# Iterative Image Reconstruction for Quanta Image Sensor by using Variance-based Motion Estimation

Kiyotaka Iwabuchi, Tomohiro Yamazaki and Takayuki Hamamoto
Graduate School of Engineering, Tokyo University of Science
6-3-1 Niijuku, Katsushika, Tokyo, 125-8585 Japan
E-mail: iwabuchi@isl.ee.kagu.tus.ac.jp TEL/FAX: +81-3-5876-1717/+81-3-5876-1614

## ABSTRACT

This paper proposes an image reconstruction method for QIS imaging of moving objects. Additionally, we also propose a new motion estimation technique that is able to estimate the motion of moving objects even in extremely low-light conditions with high estimation accuracy.

## I. INTRODUCTION

Quanta Image Sensor (QIS), which was first proposed in 2005 [1], is an imaging device that counts the number of photons arriving at each pixel [2,3]. By single-photon sensitivity, the QIS can acquire a clear image with high dynamic range even in low flux [4]. The QIS is expected to be utilized for applications in imaging life sciences, aerospace, defense, etc. Information of incident photons into a QIS is recorded as "bit-plane" images composed of binary pixel values that represent that whether a photon is detected or not. A bit-plane has only binary values; therefore, they are transformed into multi-bit images by simple summation in time (and/or space) or statistical reconstruction methods [5,6].

Using the large number of bit-planes can achieve high signal-to-noise (S/N) ratio of a reconstructed image. However, when moving objects are present, motion blur is likely to occur on the reconstructed image. To reduce the motion blur, an image reconstruction method from bit-planes with moving objects is proposed by [7]. Additionally, they also proposed a motion estimation method for QIS imaging. However, accuracy of their motion estimation is degraded as light conditions become dark. Therefore, conventional method is able to reduce motion blur clearly in limited flux only.

In this paper, we propose a reconstruction method for imaging of moving objects. Our method utilizes new motion estimation technique that can estimate the motion of moving objects with high accuracy. As a result, we can reconstruct a clear image with any motion blur, which is not possible with the conventional method. We are also able to deal with situations with several overlapping objects moving at different speeds. Generally, it is difficult to reduce the motion blur from each object. However, we can solve this problem by using our method in iterative manner.

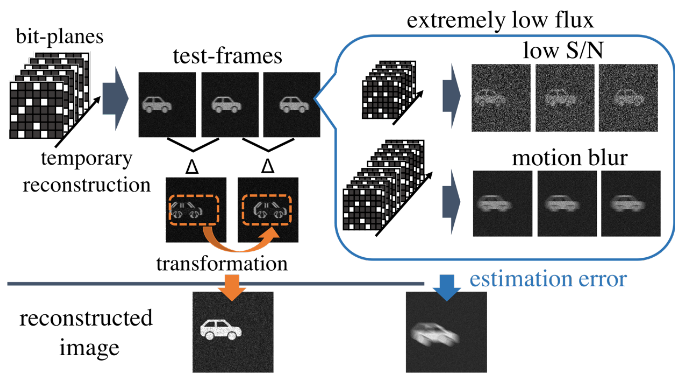


Fig. 1 Conventional motion estimation.

## II. CONVENTIONAL METHOD

In the conventional reconstruction method [7] for imaging of moving objects by QIS, the positions of bit-planes are shifted so that objects' positions aligned, which makes it possible to reconstruct free motion blur image of moving objects. Before shifting the position of bit-planes, it is required to estimate the motion of the moving objects. A bit-plane image is composed of binary pixel values; therefore, it is not easy to apply it to existing motion estimation method, whose target is multi-bit (e.g. 8bit) images. Therefore, some temporary multi-bit images ("test-frames") are reconstructed from bit-planes to apply the existing motion estimation method in [7] (Fig. 1).

Specifically, several test-frames are generated. Additionally, several "difference-frames" are generated from continuous test-frames by subtraction. To estimate the motion of moving objects, a geometric transformation that aligns regions of moving objects on two continuous difference-frames is calculated. Positions of the bit-planes eventually shifted based on the estimated transformation.

