

Comparison of super-resolution and noise reduction for passive single-photon imaging

Martin Laurenzis^{a,*}, Trevor Seets,^b Emmanuel Bacher,^a
Atul Ingle,^b and Andreas Velten^b

^aFrench-German Research Institute of Saint-Louis, Advanced Visions and Processing,
Saint-Louis, France

^bUniversity of Wisconsin, Department of Biostatistics and Medical Informatics, Madison,
Wisconsin, United States

Abstract. Single-photon sensitive image sensors have recently gained popularity in passive imaging applications where the goal is to capture photon flux (brightness) values of different scene points in the presence of challenging lighting conditions and scene motion. Recent work has shown that high-speed bursts of single-photon timestamp information captured using a single-photon avalanche diode camera can be used to estimate and correct for scene motion thereby improving signal-to-noise ratio and reducing motion blur artifacts. We perform a comparison of various design choices in the processing pipeline used for noise reduction, motion compensation, and upsampling of single-photon timestamp frames. We consider various pixelwise noise reduction techniques in combination with state-of-the-art deep neural network upscaling algorithms to super-resolve intensity images formed with single-photon timestamp data. We explore the trade space of motion blur and signal noise in various scenes with different motion content. Using real data captured with a hardware prototype, we achieved super-resolution reconstruction at frame rates up to 65.8 kHz (native sampling rate of the sensor) and captured videos of fast-moving objects. The best reconstruction is obtained with the motion compensation approach, which achieves a structural similarity (SSIM) of about 0.67 for fast-moving rigid objects. We are able to reconstruct subpixel resolution. These results show the relative superiority of our motion compensation compared to other approaches that do not exceed an SSIM of 0.5. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JEI.31.3.033042](https://doi.org/10.1117/1.JEI.31.3.033042)]

Keywords: single-photon avalanche diode; super-resolution; denoising; deblurring.

Paper 220064G received Jan. 24, 2022; accepted for publication Jun. 10, 2022; published online Jun. 25, 2022.

1 Introduction

In recent years, single photon-counting avalanche diode (SPAD) sensors have gained popularity for use in various optronic sensing applications due to their extreme sensitivity and the ability to precisely measure the time of arrival of individual photons. SPAD sensors can be integrated and manufactured inexpensively in standardized semiconductor manufacturing processes with a wide range of pixel array sizes, from single pixel detectors to megapixel SPAD arrays.^{1–6}

Often, sensors and readout circuits are fabricated side-by-side on the same chip, representing a high degree of integration and resulting in short signal propagation times. Pixel fill-factors can be improved through the use of three-dimensional (3D)-stacking and microlens arrays. SPAD sensors have an outstanding sensitivity, low dark count rates (DCR), and high time resolution. These abilities are due to the detection of single-photon impacts triggering avalanche effects and time tagging with a resolution of a few picoseconds.

Typically, SPAD sensors are used in conjunction with an active light source (e.g., a pulsed laser) to record the photon timestamps in synchronization with the pulsed illumination source. Each pixel thus observes the photon impingement time which is correlated to the scene response

*Address all correspondence to Martin Laurenzis, martin.laurenzis@isl.eu

such as in fluorescence lifetime microscopy,⁷ range imaging LiDAR,^{8–11} super-resolution ranging,¹² transient,^{13,14} and nonline-of-sight sensing.^{15–18}

The focus of this work is different. We consider the problem of passive imaging with a SPAD camera where each SPAD pixel records photon timestamps due to ambient light naturally present in the scene. These timestamps are not correlated with any active light source and are instead recorded with respect to the SPAD camera's frame start times. Although these passively acquired photon timestamps do not provide information about the 3D scene structure, by exploiting the Poisson timing statistics, it was recently shown^{19–22} that these passive timestamps provide scene intensity information.

Recent publications focused on the passive sensing capabilities of single photon counting devices by, for instance, restoring intensity images from binary photon detection^{23–28} for both static and dynamic scenes. Furthermore, the timing ability of SPAD sensors was used to determine the physical intensity by estimation of the photon flux from the photon impingement rate.^{19–22} In this approach, the time between photon events is determined from the mean event time. It was shown that with photon flux measurements SPAD sensors are able to perform sensing with high dynamic range.

Despite these promising features of extreme sensitivity and time resolution, current SPAD camera arrays are severely limited in their spatial resolution and fill factor due to manufacturing challenges. Moreover, the individual frames of passive photon timing data captured with SPAD camera are extremely noisy. Therefore, there is a need to devise efficient computational algorithms that can denoise SPAD photon frames and increase the spatial resolution of the captured images. The goal of this paper is to present a thorough analysis of the various denoising and super-resolution techniques adapted to single-photon timing information captured using low-resolution SPAD cameras.

Our proposed photon timestamp processing pipeline is summarized in Fig. 1. The scene is illuminated passively by an ambient light source. The SPAD camera captures frames of photon timestamps at a high frame rate but with at most 1 photon per pixel. These are collapsed into noisy but high-frame-rate intensity (photon flux) estimates. By a combination of denoising and motion-alignment, these frames are summed to increase the overall signal-to-noise ratio (SNR) while minimizing motion blur artifacts. Finally, the low-resolution images are up-scaled by a state-of-the-art super-resolution algorithm.

In prior work, we have demonstrated²² compensation of motion by accumulating photon information along motion trajectories in these 3D (spatiotemporal) photon timestamp data sets. Further, we have published some preliminary investigations on the application of deep neural network (DNN) image upscaling in a conference paper.²⁹ Here, we focus mainly on the comparison of the two different processing strategies that are motion compensation (MC) and DNN upscaling and perform a detailed evaluation of the denoising and upscaling steps to obtain super-resolution photon flux frames.

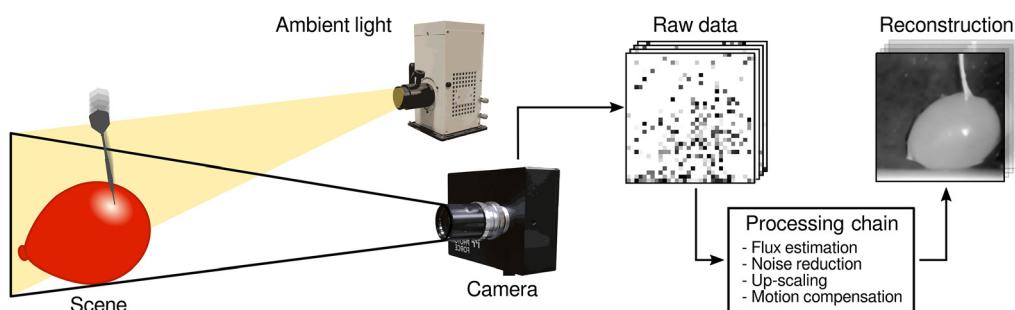


Fig. 1 A low-resolution but extremely sensitive SPAD camera captures photon timestamp frames at many kHz. These frames are temporally accumulated over short timescales to estimate noisy but blur-free photon flux (intensities). These noisy intensity frames are further denoised and accumulated over spatiotemporal motion trajectories to prevent motion-blur. Finally, super-resolved intensity images are be reconstructed from the denoised motion-blur-free image frames.

2 Related Work

2.1 Passive Single-Photon Imaging

The passive single-photon imaging aspect is related to work discussing quanta image sensors,²⁴⁻²⁷ binary single photon intensities,^{28,30} low-noise sCMOS,³¹ and EMCCD³² cameras with low light sensitivity. We consider SPAD-based imaging here because they provide much higher time resolution compared with these other sensor technologies. Moreover, SPADs can be manufactured cheaply as they are compatible with the CMOS photolithography processes.

2.2 Motion Deblurring

Motion deblurring is an ill-posed inverse problem. Conventional deblurring techniques pose this as a deconvolution problem, where the blur kernel may be assumed to be known or can be estimated from the image itself.^{33,34} Recent methods also use data driven approaches³⁵ to handle the ill-posedness.

The idea closest related to our work is burst photography where a rapid sequence of images (usually around 10) are captured and merged after MC.²⁷ Our method takes this idea to the extreme limit where the burst is composed of single-photon frames.^{21,26}

2.3 Super-Resolution and Image Upsampling

The task of image upscaling is a well-studied problem. Many methods use a subpixel movement through deliberate changes in the position of the image plane³⁶⁻³⁹ or analyze in-scene motion^{40,41} for resolution enhancement through analysis of image sequences. For single image processing, state-of-the-art methods apply data-driven approaches to train and employ deep neural networks (DNNs)⁴² to obtain super-resolution images from low-resolution datasets. Several studies were published refining, using, and comparing different approaches.⁴³⁻⁵⁰ Here, we leverage these developments and apply them to the new kind of data provided by a single-photon camera.

