Multi-Tap NIR Lock-In CMOS Imager for Non-Contact Physiological Signal Monitoring

Chen Cao, Masashi Hakamata, Keita Yasutomi, Keiichiro Kagawa and Shoji Kawahito

kawahito@idl.rie.shizuoka.ac.jp
Research Institute of Electronics, Shizuoka University, Hamamatsu, 432-8011, Japan
Keywords: CMOS image sensor, multi-tap pixel, physiological sensing

ABSTRACT

Non-contact physiological monitoring using image sensors becomes popular for the cutting-edge medical and security applications. This paper reports a multi-tap NIR lock-in CMOS imager which is able to achieve heart rate and its variability measurements under varied ambient light and artificial motion conditions with excellent SNRs.

1 Introduction

Non-contact physiological monitor systems for detecting vital indicators, such as heart rate (HR) and its variability (HRV), become popular for developing the cutting-edge products aimed to human medical and security applications. CMOS image sensors (CISs) act as a very important role as the front-end signal receiver of those systems. Diverse attempts have been devoted to CIS-based HRV sensing techniques. One of the most remarkable methods is remote photo-plethysmography (rPPG) technique, in which standard RGB CISs [1] and CISs with special color filter array (CFA) [2] were introduced and demonstrated. However, the lack of robustness on ambient disturbances (background light fluctuation, artificial motion, etc.) significantly limits the feasibilities of those techniques with respect to realistic applications. To address this issue, this paper demonstrates a multi-tap near-infrared-band (NIR-band) lock-in pixel (LIP) CIS [3] for measuring the HRV signal remotely by detecting the in-blood hemoglobin concentration variations using auxiliary NIR emitters. The reported CIS is capable of synchronizing with two NIR emitters simultaneously, thereby realizing dual NIR lock-in functionality. As a result, background light (BGL) and artificial motion can be suppressed in each HRV sensing frame. Moreover, by applying an in-pixel temporal redundant sampling (TRS) technique [4] introduced in charge domain, the signal-to-noise ratios (SNRs) of the HRV signals measured by the proposed CIS can be further enhanced.

2 Methods and Implementations

2.1 CIS Design

Fig. 1 shows a simplified block diagram of the proposed CIS design including the two-tap LIP structure and an interdigitated driving scheme for realizing dual NIR lock-in possibility. An individual LIP consists of a photodiode (PD) for photon sensing, two storage diodes (SDs) gated by switching gates (TGs) in terms of lateral electric field modulation (LEFM) mechanism [5], transfer drain (TD) gates for sweeping undesired signals. The LIP configures a four transistor (4T)-based topological circuitry including transfer gates (TXs), source followers (SFs) and select switches for reading LIP signals. The taps accumulate charges sent by PD through delivering programmable gating pulses to TGs. Those gating pulses are driven by both the upper and lower sides of the pixel array. Each side is in charge of every other column LIPs. Specifically, upper side bounds for odd columns and lower side bounds for even columns, respectively. This function is of fundamental for realizing dual NIR lock-in with LIPs having only two taps, thus releasing the pixel design difficulties when a four-tap-wise function is involved.

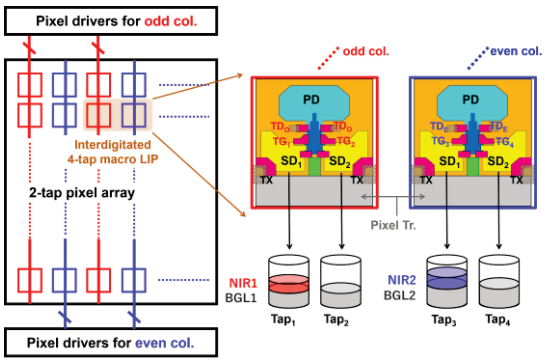


Fig. 1 Two-tap dual NIR lock-in CIS design.

