of the object can also be treated as constant within the measurement time, which is easy to achieve in practice, the half width at half maximum (HWHM) of $C(\tau)$ will be inversely proportional to the relative velocity as

$$|\Delta\mathbf{v}| = \frac{\epsilon}{\Delta\tau}, \quad (4)$$

where $\Delta\tau = \arg\frac{1}{2}C(0)$ is the HWHM of $C(\tau)$. Here $\epsilon \geq 0$ is a coefficient related to the autocorrelation of the reflectivity of the object (details in the Appendix). Considering both $|\Delta\mathbf{v}|$ and ϵ are unknown in Eq. (4), we set the illumination pattern shifting with speed $\mathbf{v}_{i1}$, $\mathbf{v}_{i2}$, and record two temporal waveforms of the bucket signal. Then two equations in the form of Eq. (4) are established and $|\Delta\mathbf{v}|$ can be calculated. To solve $\mathbf{v}_0$, the illumination shifting in two orthogonal directions need to be produced and the component values of the object's velocity in these two directions will be obtained separately, which means the velocity vector is calculated actually (details in the Appendix).

With knowledge of the calculated velocity, the spatial information of the object can be decoded and the image can be computationally reconstructed during its evolution. In addition, the illumination beam can also be steered accordingly to illuminate the moving object persistently, thus tracking and imaging are achieved simultaneously.

In our method, it is assumed that the velocity of the object can be treated as constant within t_m, which actually means

$$\int_0^{N'/f} \frac{\partial\omega(t)}{\partial t} \, dt \ll \omega(t), \quad (5)$$

with f being the sampling repetition rate and N' being the number of required samplings to obtain the velocity of the object. In practice, several hundreds of samplings

are sufficient for velocity calculation with considerable accuracy. For a GI or SPI system with sampling repetition rate of 10^5 Hz [28,29] or higher [30–32], the velocity measurement time t_m is of the order of a microsecond, within which the velocity can be taken as constant for most cases. Compared with Eq. (1), the speed of the object is no more limited or the requirement on sampling repetition rate and the relationship between the spatial resolution and the maximum displacement is relaxed tremendously. Furthermore, the image of the moving object is no longer required to be formed within the "frozen" moment, but can be gradually recovered during its evolution, which is a far longer time, thus the photon flux required to be detected is also largely reduced.

III. EXPERIMENTAL RESULTS

A. Tracking via temporal correlation of bucket results

Estimation of velocity and tracking is first verified in the experimental setup (Fig. 2). The object is moving in one dimension at a constant velocity $+4.00$ cm/s in $\mathbf{x}$ (angular speed with respect to the system is $+72.73$ mrad/s). Two
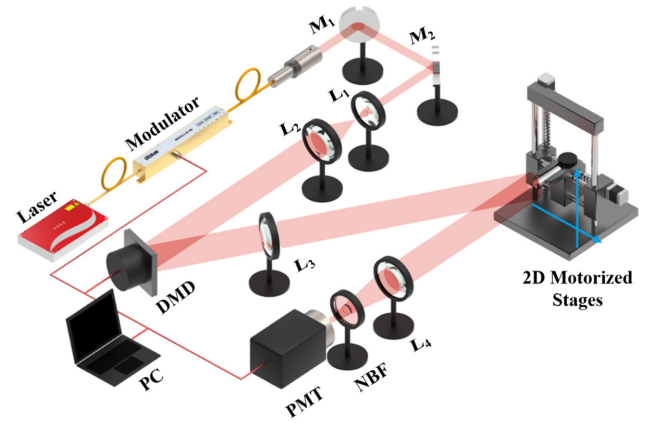

FIG. 2. Experimental illustration. A fiber laser of 1064 nm is cut into pulses of 20 ns by an intensity modulator (Photline NIR-MX-LN-10). Both M_1 and M_2 are mirrors. The beam is enlarged via a $4-f$ system consists of L_1 ($f = 83$ mm) and L_2 ($f = 500$ mm) to match the Digital Micromirrors Device (DMD, DLP 0.7 XGA). With a spherical lens L_3, the illumination pattern on DMD is projected onto the object, which is a pen cap with the end face driven by two-dimensional motorized stages (Zolix PSA 100-11-X). The size of the random pattern cell (bin 4×4 pixels) on the object plane, which is 55 cm away from L_3, is about 250.00 μm, thus the angular resolution of the system is 0.45 mrad. The size of the micromirror pixel on the object plane is about 62.5 μm, which determines the minimum shifting step of the illumination pattern. The bucket signals from the object are collected by L_4 ($f = 100$ mm) and detected by a photomultiplier tube (PMT), in front of which is located a narrow band filter (NBF). Both the intensity modulator and the PMT are triggered by the signal from the DMD.
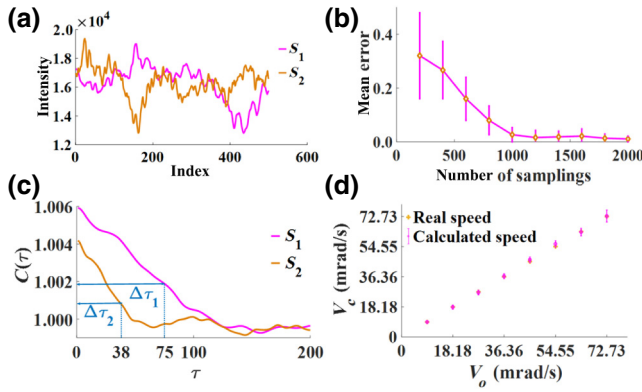
FIG. 3. Verification of the validity of velocity measurement via bucket signals. (a) Recorded bucket signals. (b) Temporal correlation of the bucket signals, the HWHM of which depends on the relative velocity between the object and the illumination pattern. (c) Mean error between calculated and real speed. (d) Real speed (V_o) and the calculated speeds (V_c) for each test when the object is moving at a different speed. For each velocity measurement, 1000 samplings are performed.

