Signal-to-Noise Ratio Enhancement in Cardiac Pulse Measurements Using Multi-Tap CMOS Image Sensors With In-Pixel Temporal Redundant Samplings

Signal-to-Noise Ratio Enhancement in Cardiac Pulse Measurements Using Multi-Tap CMOS Image Sensors With In-Pixel Temporal Redundant Samplings

Chen Cao, *Member, IEEE*, Masashi Hakamata, Keita Yasutomi, *Member, IEEE*, Keiichiro Kagawa, *Member, IEEE*, Satoshi Aoyama, *Member, IEEE*, Norimichi Tsumura, and Shoji Kawahito, *Fellow, IEEE*

***Abstract*—Insufficient signal-to-noise ratio (SNR) limits the quality of cardiac pulses measured by multi-tap CMOS image sensors (CISs). We present an in-pixel temporal redundant samplings (TRS) method for enhancing the SNR through sampling the charges originated by one individual optical scene into the multiple storage taps in turn with the identical time window, enabling magnification of the image signal in shot noise domain. Modeling indicates the linear SNR increases with a factor square root of TRS sampling number, which is also verified by characterization results utilizing a four-tap CIS. Application results show that the high frequency (HF) noise components observed in the measured cardiac pulses can be attenuated by 67.9%, 48% and 53.6% with heart rate detection accuracies up to 99.7%, 99.3% and 98.7% by applying TRS modes under given conditions of stable ambient, background light (BGL) variation and motion disturbance, respectively. Moreover, the proposed method is as well foreseeable to be extended to other multi-tap-involved imaging applications.**

***Index Terms*—CMOS image sensor (CIS), multi-tap pixel, temporal redundant samplings (TRS), signal-to-noise ratio (SNR), physiological sensing, modeling, characterization.**

I. Introduction

IN THE past decade, CMOS image sensors (CISs) with multiple in-pixel storage nodes, known as multi-tap CISs, have gained remarkable developments. The impressive functionality of one-shot multi-image capturability allows those devices to be employed by up-to-date imaging systems, particularly time-of-flight (ToF) depth finders [1]-[5], computational photographers [6]-[8], and time-resolved biomedical imagers [9]-[11]. Furthermore, recent works [12]-[14] have also realized measuring crucial vital indicators, e.g. cardiac pulse (heart rate) and its variability, remotely via facial skin by utilizing two or four-tap near infrared (NIR) lock-in pixels with lateral electric filed charge modulator (LEFM) technique [15], extending the application of multi-tap CISs to a further step toward physiological signal monitors.

Previous studies revealed that the raw physiological signal, by which the cardiac pulses can be measured, is extremely weak (less than 1% of imager detected output) compared to the large offset components corresponding to light reflection and diffusion through facial tissues. Signal-to-noise ratio (SNR) of the imager output, accordingly, is of primary importance to determine the quality of cardiac pulses obtained by physiological sensing. Typically, as regards a solid-state imager, SNR can be limited either by input-referred read noise or shot noise (given by the square root of signal) up to the light conditions. The cardiac pulse measurement works under bright scene due to necessary synchronizations between the imager and auxiliary NIR emitters. Therefore, larger charge handling capability [i.e. full well capacity (FWC)] would be preferable for the promotion of non-saturated signal level and, hence, of the SNR. Normally, regarding applications using multi-tap CISs, each individual image signal is treated by each single tap. This indicates that the achievable FWC of a multi-tap pixel is defined by the maximum charge capacity of one tap only, the area of which, however, is dramatically limited by the complicated layout created by the multi-tap configuration. As a result, to increase FWC by simply enlarging per tap area is difficult. On the other hand, the approaches for deepening the potential well (i.e. increasing the capacitance density of pinned diode) of the tap are alternatives to allow more charges per tap being accumulated, whereas the smaller potential difference causes deterioration in charge transfer efficiency between taps and floating diffusions (FDs). Therefore, the main solutions that are capable of improving charge handling capability in ordinary CMOS pixels are less adoptable from a perspective of multi-tap pixel design. In addition, pixel counting enlargement in region-of-interest (ROI) area, as mentioned in [12], presents helpfulness on high frequency (HF) noise reduction, leading to a better SNR in cardiac pulse wave. Nevertheless, larger spatial resolution, required by pixel counting increasing within a fixed scene (facial area), is a challenge for multi-tap pixels.

In this article, we present an in-pixel temporal redundant samplings (TRS) method aimed at improving SNR of multi-tap CISs without any modifications of pixel layout or fabrication process. This method samples an individual image signal by

Manuscript received xx xx, 2021. This work was supported in part by the Ministry of Education, Culture, Sports, Science and Technology (MEXT)-in-Aid for Scientific Research (S) under Grant 25220905 and Grant 18H05240; in part by the Japan Science and Technology Agency (JST) Center of Innovation (COI) under Grant JPMJCE1311; and in part by the Regional-Innovation Ecosystem Program.

C. Cao, M. Hakamata, K. Yasutomi and K. Kagawa are with the Research Institute of Electronics, Shizuoka University, Hamamatsu 432-8011, Japan. (e-mail: chencao@idl.rie.shizuoka.ac.jp).

S. Aoyama is with Brookman Technology, Inc., Hamamatsu 430-0936, Japan.

N. Tsumura is with Chiba University, Chiba 263-8522, Japan.

S. Kawahito is with the Research Institute of Electronics, Shizuoka University, Hamamatsu 432-8011, Japan, and also with Brookman Technology, Inc., Hamamatsu 430-0936, Japan.