In extremely low flux, the number of incident photon decrease; therefore, S/N ratio in reconstructed images will also decrease. This makes it difficult to estimate precise motion of moving objects. Using large number of bit-planes for reconstruction can increase the S/N ratio on test-frames. However, a large motion blur is likely to occur on test-frames, which also makes it difficult to estimate

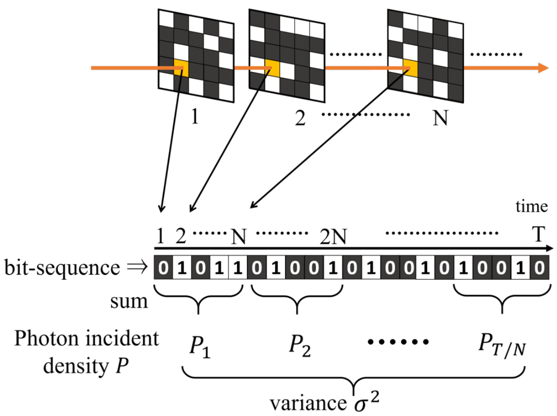



Fig. 2 Motion detection.

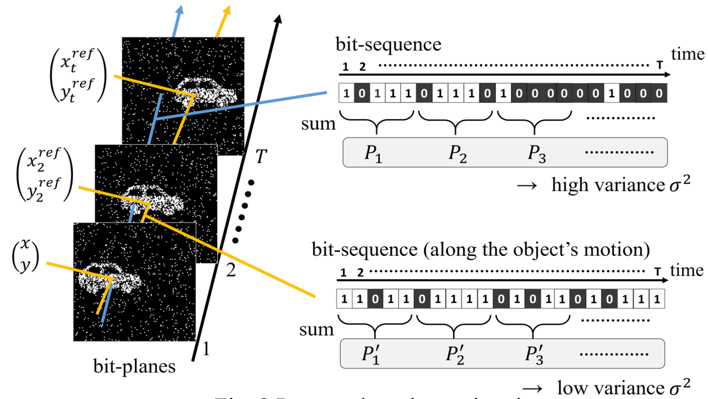



Fig. 3 Proposed motion estimation.

motion. Thus, due to this trade-off between S/N ratio and motion blur, accuracy of their motion estimation is degraded as light conditions become dark. Therefore, in the motion estimation process from bit-planes, generating temporary multi-bit images to apply them to existing motion estimation method is not the best way to accurately estimate motion.

## III. PROPOSED METHOD

For QIS imaging of moving objects, we propose an image reconstruction method with highly accurate motion estimation. Our method is able to reconstruct an image with high S/N ratio and free motion blur even in extremely low flux. In the following section, we explain our reconstruction process using $T$ frames of bit-planes.

We first detect the dynamic area in a moving object using a motion detection method from the bit-planes proposed by [8]. In [8], the dynamic area is statistically detected from the sequence of binary pixel values from each bit-plane (we term this sequence of binary pixel values as "bit-sequence"). Particularly, "photon incident density" ($P$) is calculated by summing several bit-pixels in each group of bit-planes. In Fig. 2, each $P$ is calculated from $N$ bit-pixels. Then, from some $P$ in a bit-sequence, the variance ($\sigma^2$) of each spatial pixel of bit-planes is calculated. $\sigma^2$ of each pixel represents how much the number of incident photons is changed over time. At the pixel corresponding to the moving object, $\sigma^2$ tends to shift to a high value. This is because the incident light intensity (equal to the number of incident photons) changes in the exposure time owing to the object's motion. Therefore, we regard pixels having high $\sigma^2$ as "dynamic" and generate a binary map for dynamic area. To distinguish multiple moving objects, we identify each object as an isolated dynamic area that is composed of some "dynamic" pixels; they are connected in space on the map. In the next step, we estimate the motion of objects in the units of isolated dynamic area. In our estimation process, to estimate the motion statistically, we again consider the "bit-sequence" (Fig .3).