sequences of laterally shifting random patterns are produced and 500 samplings are performed for each sequence. The bucket signals S_1 and S_2, as depicted in Fig. 3(a), is obtained under the average velocity of the shifting pattern (refreshing frequency is 1.28 kHz) being $+145.45$ mrad/s and $+218.18$ mrad/s, respectively. The $C(\tau)$ of S_2 decorrelates about twice as fast as that of S_1, as shown in Fig. 3(b). Taking the HWHM of S_1 and S_2 ($\Delta\tau_1 = 75$, $\Delta\tau_2 = 38$) into account, the velocity is calculated as $+71.75$ mrad/s according to Eq. (4). Errors of the results are mainly caused by limited number of statistics, thus can be reduced by increasing the number of samplings. With the object moving at $+72.73$ mrad/s, the relative mean error between the calculated speed and the real value is shown in Fig. 3(c), for different numbers of samplings. Tracking of moving objects at a different speed is also tested with the same number of samplings. The results are shown in Fig. 3(d), with the stars showing the real velocity and the circles showing the calculated values.

B. Simultaneous tracking and imaging of a moving object

Then the object is set moving faster than 2.5 cm/s (45.45 mrad/s) in two dimensions for 2 seconds, as depicted by the orange line in Fig. 4(a). The top left corner is the static object and the moving object recorded by a CCD camera (not shown in Fig. 2) can be seen in Video 1 within the Supplemental Material [48]. During the object's motion, four series of shifting illumination patterns are generated and projected. For each velocity calculation, bucket signals from 2000 samplings (500 samplings for each series of shifting pattern) are collected.
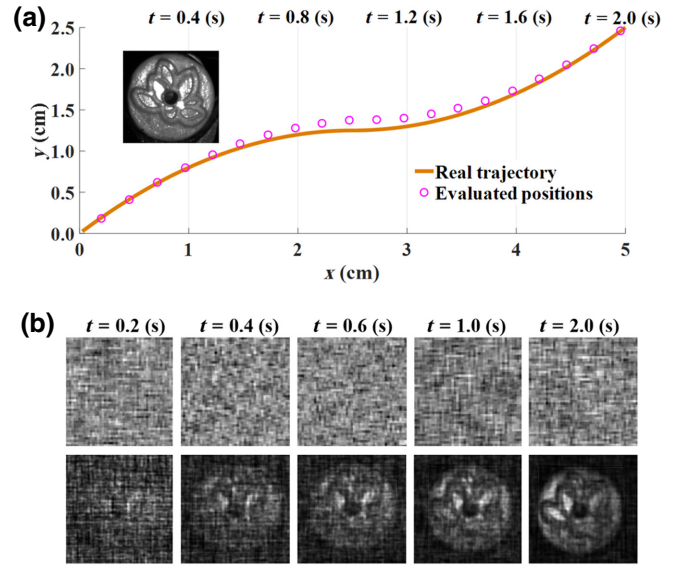

FIG. 4. Simultaneous capturing and imaging of a moving object in two dimensions. (a) Real trajectory and the evaluated positions of the moving objects within 2 seconds via our method. (b) First row is the results at every moment from traditional SPI (Video 2 within the Supplementary Material [48]), the images in the second row are obtained by our method (Video 3 within the Supplementary Material [48]).

Considering both the refreshing frequency of the DMD and the sampling frequency of the system are set as 20 kHz, the frequency of velocity measurement in our experiment is 10 Hz. According to the calculated velocity, the image of the object at the final instant ($t = 2.0$ s) is recovered with 4×10^4 samplings, as shown by the rightmost image in the second row of Fig. 4(b), the SNR of which is calculated as 4.57 (details in Appendix). With the recovered image, the final location of the object is acquired. Combining with the calculated velocity of the object at every moment the position of the object within every measurement, the trajectory is thus obtained and depicted as the red circles in Fig. 4(a).