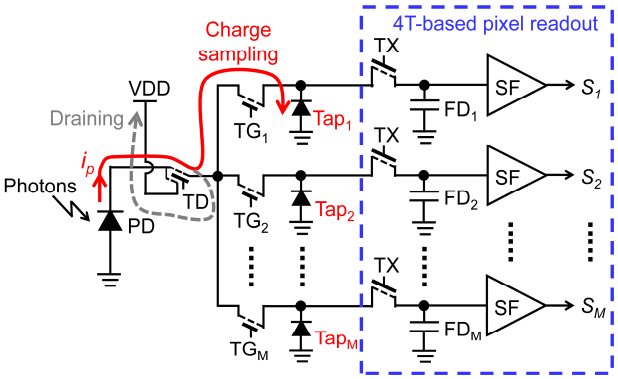


Fig. 1. Schematic of a typical multi-tap pixel with LEFM TGs and 4T-based readout topologies.

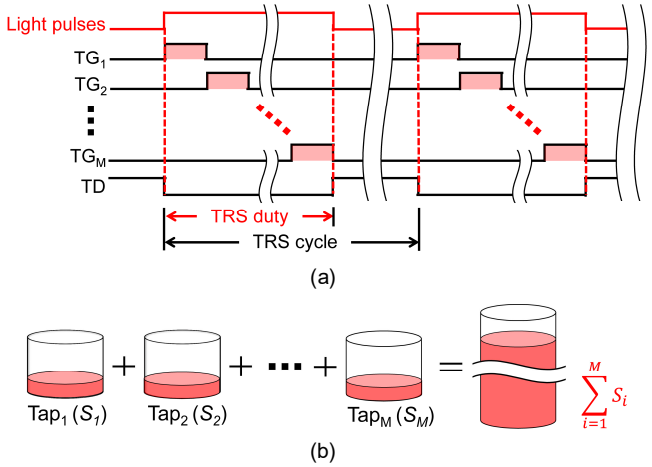


Fig. 2. Proposed TRS operational method for the pixel with *M*-tap storability. (a) Timing diagram. (b) Integration processing.

multiple times with the aid of in-pixel storage resources made by taps, leading to a magnified FWC, and, therefore, better SNR performance. The proposed technique is comprehensively verified through modeling, characterizations, and is ultimately demonstrated in real cardiac pulse measurements. The application results show apparent improvements particularly on HF noise reductions under varieties of ambient conditions.

II. Concept of the Proposed Method

A. TRS Method With Multi-Tap Pixels

Fig. 1 schematically shows a typical example of the multi-tap pixel design, which is composed of a photodiode (PD) for photon sensing, multiple storage diodes (taps) gated by switching gates (TGs) for image signals (S_1, S_2, …, S_M) samplings, a transfer drain (TD) for sweeping unwanted charges, and four transistor (4T)-based topologies with transfer gates (TXs), floating diffusions (FDs), source followers (SFs) for pixel signals reading. The taps, each of which accumulates the charges originated by PD through delivering programmable signals to TGs, enabling a basic functionality for multi-tap imaging application.

Fig. 2 conceptually depicts the proposed TRS operation offered by an *M*-tap pixel. In this method, the charges stimulated by one optical source (e.g. auxiliary light pulse) are redundantly sampled by multiple times (maximum = *M*) by switching, in turn, the TGs ON and OFF with an identical pulse width each, as shown in Fig. 2(a). Benefited from this operation, the target image signal can be equivalently amplified in charge domain by a factor proportional to TRS sampling number after integrating the sampled sub-signals tap by tap, as illustrated in Fig. 2(b). In other words, the effective FWC of the multi-tap pixel is as well gained by the same factor led by TRS, thereby the enhancement of SNR is expected.

B. SNR Modeling With TRS

The impact of the TRS operation on the enhancement of SNR can be quantitatively investigated by modeling hereafter.

Initially, the linear SNR of a multi-tap CIS operating at the conventional mode can be calculated as

$$SNR_{no-TRS} = \frac{\frac{S}{k}}{\sqrt{\left(\sqrt{\frac{S}{k}}\right)^2 + \overline{\sigma_R^2}}} = \frac{N_S}{\sqrt{N_S + \overline{\sigma_R^2}}} \tag{1}$$

where S and k are the output signal (after dark level subtraction) and conversion gain (CG) of the single tap, respectively. N_S is the number of signal electrons. $\overline{\sigma_R^2}$ is the power of input-referred read noise.

Then, when the TRS operation is supplied with a sampling number of M, the SNR calculation would be expressed as

$$SNR_{M-TRS} = \frac{\sum_{i=1}^{M} N_{Si}}{\sqrt{\left(\sqrt{\sum_{i=1}^{M} N_{Si}}\right)^2 + M \times \overline{\sigma_R^2}}} \tag{2}$$

where N_{Si} is the electron number of the i-th tap. Note that the TRS sub-signals (S_i) are recommended to be integrated by off-chip processing to avoid extra pixel or readout circuit design. This leads to an accumulation of the input-referred noise in power domain when signals are eventually integrated, as reflected by $M \times \overline{\sigma_R^2}$ in (2). Meanwhile, given that

$$\sum_{i=1}^{M} N_{Si} = M \times \overline{N_S} \tag{3}$$

where $\overline{N_S}$ denotes the average number of the sampled signal electrons by TRS, equation (2) can be simplified as

$$SNR_{M-TRS} = \frac{M \times \overline{N_S}}{\sqrt{\left(\sqrt{M \times \overline{N_S}}\right)^2 + M \times \overline{\sigma_R^2}}} = \sqrt{M} \times \frac{\overline{N_S}}{\sqrt{\overline{N_S} + \overline{\sigma_R^2}}} . \tag{4}$$

Hence, the factor of SNR enhancement triggered by M times TRS is able to be deduced by comparing with (1) as

$$\Delta SNR_{M-TRS} = \frac{SNR_{M-TRS}}{SNR_{no-TRS}} \approx \sqrt{M} . \tag{5}$$

Equation (5) indicates that SNR of an *M*-tap CIS can be maximally enhanced by a factor of $\sqrt{M}$ with the TRS