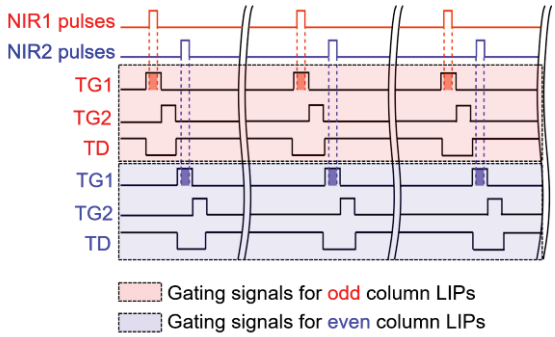


Fig. 2 Timing diagram for LIP operation.

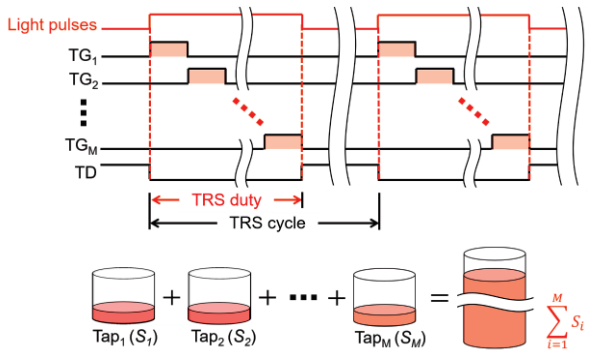


Fig. 3 TRS method for SNR enhancement [4].

2.2 Ambient disturbances suppressions

Fig. 2 illustrates the timing diagram used for LIP operation in this sensor for suppressions of BGL and motion effects. As can be seen, the timing presents a sort of double NIR pulses lock-in operation delivering two individual NIR pulse lock-in functions, respectively, to odd and even column LIPs. Each lock-in operation comprises a TG1 pulse for sampling the mixed signal with NIR and BGL into SD1 storage, a TG2 pulse for sampling BGL-only signal into SD2 storage without NIR emitting. Therefore, the difference of the charge amount between two SDs would be a pure NIR signal with BGL rejection. The TD gates are enabled between the lock-in operational periods for draining out the charges induced by the ambient light to the power lines, hence additionally offering LIPs a BGL tolerance ability during the periods other than lock-in operations. Eventually, the timing diagram shown in Fig. 2 is capable of feeding the sensor two BGL-rejected signals per each macro pixel, as schematically plotted in Fig.1, for further processing.

Subsequently, the ratio of two BGL-rejected signals from each macro pixel is carried out through off-chip processing. By applying this, the tissue-related offsets that are sensitive to artificial motion can be eliminated, remaining a motion robust signal. The mechanism behind this motion mitigation method is explained in detail in [3]. It is worth noticing that the SNR observed in the motion-mitigated signal would be degraded in comparison with that of each individual signal before motion mitigation. This is caused by the process once taking a ratio between two signals. This SNR drop can be improved by considering an in-pixel TRS operational technique, however, as a sacrifice of BGL canceling capability when using a 4-tap-wise LIP configuration in this paper. This will be briefly discussed later in section 3.

2.3 TRS method for SNR enhancement

Typically, the HRV signal is extremely weak compared to the large offset components related to light reflection and diffusion through tissues. The SNR of the CIS, thereupon, is very important to obtain a sufficiently good HRV signal quality. For a multi-tap CIS pixel usually featuring complicated layout, the full well capacity (FWC) which determined by tap property is significantly constrained due to the tap area limitation, thus weakening the SNR performance. To alleviate this issue, we proposed an in-LIP TRS method without introducing any modification of layout or fabrication process. Fig. 3 conceptually depicts the TRS operation using an *M*-tap LIP. In this method, the charges generated by an auxiliary light pulse are sampled by maximumly *M* times with identical pulse width each, resulting in a magnified FWC, and, therefore, better SNR. The SNR enhancing factor is approximately equal to $M^{1/2}$. A detailed modeling work as well as a discussion on TRS feasibility in HRV measurement have been reported in [4]. As mentioned in [4], quad-TRS (q-TRS) with 2x SNR enhancement can be applied to HRV measurement using the proposed CIS

