P21

Simulating the performance required for multi-tap charge modulation pixels in time-resolved biomedical imaging

Keiichiro Kagawa[*1*2], Nobukazu Teranishi[*1*3], Keita Yasutomi[*1], Rolf Saager[*2], Min-Woong Seo[*1], Shoji Kawahito[*1], Anthony Durkin[*2], and Bruce Tromberg[*2]

[*1]Research Institute of Electronics, Shizuoka University, 3-5-1 Johoku, Nakaku, Hamamatsu, Shizuoka 432-8011, Japan, [*2]Beckman Laser Institute, UC Irvine, 1002 Health Sciences Road, Irvine, CA 92617, USA, [*3]Laboratory of Advanced Science and Technology for Industry, University of Hyogo, 1-1-2 Koto, Kamigori, Ako-gun, Hyogo 678-1205, Japan,
E-mail: kagawa@idl.rie.shizuoka.ac.jp

Abstract

In this paper, the performance required for charge modulation pixels in typical biomedical applications based on time-resolved imaging are quantitatively estimated by simulation with modeling the detection process and compared with TCSPC as reference. Single-component fluorescence lifetime imaging (FLIM) and light scattering and absorption in tissues are discussed because quantitative measurement of metabolism based on autofluorescence of endogenous coenzymes and hemodynamics is of great interest in biomedical applications. Overall, accuracy of the deduced optical parameters is mostly determined by photon shot noise, which suggests that large full well capacity or multiple image acquisition is required. Low inter-tap crosstalk is more significant than low read noise when pixelwise signal integrity is considered. Although multiple image acquisition by shifting the tap timings can improve the performance, increase of the number of taps is more effective

1. Introduction

Time-resolved imaging provides rich information in biomedical imaging, where laser pulses with a few to tens of pico-second width excite tissues repeatedly and their temporal responses are measured. Conventionally, time-correlated single photon counting[1], gated CCD with scanning the delay of excitation light, and streak camera to observe a linear region are utilized. On the other hand, recent progress in time-resolving CMOS image sensors for time-of-flight depth acquisition and fluorescence lifetime imaging enables to realize scanningless wide-field time-resolved biomedical imaging cameras.

There are two types of time-resolving CMOS image sensors based on single photon avalanche diode[2] and charge modulation pixels[3, 4]. Lateral electric field charge modulator (LEFM)[5, 6] has opened the pico-second regime ultra-high-speed computational imaging in a range of a few nano seconds or shorter, whose charge handling speed is close to the limitation of charge transfer speed in silicon[7, 8]. The charge modulation pixels are free from the drawbacks of TCSPC such as quite a few digital elements for creating histograms and pile-up causing higher cost and limited photon rate (namely, longer measurement time), respectively. Furthermore, they are compatible with emerging efficient computational methods with coded shutters[9].

One of the drawbacks of the charge modulation pixel is small number of charge taps, typically 2 to 4, at most 8[10], due to the limitations in terms of charge transfer speed and fill factor. However, it is very important to notice the fact that impulse response of biomedical specimen is mostly smooth and are approximated by equations with the limited number of parameters, which implies that charge modulation pixels with not so many taps can offer good performance close to that of TCSPC in some situations. For example, single-component fluorescence response and reflectance of homogeneous tissues are represented by the following equations:

$$I(t) = A\exp\left(-\frac{t}{\tau_f}\right) \quad (1)$$

$$I(\rho,t) = \left(4\pi D\frac{c}{n}\right)^{-3/2} z_0 t^{-5/2}\exp\left(-\mu_a\frac{c}{n}t\right)\exp\left(-\frac{\rho^2+z_0^2}{4D\frac{c}{n}t}\right) \quad (2)$$

Note that τ_f is lifetime of fluorophore, μ_a is absorption coefficient, and μ'_s is reduced scattering coefficient. ρ means a source-detector separation. See Ref. 11 for more definitions.

2. Modeling and processing flow

Temporal performance of the multi-tap charge modulation pixel is characterized by the number of taps, steepness of the time window at the rising and falling edges, width of the time windows, and crosstalk between the taps as shown in Fig. 1. In simulation, the minimum number of the total detected electrons (MNE) for all taps that satisfies a given relative error compliance is considered as an index of performance because an error in the estimated parameters, which is defined by standard deviation and difference between the ground truth and an average of the deduced parameters, can be alleviated only by decreasing the electron shot noise for a given condition of the sensor and specimen. Smaller index leads to less optical damage on the specimen, smaller full well capacity, or shorter measurement time.