In our experiments, the angular resolution is about 0.45 mrad, considering the speed of the moving object and the sampling repetition rate of the system, only less than 200 samplings can be performed under the condition expressed by Eq. (1). Thus, the traditional SPI fails as shown in the first row of Fig. 4(a), the average SNR of which is only 0.77 and no image information can be extracted from the noise. To capture the moving object at the same speed via traditional SPI, with the same quality as our method, 4×10^4 samplings need to be achieved within the "frozen" moment. Thus, to match the condition shown in Eq. (1), the required sampling repetition frequency will be 4 MHz, which is 200 times higher than that in our system. Considering the energy of the illumination pulse in each sampling remains the same, 200 times

higher sampling repetition frequency also means that 200 times higher illumination power is required.

C. Performance at low photon flux

Collecting the photon via a photomultiplier tube (PMT), our method promises the capability of capturing a moving object under weak illumination. To verify that, the energy of the laser pulse is decreased to 3 nJ. Considering its repeat frequency, the mean laser power is about 60 μW. The field of view is about 64 mm × 48 mm and the diameter of the pen cap is about 9 mm, from which the reflected photon flux is collected by L_4 located 0.76 m away. As a comparison, a CCD camera (Allied Vision Stingray F-125 B, not shown in Fig. 2) is used to record the image of the object, and even perform tracking.

The performance of imaging via CCD is verified first. For a fair comparison between the results of CCD imaging and our method, the angular resolution of the CCD system is adjusted to be the same as the SPI system (the number of CCD pixels occupied by the image of the object is approximately equal to the number of the speckles imprinted on the object in the SPI system). Considering the angular resolution of the system (0.45 mrad) and the angular speed of the object (45.45 mrad/s), the shutter duration and the frame frequency of the CCD are set as 10 ms and 10 Hz, respectively, to avoid motion blur. Due to an insufficient number of detected photons, the quality of the results are extremely low and the average SNR is 0.93, as shown by the first row of Fig. 5(b). The position of the object at every moment is evaluated by calculating the centroid of the image (details given in the Appendix). Tracking fails as shown by EP3 in Fig. 5(a) due to the low SNR of the image. Towards a higher SNR, experiments with an eight times longer shutter are performed to obtain more photons, with the results shown as the second row of Fig. 5(b). The average SNR of the images is improved to 1.83 while motion blur appears because the displacement of the moving object within the shutter is larger than the resolution. The results of tracking the blurred images are shown by EP2 in Fig. 5(a) and the error of the evaluated position is quite large. Thus, in this case the recorded photon within the "frozen" moment is insufficient to produce a clear image while a longer shutter causes motion blur, thus this makes the resolution and SNR of the image irreconcilable, and the imaging fails, and so even does the tracking. Then photon crisis is induced under the pixelated detection.

Under the same illumination, tracking and imaging is successfully performed via our method. The results recovered from 4×10^4 samplings are obtained. The evaluated positions of the moving object at every moment are shown by EP1 in Fig. 5(a). The image with SNR of 4.32 is reconstructed at the final moment as shown by Fig. 5(b). Under this photon flux, the average current from the anode of the PMT produced by a single pulse is $\bar{i}_s = 38.59\ \mu$A, which
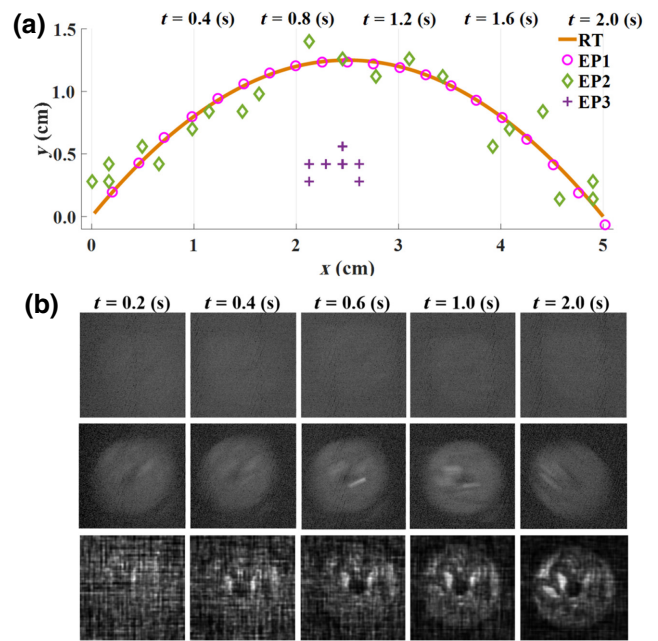