3 Methods

Here, we describe the experimental setup and the data processing steps to obtain super-resolution single-photon flux images from raw photon timestamp data frames. In Sec. 3.1, we describe our experimental setup and the structure of the recorded data. We then explain the processing pipeline, which consists of several steps, as described in Sec. 3.2. After capturing the data, we estimate an instantaneous flux value at all pixels using a the maximum likelihood estimator described in Sec. 3.2.1. We then use statistical noise filter to reduce the temporal noise described in Sec. 3.2.2. At this point, an optional align and merge motion correction step can be applied in case the scene contains moving objects or there was global motion due to camera-shake. We describe MC in Sec. 3.2.4. Next, bad pixels are replaced, and finally a super-resolution network is applied to obtain the final image (see Sec. 3.2.5). In the following section, we describe the whole process in detail.

3.1 Experimental Setup and Data Acquisition

In our experimental setup, a scene is illuminated by ambient light that is generated by an uncorrelated continuous light source. As shown in Fig. 1, the reflected light is partly captured by a lens to form an image on the camera sensor array. A computer is used to control and read-out the camera and to store and process the recorded data.

Continuous illumination was generated by a scientific light source consisting of a 100-W halogen light bulb and projection optics to form a homogeneous illumination field. The light bulb was driven by a stabilized power supply which generates a constant driving current. In our experiments, this current was set to relatively low levels between 6.5 and 7 A. The particular values were chosen to adapt to changing sensing condition and scene reflectance.

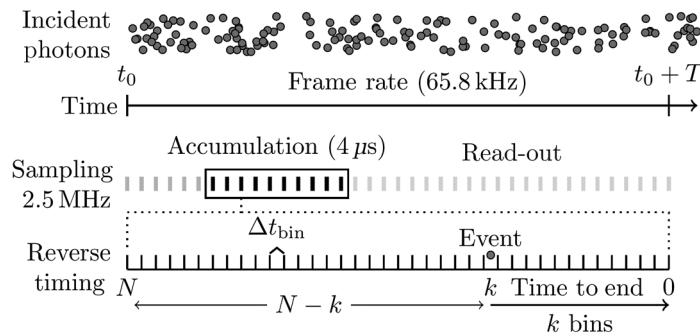


Fig. 2 Each pixel in our SPAD array (PhotonForce PF32 camera) captures a stream of incident photon timestamps; these timestamp frames are read-out at 65.8 kHz. The internal sampling rate for each pixel is 2.5 MHz and we accumulate 10 sampling intervals which amounts to a 4 μ s exposure time for each raw photon frame.

The reflected light was received by a PF32 camera (PhotonForce Ltd., Edinburgh, United Kingdom) with 16 mm focal length lens and an aperture of $f/1.4$. The PF32 camera has a silicon detector array of 32×32 single photon-counting avalanche diodes (SPAD). These SPAD sensors had an active area with a diameter of $6.95 \mu\text{m}$, a pitch of $50 \mu\text{m}$, and an optical fill factor of 1.5%. Moreover, the photon detection efficiency at 500 nm reaches a peak of 28%.⁵¹

The camera was read out with a frame rate of 65.8 kHz. However, on chip-level, the sensor array was triggered, quenched, and read at a frequency of 2.5 MHz. Furthermore, the camera was operated with an exposure time of 4 μ s to increase the overall photon detection probability by accumulating 10 sampling cycles in a single frame. Further camera details can be found in the manufacturer documentation.⁵¹

Each sensor element can detect the advent of a single photon event. The time of an event was measured within a sampling period of 57.3 ns with a 10-bit resolution of 1024 time bins and a bin width Δt_{bin} of 56 ps. The camera uses a reverse timing to record the event time as the waiting time until the end of the measurement cycle. The sensing process is shown in Fig. 2.

During the experiments, we captured two scenes with different motion content: “rotating fan” and “bursting balloon.” For each scene, we recorded data sets of 32×32 pixel positions and 5e5 samples (frames). For ease of use, recording and data analysis were performed on two different computers. In principle, however, both processes can be integrated to run on a single machine. Our implementation uses a combination of MATLAB⁵² (for capturing data from the camera over USB3.0) and Python⁵³ with OpenCV^{54,55} (for data analysis).

3.2 Data Processing Chain

Details of the data processing pipelines are shown in Figs. 3(a) and 3(b). Without MC, the pipeline follows four step process as illustrated in Fig. 3(a) as a row of gray-scale images. This data processing chain first retrieves the position of all detected photons in the $32 \times 32 \times 500,000$ dataset, pToF, estimates the photon flux for each single photon (estimate Φ) and noise is reduced by applying a filter (see Sec. 3.2.2). We tested different linear and statistical noise filter. Then, we correct the pixel values due to a previous determined dark current map (DCR pattern) and apply a DNN approach to upscale the resulting data frames.

With MC applied, see Fig. 3(b), the data processing pipeline is bypassed through additional processing steps [in Fig. 3(a) at * to ‡] and we compensate motion through an align and merge process. Again, we scale the dataset to super-resolution, but we distinguish the opportunity to scale either before or after MC.

Details of each processing step are now described below.

3.2.1 Photon flux estimation

In recent papers,^{19–22} the method of photon flux measurement was introduced using the timing capability of single photon-counting avalanche diodes to estimate the photon flux as the rate at

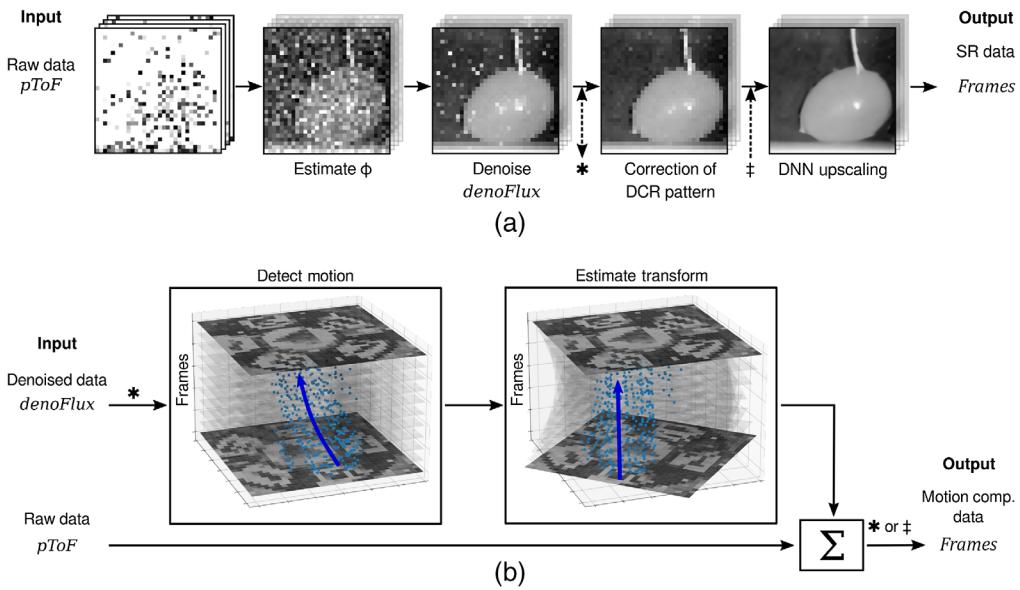


Fig. 3 The reconstruction algorithms follow a pipeline of different processing steps for (a) pixelwise estimation of the photon flux (Φ), denoising ($denoFlux$), correction of photon flux estimation errors associated with dark current noise (DCR pattern), and upscaling by a DNN (upsampling). Optionally, we can (b) compensate motion by estimating frame-to-frame motion and merging photon counts along predetermined motion paths.

which photons arrive at the detector. Equation (1) states the photon flux estimate, $\hat{\phi}$, over m sampled frames which is given as the reciprocal of the mean photon arrival time (\bar{t}) within the m sampled frames. Here, t_i is the measured event time in the i 'th sampling frame, while $n_i \in [0,1]$ indicates if a photon was detected. If no photon is detected, t_i returns the measuring cycle time and increases the overall waiting time between two photon events. We define n as the number of detected photon events within the m sampling approaches:

$$\hat{\phi} = \frac{1}{\bar{t}}, \quad \text{with } \bar{t} = \frac{\sum_{i=1}^m t_i}{\sum_{i=1}^m n_i} = \frac{\sum_{i=1}^m t_i}{n}, \quad (1)$$

$$\text{with } t_i = \begin{cases} b_i \Delta t_{\text{bin}} & (\text{forward sampling}), \\ (N_{\text{bins}} - b_i) \Delta t_{\text{bin}} & (\text{reverse sampling}). \end{cases} \quad (2)$$

For an event read-out with forwardsampling, the event time t_i equals simply the the bin number (b_i) times the sampling bin width (Δt_{bin}), see Eq. (2). In the case of reverse sampling, first, we have to convert the bin number by subtraction from the total number of bins N_{bins} , as given in Eq. (2).