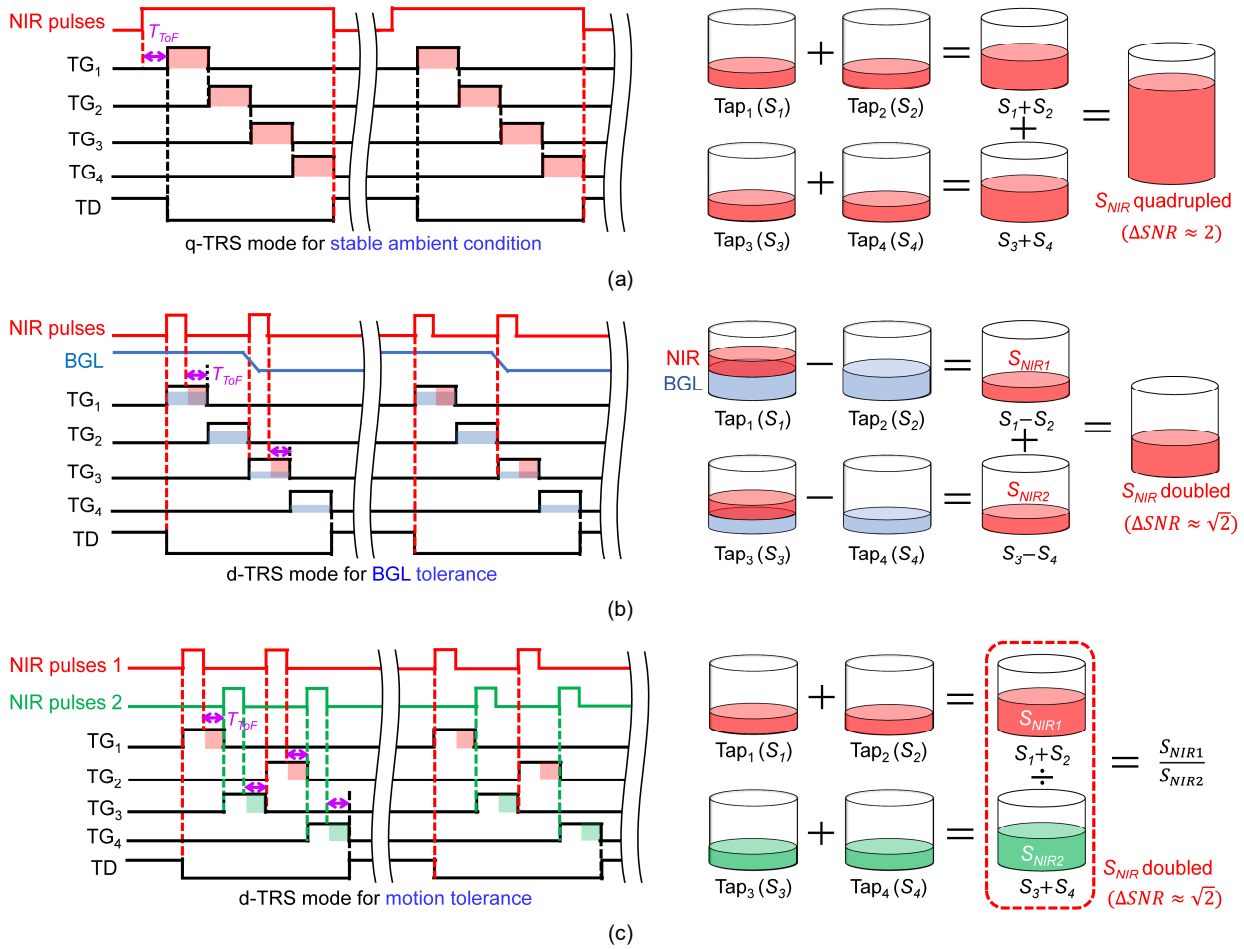


Fig. 3. Implementations of TRS modes based on a 4-tap pixel for cardiac pulse measurement application. (a) q-TRS mode under stable ambient conditions. (b) d-TRS mode for BGL rejection. (c) d-TRS mode for motion tolerance.

operation. Note that $\overline{N_S}$ should be ideally equal to the electron number from each single tap when properties of the taps are perfectly matched with each other. However, unfortunately, the nonideal factors (e.g. nonlinear responses, modulated charge transfer uncertainties, CG mismatches, etc.) induced by inherent device defects or process imperfections are likely to produce deviations on tap outputs, causing an inaccuracy of the SNR enhanced factor led by TRS.

The modeling clarifies that the TRS efficiency strongly relies on the available tap quantities. In other words, the more taps involved in the TRS operation, the larger the SNR enhanced factor is attained. On the other hand, the superiority of the imaging functionalities benefited from multi-tap configuration would be, however, weakened due to the decrease of the tap resources. To deal with this issue, the tradeoffs between the tap resources and application demands should be flexibly considered to maximize the TRS efficiency while preserving the necessary functionalities in the cases of real applications.

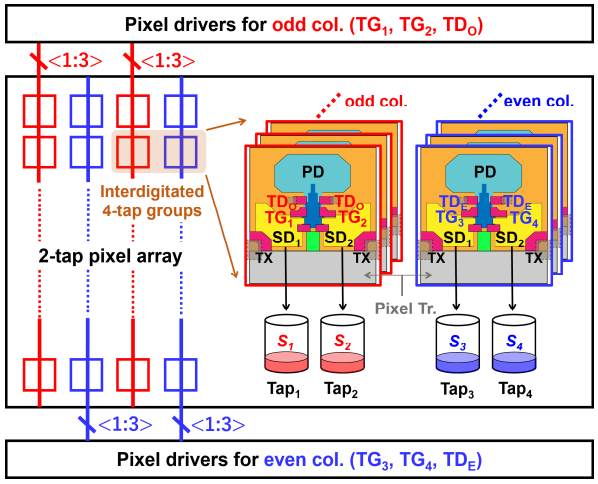


Fig. 4. Simplified block diagram of the 2-tap-pixel based interdigitated 4-tap driving scheme proposed in [14] with in-pixel layout.