FIG. 5. Capturing a moving object at low photon flux. (a) Real trajectory and the evaluated positions via CCD with a shutter time of 10 ms (EP3), 80 ms (EP2), and our method (EP1). (b) First two rows show the recorded results via CCD, the shutter duration to record the images in the first (Video 4 within the Supplementary Material [48]) and the second (Video 5 within the Supplementary Material [48]) row is 10 ms and 80 ms, respectively. The last row shows the recovered images by our method (Video 6 within the Supplementary Material [48]), in which the rightmost image possesses better resolution and SNR than results from CCD.

is calculated by subtracting the noise current produced by background photon and dark current from the out current. The gain γ is about 2×10^3, and the average number of detected photons for a single pulse $\bar{n}$ can be estimated as 2412 ($\bar{n} = \bar{i}_s \tau_p / \gamma e$, in which τ_p is the duration of the pulse and e is the charge of a single electron). As the average number of speckles imprinted on the object is about 1018 (the ratio between the area of the object and the average area of the speckle), the average number of detected photons is about 2.4 photons/speckle, under each illumination pulse. As the small number of photons are collected by a single pixel detector and integration over time is no longer required, the moving object is captured by our method successfully. Moreover, because the velocity is recovered with a correlation method, uncorrelated detection noise is also suppressed to some degree [49–51].

In order to suppress the effect of the different capabilities of detection devices on the performance of the two kinds of imaging scheme, let us conduct a more fair analysis apart from the above experiments. In SPI, the number of the speckles imprinted on the object plays the same role as the pixel number occupied by the image on the array sensor in focal plane imaging (FPI), both of which determine

the resolution of the reconstructed or the recorded image. Thus, photon/speckle in SPI corresponds to photon/pixel in FPI. Therefore, if an array sensor with the same sensitivity as the PMT is used to perform FPI with the same resolution, in the same experimental setup, 2.4 photons/pixel will be captured to form the image. Considering the shot noise of photon number, the SNR of the image will be extremely low and tracking will be very difficult. This is the case when the shutter of the array sensor is set as the same width as the single-pulse laser. When the shutter is longer and the signal from a lot of illumination pulses is integrated, motion blur will also appear if the object moves too quickly.

Compared with FPI, in which the signal photons are required to be recorded within the "frozen" moment, our method can gradually integrate sufficient photons to reconstruct the image of the moving object during its evolution. Therefore, under the same weak illumination, more photons can be collected and images with higher quality can be obtained via our method. Thus, our method is more photon efficient and can work under weaker illumination.

IV. DISCUSSION

In this scheme, the error of the calculated velocity can also be reduced by setting a smaller shifting step of the illumination pattern, the minimum of which is limited by the pixel pitch of the DMD in our method. To obtain a calculated velocity with a smaller error, an interpolation method [52] is implemented for the bucket signal. Alternatively, a continuously shifting pattern can be effectively produced via a memory effect [46], then the influence of the shifting step of the illumination pattern on the calculation error can be neglected. In addition, as shown by Eq. (4), a more accurate $\delta\tau$ will generate a velocity with smaller error, thus a bucket detector with higher sampling repetition frequency is instrumental in achieving a better numerical accuracy. The error of the calculated velocity will be accumulated over time, which will consequently deteriorate quality of the reconstructed image. Towards this issue, a better value of velocity is searched within the neighborhood of the calculated value. As smaller velocity error implies higher-quality imaging, the quality of the final image is set as the target function for finding the velocity of smaller error [53–55]. In the case that the trajectory of the object is a more meandering curve, spline function interpolation [56,57] can also be used to reduce the accumulation of the velocity error.

The real-time performance of the tracking and imaging method is significant to capture the moving object continuously and steadily. Thus, in our experiment, the image of the object is reconstructed via intensity correlation shown as Eq. (A7) [16], which is a linear algorithm that can achieve real-time imaging. Given longer time for data processing, the quality of image can be improved

further using nonlinear algorithms [33–35,58] or machine learning [37,38].

In practice, longitudinal motion can be detected with a time-of-flight method. Thus, the image of the object along a three-dimensional trajectory can also be captured based on our method. Except for the translating motion, rotation or the relative motion between multiple objects exposed in the illumination pattern can also be reflected in the bucket signals. Thus, we believe that technology to capture rotating or multiple moving objects can also be inspired by our method. In addition, in the situation where the object is moving in a complex or bright background, the influence of the bucket signals from the background can be suppressed via background subtraction [32], then our method will also be effective in performing tracking and imaging.

Motion blur in photoelectric imaging can be essentially attributed to the coupling between the motion and spatial information of the moving object. To achieve information decoupling, the basic concept of the traditional method is to finish the acquisition of the spatial information before the motion information of the moving object increases. As a consequence, the response speed of the optoelectronic devices is required to improved endlessly subject to the object moving faster. In our method, both the motion and spatial information of the object is encoded via the operation of sequential illumination and single pixel detection. The information decoupling is achieved by data processing, thus the requirement on high performance of the device is greatly relaxed.