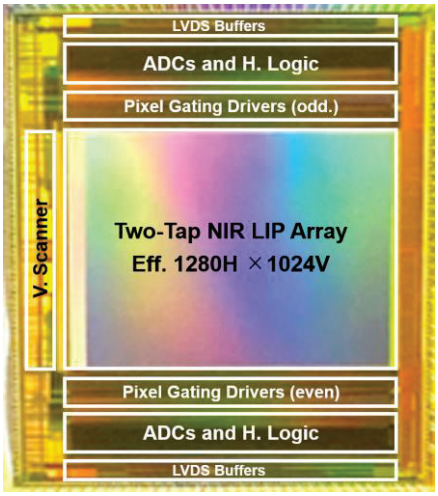
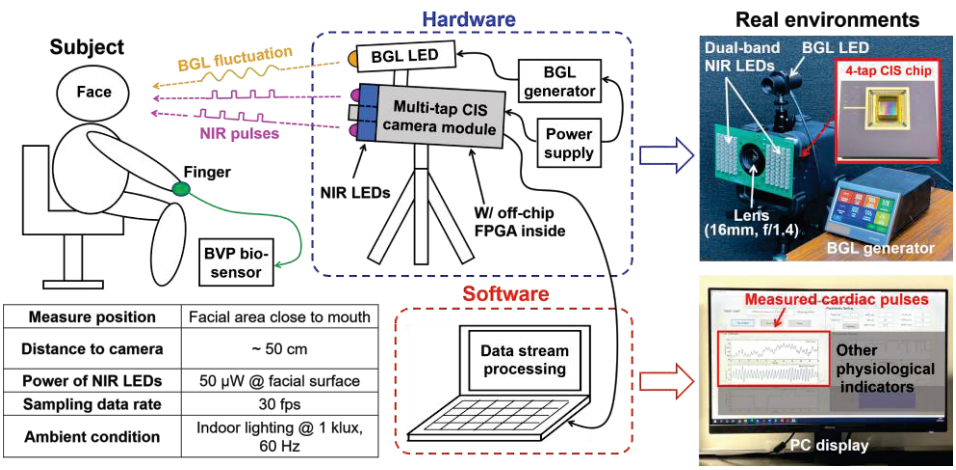


Fig. 4 Micrograph of sensor chip.

TABLE I
Summary of Sensor Characteristics

Parameter	Value
Technology	0.11-μm 1P4M CIS
Effective pixel number	1280H ×1024V
Pixel pitch	7.1-μm
Tap mode	2-tap / 4-tap
NIR pulses sync. ability	Dual-band
ADC resolution	18-bit
Conversion gain	110 μV/e-
Dark current @300K	18.2 pA/cm^2
Random noise @ 300K	0.67 e-rms @peak, analog gain x64
Full well capacity @SDs	4.2 ke-
Frame rate	30 fps



Measure position	Facial area close to mouth
Distance to camera	~ 50 cm
Power of NIR LEDs	50 μW @ facial surface
Sampling data rate	30 fps
Ambient condition	Indoor lighting @ 1 klux, 60 Hz

Fig. 5 Experimental setup with photographs of real hardware and software environments [4].

under stable ambient conditions, whereas double-TRS (d-TRS) is appropriate for giving a $2^{1/2}$x SNR up under the case when either BGL or motion tolerance is necessary.

2.4 Implementations

A prototype CIS chip has been fabricated by a 0.11-μm 1-poly 4-metal (1P4M) CIS technology. Fig. 4 shows the die micrograph of the imager chip. In addition to the LIP array and the pixel gating drivers, column ADCs with noise canceller function are also implemented on both sides of LIP array for quantizing the pixel signals with a relatively high speed. The quantized data are ultimately buffered out through low voltage differential signaling (LVDS) pairs after serialized by horizontal logic. The characterization results of the chip are summarized in Table I. The imager features a SXGA resolution and an ultra-low random noise (0.67 e-rms) at video data rate of 30 fps.

