Sub-pixel Architecture of CMOS Image Sensor Achieving over 120 dB Dynamic Range with less Motion Artifact Characteristics

T. Asatsuma[1], Y. Sakano[1], S. Iida[1], M. Takami[2], I. Yoshiba[1], N. Ohba[2], H. Mizuno[1], T. Oka[1], K. Yamaguchi[1], A. Suzuki[1], K. Suzuki[1], M. Yamada[1], Y. Tateshita[1], and K. Ohno[1]
[1]Sony Semiconductor Solutions, Kanagawa, Japan. [2]Sony Semiconductor Manufacturing, Kumamoto, Japan.
E-mail: Tomohiko.Asatsuma@sony.com

Abstract—We developed a CMOS image sensor with a 3 μm pixel pitched sub-pixel architecture. This sensor architecture shows a high dynamic range of over 120 dB with less motion artifact characteristics.

I. INTRODUCTION

In recent years, real-time sensing has been creating new businesses and social changes, specifically in the internet of things and automotive fields. The accurate perception of moving objects and detection with high color reproducibility for all light conditions is a necessity.

A conventional high dynamic range (HDR) technique uses the multiple exposure method [1–2]. However, this technique causes motion artifacts depending on the sampling time difference of the moving objects.

We have developed a CMOS image sensor with a new architecture. The characteristic of this sensor is that it has been designed with a sub-pixel architecture that contains a single large photodiode, a single small photodiode, and an in-pixel floating capacitor.

II. SENSOR ARCHITECTURE

Fig. 1 shows the block diagram of the image sensor with a pixel array of 1920 × 1200 pixels, read-out circuits (load MOS transistors, column ADCs, DAC), driver circuits (row driver, row decoder), image signal processors, and other circuits (PLL, regulator, MIPI I/F, CPU, etc.).

Fig. 2 shows the pixel schematic of the sub-pixel architecture. This circuit employs a single large photodiode (SP1), a single small photodiode (SP2), an in-pixel floating capacitor (FC), and seven transistors. The SP1 has a high sensitivity (Green) of 36000e-/lx-s, which is 10 times that of SP2. The linear full-well capacities (FWCs) of SP1 and SP2 are 10000e- and 78500e-; these values are attributed to the FC, respectively. The seven transistors are as follows: transfer gate of SP1 (TGL), transfer gate of SP2 (TGS), floating diffusion gate (FDG), floating capacitor gate (FCG), reset transistor (RST), select transistor (SEL), and the source follower amplifier (AMP). A floating diffusion (FD) is separated as FD1, FD2, and FD3 by FDG and FCG. The two electrodes of the FC are connected to FD3 and a counter electrode with a supply voltage of FCVDD.