From Eq. (1), we can estimate the photon flux related to a single photon event ($n \rightarrow 1$) called the instantaneous photon flux $\hat{\phi}_{\text{inst.}}$. The instantaneous flux estimator $\hat{\phi}_{\text{inst.}}$ is no longer defined as the mean detection rate over a certain number of sampling cycles, but as the flux which persists in the period between a previous p and the current detection event $p + \delta$, see Eqs. (3) and (4):

$$\hat{\phi}_{\text{inst.}}(u, v, k) \underset{n \rightarrow 1}{\approx} \frac{1}{t_{\text{inst.}}(u, v, k)}, \quad (3)$$

$$\text{with } t_{\text{inst.}}(u, v, k) = \sum_{i=p+1}^{p+\delta} t_i(u, v), \quad \delta \geq 1 \text{ and } k \in]p, p + \delta]. \quad (4)$$

For each detector position (u, v) , we can extrapolate the photon flux to any instant $S = k \in]p, p + \delta]$ from the later single photon detection $(p + \delta)$. Thus, we can obtain an estimation of

the instantaneous photon flux for the whole detector array even if we do not have a photon detection at every sensor element that instant.

3.2.2 Noise characteristics and noise reduction

Uncorrelated single photon detection is strongly affected by noise effects²⁰ due to the Poisson nature of the detection process, as illustrated in Fig. 4. Here, the noise characteristics are shown for different level of photon flux. The target consisted of six patches ($P_1 \dots P_6$) covered with a diffuse reflecting coating (Permaflect[®], Labsphere Inc.) of different reflectively ($R \in [94\%, \dots 5\%]$). Therefore, these patches represent a wide range of shades of gray that include all coatings available at the time of the experiment. Figure 4(c) shows a histogram of $\hat{\phi}_{\text{inst}}$ for the measurements of a single pixel on each patch. The obtained count distributions are specific for the camera used. Other SPAD cameras may have different active areas, DCR, quantum efficiency, and reception optics which will impact the absolute values and the shape of the count distribution.

We can fit these distribution functions by skewed normal distribution⁵⁶ and observe that different photon flux levels can be distinguished and the shift between the different distribution functions matches with the intensities or photon flux level coming from the target surface. Further, in each sampling cycle, we can estimate to have between 1.47 and 0.08 photons, $\bar{N}_{\text{Photon}} = \Delta t_{\text{bin}} N_{\text{bins}} \bar{\phi}$, impinging each sensor element. The real number may be higher due to the sensor's quantum efficiency.

However, our photon flux estimation is strongly affected by noise. Thus, further processing is needed to reduce the noise and to give a better representation of the photon flux in a single instant. In our study, we have evaluated different noise filters which are pixelwise applied to the estimated photon flux data:

- Gaussian average (GA): we calculate the convolution with a Gaussian kernel with radius ω and window 8ω .
- Total variance regularization (TV): TV⁵⁷ is an edge-preserving denoising method.⁵⁸ In our analysis, we use the direct one-dimensional TV algorithm.⁵⁹⁻⁶¹
- Change point fitting (CPF): As in Ref. 22, we use statistical change point detection⁶² to analyze photon flux data with either the kernel based^{63,64} or pruned exact linear time (PELT)⁶⁵ algorithm. Then, the photon flux is fitted by the mean photon flux between two consecutive change points.

We also evaluated further noise reduction approaches, such as moving average, bilateral, and median filter, but we exclude these results as they do not significantly contribute to the overall discussion.

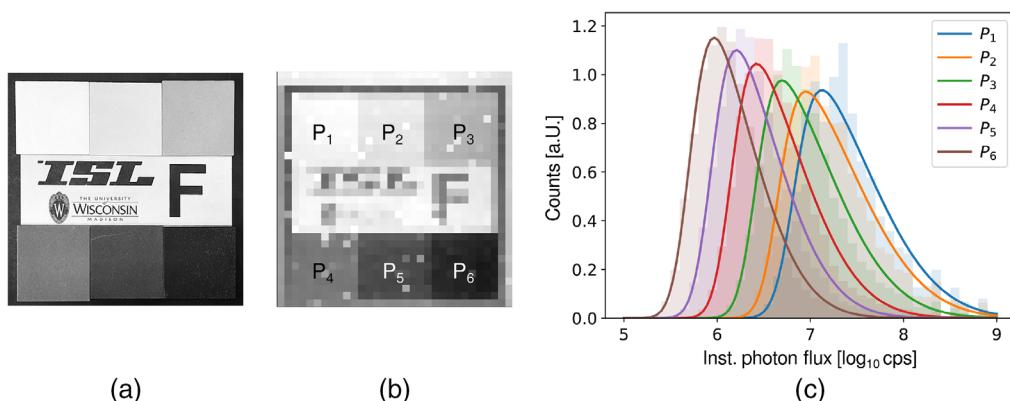


Fig. 4 Single photon flux noise characteristics: (a) gray-scale target, (b) measured mean photon flux, and (c) single pixel instantaneous photon flux distributions (500,000 samples).

3.2.3 Correction of pixel values

After denoising temporally each pixel, we noticed a spatial pattern of high estimated photon flux values regardless of the scene. Therefore, we measured the DCR of the SPAD array and identified the pixels which have a significant higher probability to fire without any impinging photons. We assumed that pixels with a dark rate count of more than 5% are bad pixels that will not give accurate flux measurements. To get rid of these hot pixels, we apply a 3×3 median filter to each frame and replace the wrong pixel values by the median value of its neighboring pixels.

3.2.4 Motion compensation

Motion in the observed scene can lead to significant blurring of edges and contours in the restored image frames; therefore, we want to reduce these effects and follow an approach similar to Ref. 22 for MC by aligning photon detection events through the data volume. The process is illustrated in Fig. 3(b).

Into the MC algorithm, we pass in the denoised fluxes denoFlux (we use CPF for the denoise method) and the unprocessed data pToF. We then find motion between successive frames using either dense optical flow (openCV⁵⁴ function calcOpticalFlowFarneback) or a euclidean motion model (openCV⁵⁴ function findTransformECC). Then, we linearly interpolate the found motion and apply it to the raw photon timestamps pToF this aligns the measurements to the next frame in the sampled data set.

Then, we merge the aligned measurements into frames, we hold the counts and the average time step separately to correctly estimate flux in later steps. Next, we estimate flux to get a new set of motion compensated frames (frames). We then recursively find motion, align, and merge frames, each step halving the number of frames, until we are left with one frame. This process can be done with a moving window to get a full-frame rate video.

3.2.5 Upscaling to super-resolution

It is known that image upscaling can be done by application of convolutional neural networks. There are several approaches published and their codes and even trained network are available. In our approach, we used the upscaling function of OpenCV⁵⁴ in Python. The upscaling algorithm is based on a pretrained DNNs with a scaling factor of $4\times$, to scale 32×32 low-resolution to 128×128 super-resolution frames. In our algorithm, we use the enhanced deep super-resolution network approach from Lim et al.⁴⁸

We also investigated upscaling prior to an MC step as in Ref. 22. In this approach, the raw data are nearest-neighbor upscaled before applying align and merge MC. This gives sub-pixel resolution because each SPAD pixel will sample along a continuous motion curve. The nearest-neighbor upsampling allows photon data to be combined along these motion curves better than working in the captured resolution. This technique allows SPADs to get around some of the issues with low fill factors for scenes with motion.