To experimentally measure the HRV signals using the proposed CIS under ambient disturbances, a camera system environment is established. Fig. 5 shows the experimental setup, in which a camera module and a customized software for HRV data streaming are included. Two light emitting diode (LED) families centered at wavelengths of 880 and 760 nm, as well as a white LED for arbitrary BGL generation are, respectively, mounted to the front and top of the camera. Synchronizations between the LEDs and the CIS are offered by an off-chip field programmable gate array (FPGA). The subject is located at an approximately 50-cm distance from camera lens, and the measured position is the facial skin close to the mouth. A commercial blood volume pulse (BVP) biosensor is introduced to measure the HRV signal simultaneously via fingertips for comparisons. The mean power of the pulsed NIR LEDs is measured to be around 50 μW at the facial surface, which is assumed safe enough to human eyes. The experiments are performed at indoor circumstance with a baseline lighting condition of 1000 lux and 60 Hz by a fluorescent lamp.

3 Results and Discussion

Fig. 6 shows the HRV signals measured by the proposed CIS under the conditions of 2.4-Hz sinusoidal BGL fluctuation [Fig. 6(a)] and about 0.5-Hz head motion [Fig. 6(b)], respectively. A region of interest (ROI) with a size of 200 by 200 pixels is taken account of for calculation. Fast Fourier transforms (FFTs) are applied to both the waveforms with and without TRS, respectively, for frequency domain analyses. Compared to the results reported in [4], the measured data before ambient disturbances rejection are additionally given in Fig. 6 for comparison. As can be seen in the top graphs of Fig. 6(a) and (b), the output waveforms before ambient disturbances rejection behave completely ambient dependencies without any HRV signal indications. However, relatively clear HRV effective signals, though a bit noisy, are extracted by applying ambient rejection techniques with NIR lock-in operations launched by the proposed CIS, as shown in the middle graphs of Fig. 6. Hence, measuring accuracy of 96.5% is attained for both the BGL-cancelled and the motion-mitigated waveforms. Furthermore, the high frequency (HF) noise levels of the ambient rejected data can be reduced, as shown in the bottom graphs of Fig. 6, by introducing proper TRS operations. The SNRs are, accordingly, enhanced, thereby offering more accuracies to the resulted data. FFT results placed at the side of the graphs quantify the HF noise reduction effects led by TRS operations by plotting normalized power spectrum densities (PSD) versus frequencies. The PSD integrations (P_{int}) obtained from the noise components with frequencies larger than 5-Hz (regarded as HF noises) can be decreased by 48% and 53.6%, respectively, after applying TRS to BGL cancelling and motion mitigation cases.

Note that, TRS technique is only applicable for simple BGL rejection or motion rejection in this work due to the limited tap resource. We plan to improve LIP design by adopting more storage taps in future works. Then, TRS