In optical sensing, sufficient photons are indispensable to record or computationally recover the commensurate spatial information of an image [1]. On the issue of capturing a fast-moving object, a conventional way is to record countless clear images in sequential transient moments. Thus, the required photon flux will be increased with the improving temporal resolution, which makes the performance of this method limited under weak illumination. As a different tack, in our method, the detected signal is used to acquire the velocity of the object directly. Motion information contained in a moving object is usually far less than in a spatial-resolving image [44,45], which means the required photon flux is smaller and the photon crisis can be alleviated essentially. These significant enhancements make our method promising in autonomous vehicles, medical imaging, and remote sensing applications.

APPENDIX

In SPI, the image of a static object is computationally retrieved by correlation between fluctuations of the bucket

signals and the illumination patterns as

$$G(\mathbf{r}_r) = \langle I(\mathbf{r}_r)B\rangle - \langle I(\mathbf{r}_r)\rangle\langle B\rangle, \qquad (A1)$$

in which $\langle\cdot\rangle$ represents the ensemble average of sampling for imaging, $I(\mathbf{r}_r)$ is the recorded pattern, and B is the bucket signal from the static object,

$$B = \int T(\mathbf{r}_o)I(\mathbf{r}_o)\, d\mathbf{r}_o. \qquad (A2)$$

Under the illumination of laterally shifting illumination pattern, the bucket result from the moving object can be expressed as

$$\begin{aligned}
B(t) &= \int T(\mathbf{r}_o - \boldsymbol{v}_o t)I(\mathbf{r}_o - \boldsymbol{v}_i t)\, d\mathbf{r}_o \\
&= \int T(\mathbf{r}_o)I(\mathbf{r}_o - \Delta\boldsymbol{v}t)\, d\mathbf{r}_o \\
&= \int T(\mathbf{r}_o)I(\mathbf{r}_o - \Delta\mathbf{r})\, d\mathbf{r}_o, \qquad (A3)
\end{aligned}$$

with $\Delta\mathbf{r}(t)$ representing corresponding relative displacement. Substituting Eqs. (A3) into (3), we can obtain

$$C(\tau) = \langle C_T[\Delta\mathbf{r}(t)]\, C_I[\Delta\mathbf{r}(t)]\rangle_{t_m}, \qquad (A4)$$

with $C_I[\Delta\mathbf{r}(t)] = \int[I(\mathbf{r}_o) - \langle I(\mathbf{r}_o)\rangle][I[\mathbf{r}_o - \Delta\mathbf{r}(t)] - \langle I[\mathbf{r}_o - \Delta\mathbf{r}(t)]\rangle]\, d\mathbf{r}_o$ and $C_T[\Delta\mathbf{r}(t)] = \int T(\mathbf{r}_o)T[\mathbf{r}_o - \Delta\mathbf{r}(t)]\, d\mathbf{r}_o$ representing the autocorrelation function of the illumination pattern and that of the reflectivity of the object, respectively. With the assumption that the velocity of the moving object can be taken as constant within the measurement time, the HWHM of $C(\tau)$ will be inversely proportional to the relative velocity between the object and the illumination pattern. Therefore, the relative velocity between the object and illumination pattern can be expressed as

$$|\Delta\boldsymbol{v}| = \frac{\varepsilon}{\Delta\tau} + \zeta, \qquad (A5)$$

where ζ is a hypothetical constant. Considering that an infinite $\Delta\tau$ means the illumination pattern and the object remain relatively static, thus $\zeta = 0$ and Eq. (4) can be obtained. To solve $\mathbf{v}_0$, we set the illumination pattern shifting in two orthogonal directions with speed of $\mathbf{v}_{x1}, \mathbf{v}_{x2}, \mathbf{v}_{y1}$, and $\mathbf{v}_{y2}$ separately. Decomposing the velocity of the object

as $\mathbf{v}_0 = \mathbf{v}_x + \mathbf{v}_y$, equations can be established as

$$\begin{aligned}
\frac{\varepsilon_x}{\Delta\tau_{x\beta}} &= \sqrt{v_{x\beta}^2 + v_o^2 - 2v_{x\beta}v_x}, \\
\frac{\varepsilon_y}{\Delta\tau_{y\beta}} &= \sqrt{v_{y\beta}^2 + v_o^2 - 2v_{y\beta}v_y}, \\
v_o^2 &= v_x^2 + v_y^2 \qquad (A6)
\end{aligned}$$

in which $v_{x\beta}(v_{y\beta})$ $(\beta = 1, 2)$ represents the shifting speed of the illumination pattern in the $\mathbf{x}(\mathbf{y})$ direction. Here $\Delta\tau_{x\beta}(\Delta\tau_{y\beta})$ is the HWHM of $C(\tau)$ under the corresponding illumination and $\varepsilon_x(\varepsilon_y)$ is the autocorrelation of the reflectivity of the object in the corresponding direction. With the operation that speed of the illumination pattern is larger than that of the moving object, the unknown $\varepsilon_{x,y}$ can be treated as constants and will be canceled. Thus, v_x and v_y can be solved, which means the velocity vector $\mathbf{v}_0$ is calculated.

