https://doi.org/10.2352/ISSN.2470-1173.2017.11.IMSE-190
© 2017, Society for Imaging Science and Technology

RTS and photon shot noise reduction based on maximum likelihood estimate with multi-aperture optics and semi-photon-counting-level CMOS image sensors

Haruki Ishida[a], Keiichiro Kagawa[b], Min-Woong Seo[b], Takashi Komuro[c], Bo Zhang[b], Taishi Takasawa[b], Keita Yasutomi[b] and Shoji Kawahito[b];
[a]*Department of Engineering, Shizuoka University, 3-5-1, Johoku, Nakaku, Hamamatu, 432-8011, Japan*
[b]*Research Institute of Electronics, Shizuoka University, 3-5-1, Johoku, Nakaku, Hamamatu, 432-8011, Japan*
[c]*Graduate School of Science and Engineering, Saitama University, 255, Simookubo, Sakuraku, Saitama, 338-8570, Japan*

Abstract
This paper presents an effective noise reduction method for both RTS noise and photon shot noise based on a multi-camera configuration composed of multi-aperture optics and a semi-photon-counting-level CMOS image sensor. We model noise characteristics of the CMOS image sensor with a probability distribution based on measured read noise including RTS noise and shot noise. We estimate the average number of photons by the maximum likelihood estimation with all corresponding pixels in the multi-aperture optics. We compared the noise reduction performance of several schemes when the incident photon number was assumed to be 2 for each aperture. 2 out of 9 apertures showed RTS noise. Simple averaging gave effective noise of 0.68 e^-_{RMS}, where RTS noise still exist. Selective averaging, which minimizes synthetic sensor noise, gave that of 0.60 e^-_{RMS}. Although RTS noise was removed, photon shot noise was less suppressed because only 7 apertures were considered in averaging. On the other hand, with the maximum likelihood estimation, the effective noise became 0.48e^-_{RMS}, and both RTS noise and photon shot noise were reduced.

Introduction
Low light imaging is required in various fields such as bio-imaging, surveillance cameras, and astronomical measurement, where fast optics and ultra-low-noise image sensors are inevitable. Recently, semi-photon-counting-level CMOS image sensors based on high conversion gain floating diffusions have emerged [1-3]. However, extremely large noise called random telegraph signal (RTS) noise is becoming more fatal as the source-follower transistor scales down [4,5]. In the low light imaging, RTS noise becomes more visible and degrades the image quality because a large digital gain that enhances visibility of noise is necessary. In addition, fast optics with an F-number much smaller than unity makes physical embodiment very hard. Because, not only it becomes extremely huge and heavy, also spatial resolution decreases due to aberration. Moreover, a depth of field becomes shallow. In our previous research, we proposed a selective averaging method based on the redundancy of the multi-aperture optical system to eliminate the RTS noise and demonstrated its effectiveness [6]. However, this method suffers from photon shot noise when semi-photon-counting-level CMOS image sensors are used. Because the number of pixels used in averaging is smaller than that of the apertures, photon shot noise can be relatively larger than in the simple averaging where all the apertures are considered. Reduction of both RTS noise and photon shot noise is necessary.

In this work, we use a multi-aperture optical system in order to realize a fast lens which is much smaller than F/1. Moreover, we model the sensor noise of each pixel as a probability distribution of read noise based on measurement, where RTS noise is included. By using the maximum likelihood estimation [7], we attempt to decrease both RTS noise and photon shot noise. In the simulation, we confirmed that maximum likelihood estimation shows the best noise reduction capability compared with simple averaging and selective averaging.

Semi-photon-counting-level CMOS image Sensor
The fast lens and low-noise image sensors are essential for low light imaging. In general, in the CMOS image sensors, 4-transistor pixels with the pinned photodiode is used for high image quality. In this pixel structure, the parasitic capacitances generated between the floating diffusion (FD) and a transfer gate (TG) node and between the FD and a reset gate (RG) node are added to the capacitance of the FD. It is necessary to reduce these parasitic capacitances around the FD node to achieve ultra-high sensitivity. However, it is difficult in most cases. Figure 1 shows a cross-sectional view of the high conversion gain (HCG) pixel structure which is developed in our laboratory.