4 Results

4.1 Simulated Results: Pixelwise Fitting

To compare the different three pixelwise noise reduction approaches (see Sec. 3.2.2), we simulate a simple square light pulse incident on a single pixel SPAD to represent an object moving across a background. We are interested in the behavior of our filters for different motion speeds and contrasts. Therefore, we vary the pulse width as an analogy to speed and the pulse contrast—the ratio between the tallest and lowest point (set at 10^7) in the pulse.

We simulate photon detections of a 1024-bin SPAD sensor with 56 ps time resolution over 5000 sampled frames, for details see Appendix A, and denoise the photon timestamps with the three noise reduction methods. An example of a simulated pulse and the results of each fitting method is shown in Fig. 5(a). For each method, we try a variety of hyperparameters and use the best hyperparameter for each contrast/duration pair to calculate the SNR for each fitting method.

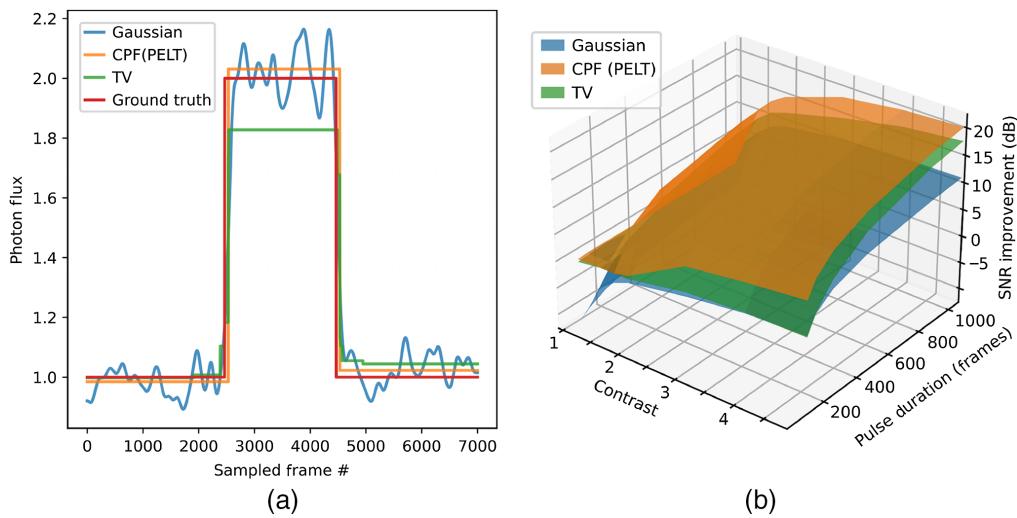


Fig. 5 Performance of different denoising strategies on simulated photon timestamp data. (a) Example for ground truth (red) and denoised single pixel signal and (b) SNR improvement Δ_{SNR} [Eq. (6)] for different algorithms.

Let $\Phi := (\phi_1, \dots, \phi_k, \dots)$ be the vector of the true photon flux incident on a pixel at sampling frame k , and $\hat{\Phi}^{\text{fit}} := (\hat{\phi}_1^{\text{fit}}, \dots, \hat{\phi}_k^{\text{fit}}, \dots)$ be the vector of estimated fluxes found from a given fitting method. We define SNR as²⁸

$$\text{SNR}(\Phi, \hat{\Phi}^{\text{fit}}) = 10 \log_{10} \left(\frac{\|\Phi\|_2^2}{\|\Phi - \hat{\Phi}^{\text{fit}}\|_2^2} \right), \quad (5)$$

where $\|\cdot\|_2^2$ represents the l^2 norm.

To better understand how the fitting methods deal with the dynamics of the scene, we compare the SNR from a fitted signal, $\text{SNR}(\Phi, \hat{\Phi}^{\text{fit}})$, to the SNR given by averaging over all frames, $\text{SNR}(\Phi, \hat{\Phi}^{\text{Avg}})$. Where $\hat{\Phi}^{\text{Avg}}$ is a constant vector representing the average photon flux measured and is calculated using all sampling frames in Eq. (1). By comparing $\text{SNR}(\Phi, \hat{\Phi}^{\text{fit}})$ to $\text{SNR}(\Phi, \hat{\Phi}^{\text{Avg}})$, we can calculate how much SNR improvement we get from fitting to motion in the scene. The SNR improvement, Δ_{SNR} , is given as

$$\Delta_{\text{SNR}} = \text{SNR}(\Phi, \hat{\Phi}^{\text{fit}}) - \text{SNR}(\Phi, \hat{\Phi}^{\text{Avg}}). \quad (6)$$

The SNR improvement results are shown in Fig. 5(b). The CPF did the best for all tested scenarios, with TV fitting close behind. Both TV and CPF properly dealt with the constant case (contrast = 1) by giving the same SNR as the full time average.

Gaussian averaging did worse than $\hat{\Phi}^{\text{Avg}}$, for short durations (fast objects) and low contrast. The GA struggles in these tough scenarios because it only can look at a small number of samples which can cause inaccurate estimations at the level of the pulse, whereas both CPF and TV fitting can adaptively average more samples.

The run times for each method are dependent on the choice of hyperparameter. In general, Gaussian averaging took about 10 ms, TV took about 100 ms, and CPF took about 1s.

4.2 Experimental Results

In Fig. 6, the results of the pixelwise data processing and the MC are summarized and depicted as a single frame for each scene and algorithm. The first and second rows show the results of the “rotating fan” scene, and the third and fourth show the “bursting balloon” scene, respectively, at single (32 × 32) and super-resolution (128 × 128), with each column representing a different

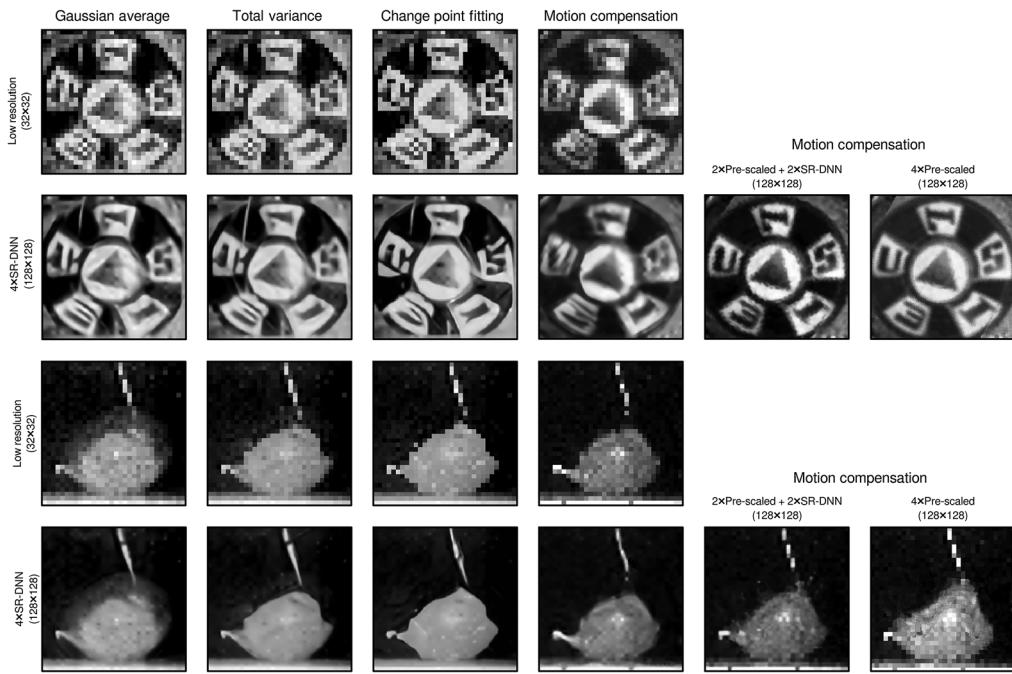


Fig. 6 Reconstructing various scenes with native (32×32) and super-resolution (128×128) using pixelwise (GA, TV, CPF) and MC algorithms. We reconstruct the subpixel details of the letters on the rotating fan blades (prescaled MC) and observe the fast motion of the bursting balloon (e.g., CPF).

reconstruction pipeline. In all frames, the logarithmic gray scale illustrates a photon flux of about $10^6 - 10^8$ cps.