C. Feasibility of TRS in Cardiac Pulse Measurement

Cardiac pulses have been reported to be remotely measurable using multi-tap CISs with NIR light lock-in techniques under various ambient conditions, e.g. stable or unstable background light (BGL) [12], [13], artificial motion [14], etc. In terms of the pixel operational differences with respect to those different conditions, the proposed TRS method can be implemented by three operational modes with, for instance, a four-tap pixel configuration, as illustrated in Fig. 3. Fig. 3(a) shows the basic TRS implementation allowing cardiac pulses to be measured at

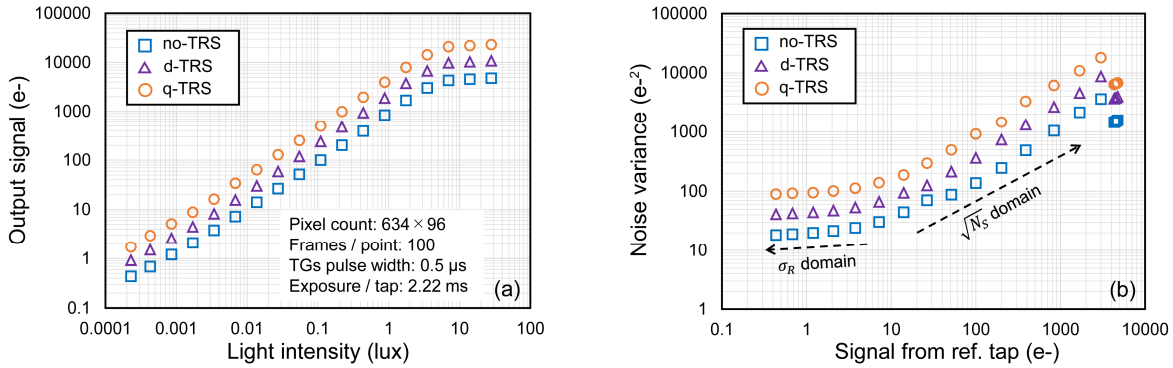


Fig. 5. Photo response characteristics of 4-tap TRS variations measured by 3746-K light source with a collimator at 25 ℃. (a) Responsivity. (b) PTC.

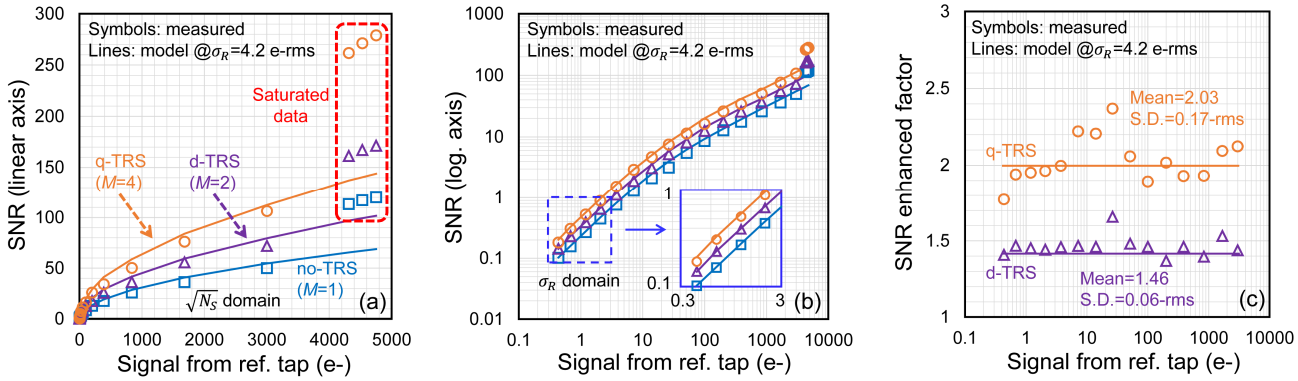


Fig. 6. Comparisons of the measured SNR characteristics versus TRS variations with the model predictions. (a) SNR on linear coordinates. (b) SNR on logarithmic coordinates with zoomed-in inset for highlighting read noise domain. (c) SNR enhanced factors.

a stable ambient environment, where the BGL is dark or exhibits comparatively constant brightness[1], and the subject keeps still. Under this case, the tap resources can be maximally used for launching a quad-TRS-wise (q-TRS) operation, by which the amplitude of the effective NIR signal is quadrupled, thus resulting SNR extension by a factor of 2. Note that a time delay (T_{ToF}) is necessary to be introduced between the raising edges of NIR pulses and TG_1 pulses for avoiding the charge sampling delay due to ToF effects caused by remotely sensing. Fig. 3(b) shows the TRS strategy for cardiac pulse measuring against fluctuated ambient brightness when BGL tolerance is essential. Since a couple of taps has to be contributed beforehand to the BGL samplings with and without NIR signal, respectively, the TRS operation can only be permitted after BGL subtractions are applied. Thus, dual-TRS (d-TRS) is feasible to be adopted to double the BGL-rejected signal with, accordingly, an SNR enhanced factor of $\sqrt{2}$. Similar to q-TRS, T_{ToF} should be set as well in d-TRS between the falling edges of NIR pulses and TG pulses with NIR duties to prevent the NIR signal from being delayed to BGL-only signal sampling phases. Fig. 3(c) shows the TRS mode for cardiac pulses measuring with dual-band lock-in technique dedicated to artificial motion mitigation. As presented in [14], motion effect can be reduced in this application by dividing the two lock-in signals generated by dual-NIR-band light pulses to eliminate the offset signal component which is very sensitive to the subject movements.