According to the calculated velocity vector, with sufficient samplings, the image of the object at instant t can be reconstructed as

$$\begin{aligned}
G[\mathbf{r}(t)] &= \langle S[\Delta\boldsymbol{v}(t)]I[\mathbf{r}(t)]B(t)\rangle \\
&\quad - \langle S[\Delta\boldsymbol{v}(t)]I[\mathbf{r}(t)]\rangle\, \langle B(t)\rangle, \qquad (A7)
\end{aligned}$$

where $S[\Delta\mathbf{v}(t)]$ represents the shifting operation on the reference patterns according to the relative velocity at every moment, thus the illumination pattern can be treated as relatively static with the object in the data processing and the image of the moving object can be obtained.

In order to indicate the quality of the images recorded or reconstructed in a quantitative way, the SNR of image is calculated as

$$\mathrm{SNR} = \frac{\frac{1}{RV}\sum_{k=1}^{R}\sum_{l=1}^{V} M_i(k,l)}{\sqrt{\frac{1}{RV}\sum_{k=1}^{R}\sum_{l=1}^{V}[M_i(k,l) - M_o(k,l)]^2}}, \qquad (A8)$$

where M_i is the images retrieved and M_o is the image shown in Fig. 4(a), which is treated as the ground truth. Here R and V represent the pixel number of the image in two dimensions. The numerator in the right is the mean value of the retrieved image (signal) and the denominator is the standard deviation between the signal and the ground truth, which is noise. Thus, a larger SNR means a higher quality of the image and no image information can be extracted from noise when SNR ≤ 1.

To achieve tracking with sequential images, the position of the object at different moments is evaluated by calculating the centroid of the image, as

$$\begin{cases} C_k = \dfrac{\sum_{k=1}^{R} kM_i(k,l)}{\sum_{k=1}^{R} M_i(k,l)} \\[3mm] C_l = \dfrac{\sum_{l=1}^{V} lM_i(k,l)}{\sum_{l=1}^{V} M_i(k,l)} \end{cases} \qquad (A9)$$

in which C_k and C_l are the coordinates of the centroid in two dimensions.

———————————

[1] S. D. Johnson, P.-A. Moreau, T. Gregory, and M. J. Padgett, How many photons does it take to form an image?, Appl. Phys. Lett. **116**, 260504 (2020).

[2] Z.-P. Li, X. Huang, Y. Cao, B. Wang, Y.-H. Li, W. Jin, C. Yu, J. Zhang, Q. Zhang, and C.-Z. Peng, *et al.*, Single-photon computational 3D imaging at 45 km, Photonics Res. **8**, 1532 (2020).

[3] Z.-P. Li, J.-T. Ye, X. Huang, P.-Y. Jiang, Y. Cao, Y. Hong, C. Yu, J. Zhang, Q. Zhang, and C.-Z. Peng, *et al.*, Single-photon imaging over 200 km, Optica **8**, 344 (2021).

[4] P. Fuller, An introduction to high speed photography and photonics (2009).

[5] M. El-Desouki, M. Jamal Deen, Q. Fang, L. Liu, F. Tse, and D. Armstrong, CMOS image sensors for high speed applications, Sensors **9**, 430 (2009).

[6] K. Nakagawa, A. Iwasaki, Y. Oishi, R. Horisaki, A. Tsukamoto, A. Nakamura, K. Hirosawa, H. Liao, T. Ushida, and K. Goda, *et al.*, Sequentially timed all-optical mapping photography (STAMP), Nat. Photonics **8**, 695 (2014).

[7] L. Gao, J. Liang, C. Li, and L. V. Wang, Single-shot compressed ultrafast photography at one hundred billion frames per second, Nature **516**, 74 (2014).

[8] M. Lesaffre, N. Verrier, and M. Gross, Noise and signal scaling factors in digital holography in weak illumination: Relationship with shot noise, Appl. Opt. **52**, A81 (2013).

[9] P. Magnan, Detection of visible photons in CCD and CMOS: A comparative view, Nucl. Instrum. Methods Phys. Res. Sect. A: Accelerators, Spectrometers, Detectors and Associated Equipment **504**, 199 (2003).

[10] D. Shin, A. Kirmani, V. K. Goyal, and J. H. Shapiro, in *2014 IEEE International Conference on Image Processing (ICIP)* (IEEE, 2014), p. 46.

[11] D. Shin, A. Kirmani, V. K. Goyal, and J. H. Shapiro, Photon-efficient computational 3-D and reflectivity imaging with single-photon detectors, IEEE Trans. Comput. Imaging **1**, 112 (2015).

[12] R. S. Aspden, D. S. Tasca, R. W. Boyd, and M. J. Padgett, EPR-based ghost imaging using a single-photon-sensitive camera, New J. Phys. **15**, 073032 (2013).