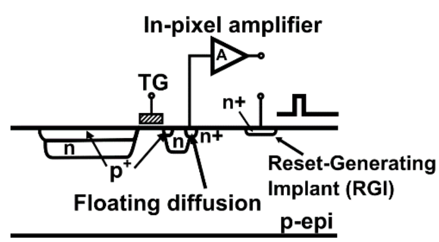

Figure 1. High conversion gain pixel structure.

This pixel structure enables to drastically reduce the parasitic capacitance using two in-pixel technologies. The first method reduces the influence of the parasitic capacitance between the FD and TG node [8]. By inserting p^+ layer at the boundary between the FD and TG node and making it fully depleted, the parasitic capacitance between the FD and TG nodes can be minimized. The second method completely removes the parasitic capacitance between the FD and RG node because this HCG pixel structure does not use a reset transistor [3]. The pixel reset operation is implemented by an implanted n^+ layer located close to the FD node. We call this implanted n^+ Reset-Generating-Implant (RGI). When a high voltage is applied, a barrier between the FD node and RGI is lowered by punch-through. Then, the FD is soft-reset. As a result, the effective capacitance of the FD drastically decreased and high conversion gain of 220μV/e^- was achieved using 0.11μm CIS process [3]. Furthermore, by combining this pixel with correlated multiple sampling (CMS) [9], we realized extremely low read noise of 0.27 e^-_{RMS} of semi-photon-counting-level.

Multi-aperture camera

No matter how the image sensor becomes high sensitivity, if signal amount is small, it suffers from large shot noise. Therefore, ultra-high sensitivity low noise cameras require a fast lens with an F-number much smaller than unity. However, such cameras will be huge and heavy because a small F-number lens has an extremely large pupil. Furthermore, it is required to correct huge aberration. Therefore, it is almost impossible to embody with single-aperture optics. Figure 2 shows a multi-aperture camera system structure. Multi-aperture cameras have some lenses and corresponding CMOS image sensors. The synthetic F-number (F_M) in the multi-aperture camera system is represented by the following equation.

$$F_M = F_0/\sqrt{M} \tag{1}$$

Where F_0 is the F-number of the elemental lens and M is the number of the apertures. Therefore, we can realize a fast lens whose F-number is smaller than 1 by using multiple moderately fast lens. In addition, the multi-aperture camera has functionality such as estimating a three-dimensional shape from disparity information [10], increasing signal amount, and reducing noise by merging redundant images obtained from multiple apertures. It is difficult to reduce the influence of RTS noise with the image sensor alone. However, in this study, if there are pixels which generate RTS noise, we can remove the noise by use of the correlation among the corresponding pixels.

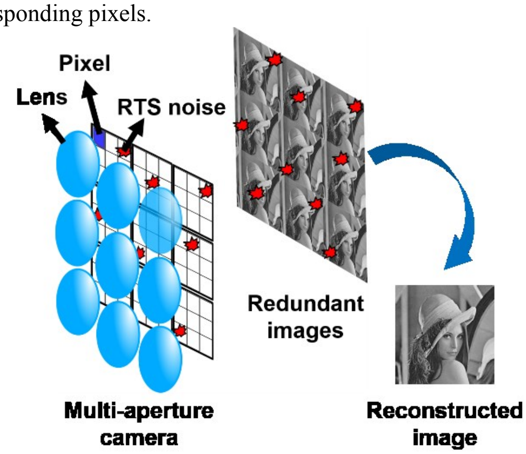

Figure 2. Structure of multi-aperture camera system.

Modeling of RTS noise and maximum likelihood estimation

In the CMOS image sensors, there are pixels that generate RTS noise. Since RTS noise appears as very large noise at low frequencies, simple averaging has a less noise reduction effect. Previously, we proposed an averaging method that we call selective averaging. This method adaptively removes pixels that generate RTS noise due to minimizing synthetic sensor noise. Here, synthetic sensor noise is calculated by using variance of the each pixel value in many frames. Although this method has a large noise reduction effect for RTS noise, the number of apertures that is used for averaging decreases, so that the noise reduction effect for photon shot noise becomes impaired.