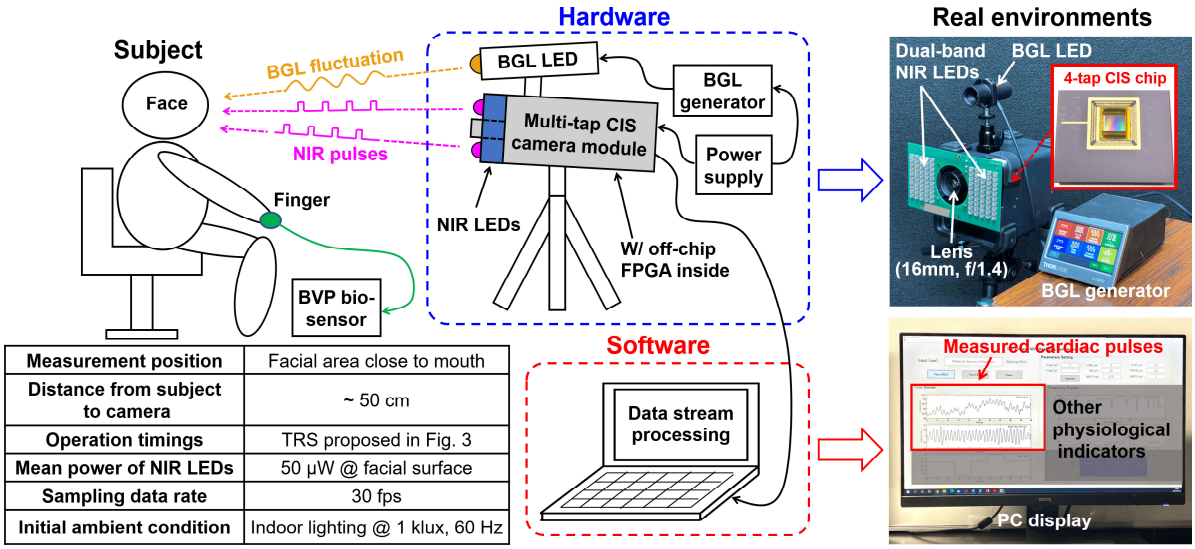
[1]Indoor lighting is also allowable as long as the light meets an integral multiple frequency with CIS frame rate.

Hence, in this mode, the four tap resources can be divided into two groups, each possesses double taps bound for an individual NIR wavelength. Therefore, d-TRS is able to be provided by the couple of taps belonging to each group for gaining SNR enhancement before the ratio of dual-band signals is taken. Correspondingly, T_{ToF} still needs to be added to the concerned pulse edges, as marked in Fig. 3(c).

It is of importance to notice that the aforementioned TRS modes, as well as the conventional operations, can be simply switchable according to the real application situations by means of reprogramming the control sequences supplied to the light sources or TG gates. This implies that the presented TRS method features good feasibilities in multi-tap based cardiac pulse measurements. Furthermore, promising compatibilities of the proposed concept with other multi-tap imaging applications are also well predictable.

III. Characterizations and Discussion

To verify the effectiveness of the proposed method, the CIS with interdigitated four-tap pixels, the specifications of which were reported in [14], has been used for TRS (specifically, the variations of no-TRS, d-TRS, q-TRS) characterizations. A simplified block diagram of the interdigitated four-tap driving scheme as well as the pixel layout is reviewed in Fig. 4. In this section, the pixels are operated at four-tap modulation similar to the one shown in Fig. 3(a). The pulse width of TGs is set to 0.5 μs each. The TRS variations are evaluated by using respectively the output data from one tap (no-TRS), summations of two taps (d-TRS) and four taps (q-TRS).



Measurement position	Facial area close to mouth
Distance from subject to camera	~ 50 cm
Operation timings	TRS proposed in Fig. 3
Mean power of NIR LEDs	50 μW @ facial surface
Sampling data rate	30 fps
Initial ambient condition	Indoor lighting @ 1 klux, 60 Hz

Fig. 7. Experimental setup of the cardiac pulse measurements involved in this work with the table (left bottom) listing necessary information and the photographs showing real environments of hardware (right top) and software (right bottom) configurations.

Fig. 5(a) shows the averaged outputs as a function of light intensity. The output digital numbers (DN) are converted to electrons (as N_S presented in Section II-B) for compensating the CG mismatches between each tap. As can be seen, fairly similar linearity curves with approximately uniform signal differences at each illuminance are obtained. The linear FWCs are measured as 3000 e^-, 6700 e^-, 14300 e^- correlated to no-TRS, d-TRS, q-TRS, respectively, indicating that charge handling capability of a multi-tap pixel increases proportionally versus TRS sampling number. It is observed that the FWCs obtained by d-TRS and q-TRS are not precisely double and quadruple the one measured with no-TRS. This can be interpreted by the linearity variation between each tap. Fig. 5(b) plots the photon transfer curves (PTCs) of three TRS variations. To facilitate comparisons, the output from one tap (SD_1 of odd column pixels in Fig. 4) is referred to as a reference signal, which is common to all curves. One can see from Fig. 5(b) that a proportional behavior to TRS sampling number appears not only at the noise variances dominated by shot noise, but also at the ones dominated by read noise (σ_R). This can be explained by the read noise power accumulation caused by the TRS signals integration at the off-chip stage, as mentioned in Section II-B.