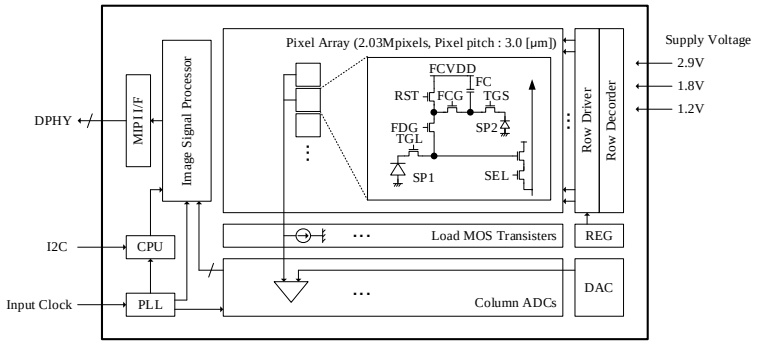


Fig. 1. Sensor block diagram

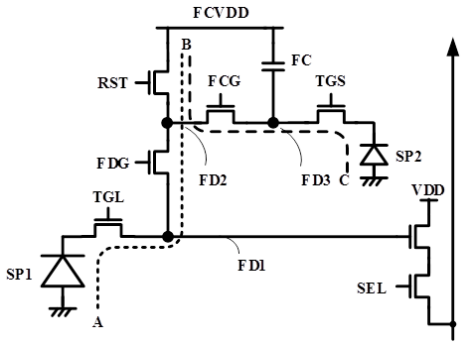


Fig. 2. Pixel schematic

Fig. 3 shows the pixel top view. One pixel consists of one each of a large and small on-chip microlens (OCL). The OCL of SP2 is located in the gap section of the OCL of SP1. This results in a sensitivity ratio of 10:1 for SP1:SP2, where the sensitivity of SP1 is maintained. Fig. 4 shows the cross-sectional view of pixels. The OCL thickness of SP1 is ~4 times the thickness of SP2 for a high sensitivity ratio for SP1:SP2. Deep trench isolations are employed in the silicon substrate. As shown in Fig. 5, the optical crosstalk and leakage of electrical charges from SP1 to SP2 are prevented by the refractive index difference and physical isolation, respectively. The color ratio (Red/Green) of SP2 at a wavelength of 530 nm is reduced by 59%, thereby achieving the same value as the color ratio of SP1.

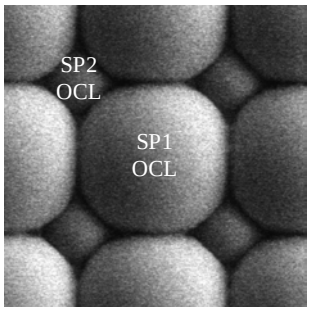


Fig. 3. Pixel top view

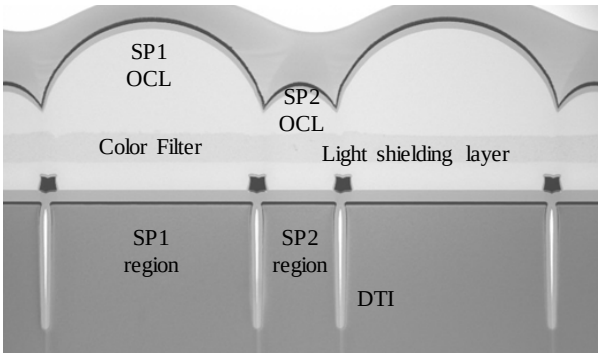


Fig. 4. Pixel cross-sectional view

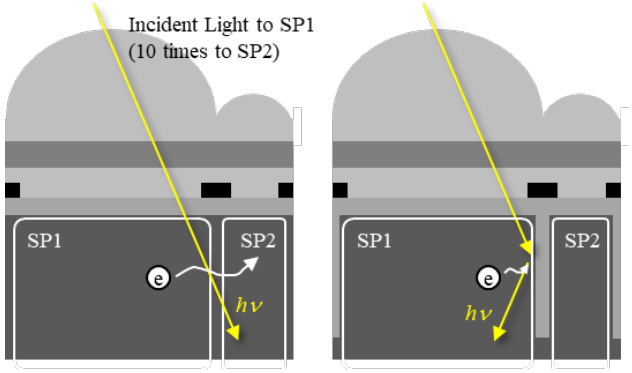


Fig. 5. Sensor structure for suppressing electrical and optical crosstalk

Fig. 6 shows the pixel-driving sequence. The signals from SP1 and SP2 are output serially. Further, the electrical charges accumulated in SP1 are converted to signal voltage in two modes by switching the FDG: high conversion gain (HCG) and low conversion gain (LCG). As such, three types of signals are read out in a single exposure. First, an exposure of SP1 and SP2 begins by resetting SP1, SP2, and FC. Then, LCG reset level 2 and HCG reset level 1 are sampled. Subsequently, HCG signal level 1 is sampled after switching TGL, and LCG signal level 2 is sampled after switching TGL once again. Two signals are read out by performing corelated double sampling for each reset and signal level: SP1H and SP1L. Subsequently, the signal that comes from SP2 (SP2L) is read out by performing delta reset sampling (DRS), in which signal level 3 is sampled first, followed by reset level 3. As signal charges are accumulated in FD3, FD3 cannot be reset prior to sampling signal level 3. The disadvantage of DRS is that kTC noise cannot be removed; however, it can be suppressed by ensuring sufficient capacitance of the FC.