The processing flow to find the MNE is shown in Fig. 2. How to generate noisy signals detected by the sensor taps and to deduce the optical parameters are described in Fig. 3. As shown in Fig. 4, a virtual response of sensor is firstly calculated by convolving an impulse excitation light with the responses of a specimen and a sensor. Here, sensor's temporal response is modeled by a 1st order lag. The number of electrons in each tap is figured out by integrating the virtual response in its time window. Fig. 5 shows examples of the responses of fluorophors with different lifetimes and their virtual responses assumed that sensor response, τ_f, is 0.2ns. Comparison of Fig. 5(b) and 5(c) suggests that modeling of the sensor response with a 1st order lag gives a good approximation of the virtual sensor response.

To measure temporal responses with a few taps, alignment of the time windows is important (Fig. 6). For a given range of fluorescence lifetime in FLIM, there are two options; the

alignment is optimized for the shortest or the longest lifetime. If the alignment is optimized for the shortest lifetime, in measurement of the longest lifetime, most electrons are detected by the 2nd time window, so that the 1st time window will suffer from huge photon shot noise, and vice versa. Fig. 7 compares MNEs in a given range of τ_f for two alignment cases, which suggests that the accuracy is not determined by the lifetime itself, but mostly determined by the alignment.

3. Simulations

Single-component fluorescence lifetime imaging and light scattering and absorption in tissues are discussed in the followings because quantitative measurement of metabolism based on autofluorescence of endogenous coenzymes and hemodynamics is of great interest in biomedical applications.

3.1 FLIM

Fig. 8 compares MNEs for different error compliances. For τ_f of 0.1-0.5ns (a half decades), MNEs are as follows: 2,000e- for ±10% error, 8,000e- for ±5%, and 200,000e- for ±1%. These results suggest that larger full well capacity is required to satisfy a smaller error compliance. MNE becomes smaller as the sensor response becomes faster, namely, as τ_i becomes smaller. Measurement of whole one decade (0.05-0.5ns) requires about 10 times more electrons than that for the half decade because, as mentioned above, penalty in MNE due to the time windows alignment becomes more significant.

The number of taps is discussed in Fig. 9. Fig. 9(a) shows an ideal case without any read noise and inter-tap crosstalk. The MNE decreases as the number of taps increases. However, if the sensor response is fast enough, the difference between 4 and 8 taps is not significant. Note that TCSPS gives the lower limit of MNE. If read noise and inter-tap crosstalk are finite, MNEs for 2 taps increases apparently (Fig. 9(b)). The difference between 2 taps and 4 taps becomes smaller probably because more total sensor noise is introduced by additional taps and less electrons are accumulated in each tap when more taps are prepared under the constraint that the number of total electrons is constant.

The effect of read noise and inter-tap crosstalk is shown in Figs. 10 and 11. Read noise is not very significant because measurement is almost shot noise limited. Inter-tap crosstalk is also not so influential, but more significant than read noise because it introduces more shot noise.

Two lifetime ranges are compared in Fig. 12. MNE is basically determined by the ratio of the shortest and longest lifetimes in a given range, which is independent of the absolute lifetime. However, for approximately $\tau_i > 2\tau_f$, MNE visibly increases.

3.2 Tissues

Fig. 13 shows some examples of temporal reflectance of tissues for two source-detector separations, 10mm and 20mm. As reference, MNEs for TCSPC are shown in Fig. 14. For large μ_a and small μ's, MNE increases because they cause a faster tissue response. For small μ_a, MNE also increases because the responses for different μ_a's become similar and more accuracy in measurement is necessary. Fig. 15 compares the number of taps and the number of time window shifting steps. As the number of taps increases, MNE decreases. Multiple measurement with shifting the tap timings decreases the MNE because that increases the number of measurements. However, it is less effective than increasing the number of taps. This is because with a wider time window width, information such as the position of the peak and gradual change of the gradient is more lost. TCSPC gives the smallest MNE because there are no read noise and crosstalk and the sampling pitch is small enough compared with the tissue response.