In the normal bit-sequence, bit values show high variance in terms of time. However, in the specific bit-sequence achieved by shifting the position of the target pixels at each bit-plane along the object's movement, the binary values show low variance. Therefore, we estimate the motion by detecting a shift amount that minimizes the variance of bit values in a bit-sequence. Particularly, we estimate the optimal parameters of (1) by search. These parameters minimize the variance of the bit-sequence from target pixels $(x^{ref}, y^{ref})$ from (1).

$$\begin{pmatrix} x_t^{ref} \\ y_t^{ref} \end{pmatrix} = \Delta\lambda_t \begin{pmatrix} \cos\Delta\theta_t & -\sin\Delta\theta_t \\ \sin\Delta\theta_t & \cos\Delta\theta_t \end{pmatrix} \begin{pmatrix} x - X_O \\ y - Y_O \end{pmatrix} + \begin{pmatrix} \Delta x_t \\ \Delta y_t \end{pmatrix} + \begin{pmatrix} X_O \\ Y_O \end{pmatrix}$$

$$\begin{aligned} \Delta\lambda_t &= 1 + d\lambda(t-n) \quad & \Delta\theta_t &= d\theta(t-n) \\ \Delta x_t &= dx(t-n) \quad & \Delta y_t &= dy(t-n) \end{aligned} \qquad (1)$$

where $d\lambda, d\theta, dx\ and\ dy$ represent the scaling factor, rotation factor, horizontal translations and vertical translations, respectively. These are relative parameters based on the $n^{\text{th}}$ bit-plane, which is the reference frame for arranging the position of a moving object in the bit-planes. Here, $(X_O, Y_O)$ is the centroid position of the moving object, which can be estimated from the dynamic area map. $t$ is the frame number within $T$ bit-planes. We estimate the parameters in the units of isolated dynamic area. The estimated parameters of each area minimize the sum total of the variance within each isolated dynamic area. Our method can estimate motion without generating any temporary reconstructed images. Therefore, we can avoid the trade-off (between S/N ratio and motion blur) that can reduce estimation accuracy.

In summary, we reduce the motion blur by motion compensation that replaces the binary pixel values' constitution of each bit-plane by using estimated motion parameters. When we replace the pixel values of the $t^{\text{th}}$

bit-plane, we replace the pixel value of $(x, y)$ by that of $(x_t^{ref}, y_t^{ref})$ from (1). This replacement aligns the object's position among the continuous bit-planes. In the image reconstruction using new bit-planes with motion compensation, we regard the area with the moving object as "static" area. Therefore, we can reconstruct a clear image without motion blur.

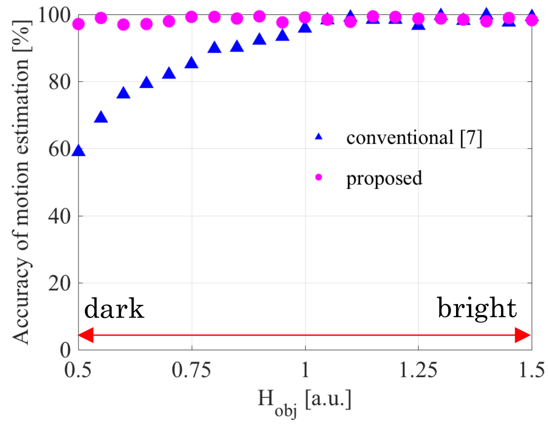
If several overlapping objects are moving at different velocities, it is difficult to reduce every motion blur because the objects are detected as a "single object." In this case, we cannot estimate the motion of each object, which is necessary for precise motion compensation. Our method can also deal with this problem by iterative process. Particularly, we feedback the bit-planes as revised input bit-planes and iterate our method from the point of detecting the dynamic area. When bit-planes continue to have insufficient area about motion compensation after one process, we can detect this area as dynamic area in next process. Therefore, we can gradually reduce the motion blur of each object by using an iterative approach with our process.