In the first column, the application of the GA filter is shown. We had optimized the Gaussian radius to trade the noise reduction ability against the tendency of motion blur. In our analysis, we set this radius to $\omega = 15$ samples and the convolution kernel size to $8\omega = 120$. Although both a significant remaining noise component as well as an incipient motion blur can be observed, this filter is quite effective. Shapes and letters can be recognized, but the ripping edge of the bursting balloon is smeared due to significant motion blur. No clear contours can be seen here.

The results of the total variation (TV) filter are shown in the second column. We observe a significant lower effect of noise and motion blur through out all scenes. Letters and shapes are reconstructed in good quality. Only a slight amount of motion blur can be observed at the balloon's ripping edge. But, in contrast to the GA filter, now, the motion blur effect is much lower and the edge can be observed clearly. In the third column, the CPF results show very low noise effects and so far the highest contrast between bright and dark surfaces (letters and background).

The results of the MC algorithm are presented in the last three columns. In the fan scene, we used the euclidean motion model to compensate motion, while for the balloon we used optical flow analysis. Further, we tried out different scaling methods to obtain super-resolution by combining prior and posterior scaling of the data sets with linear and SR-DNN methods.

A first MC approach works similarly to the pixelwise algorithms above. MC was used on the datasets with native resolution and resulting frames were later scaled to $4\times$ resolution by employing the pretrained DNN. The results of the two processing steps are shown in the fourth column. In both scenarios, the impact of noise is significantly reduced. Further, shapes and the balloon's ripping edge are reconstructed well. However, the resolution of the letters on the fan appears to be similar to the results of the previous pixelwise processing algorithms.

In a further MC approach, fifth column, we combined a linear scaling ($2\times$) of the datasets before application of MC and $2\times$ DNN scaling. With this approach, it is possible to obtain effective noise reduction and a high contrast. Further, due to the prior scaling of the datasets, it is possible to reveal subpixel information in areas of continuous motion that is, for instance, the rotating fan. In these areas, we can identify details of the letters typeface⁶⁶ such as the serifs

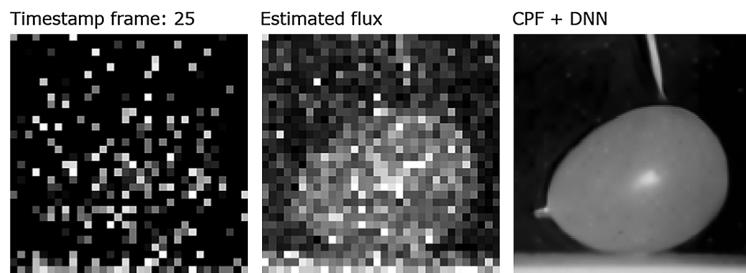


Fig. 7 [Video 1](#) showing reconstruction of “bursting balloon” scene with timestamp frames, estimated flux, and super-resolution reconstruction with CPF + DNN ([Video 1](#), MPEG, 1.4 MB [URL: <https://doi.org/10.1117/1.JEI.31.3.033042.1>]).

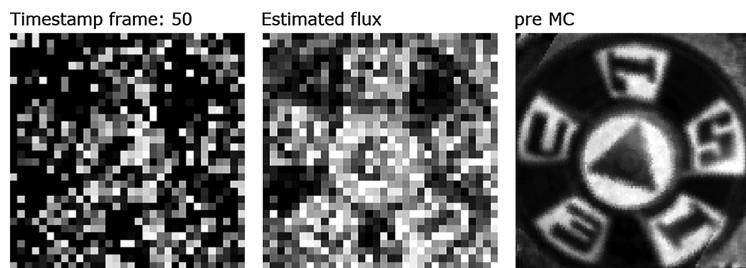


Fig. 8 [Video 2](#) showing reconstruction of “rotating fan” scene with timestamp frames, estimated flux, and super-resolution reconstruction with prescaled MC ([Video 2](#), MPEG, 1 MB [URL: <https://doi.org/10.1117/1.JEI.31.3.033042.2>]).

(in “I” and “L”), hairlines (in “S” and “L”), and bows (in “U” and “W”) which were not visible before. On the other hand, in the balloon scene, we observe pixelated representation of areas with no or slow motion (e.g., balloon surface and dropping dart). Here, the pixelwise processing methods obtain much better resolutions.

Finally, in a the third MC approach (sixth column), we employed only the prescaling (4 \times) of the datasets before application of MC. Again, it is possible to obtain effective noise reduction and a high contrast. In the fan scene, again, we obtain very detailed subpixel resolution in areas with continuous motion. For instance, the representation of the typefaces appears to be much clearer than in the aforementioned approaches. However, the algorithm results are very pixelated in areas with no or few motion (e.g., balloon body). Moreover, the ripping edge, although having huge amount of motion, is reconstructed pixelated and does not appear clearly in the resulting frame.

Figures 7 and 8 show still images of videos illustrating our processing pipelines. [Video 1](#) (Fig. 7) depicts the “bursting balloon” scene and shows raw timestamp frames, pixelwise estimation of the instantaneous photon flux, and the super-resolution reconstruction using CPF denoising and DNN upscaling. [Video 2](#) (Fig. 8) presents reconstruction of the “rotating fan” scene with prescaled MC.

5 Discussion

In our analysis, we have seen that we can apply various noise reduction and super-resolution algorithms to passive single-photon timing data to obtain a super-resolution reconstruction of the instantaneous photon flux. However, the quality of the reconstruction depends on the amount of motion contained in the scene and the reconstruction filter applied. Nevertheless, it is possible to obtain a reconstruction frame rate identical to the sampling frame rate. In our experimental datasets, we were able to reconstruct an original frame rate of 65.8 kHz. This rate is limited by the achievable frame rate of the hardware used.

For further discussion, we focus on the super-resolution results obtained in the rotating fan scene. An overview can be seen in Fig. 9, which shows two sets of magnified sections

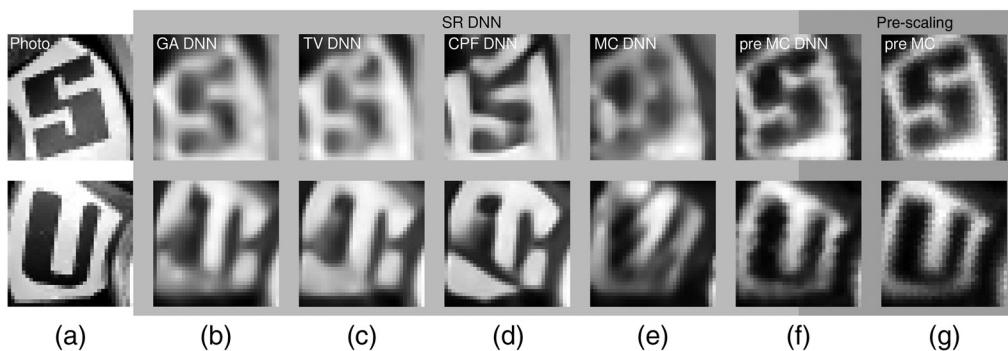


Fig. 9 Effect of MC and different denoising algorithms on super-resolution results. (a) Ground-truth photograph of the object. (b)–(e) Super-resolution reconstruction using DNN scaling after different noise reduction filters Gaussian averaging, total variation, change-point fitting, and MC. (f)–(g) Prior linear data scaling with later MC.

representing the fan blades with the letter “S” and “U.” It is obvious that very different reconstruction qualities are achieved with the different algorithms. However, the algorithms with 4× SR-DNN upscaling result in blurred reconstructions, while applying MC to prescaled datasets achieves sharp reconstruction of subpixel information.

To compare the reconstruction performance of the different algorithms, we calculated the structural similarity (SSIM)^{67,68} and mean square error (MSE) by comparison to a scene photo (see Appendix B, Fig. 11). Both evaluation metrics were applied to the whole images and to 35 × 35 cut-out sections. For the comparison, we normalized the data using a min-max normalization such as the MSE range from zero to one ($\text{MSE} \in [0,1]$). By definition, the SSIM value can range from 1 (good similarity) to -1 . Due to the fact that we do not have a real ground truth, but a very similar photo, we do not expect to obtain an SSIM of 1. For details of the SSIM analysis process, see Appendix B.