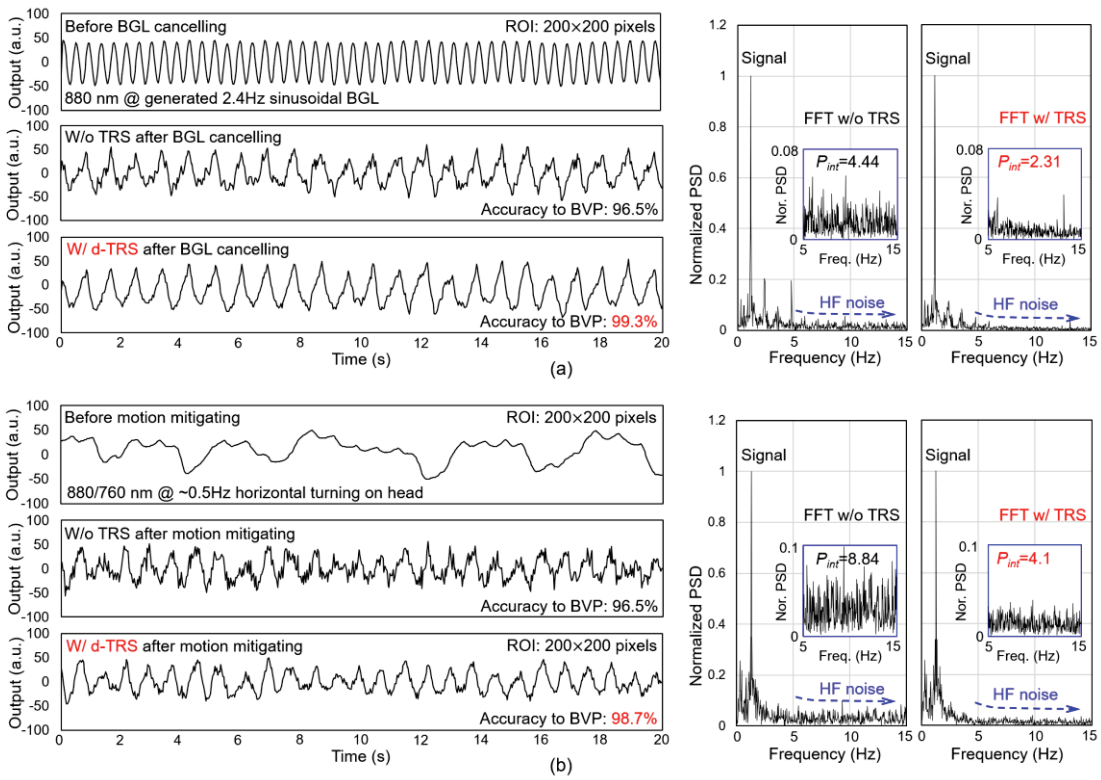


Fig. 6 HRV signals extraction under the conditions of (a) BGL disturbance, and (b) motion disturbance, respectively. FFTs are applied to quantify the TRS effects on HF noises reduction.

method is expected to be combined with simultaneous BGL and motion rejections to gain SNR with stronger functionalities.

4 Conclusions

In this work, we have presented a two-tap NIR-band LIP CIS for remote HRV signal monitoring. The presented CIS is capable of synchronizing with dual NIR-band pulses owing to the LIP design combined with a unique LIP driver configuration. This gives the sensor a strong robustness against ambient disturbances during HRV measuring. The effectiveness of the presented technique has been verified by experimental results. Fairly good accuracies are acquired when comparing our results with those of the reference biosensor. An in-pixel TRS method has also been demonstrated to improve the SNR of the resulted HRV signals.

Acknowledgements

This work was supported in part by the Japan Science and Technology Agency (JST) Center of Innovation (COI) under Grant JPMJCE1311. The authors would like to acknowledge DB HiTek for chip fabrication.

References

[1] W. Verkruysse, L. Svaasand, and J. Nelson, “Remote plethysmographic imaging using ambient light,” Opt. Express., vol. 16, no. 26, pp. 21434–21445 (2008).

[2] D. McDuff, S. Gontarek, and R. W. Picard, “Improvements in Remote Cardiopulmonary Measurement Using a Five Band Digital Camera,” IEEE Trans. Biomed. Eng., vol. 61, no. 10, pp.2593–2601 (2014).

[3] 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 (2021).

[4] C. Cao, M. Hakamata, K. Yasutomi, K. Kagawa, S. Aoyama, N. Tsumura, and S. Kawahito, “Signal-to-Noise Ratio Enhancement in Cardiac Pulse Measurements Using Multitap CMOS Image Sensors With In-Pixel Temporal Redundant Samplings,” IEEE Trans. Electron Devices, vol. 69, no. 6, pp. 2851–2857 (2022).

[5] 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,” Proc. Int. Image Sensor Workshop (IISW), pp. 361–364 (2013).