On the other hand, by applying maximum likelihood estimation to multi-aperture camera, we can expect a reduction effect for both of them. Maximum likelihood estimation is a method that estimates parameters based on statistics. In this study, we consider that the pixel outputs of each pixel obtained from the image sensors are the result that gives the maximum probability. Maximum likelihood estimation is performed in three steps. Step-1) we model sensor noise for each pixel as a conditional probability density distribution $p(x|\lambda)$, where x is observed value and λ is the number of electrons that we want to find. Step-2) we calculate the likelihood function from the modeled probability density function. The likelihood function is the product of the probability density for each pixel output. Step-3) step-2 is repeated by changing λ to find the optimal λ that gives the maximum likelihood. Figure 3 shows the flow chart of the proposed method. Firstly, we capture the dark images of many frames with a multi-aperture camera. The sensor noise of each pixel is modeled as a probability density distribution of the dark read noise based on measurement, where RTS noise is included. Photon shot noise is modeled as a Poisson distribution, and convolved with the read noise distribution. We fitted pixel by pixel of each aperture with the probability density function expressed the following equation.

$$p(x|\lambda) = \sum_{s=0}^{m} \sum_{k=0}^{n} \alpha_s \cdot \frac{\lambda^k}{k!} exp(-\lambda) \cdot \frac{1}{\sqrt{2\pi}\sigma} exp\left\{-\frac{(x-k-r_s)^2}{2\sigma^2}\right\} \tag{2}$$

$$L(\lambda) = \prod_{j=1}^{N} p(x|\lambda) \tag{3}$$

Where s is number of levels of RTS noise, α_s is the peak value of the noise histogram at each noise level, λ is average number of incident photons, σ is standard deviation of the sensor's read noise without RTS noise and photon shot noise, x is the pixel output of the semi-photon-counting-level CMOS image sensor, which is a measured value at the time of shooting, and r_s is the position of each noise level of RTS noise. We assume that the same amount of photons are incident in average to the corresponding pixels of the multi-aperture image. Eq. 2 can represent the basic sensor noise without RTS noise, RTS noise, and photon shot noise. After shooting, we change the value of λ in Eq. 2 within a certain range and calculate the likelihood function $L(\lambda)$ by the Eq. 3. Then, λ which maximizes the likelihood function $L(\lambda)$ becomes the estimated value of the pixel in the reconstructed image.

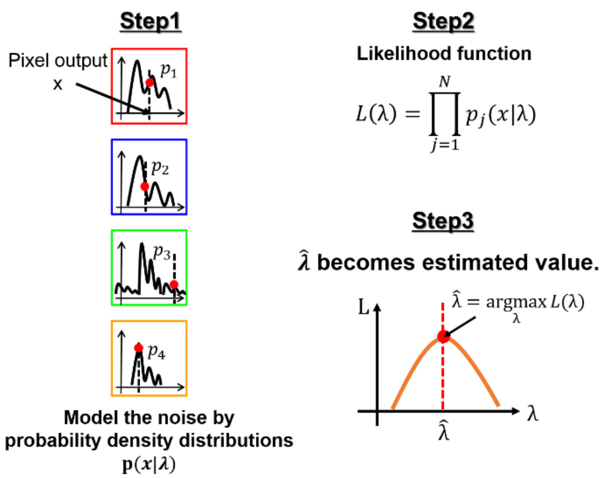

Figure 3. Flow chart of the proposed method.

Simulation results by using measured noise

To confirm the effectiveness, we conducted a simulation with 5000 dark images measured by a semi-photon-counting-level CMOS image sensor. In order to suppress the dark current shot noise, the measurement was carried out with cooling the sensor at about -10℃. We assumed a 3×3 multi-aperture camera, and one pixel per aperture was considered for preliminary verification. When the quantum efficiency is 100%, an average incident photon number λ is equivalent to the number of electrons of the sensor. When the photon shot noise with an average incident photon number λ of 2.0 e^- was added to the measured readout noise, we simulated the noise which is generated pixel by pixel. 2 out of 9 pixels showed RTS noise, and the average read noise for these pixels were 3.41e^-_{RMS}. That for the other pixels was 0.31 e^-_{RMS}. Figures 4 (a), (b) show an example of the probability density distribution such as with and without RTS noise which is real measured by semi-photon-counting-level CMOS image sensor. Fig. 4 (a) shows pixel output which does not include RTS noise and Fig. 4 (b) shows pixel output which include RTS noise. Then, some parameters such as σ, α_s and r_s are extracted from these histograms. Figures 4 (c) and (d) are probability density distributions modeled by Eq. 2. In order to calculate the likelihood, we calculate the probability density for each λ. To estimate the pixel value by the maximum likelihood estimation (MLE), we changed the average incident photon number λ with a search step width $\Delta\lambda = 0.1$ e^- in the range of $0 \leq \lambda \leq 4$ to calculate the likelihood L. λ that gives the maximum likelihood is the estimated value of the pixel of the reconstructed image. Figure 5 shows the relationship between λ and the likelihood function $L(\lambda)$ in one frame. In the frame without influence of RTS noise (Fig. 5 (a)), the result of the maximum likelihood estimation almost agrees with the simple averaging. On the other hand, the frame that the influence of RTS noise is large (Fig. 5 (b)) differs greatly each other. In the simple averaging, the estimated value is largely deviated from 2.0 e^- because the influence of RTS noise is not negligible. On the other hand, in maximum likelihood estimation, the level of RTS noise is probabilistically considered. Therefore, the estimated value is little affected by the RTS noise.