For the different algorithms, we found SSIM ranging from 0.32 to 0.67 while the MSE values are almost constant at 0.09 ± 0.04 . The algorithms based on DNN upscaling after noise processing (GA DNN, TV DNN, CPF DNN, and MC DNN) result in $\text{SSIM} \in [0.32, 0.5]$ with very similar values for the different letters and the whole image. Significantly, better SSIM values are obtained for MC employing prescaled data sets, $\text{SSIM} \in [0.46, 0.67]$. The obtained values are summarized in Fig. 10.

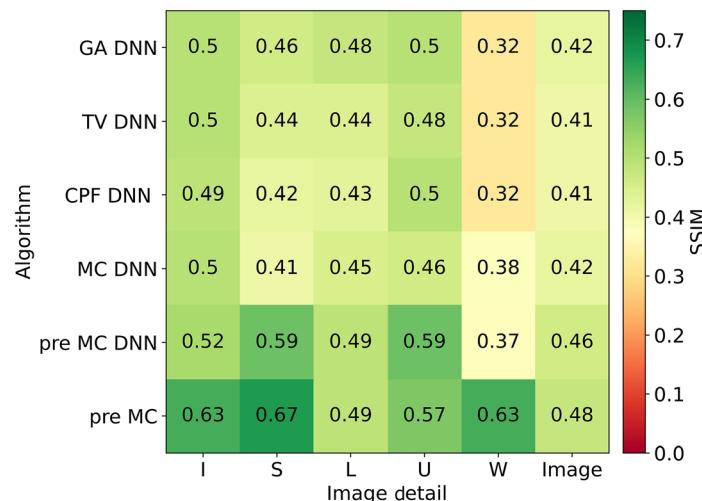


Fig. 10 Image quality metric (SSIM) computed for different denoising methods of different sections of the “rotating fan” scene for different reconstruction algorithms: DNN upscaling with Gaussian averaging (GA DNN), total variation (TV DNN), change-point fitting (CPF DNN) and motion compensation (MC DNN), and prescaled motion compensation (pre MC DNN, pre MC).

6 Conclusion

In this paper, we demonstrated the use of a commercially available single photon counting camera to measure the event times of passively acquired photons in the absence of any active (time-correlated) light source. Reconstruction of the instantaneous photon flux at high sampling rates was possible with an impingement of just $O(1)$ photons per sample. However, since this initial estimate is strongly affected by noise and the spatial resolution of the sensor array is low, we applied and compared various denoising and super-resolution strategies. In doing so, we were able to reconstruct super-resolved images at a frame rate of 65.8 kHz, limited only by the native frame rate of the camera.

Further, we have shown that SPAD sensors are more than just photon counters for range-finders and time-correlated measurements. They can be used to measure real physical quantities of light such as photon flux, which expresses the light intensity and thus the radiated energy flux with physical accuracy. We believe that photon flux measurements can provide valuable information, for example, in high-speed imaging which will have implications for a variety of applications including scientific imaging and consumer machine vision applications.

In general, our algorithms do not use any assumptions about the physical properties of the sensor and can therefore also be used for other sensors. Nevertheless, the use of other sensors could lead to different results. For instance, the use of sensors with higher fill factor could lead to higher spatial resolution, as we show in [Appendix C](#). Furthermore, higher count rates could be achieved with such sensors.

Finally, with our work, we hope to encourage hardware developers to implement the photon flux estimation, motion correction, denoising, and super-resolution pipelines on the sensor chip level. With a specialized photon flux hardware, we could dramatically reduce the amount of data postprocessing required. In addition, such devices could also operate at much higher frame rates up to several MHz. Although future photon counting devices will have much higher frame rates and higher spatial resolution, our concept of MC will still be an interesting approach to reduce the effects of motion blur and to achieve sub-pixel resolution.

7 Appendix A: Simulations

To quantitatively compare the pixelwise fitting algorithms, we simulate a single pixel with incident flux given by a single pulse wave at a randomly generated time. The pulse width represents the motion speed while the ratio between the pulse height and background level represents the contrast. We use 10^7 photons per second for the background. Our simulated pixel uses 1024 bins each bin having width of 56 ps and measures for a total of 5000 frames.

We do the following to simulate the photon detection data. In each bin, the probability of detecting a photon is given from Poisson statistics to be $p = 1 - e^{-\phi\Delta t_{\text{bin}}}$ where ϕ is the flux during that bin (we assume that the pulse changes at a bin edge) and Δt_{bin} is the bin width. We then pull a Bernoulli sample for each bin and keep the first success as the arrival time of the first photon in a given frame. Doing this for many frames, we generate simulated SPAD data for the pulse waveform.

We fit the simulated data using each of the three fitting methods, with different parameters where possible, and compare the performance of each fit. We use the following parameter for each fitting method parameter:

- Gaussian averaging: $\sigma = [15, 25, 45, 65]$
- CPF (PELT): $\lambda = [2, 4, 6, 8, 10]$
- TV: $\lambda = [1000, 4000, 6000, 8000, 10000]$

We measure the MSE between the ground truth and fitted signal. We calculate SNR as the ratio between the signal energy (L^2 norm of the signal) and root MSE.

8 Appendix B: Determine the Structural Similarity

The SSIM index was developed by Wang and Bovik⁶⁷ to have an universal metric that describes the similarity of an image (e.g., reconstructed image) (X) and a reference (Y). Unlike other error

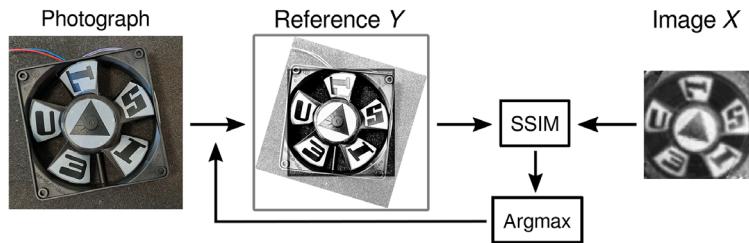


Fig. 11 A scene photograph is adapted as a reference (Y) for determining SSIM to a reconstructed image (X).

metrics (such as root-mean-square-error, MSE, or peak-signal-to-noise) which evaluate the absolute difference between pixel values the SSIM characterizes the image quality in means of the human perception or the human visual system and incorporate perceptual quality measures. SSIM [see Eq. (7)] evaluates the luminance, contrast, and structure:⁶⁸

$$\text{SSIM} = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)}. \quad (7)$$

Here, μ_i denotes the mean intensity of image i and σ_i is its standard deviation. Then, the structural correlation between the two images is evaluated by $\sigma_{XY} = \frac{1}{N-1} \sum_{i=1}^N (X_i - \mu_X)(Y_i - \mu_Y)$. Furthermore, C_1 and C_2 are constants to bias low values and to stabilize the division.

To use the SSIM on our results, we took a photograph of the scene (only the “fan scene”) as a reference image and used Euclidean transform (translation $\vec{T} = (t_u, t_v)$, rotation α_R , and scaling (S) where S is a constant factor for all results) of the reference image to enable an overlay with the reconstructed image. Both images are cropped to the same size and the intensity as well as the photon flux is normalized to a min-max span, Eq. (8):

$$I_{\text{norm.}} = \frac{I - \min(I)}{\max(I) - \min(I)}. \quad (8)$$

The overlay process is illustrated in Fig. 11. The parameter set $p = (\vec{T}, \alpha_R, S)$ was optimized by maximizing the SSIM value, Eq. (9):

$$\arg \max_p \text{SSIM}(X, Y, p) \quad \text{with } p = (\vec{T}, \alpha_R, S). \quad (9)$$

9 Appendix C: Impact of Sensor Fill Factor

The sensor fill factor has enormous impact on the super-resolution reconstruction capabilities of our MC algorithm. To illustrate this impact, we conducted a simulation using scene images rotated by a known angle. Under this conditions, we are able to ideally compensate this motion and reconstruction a super-resolution image of the scene.

In our simulation, we used the following procedure:

1. Rotate scene: The scene image is rotated by an angle $\phi \in [0 \text{ deg}, 90 \text{ deg}]$.
2. Create image: From the rotated scene, we generate a low-resolution sensor image.
 - (a) Crop scene: Scene image is cropped to the sensor field of view (320×320 pixel).
 - (b) Resample image: Reduce resolution to the sensor resolution (32×32 pixel).

Resampling was done for two different sensor fill factors:

- i. Low fill factor of 1% was achieved by sparse sampling using nearest-neighbor interpolation.
- ii. High fill factor of 100% was obtained by integration over the sensing area using bilinear interpolation.