4. Conclusions and discussions

Overall, in FLIM, if error compliance is not severe (+/-10%), the number of electrons is less than 10,000. However, for more severe error compliance in FLIM and optical scattering cases, the number of electrons is more than tens of thousands. Therefore, large full well capacity is required. Although multiple image acquisition by shifting the tap timings can alleviate the number of electrons, increase of the number of taps is more effective. When the sensor response is slower or approximately equal to specimen's response, the number of necessary electrons significantly increases.

In this paper, the sensor performance required for two typical biomedical imaging cases were evaluated with a straightforward sampling method using the taps with the same pitch and width. However, more efficient computational sampling methods such as compressive sensing with coded shutters should be discussed for the potential of the charge modulation pixel to compete with or overcome TCSPC's performance.

Acknowledgements

This work is partially supported by Grant-in-Aid for Scientific Research (B) Number 15H03989 and (S) Number 2522905, and MEXT/JST COI-STREAM program.

References

[1] W. Becker, *Advanced time-correlated single photon counting techniques*, Springer (2005).
[2] J. Pavia *et al*., IEEE J. Solid-State Circuits **50** (2015) 2406.
[3] A. Esposito *et al*., Optics Express **13** (2005) 9812.
[4] A. Payne *et al*., ISSCC Dig. Tech. Papers (2014) 134.
[5] S. Kawahito *et al*., Proc. Image Sensor Workshop (2013) 1417.
[6] M. -W. Seo *et al*., ISSCC Dig. Tech. Papers (2015) 189.
[7] C. Canali *et al*., IEEE Trans. on Electron Devices **22** (1975) 1045.
[8] V. Dao *et al*., Proc. Image Sensor Workshop (2015) 6-02.
[9] F. Mochizuki *et al*., ISSCC Dig. Tech. Papers (2015) 116.
[10] Y. Shirakawa *et al*., Proc. Int'l Workshop on Image Sensors and Imaging Systems, poster (2016) 17.
[11] M. Patterson *et al*., Appl. Opt. **28** (1989) 2331.
[12] T. Kasugai *et al*., Proc. IS&T Int'l Symp. on Electronic Imaging 2016 (2016) IMSE-D48.

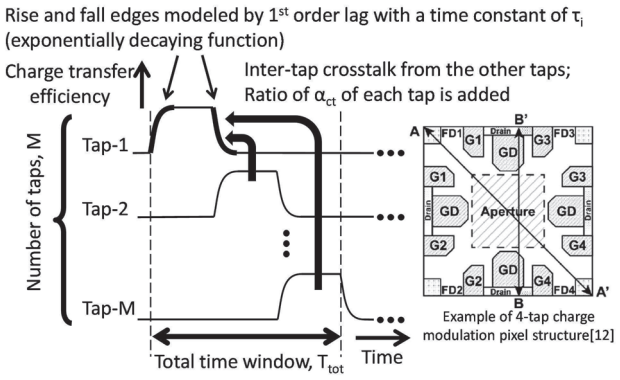


Fig. 1. Specifications of charge modulation pixel operation.

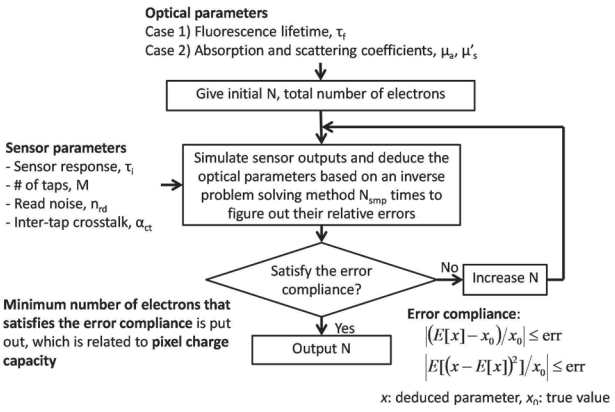


Fig. 2. Processing flow to find the minimum electron number.