Figure 6 shows the results of the same processing for 100 frames. With simple averaging, the effective noise became 0.68 e^-_{RMS}. With the selective averaging which minimizes only the synthetic sensor's read noise, the effective noise was 0.60 e^-_{RMS}. The maximum likelihood estimation showed the smallest noise, 0.48e^-_{RMS}, because this method reduced photon shot noise as well as RTS noise.

Conclusion

We applied maximum likelihood estimation to multi-aperture camera using semi-photon-counting-level CMOS image sensors. We modeled the noise of CMOS image sensors as conditional probability density distributions and attempted to reduce both of RTS and photon shot noise. We assumed a multi-aperture camera with 9 apertures and simulated the noise reduction capability using measured noise and calculated photon shot noise. In the simulation, we confirmed that the maximum likelihood estimation showed better noise reduction capability compared with the simple averaging and the selective averaging. With the simple averaging, the effective noise became 0.68 e^-_{RMS}. With the selective averaging method, the effective noise was 0.60 e^-_{RMS}. The maximum likelihood estimation showed the smallest noise, 0.48e^-_{RMS}. In the proposed method, the noise reduction capability was improved about 29% compared with simple averaging. This noise level is approximately equal to the theoretical value 0.47 e^-_{RMS} when considering only photon shot noise.

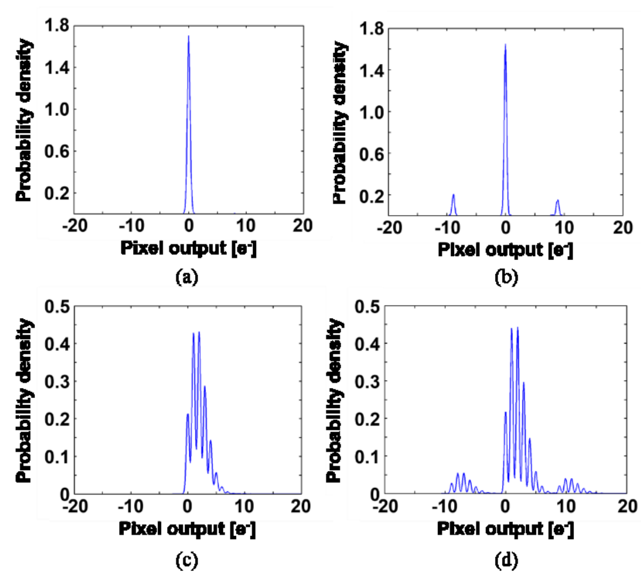

Figure 4. Probability density distributions of pixel output. Measured pixel outputs of the pixel that (a) does not generate and (b) generate RTS noise. The simulated values of the pixel that (c) does not generate and (d) generate RTS noise.

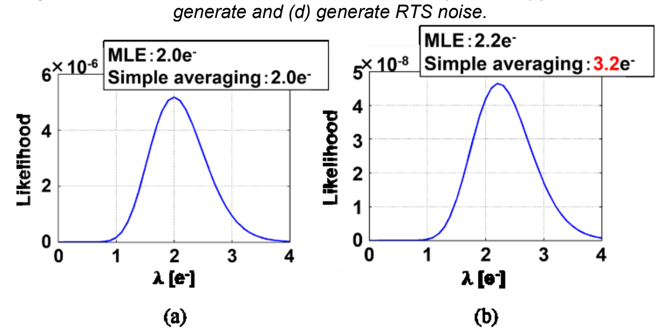

Figure 5. Examples of likelihood. (a) A frame without influence of RTS noise. (b) A frame strongly affected by RTS noise.