Based upon the data harvested from Fig. 5, the SNR characteristics versus TRS variations are further reported in Fig. 6 where the results between measured data and model estimations are also compared. The σ_R for model calculations, as discussed in Section II-B, is determined here as 4.2 e-rms. This is obtained by the mean standard deviation of the dark outputs from the reference tap within 100 frames at 25 °C. Fig. 6(a) and (b) respectively plot SNR results on linear and logarithmic coordinates for highlighting the properties in shot noise and read noise domains simultaneously. As can be observed, good agreements between measured data and model predictions are exhibited with averaged deviations of 8.24%, 8.13%, and 9.98%, respectively for no-TRS, d-TRS, and q-TRS. The enhanced factors of SNR (ΔSNR) led by d-TRS and q-TRS are plotted afterwards in Fig. 6(c). It is shown that d-TRS offers ΔSNR a mean of 1.46 with standard deviation (S.D.) of 0.06-rms, whereas those for q-TRS are measured as 2.03 and 0.17-rms, respectively. Both mean ΔSNRs are extremely approach to $\sqrt{M}$ (model prediction) with respect to varied TRS modes, indicating the effectiveness of the modeling work, and, therefore, of the proposed method. Moreover, the phenomena of ΔSNR fluctuations occurred in both TRS modes [see Fig. 6(c)] are attributed to the influences of nonideal factors other than CG mismatches (already compensated) on tap outputs, as mentioned in Section II-B. It is possible to further compensate those nonideal factors by increasing the number of frames for each data point to acquire more stable ΔSNR trends, but beyond the scope of this article.

Additionally, based on the characterization results reported in this section, one can observe that the FWC (3000 e^-) and dark noise (4.2 e-rms) show obvious differences from the ones (4200 e^- and 0.67 e-rms, respectively) achieved in [14] by the same CIS. The mechanisms behind these deviations are worth discussing hereafter. As regards the FWC characterization, constant biases were applied to the modulation gates (i.e. TG_1 ON, TG_2 and TD OFF) in case of the previous work, thus allowing the potential well of SD_1 to be extended towards the modulation channel between TGs (see Fig. 4), resulting in an expansion of the effective storage area. However, with the TRS operation used in this article, the aforementioned extra potential well is disappeared when TGs turned OFF simultaneously at the end of each modulation cycle, leading to lower charge handling capabilities at SD nodes. The deviation of the noises can be attributed to the difference of the noise characterization methods. Specifically, this article measures the mean noise over all pixels for SNR calculations, whereas the previous work picked up the noise from the peak probability appeared in noise histogram. With the latter method, the low frequency components (1/f, random telegraph noises, etc.) behaving larger noise levels but lower probabilities are less considered, thereupon a better noise performance tends to be attained. Furthermore, the TRS operation with fast gating signals produce possibly additional noise causes in comparison with the constant state measurements applied in [14], such as, e.g.

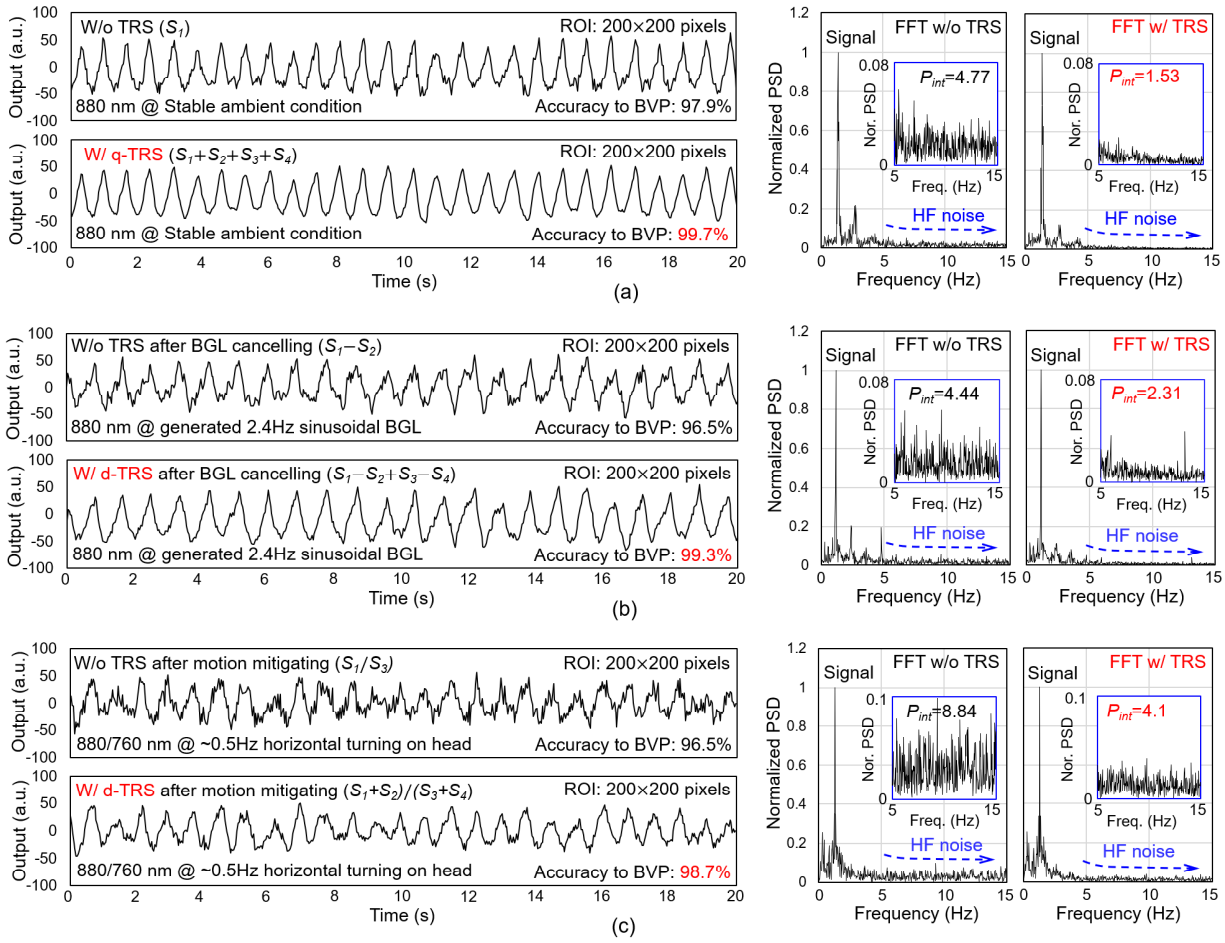


Fig. 8. Detrended cardiac pulses extracted by sensor outputs with and without in-pixel TRS under the conditions of (a) stable ambient environment, (b) BGL disturbance, and (c) motion disturbance, respectively. FFTs (normalized) are applied to each corresponding pulse wave for frequency domain analyzing to characterize the effects of TRS on HF noises reduction in the measured cardiac pulses.