[13] P. A. Morris, R. S. Aspden, J. E. Bell, R. W. Boyd, and M. J. Padgett, Imaging with a small number of photons, Nat. Commun. **6**, 1 (2015).

[14] R. S. Aspden, N. R. Gemmell, P. A. Morris, D. S. Tasca, L. Mertens, M. G. Tanner, R. A. Kirkwood, A. Ruggeri, A. Tosi, R. W. Boyd, G. S. Buller, R. H. Hadfield, and M. J. Padgett, Photon-sparse microscopy: Visible light imaging using infrared illumination, Optica **2**, 1049 (2015).

[15] X. Liu, J. Shi, X. Wu, and G. Zeng, Fast first-photon ghost imaging, Sci. Rep. **8**, 1 (2018).

[16] K. W. C. Chan, M. N. O'Sullivan, and R. W. Boyd, High-order thermal ghost imaging, Opt. Lett. **34**, 3343 (2009).

[17] J. H. Shapiro and R. W. Boyd, The physics of ghost imaging, Quantum Inf. Process. **11**, 949 (2012).

[18] Y.-K. Xu, S.-H. Sun, W.-T. Liu, G.-Z. Tang, J.-Y. Liu, and P.-X. Chen, Detecting fast signals beyond bandwidth of detectors based on computational temporal ghost imaging, Opt. Express **26**, 99 (2018).

[19] S. Li, F. Cropp, K. Kabra, T. Lane, G. Wetzstein, P. Musumeci, and D. Ratner, Electron Ghost Imaging, Phys. Rev. Lett. **121**, 114801 (2018).

[20] W. Li, Z. Tong, K. Xiao, Z. Liu, Q. Gao, J. Sun, S. Liu, S. Han, and Z. Wang, Single-frame wide-field nanoscopy based on ghost imaging via sparsity constraints, Optica **6**, 1515 (2019).

[21] H. Yu, R. Lu, S. Han, H. Xie, G. Du, T. Xiao, and D. Zhu, Fourier-Transform Ghost Imaging with Hard x Rays, Phys. Rev. Lett. **117**, 113901 (2016).

[22] A.-X. Zhang, Y.-H. He, L.-A. Wu, L.-M. Chen, and B.-B. Wang, Tabletop x-ray ghost imaging with ultra-low radiation, Optica **5**, 374 (2018).

[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**, 285 (2014).

[24] L. Olivieri, J. S. T. Gongora, L. Peters, V. Cecconi, A. Cutrona, J. Tunesi, R. Tucker, A. Pasquazi, and M. Peccianti, Hyperspectral terahertz microscopy via nonlinear ghost imaging, Optica **7**, 186 (2020).

[25] R. I. Stantchev, B. Sun, S. M. Hornett, P. A. Hobson, G. M. Gibson, M. J. Padgett, and E. Hendry, Noninvasive, near-field terahertz imaging of hidden objects using a single-pixel detector, Sci. Adv. **2**, e1600190 (2016).

[26] R. I. Stantchev, X. Yu, T. Blu, and E. Pickwell-MacPherson, Real-time terahertz imaging with a single-pixel detector, Nat. Commun. **11**, 1 (2020).

[27] E. Hahamovich, S. Monin, Y. Hazan, and A. Rosenthal, Single pixel imaging at megahertz switching rates via cyclic Hadamard masks, Nat. Commun. **12**, 1 (2021).

[28] Z.-H. Xu, W. Chen, J. Penuelas, M. Padgett, and M.-J. Sun, 1000 fps computational ghost imaging using LED-based structured illumination, Opt. Express **26**, 2427 (2018).

[29] W. Zhao, H. Chen, Y. Yuan, H. Zheng, J. Liu, Z. Xu, and Y. Zhou, Ultrahigh-Speed Color Imaging with Single-Pixel Detectors at low Light Level, Phys. Rev. Appl. **12**, 034049 (2019).

[30] Y. Kohno, K. Komatsu, R. Tang, Y. Ozeki, Y. Nakano, and T. Tanemura, Ghost imaging using a large-scale silicon photonic phased array chip, Opt. Express **27**, 3817 (2019).

[31] K. Nitta, Y. Yano, C. Kitada, and O. Matoba, Fast computational ghost imaging with laser array modulation, Appl. Sci. **9**, 4807 (2019).

[32] G. Zhou, Z. H. Lim, Y. Qi, and G. Zhou, Single-pixel MEMS imaging systems, Micromachines **11**, 219 (2020).

[33] O. Katz, Y. Bromberg, and Y. Silberberg, Compressive ghost imaging, Appl. Phys. Lett. **95**, 131110 (2009).

[34] O. S. Magana-Loaiza, G. A. Howland, M. Malik, J. C. Howell, and R. W. Boyd, Compressive object tracking using entangled photons, Appl. Phys. Lett. **102**, 231104 (2013).