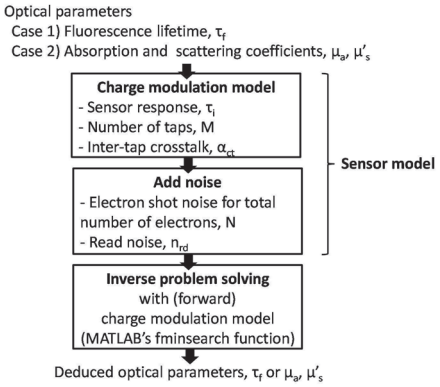


Fig. 3. Processing flow of generating noisy sensor output and deduction of optical parameter(s).

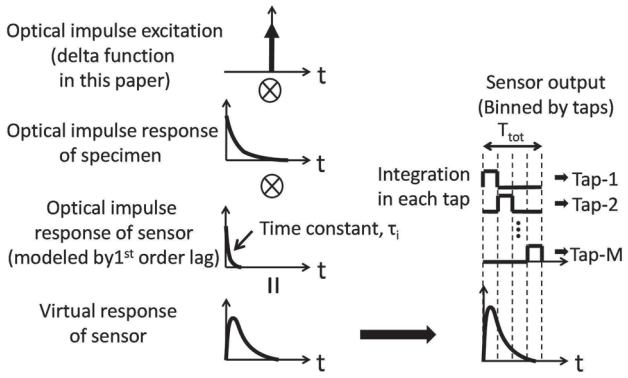


Fig. 4. Charge modulation pixel model.

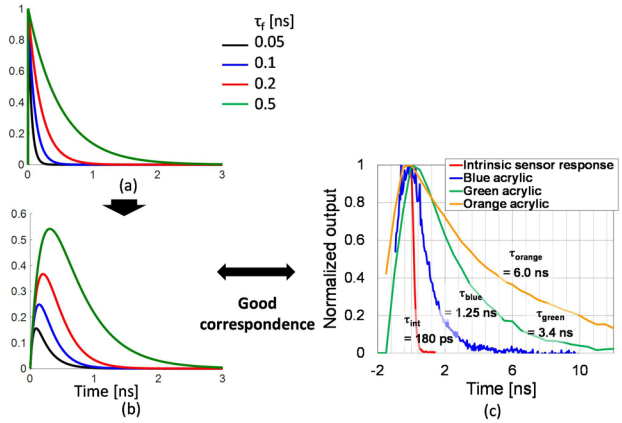


Fig. 5. Example waveforms in FLIM: (a) responses of specimen (no sensor response included), (b) virtual sensor response for τ_i=0.2ns, and (c) measured signals.

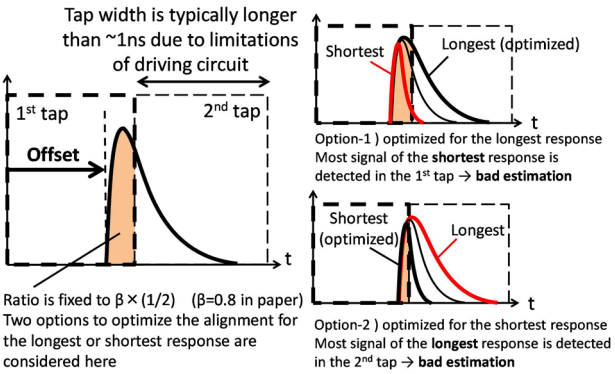


Fig. 6. Importance of tap alignment (e.g. 2 taps).

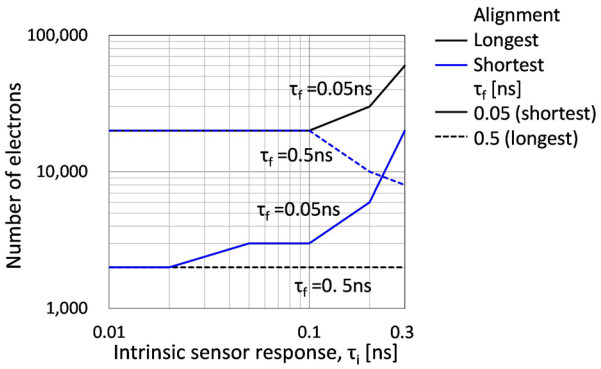


Fig. 7. Effect of tap alignment (T_{tot}= 4ns, τ_f=0.05-0.5ns, M=2, n_{rd}= 2e-, α_{ct}= 3%, err= +/-5%).