## IV. SIMULATION

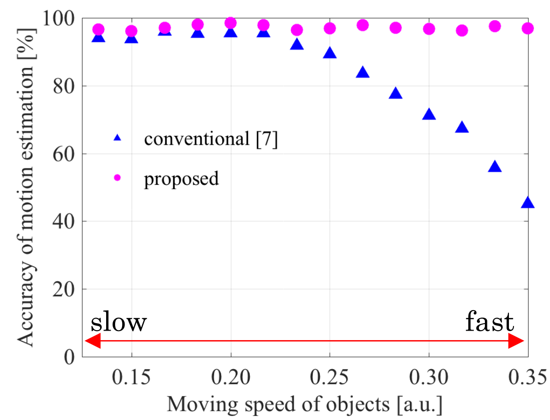
We present the simulation results of the comparison between our method and the conventional method [7].

We generated the bit-plane images by calculating incident photons at each pixel by using a random function that follows a Poisson distribution. Additionally, we assumed that the QIS pixel outputs "1" (if there are one or more photons incident) or "0" (if there are no photons incident) at 10 kfps. We used 240 bit-planes, where the image size of a single bit-plane is $256 \times 256$. In the conventional method, we assumed that a single test-frame is generated by summation of bit values from 80 bit-planes. The spatial kernel size $(N_x \times N_y)$ for generating a pixel of single test-frame is (4×4). Then we used 3 test-frames generated from 240 bit-planes for the motion estimation by using MATLAB's function: "imregtform" as well as [7] used. In our method, we calculated "photon incident density"$(P)$ in the units of (4×4) pixels block in space and 10 bit-planes in time when we estimated the motion of objects. We employed the same ideal dynamic area for each estimation method.

First, we evaluated the accuracy of motion estimation in each method, and then compared them. We calculated the estimation accuracy from the difference between "true movement" and "estimated movement." However, the pattern of photon detection in bit-planes is random and different every time. Result of motion estimation depends on the pattern of photon detection in bit-planes.




(a)

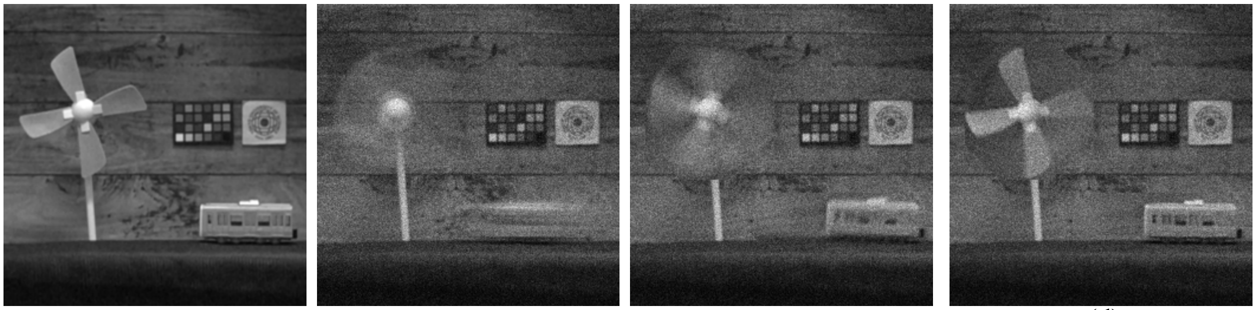



(b)

Fig. 4 Accuracy of motion estimation: (a) and (b) show estimation accuracy as flux and speed of moving object.

Therefore, we generate 10 sets of 240 bit-planes for one scene that have different patterns of photon detection. Then, we estimate the translational movement of an object 10 times by using each motion estimation method. We determine the average value of the estimation as estimation accuracy. Fig. 4 (a) shows the accuracy of motion estimation as flux of moving objects. Particularly, $H_{obj}$ represents the number of incident photons during the exposure time of the single bit-plane. The flux rate of background to object is 80%. An object moves 40 pixels during continuous 240 bit-planes. In Fig. 4 (a), it can be seen that our method has improved accuracy of motion estimation compared to [7]'s especially dark environment.