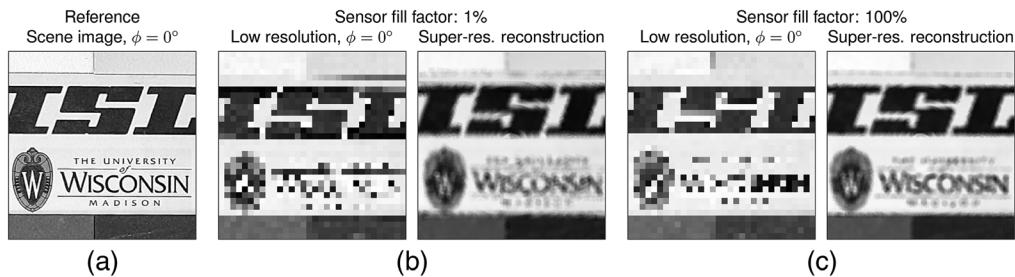


Fig. 12 The sensor fill factor has impact on the super-resolution reconstruction performance. We simulated the reconstruction of a high-resolution input image (a) from low-resolution images. The input scene was rotated at 30 different rotation angles ϕ , $\phi \in [0 \text{ deg}, 90 \text{ deg}]$. The low-resolution images were sampled from the input image using a fill factor of (b) 1% and (c) 100%. Super-resolution reconstructions were calculated from the image stack using the prescaled motion compensation approach (preMC).

3. Prescaled MC was obtained by

- (a) Prescaling the image to super-resolution (320×320 pixel) using nearest-neighbor interpolation.
- (b) MC was achieved by reversing previous scene rotation ($-\phi$).
- (c) Stacking: The result is stacked into the data volume.

4. Reconstruction: The final motion compensated image was obtained by integration over the aligned image stack.

The simulation process is shown in Fig. 12. In Fig. 12(a), we show a scene image at a rotation angle of $\phi = 0$ as a reference. The simulation with a low sensor fill factor (1%) is shown in Fig. 12(b) while the high sensor fill factor (100%) is illustrated in Fig. 12(c). In both cases, we show an example frame with low sampling resolution at a rotation angle $\phi = 0 \text{ deg}$ and the super-resolution reconstruction.

The main difference between the two sampling methods can be seen in the fact that with a low sensor fill factor the scene content is sparsely sampled while a high fill factor integrates more spatial details. Thus, compared with the low sensor fill factor simulation, the scene reconstruction obtained from high fill factor sampling is much sharper and shows more details. In addition, in the low sensor fill factor simulation, we can observe more reconstruction artifacts. Therefore, from this simulation, we can conclude that higher sensor fill factor will enable better super-resolution reconstruction.

Acknowledgments

We would like to acknowledge financial funding through the General Funding of the French-German Research Institute of Saint-Louis (ISL) by the République Française (France) and the Bundesrepublik Deutschland (Germany). Support for this research was provided by the University of Wisconsin-Madison, Discovery to Product (D2P) with funding from the State of Wisconsin. This material was based upon work supported by the Department of Energy/National Nuclear Security Administration under Award Number DE-NA0003921, the National Science Foundation (GRFP DGE-1747503, CAREER 1846884). The authors declare no conflict of interests. Disclaimer: This report was partly prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement,

recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Code, Data, and Materials Availability

Supplementary material such as code, data, and other materials are available from the authors upon request. This material or parts of it may be published on online repositories (such as GitHub) in the future.

References

1. E. Charbon, “Single-photon imaging in complementary metal oxide semiconductor processes,” *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* **372**(2012), 20130100 (2014).
2. J. Richardson et al., “A 32×32 50ps resolution 10 bit time to digital converter array in 130nm cmos for time correlated imaging,” in *IEEE Custom Integr. Circuits Conf.*, IEEE, pp. 77–80 (2009).
3. C. Veerappan et al., “A 160×128 single-photon image sensor with on-pixel 55ps 10b time-to-digital converter,” in *IEEE Int. Solid-State Circuits Conf.*, IEEE, pp. 312–314 (2011).
4. O. Kumagai et al., “A 189×600 back-illuminated stacked spad direct time-of-flight depth sensor for automotive lidar systems,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC)*, IEEE, Vol. 64, pp. 110–112 (2021).
5. F. Villa et al., “Cmos imager with 1024 spads and tdc for single-photon timing and 3-D time-of-flight,” *IEEE J. Sel. Top. Quantum Electron.* **20**(6), 364–373 (2014).
6. F. Guerrieri et al., “Two-dimensional spad imaging camera for photon counting,” *IEEE Photonics J.* **2**(5), 759–774 (2010).
7. C. Bruschini et al., “Single-photon avalanche diode imagers in biophotonics: review and outlook,” *Light: Sci. Appl.* **8**(1), 1–28 (2019).
8. M. A. Albota et al., “Three-dimensional imaging laser radar with a photon-counting avalanche photodiode array and microchip laser,” *Appl. Opt.* **41**(36), 7671–7678 (2002).
9. A. Kirmani et al., “First-photon imaging,” *Science* **343**(6166), 58–61 (2014).
10. A. Gupta, A. Ingle, and M. Gupta, “Asynchronous single-photon 3D imaging,” in *Proc. IEEE/CVF Int. Conf. Comput. Vision*, pp. 7909–7918 (2019).
11. R. Tobin et al., “Three-dimensional single-photon imaging through obscurants,” *Opt. Express* **27**(4), 4590–4611 (2019).
12. R. Tobin et al., “Robust real-time 3d imaging of moving scenes through atmospheric obscurant using single-photon lidar,” *Sci. Rep.* **11**(1), 1–13 (2021).
13. M. O’Toole et al., “Reconstructing transient images from single-photon sensors,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 1539–1547 (2017).
14. Q. Sun et al., “Depth and transient imaging with compressive spad array cameras,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 273–282 (2018).
15. A. Kirmani et al., “Looking around the corner using transient imaging,” in *IEEE 12th Int. Conf. Comput. Vision*, IEEE, pp. 159–166 (2009).
16. M. Buttafava et al., “Non-line-of-sight imaging using a time-gated single photon avalanche diode,” *Opt. Express* **23**(16), 20997–21011 (2015).
17. D. Faccio, A. Velten, and G. Wetzstein, “Non-line-of-sight imaging,” *Nat. Rev. Phys.* **2**(6), 318–327 (2020).
18. M. Laurenzis et al., “Multiple-return single-photon counting of light in flight and sensing of non-line-of-sight objects at shortwave infrared wavelengths,” *Opt. Lett.* **40**(20), 4815–4818 (2015).
19. A. Ingle, A. Velten, and M. Gupta, “High flux passive imaging with single-photon sensors,” in *Proc. IEEE/CVF Conf. Comput. Vision and Pattern Recognit.*, pp. 6760–6769 (2019).
20. M. Laurenzis, “Single photon range, intensity and photon flux imaging with kilohertz frame rate and high dynamic range,” *Opt. Express* **27**(26), 38391–38403 (2019).
21. A. Ingle et al., “Passive inter-photon imaging,” in *Proc. of the IEEE/CVF Conf. on Comput. Vision and Pattern Recognit.*, pp. 8585–8595 (2021).