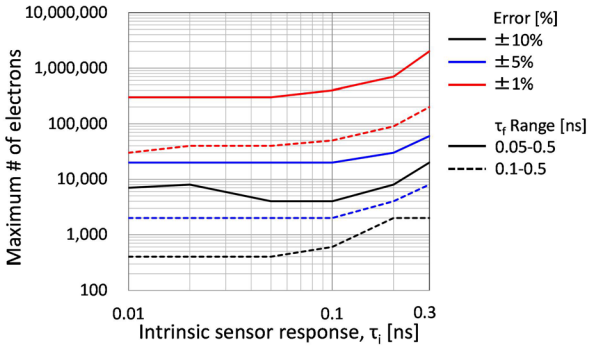
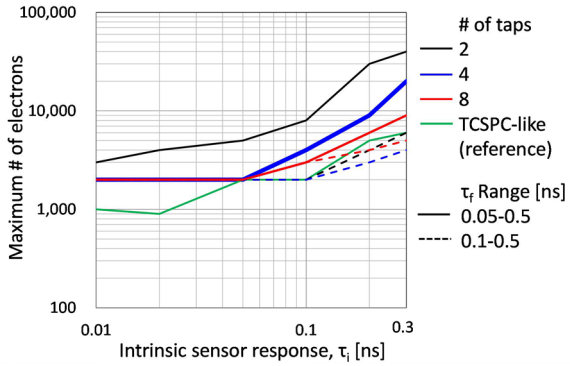
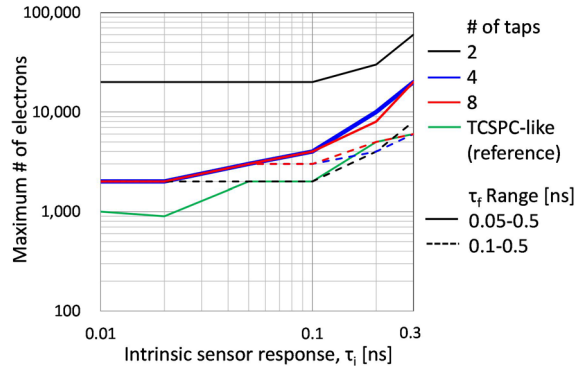


Fig. 8. Error compliance (T_{tot}=4ns , τ_f=0.05-0.5ns, M= 2, n_{rd}= 2e-, α_{ct}= 3%, alignment: longest).



(a)



(b)

Fig. 9. Number of taps: (a) without read noise and crosstalk (T_{tot} = 4ns, τ_f=0.05-0.5ns, n_{rd}= 0e-, α_{ct} = 0%, alignment: longest, err= +/-5%) and (b) with read noise and crosstalk (n_{rd}= 2e-, α_{ct} = 3%).

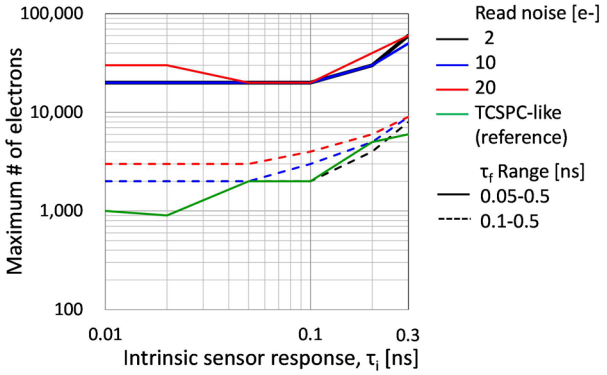


Fig. 10. Effect of read noise (T_{tot} = 4ns , τ_f=0.05-0.5ns, M= 2, α_{ct}= 3%, alignment: longest, err= +/-5%).