Fig. 4 (b) also shows the estimation accuracy as the moving speed of an object. $H_{obj}$ is 0.5 and fixed. In Fig. 4 (b), the value of horizontal axis (moving speed of objects) represents how large the object moves during continuous 240 bit-planes. The estimation accuracy of [7] degrades as the object moves to a large extent. This is because the motion blur in test-frames become large. Moreover, by avoiding the trade-off between S/N ration



(a) PNSR | (b) 22.81 dB | (c) 21.80 dB | (d) 24.23 dB

Fig. 5 Reconstructed images: (a) is reference frame, which is 1st frame of gray scale images for generating bit-planes, (b) is a simple reconstructed image without motion compensation, (c) is [7]'s reconstructed image, (d) is ours.

and motion blur, our method can estimate the motion of the object accurately in case of large moving object.

Further, we show a reconstructed image of each method in Fig. 5. In this simulation, we use 240 bit-planes generated from the light intensity information of 8-bit grayscale images captured by high speed camera at 1000 fps. We set the number of incident photons as 0.5 photon/pixel/bit-plane on average. Fig. 5(a) is 1st frame of reference images for generating bit-planes. Fig. 5(b) is a simple reconstructed image without motion compensation. Figs. 5(c) and (d) are those of [7] and our reconstructed images, respectively. In Fig. 5(c), the reduction of motion blur is insufficient owing to incorrect motion estimation. In contrast, our method satisfactorily reduced motion blur. The details of the objects can be seen in Fig. 5(d). Additionally, the image has the highest PSNR value.

Through an example it can been seen that we can reduce motion blur by an iterative process overlapping objects are present. Fig. 6(a) is a simple reconstructed image without motion compensation. Figs. 6(b) and (c) are our method's reconstructed images. They show the result of motion blur reduction once and twice, respectively. Fig. 6(b) shows the motion blur of a "Fast Object" because it overlaps with a "Slow Object". However, in Fig. 6(c), we succeeded to reduce the motion blur of the "Fast Object." We can treat the "Fast Object" mainly in the second process.

## V. CONCLUSION

In this paper, we proposed an efficient motion blur reduction method with variance-based motion estimation, which is appropriate to bit-plane images captured by the QIS. Our motion estimation achieved high estimation accuracy even in extremely low flux by avoiding the trade-off between S/N ratio and motion blur. As a result, we can obtain highly clear reconstructed images compared to conventional method. We also deal with the situation with several overlapping objects moving at different speeds by the iterative process of our method.


(a) (b) (c)

Fig. 6 Effect of iteration: (a) without motion compensation, (b) without iteration, (c) with iteration(twice).

## REFERENCES


[1] Fossum, E. "What to do with sub-diffraction-limit (SDL) pixels?– A proposal for a gigapixel digital film sensor (DFS)," in *Proc. IEEE Workshop CCDs Adv. Image Sensors*, 2015.

[2] Jiaju, M. *et al.*, "Photon-number-resolving megapixel image sensor at room temperature without avalanche gain," *Optica*, vol. 4, no. 12, pp. 1474-1481, 2017.

[3] Dutton, N. A. W. *et al.*, "A SPAD-Based QVGA Image Sensor for Single-Photon Counting and Quanta Imaging," *IEEE Transactions on Electron Devices*, vol. 63, no. 1, pp. 189-196, 2016.

[4] Dutton, N. A. W. *et al.*, "Extending the Dynamic Range of Over sampled Binary SPAD Image Sensors," in *Proc. International Image Sensor Workshop,* 2017.

[5] Yang, F. *et al.*, "Bits From Photos: Oversampled Image Acquisition Using Binary Poisson Statics," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1421-1436, 2012.

[6] Elgendy, O. A. *et al.*, "Image reconstruction and threshold design for Quanta Image Sensors," in *Proc. IEEE International Conference on Image Processing*, 2016

[7] Gyongy, I. *et al.*, "Object Tracking and Reconstruction with Quanta image Sensor," in *Proc. International Image Sensor Workshop,* 2017.

[8] Gyongy, I. *et al.*, "Bit-plane processing techniques for low-light, high speed imaging with SPAD-based QIS," in *Proc. International Image Sensor Workshop,* 2015.