22. T. Seets et al., "Motion adaptive deblurring with single-photon cameras," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vision*, pp. 1945–1954 (2021).
23. Y. Altmann et al., "A bayesian approach to denoising of single-photon binary images," *IEEE Trans. Comput. Imaging* **3**(3), 460–471 (2017).
24. E. R. Fossum, "The quanta image sensor (QIS): concepts and challenges," in *Imaging Syst. and Appl.*, Optical Society of America, p. JTUE1 (2011).
25. E. R. Fossum, "Modeling the performance of single-bit and multi-bit quanta image sensors," *IEEE J. Electron Devices Soc.* **1**(9), 166–174 (2013).
26. Y. Chi et al., "Dynamic low-light imaging with quanta image sensors," *Lect. Notes Comput. Sci.* **12366**, 122–138 (2020).
27. S. Ma et al., "Quanta burst photography," *ACM Trans. Graphics* **39**(4), 79–1 (2020).
28. F. Yang et al., "Bits from photons: oversampled image acquisition using binary poisson statistics," *IEEE Trans. Image Process.* **21**(4), 1421–1436 (2011).
29. M. Laurenzis et al., "Passive imaging of single photon flux: strategies for de-noising, motion blur reduction and super-resolution up-scaling," *Proc. SPIE* **11868**, 1186805 (2021).
30. L. Carrara et al., "A gamma, x-ray and high energy proton radiation-tolerant cis for space applications," in *IEEE Int. Solid-State Circuits Conf.-Digest of Tech. Pap.*, IEEE, pp. 40–41 (2009).
31. M. Kobayashi et al., "A 1.8 e⁻ _{rms} temporal noise over 110-db-dynamic range 3.4 μm pixel pitch global-shutter CMOS image sensor with dual-gain amplifiers SS-ADC, light guide structure, and multiple-accumulation shutter," *IEEE J. Solid-State Circuits* **53**(1), 219–228 (2017).
32. O. Daigle et al., "Extreme faint flux imaging with an emccd," *Publ. Astron. Soc. Pac.* **121**(882), 866 (2009).
33. L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," *Lect. Notes Comput. Sci.* **6311**, 157–170 (2010).
34. J. Pan et al., "Kernel estimation from salient structure for robust motion deblurring," *Signal Process.: Image Commun.* **28**(9), 1156–1170 (2013).
35. A. Chakrabarti, "A neural approach to blind motion deblurring," *Lect. Notes Comput. Sci.* **9907**, 221–235 (2016).
36. S. Kim, N. K. Bose, and H. M. Valenzuela, "Recursive reconstruction of high resolution image from noisy undersampled multiframes," *IEEE Trans. Acoust. Speech Signal Process.* **38**(6), 1013–1027 (1990).
37. S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Process. Mag.* **20**(3), 21–36 (2003).
38. M. Ben-Ezra, A. Zomet, and S. K. Nayar, "Jitter camera: high resolution video from a low resolution detector," in *Proc. 2004 IEEE Comput. Soc. Conf. Comput. Vision and Pattern Recognit., CVPR 2004.*, IEEE, Vol. 2, p. II-II (2004).
39. M. Ben-Ezra, A. Zomet, and S. K. Nayar, "Video super-resolution using controlled subpixel detector shifts," *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(6), 977–987 (2005).
40. B. Basile, A. Blake, and A. Zisserman, "Motion deblurring and super-resolution from an image sequence," *Lect. Notes Comput. Sci.* **1065**, 571–582 (1996).
41. S. K. Nayar and M. Ben-Ezra, "Motion-based motion deblurring," *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(6), 689–698 (2004).
42. C. Dong et al., "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.* **38**(2), 295–307 (2016).
43. C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," *Lect. Notes Comput. Sci.* **9906**, 391–407 (2016).
44. W.-S. Lai et al., "Fast and accurate image super-resolution with deep Laplacian pyramid networks," *IEEE Trans. Pattern Anal. Mach. Intell.* **41**(11), 2599–2613 (2019).
45. Y. Zhang et al., "Residual dense network for image super-resolution," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit. (CVPR)* (2018).
46. C. Ledig et al., "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit. (CVPR)* (2017).
47. X. Wang et al., "ESRGAN: enhanced super-resolution generative adversarial networks," in *Proc. Eur. Conf. Comput. Vision (ECCV) Workshops* (2018).

48. B. Lim et al., “Enhanced deep residual networks for single image super-resolution,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit. Workshops*, pp. 136–144 (2017).
49. W. Shi et al., “Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 1874–1883 (2016).
50. M. Ayazoglu, “Extremely lightweight quantization robust real-time single-image super resolution for mobile devices,” in *Proc. IEEE/CVF Conf. Comput. Vision and Pattern Recognit.*, pp. 2472–2479 (2021).
51. Photon Force Ltd., “Pf32 user manual, v1.5.14,” 2021, <https://www.photon-force.com/> (accessed 2021-04-28).
52. MATLAB, *Version 9.9.0.1495850 (R2020b)*, The MathWorks Inc., Natick, Massachusetts (2020).
53. G. Van Rossum and F. L. Drake Jr., *Python Tutorial*, Vol. **620**, Centrum voor Wiskunde en Informatica, Amsterdam (1995).
54. G. Bradski, “The OpenCV library,” *Dr. Dobb’s J.: Software Tools for the Professional Programmer* **25**(11), 120–123 (2000).
55. OpenCV, “Open source computer vision library,” 2015, <https://github.com/opencv/opencv> (accessed 22 June 2022).
56. A. Azzalini and A. Capitanio, “Statistical applications of the multivariate skew normal distribution,” *J. R. Stat. Soc.: Ser. B (Stat. Methodol.)* **61**(3), 579–602 (1999).
57. L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Phys. D: Nonlinear Phenom.* **60**(1–4), 259–268 (1992).
58. D. Strong and T. Chan, “Edge-preserving and scale-dependent properties of total variation regularization,” *Inverse Prob.* **19**(6), S165 (2003).
59. L. Condat, “A direct algorithm for 1-D total variation denoising,” *IEEE Signal Process. Lett.* **20**(11), 1054–1057 (2013).
60. Y. Liu, “1D-MCTV-denoising,” 2018, <https://github.com/MrCredulous/1D-MCTV-Denoising> (accessed 28 April 2021).
61. H. Du and Y. Liu, “Minmax-concave total variation denoising,” *Signal, Image Video Process.* **12**(6), 1027–1034 (2018).
62. C. Truong, L. Oudre, and N. Vayatis, “Selective review of offline change point detection methods,” *Signal Process.* **167**, 107299 (2020).
63. A. Celisse et al., “New efficient algorithms for multiple change-point detection with reproducing kernels,” *Computa. Stat. Data Anal.* **128**, 200–220 (2018).
64. S. Arlot, A. Celisse, and Z. Harchaoui, “A kernel multiple change-point algorithm via model selection,” *J. Mach. Learn. Res.* **20**(162), 1–56 (2019).
65. R. Killick, P. Fearnhead, and I. A. Eckley, “Optimal detection of changepoints with a linear computational cost,” *J. Am. Stat. Assoc.* **107**(500), 1590–1598 (2012).
66. “Learn: anatomy of a typeface,” <http://typedia.com/learn/only/anatomy-of-a-typeface/> (accessed June 2021).
67. Z. Wang and A. C. Bovik, “A universal image quality index,” *IEEE Signal Process. Lett.* **9**(3), 81–84 (2002).
68. Z. Wang et al., “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. Image Process.* **13**(4), 600–612 (2004).

Martin Laurenzis received his PhD in electrical engineering and information technologies from the RWTH Aachen University, Germany, in 2005, and his MSc degree in physics from the Technical University of Dortmund, Germany, in 1999. Since 2004, he has been with the French-German Research Institute of Saint-Louis, France. As a principal scientist in the Advanced Visionics and Processing group, he is responsible for international and national research projects. His main research interests are in the fields of active and passive imaging as well as in computational imaging. He is the member of SPIE, the European Optical Society (EOS), and the German Society for Applied Optics (DGaO).

Trevor Seets received his BS degree in electrical engineering from the University of Wisconsin-Madison in 2019. He is now a graduate student under Professor Andreas Velten at Madison. His research interests include computational optics, statistical signal processing, and imaging.

Emmanuel Bacher is a research engineer at the Franco-German Research Institute of Saint-Louis (ISL), France. He received his Higher National Diploma in electrical and industrial computer science from the University of Haute-Alsace in Mulhouse, France, in 1993, and his MSc degree in experimental physics from the Conservatoire national des arts et métiers, France, in 2001. He is a team member in the Advanced Visionics and Processing group and is responsible for conducting experimental trials, automating experimental equipment, and designing integrated electronic circuits.

Atul Ingle received his PhD in electrical engineering from the University of Wisconsin-Madison in 2015. He was a visiting R&D scientist at Philips Healthcare in Andover, his MA degree in 2013 and 2014, and a research scientist at Fitbit, Inc., in Boston, Massachusetts, in 2016 to 2017. He is currently a postdoctoral researcher in the Departments Biostatistics and Computer Sciences at UW-Madison. His research interests include computational imaging, computer vision, statistical signal processing, and medical imaging.

Andreas Velten is an assistant professor at the Department of Biostatistics and Medical Informatics and the Department of Electrical and Computer Engineering at the University of Wisconsin-Madison and directs the Computational Optics Group. He obtained his PhD in physics from the University of New Mexico in Albuquerque and was a postdoctoral associate of the Camera Culture Group at the MIT Media Lab. His research focuses on applied computational optics and imaging.