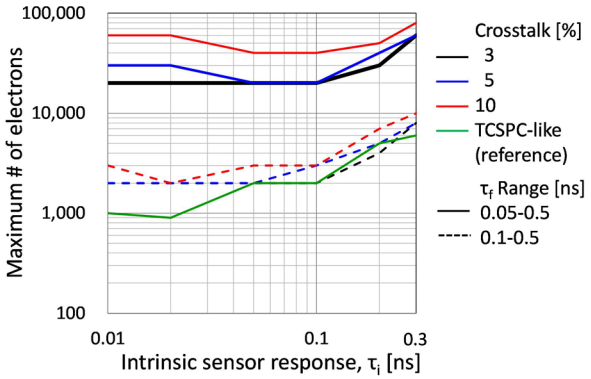


Fig. 11. Effect of inter-tap crosstalk (T_{tot}= 4ns, τ_f=0.05-0.5ns, M= 2, n_{rd}= 2e-, alignment: longest, err= +/-5%).

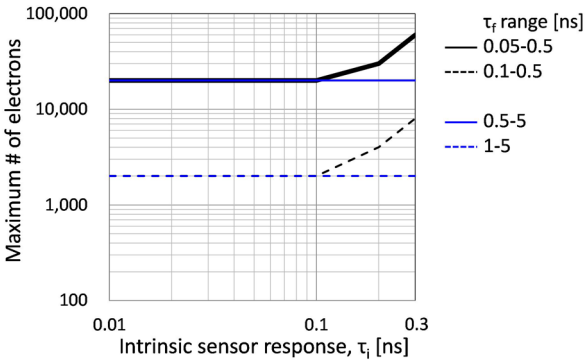


Fig. 12. Lifetime range (T_{tot}= 4/8ns, M= 2, n_{rd}= 2e-, α_{ct}= 3%, alignment: longest, err= +/-5%).

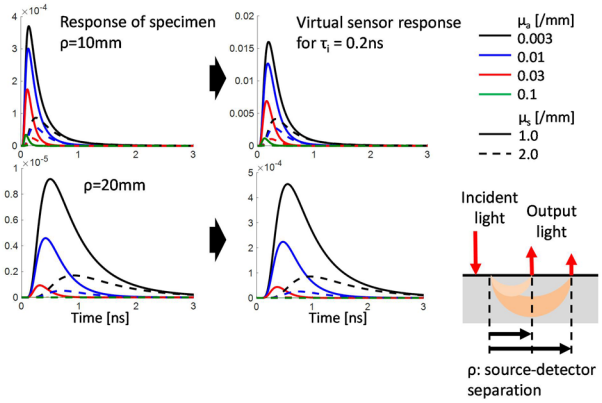


Fig. 13. Example waveforms in tissue scattering.

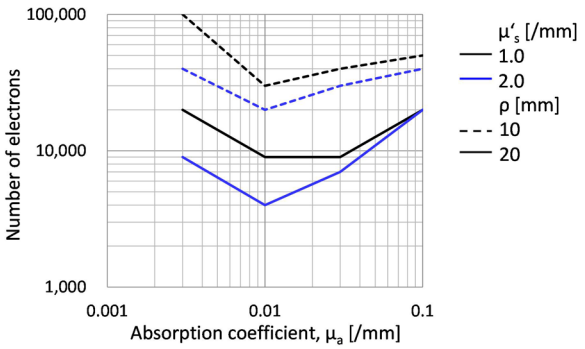


Fig. 14. TCSPC-like condition as reference (T_{tot}= 2.5/5ns, ρ=10/20mm, M= 512, n_{rd}= 0e-, α_{ct}= 0%, τ_i= 0.2ns, err=+/-5%).

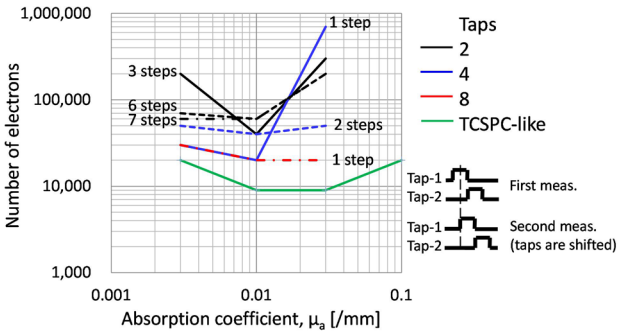


Fig. 15. μ_a, numbers of taps and shift steps (T_{tot}= 5ns, μ'_s=1.0, ρ=20mm, n_{rd}= 2e-, α_{ct}= 3%, τ_i= 0.2ns, alignment: shortest, err= +/-5%).