coupling noise between pixel array and analog-to-digital converter (ADC) or other readout circuitries with high speed operating clocks.

IV. Application Results

This section is dedicated to show the evaluation results of the cardiac pulse measurements with applying the TRS modes introduced in Section II-C. Fig. 7 shows the experimental setup, in which a camera module with the utilized four-tap CIS, and a graphical user interface (GUI) for cardiac pulses data streaming are chiefly included. Moreover, two NIR light emitting diode (LED) families centered at wavelengths of 880 nm and 760 nm, as well as a BGL LED stimulated by a functional generator are respectively mounted to the front and top of the camera module for experimental requirements. Synchronizations between the LEDs and the CIS are supported by an off-chip field programmable gate array (FPGA). The subject is located at an approximately 50 cm distance from the camera lens, and the position of the measurement is on the facial skin close to the mouth. A commercial blood volume pulse (BVP) bio-sensor (Procomp2 [16], Thought Technology Ltd.) is introduced to also measure the cardiac pulses simultaneously via fingertips for data comparisons. The measurement is performed indoors with an initial ambient illuminance of 1000 lux. This indoor lighting features an AC frequency of 60 Hz, which is double faster than that of the CIS sampling rate (30 Hz), thus feeding only a constant offset to the output signals by the CIS.

The cardiac pulses measured with and without in-pixel TRS under the given conditions of stable ambient environment, 2.4 Hz BGL fluctuation, and approximately 0.5 Hz head motion are respectively reported in Fig. 8(a)-(c), where Fast Fourier Transforms (FFTs) are applied to each waveform for further analyses in frequency domain. It is worth noting that detrending method [16] has been performed to the raw cardiac pulses outputted by the imager for avoiding the influences of low frequency fluctuations (induced by the involuntary motions such as respiration, etc.) on the readabilities of FFT results. Hence, *y*-axes of the waveforms are displayed by arbitrary unit (a.u.), as illustrated in Fig. 8. From Fig. 8(a), one can observe that the cardiac pulse wave with q-TRS method [see Fig. 3(a)] is clearly improved compared to that without TRS. An excellent heart rate detection accuracy up to 99.7% is achieved comparing with the reference data provided by the bio-sensor. FFT results reveal in depth that the integration of power spectrum density (PSD) (P_{int}) versus the HF noise component (typically larger than 5 Hz, according to the frequency range other than the fundamental signal and two harmonics [18]) can be reduced by 67.9% with q-TRS mode. This is definitely due

to the merit of sensor SNR enhancement led by the proposed method. Similarly, as shown in Fig. 8(b) and (c), the measured heart rate accuracies can be promoted to 99.3% and 98.7% under the conditions of BGL and motion disturbances, respectively, by applying the correlated d-TRS modes [see Fig. 3(b) and (c)]. Correspondingly, the PSD integrations obtained from FFT analyses are reduced by 48% and 53.6%, respectively, versus HF noise components.

In conclusion, the validity of the proposed method toward the real application has been demonstrated. Particularly, the HF noise components coupled on the effective cardiac pulse signals can be significantly attenuated by the proposed in-pixel TRS operations.

V. Conclusion

In this work, an in-pixel TRS technique for improving the SNR performance of multi-tap CISs is presented, modeled and characterized. By means of this method, the image signal is allowed to be sampled multiple times in temporal domain by the available tap resources, resulting in a significant promotion of in-pixel charge handling capability. The positive impact of the presented method on application has been demonstrated through cardiac pulse measurements under various ambient conditions using a four-tap CIS combining with NIR lock-in techniques. The concept we proposed can be further introduced to other multi-tap imaging applications by taking account of specific tradeoffs between the tap resources and application requirements.

References

[1] D. Stoppa, N. Massari, L. Pancheri, M. Malfatti, M. Perenzoni, and L. Gonzo, "A Range Image Sensor Based on 10-µm Lock-In Pixels in 0.18-µm CMOS Imaging Technology," *IEEE J. Solid-State Circuits*, vol. 46, no. 1, pp. 248–258, Jan. 2011, doi: 10.1109/JSSC.2010.2085870.

[2] T.-Y. Lee, Y.-J. Lee, D.-K. Min, S.-H. Lee, W.-H. Kim, J.-K. Jung, I. Ovsiannikov, Y.-G. Jin, Y. Park, and E. R. Fossum, "A Time-of-Flight 3-D Image Sensor With Concentric-Photogates Demodulation Pixels," *IEEE Trans. Electron Devices*, vol. 61, no. 3, pp. 870–877, Mar. 2014, doi: 10.1109/TED.2014.2301875.

[3] E. Tadmor, D. Cohen, G. Yahav, G. Tennenholtz, G. Lehana, A. Lahav, A. Birman, A. Fenigstein, and A. Fish, "Development of a ToF Pixel With VOD Shutter Mechanism, High IR QE, Four Storages, and CDS," *IEEE Trans. Electron Devices*, vol. 63, no. 7, pp. 2892–2899, Jul. 2016, doi: 10.1109/TED.2016.2560941.