[35] M. Aβmann and M. Bayer, Compressive adaptive computational ghost imaging, Sci. Rep. **3**, 1 (2013).

[36] D. B. Phillips, M.-J. Sun, J. M. Taylor, M. P. Edgar, S. M. Barnett, G. M. Gibson, and M. J. Padgett, Adaptive foveated single-pixel imaging with dynamic supersampling, Sci. Adv. **3**, e1601782 (2017).

[37] Y. He, G. Wang, G. Dong, S. Zhu, H. Chen, A. Zhang, and Z. Xu, Ghost imaging based on deep learning, Sci. Rep. **8**, 1 (2018).

[38] F. Wang, H. Wang, H. Wang, G. Li, and G. Situ, Learning from simulation: An end-to-end deep-learning approach for computational ghost imaging, Opt. Express **27**, 25560 (2019).

[39] S. Sun, H. Lin, Y. Xu, J. Gu, and W. Liu, Tracking and imaging of moving objects with temporal intensity difference correlation, Opt. Express **27**, 27851 (2019).

[40] W. Jiang, X. Li, X. Peng, and B. Sun, Imaging high-speed moving targets with a single-pixel detector, Opt. Express **28**, 7889 (2020).

[41] S. Sun, J.-H. Gu, H.-Z. Lin, L. Jiang, and W.-T. Liu, Gradual ghost imaging of moving objects by tracking based on cross correlation, Opt. Lett. **44**, 5594 (2019).

[42] S. Sun, W.-T. Liu, J.-H. Gu, H.-Z. Lin, L. Jiang, Y.-K. Xu, and P.-X. Chen, Ghost imaging normalized by second-order coherence, Opt. Lett. **44**, 5993 (2019).

[43] D. Shi, K. Yin, J. Huang, K. Yuan, W. Zhu, C. Xie, D. Liu, and Y. Wang, Fast tracking of moving objects using single-pixel imaging, Opt. Commun. **440**, 155 (2019).

[44] Z. Zhang, J. Ye, Q. Deng, and J. Zhong, Image-free real-time detection and tracking of fast moving object using a single-pixel detector, Opt. Express **27**, 35394 (2019).

[45] Q. Deng, Z. Zhang, and J. Zhong, Image-free real-time 3-D tracking of a fast-moving object using dual-pixel detection, Opt. Lett. **45**, 4734 (2020).

[46] M. I. Akhlaghi and A. Dogariu, Tracking hidden objects using stochastic probing, Optica **4**, 447 (2017).

[47] S. Ota, R. Horisaki, Y. Kawamura, M. Ugawa, I. Sato, K. Hashimoto, R. Kamesawa, K. Setoyama, S. Yamaguchi, and K. Fujiu, *et al.*, Ghost cytometry, Science **360**, 1246 (2018).

[48] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevApplied.17.024050 for the videos with descriptions.

[49] C. Deng, L. Pan, C. Wang, X. Gao, W. Gong, and S. Han, Performance analysis of ghost imaging lidar in background light environment, Photonics Res. **5**, 431 (2017).

[50] L. Pan, C. Deng, Z. Bo, X. Yuan, D. Zhu, W. Gong, and S. Han, Experimental investigation of chirped amplitude modulation heterodyne ghost imaging, Opt. Express **28**, 20808 (2020).

[51] D. Li, D. Yang, S. Sun, Y.-G. Li, L. Jiang, H.-Z. Lin, and W.-T. Liu, Enhancing robustness of ghost imaging against environment noise via cross-correlation in time domain, Opt. Express **29**, 31068 (2021).

[52] Q. Wang and R. K. Ward, A new orientation-adaptive interpolation method, IEEE Trans. Image Process. **16**, 889 (2007).

[53] E. Li, Z. Bo, M. Chen, W. Gong, and S. Han, Ghost imaging of a moving target with an unknown constant speed, Appl. Phys. Lett. **104**, 251120 (2014).

[54] X. Li, C. Deng, M. Chen, W. Gong, and S. Han, Ghost imaging for an axially moving target with an unknown constant speed, Photonics Res. **3**, 153 (2015).

[55] S. Jiao, M. Sun, Y. Gao, T. Lei, Z. Xie, and X. Yuan, Motion estimation and quality enhancement for a single image in dynamic single-pixel imaging, Opt. Express **27**, 12841 (2019).

[56] M. P. Wernet and A. Pline, Particle displacement tracking technique and Cramer-Rao lower bound error in centroid estimates from CCD imagery, Exp. Fluids **15**, 295 (1993).

[57] F. Pereira, H. Stüer, E. C. Graff, and M. Gharib, Two-frame 3D particle tracking, Meas. Sci. Technol. **17**, 1680 (2006).

[58] L. Bian, J. Suo, Q. Dai, and F. Chen, Experimental comparison of single-pixel imaging algorithms, JOSA A **35**, 78 (2018).