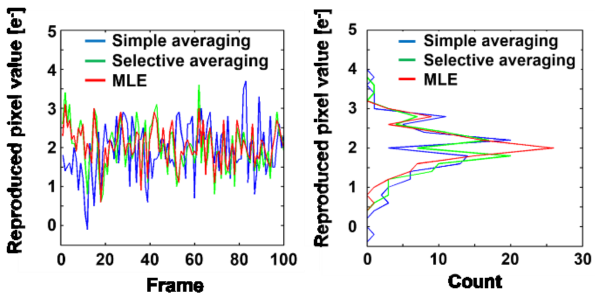

Figure 6. The noise reduction capability by maximum likelihood estimation (MLE). (a) Estimated value. (b) Noise histogram.

References

[1] J. Ma, E. R. Fossum: "Quanta Image Sensor Jot with Sub 0.3e-r.m.s. Read Noise and Photon Counting Capability," IEEE Electron Device Lett., vol. 36, no. 9, pp. 926-928 (2015).

[2] S. Wakashima, F. Kusuhara, R. Kuroda, and S. Sugawa: "A linear response single exposure CMOS image sensor with 0.5e- readout noise and 76ke- full well capacity," in symp. VLSI Circuits Digest of Technical papers, pp. 88-89 (2015).

[3] M. -W. Seo, S. Kawahito, K. Kagawa, K. Yasutomi: "A 0.27e-rms Read Noise 220- V/e- Conversion Gain Reset-Gate-Less CMOS Image Sensor With 0.11-μm CIS Process," IEEE Electron Device Lett., vol. 36, no. 12, pp. 1344-1347 (2015).

[4] C. Leyris, F. Martinez, M. Valenza, A. Hoffmann, J. C. Vildeuil, F. Roy: "Impact of Random Telegraph Signal in CMOS Image Sensors for Low-Light Levels," IEEE Solid-State Circuits Conference, pp. 376-379 (2006).

[5] H. Shimizu, H. Awano, M. Hiromoto, T. Sato: "Automation of Model Parameter Estimation for Random Telegraph," IEICE Trans Fundamentals, vol. E97-A, no.1, pp. 1-9 (2014).

[6] R. A. Fisher: "On the mathematical foundations of theoretical statistics," Philosophical Transactions of the Royal Society, vol. 222, pp. 309-368 (1922).

[7] B. Zhang, K. Kagawa, T. Takasawa, M. -W. Seo, K. Yasutomi, S. Kawahito: "RTS Noise and Dark Current White Defects Reduction Using Selective Averaging Based on a Multi-Aperture System," Sensors, vol. 14, no. 1, pp. 1528-1543 (2014).

[8] M. -W. Seo, T. Sawamoto, T. Akahori, T. Iida, T. Takasawa, K. Yasatomi, S. Kawahito: "A Low Noise Wide Dynamic Range CMOS Image Sensor With Low-Noise Transistors and 17b Column-Parallel ADCs," IEEE Sensors J., vol. 13, no. 8, pp. 2922-2929 (2012).

[9] M. -W. Seo, S. –H. Suh, T. Iida, T. Iida, T. Takasawa, K. Isobe, T. Watanabe, S. Itoh, K. Yasatomi, S. Kawahito: "A Low Noise High Intrascene Dynamic Range CMOS Image Sensor with a 13 to 19b Variable-Resolution Column-Parallel Folding-Integration/Cyclic ADC," IEEE Solid-State Circuits, vol. 47, no. 1, pp. 272-283 (2012).

[10] K. Kugenuma, T. Komuro, B. Zhang, K. Kagawa, S. Kawahito: "Depth Map Estimation and Restoration of an Improved-quality Image from Multi-aperture Images," ITE Technical Report, vol. 40, no. 40, pp. 3-4 (2016).

Author Biography

Haruki Ishida received his Bachelor of Engineering from the Shizuoka University (2016). Currently, he is a master course student at Department of Engineering, Shizuoka University.