[4] K. Yasutomi, Y. Okura, K. Kagawa, and S. Kawahito, "A Sub-100 µm-Range-Resolution Time-of-Flight Range Image Sensor With Three-Tap Lock-In Pixels, Non-Overlapping Gate Clock, and Reference Plane Sampling," *IEEE J. Solid-State Circuits*, vol. 54, no. 8, pp. 2291–2303, Aug. 2019, doi: 10.1109/JSSC.2019.2916310.

[5] Y. Shirakawa, K. Yasutomi, K. Kagawa, S. Aoyama, and S. Kawahito, "An 8-Tap CMOS Lock-In Pixel Image Sensor for Short-Pulse Time-of-Flight Measurements," *Sensors*, vol. 20, no. 4, p. 1040, Apr. 2020, doi: 10.3390/s20041040.

[6] G. Wan, X. Li, G. Agranov, M. Levoy, and M. Horowitz, "CMOS Image Sensors With Multi-Bucket Pixels for Computational Photography," *IEEE J. Solid-State Circuits*, vol. 47, no. 4, pp. 1031–1042, Apr. 2012, doi: 10.1109/TIP.2015.2436342.

[7] Y. Luo, D. Ho, and S. Mirabbasi, "Exposure-Programmable CMOS Pixel With Selective Charge Storage and Code Memory for Computational Imaging," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 5, pp. 1555–1566, May 2018, doi: 10.1109/TCSI.2017.2763822.

[8] N. Sarhangnejad, N. Katic, Z. Xia, M. Wei, N. Gusev, G. Dutta, R. Gulve, P. Z. X. Li, H. F. Ke, H. Haim, M. M. Garcia, D.Stoppa, K. N. Kutulakos, and R. Genov, "Dual-Tap Computational Photography Image Sensor With Per-Pixel Pipelined Digital Memory for Intra-Frame Coded Multi-Exposure," *IEEE J. Solid-State Circuits*, vol. 54, no. 11, pp. 3191–3202, Nov. 2019, doi: 10.1109/JSSC.2019.2932623.

[9] Z. Li, S. Kawahito, K. Yasutomi, K. Kagawa, J. Ukon, M. Hashimoto, and H. Niioka, "A time-resolved CMOS image sensor with draining-only modulation pixels for fluorescence lifetime imaging," *IEEE Trans. Electron Devices*, vol. 59, no. 10, pp. 2715–2722, Oct. 2012, doi: 10.1109/TED.2012.2209179.

[10] M.-W. Seo, K. Kagawa, K. Yasutomi, Y. Kawata, N. Teranishi, Z. Li, I. A. Halin, and S. Kawahito, "A 10 ps time-resolution CMOS image sensor with two-tap true-CDS lock-in pixels for fluorescence lifetime imaging," *IEEE J. Solid-State Circuits*, vol. 51, no. 1, pp. 141–154, Jan. 2016, doi: 10.1109/JSSC.2015.2496788

[11] M.-W. Seo, Y. Shirakawa, Y. Kawata, K. Kagawa, K. Yasutomi, and S. Kawahito, "A time-resolved four-tap lock-in pixel CMOS image sensor for real-time fluorescence lifetime imaging microscopy," *IEEE J. Solid-State Circuits*, vol. 53, no. 8, pp. 2319–2330, Aug. 2018, doi: 10.1109/JSSC.2018.2827918.

[12] C. Cao, Y. Shirakawa, L. Tan, M.-W. Seo, K. Kagawa, K. Yasutomi, T. Kosugi, S. Aoyama, N. Teranishi, N. Tsumura, and S. Kawahito, "A Time-Resolved NIR Lock-In Pixel CMOS Image Sensor with Background Cancelling Capability for Remote Heart Rate Detection," *IEEE J. Solid-State Circuits*, vol. 54, no. 4, pp. 978–991, Apr. 2019, doi: 10.1109/JSSC.2018.2885528.

[13] C. Cao, M. Oishi, L. Tan, K. Kagawa, K. Yasutomi, S. Aoyama, N. Teranishi, N. Tsumura, and S. Kawahito, "A Sub-Electron Temporal Noise High Modulation Contrast NIR Lock Lock-In Pixel CMOS Image Sensor for Non-Contact Physiological Measurement," in *Proc. Int. Image Sensor Workshop (IISW)*, 2019, pp. 286–289.

[14] C. Cao, J. K. Dutta, M. Hakamata, K. Yasutomi, K. Kagawa, S. Aoyama, N. Tsumura, and S. Kawahito, "A Dual NIR-Band Lock-In Pixel CMOS Image Sensor With Device Optimizations for Remote Physiological Monitoring," *IEEE Trans. Electron Devices*, vol. 68, no. 4, pp. 1688–1693, Apr. 2021, doi: 10.1109/TED.2021.3057035.

[15] S. Kawahito, G. Baek, Z. Li, S.-M. Han, M.-W. Seo, K. Yasutomi, and K. Kagawa, "CMOS lock-in pixel image sensors with lateral electric field control for time-resolved imaging," in *Proc. Int. Image Sensor Workshop (IISW)*, 2013, pp. 361–364.

[16] M. P. Tarvainen, P. O. Ranta-aho, and P. A. Karjalainen, "An Advanced Detrending Method With Application to HRV Analysis," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 2, pp.172–175, Feb. 2002, doi: 10.1109/10.979357.

[17] *ProComp2 User Guide and Hardware Manual*. [Online]. Available: https://thoughttechnology.com/content/docs/manual/SA7420Rev7%20ProComp2%20Manual.pdf.

[18] A. Caizzone, A. Boukhayma, and C. Enz, "A 2.6µW monolithic CMOS photoplethysmographic sensor operating with 2µW LED power," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 6, pp. 1243–1253, Dec. 2019, doi: 10.1109/TBCAS.2019.2944393.