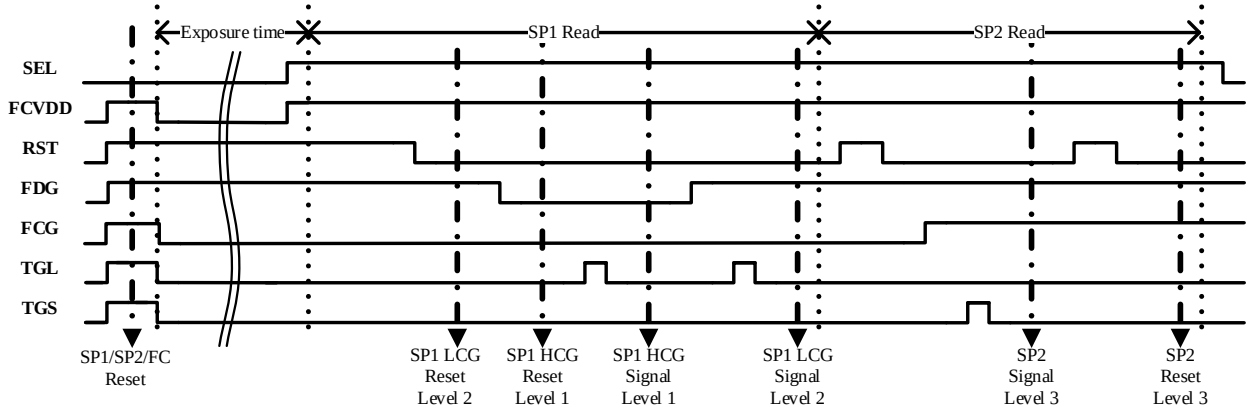


Fig. 6. Sensor driving sequence

Fig. 7 shows the potential diagram of SP1. The cross section shows the path of dotted line A-B in Fig. 2. At the start of the exposure period, switching the TGL, FDG, and RST resets the electrical charges of SP1 (Fig. 7-a). FD1 and FD2 are always reset during the exposure period (Fig.7-b), and then, LCG reset level 2 is sampled when RST is turned off (Fig. 7-c). Subsequently, HCG reset level 1 is sampled when the FDG is turned off (Fig. 7-d). Then, HCG signal level 1 is sampled after TGL is switched and the electrical charges accumulated in SP1 are transferred to FD1 (Fig. 7-e). After FD1 is connected to FD2 by turning the

FDG on, LCG signal level 2 is sampled when TGL is switched once again and the remaining charges in SP1 are fully transferred to FD1 and FD2 (Fig. 7-f). Thus, SP1H and SP1L can be read out from the electrical charges accumulated in SP1.

Fig. 8 shows the potential diagram of SP2. The cross section shows the path of dotted line C-B in Fig. 2. At the start of the exposure period, switching the TGS, FCG, and RST resets the electrical charges of SP2 and FD3 (Fig. 8-a). During the exposure period, the electrical charges that come from SP2 are accumulated in SP2 and FD3 (Fig. 8-b). After FD3 is connected to FD1 and FD2 by turning the FCG on, the signal level of SP2 is sampled when TGS is switched and the electrical charges accumulated in SP2 are fully transferred to FD3 (Fig. 8-c). Again, RST is turned on and the electrical charges of FD1, FD2, and FD3 are reset and sampled as the reset level of SP2 when RST is turned off (Fig. 8-d). Thus, SP2L can be read out from the electrical charges accumulated in SP2.

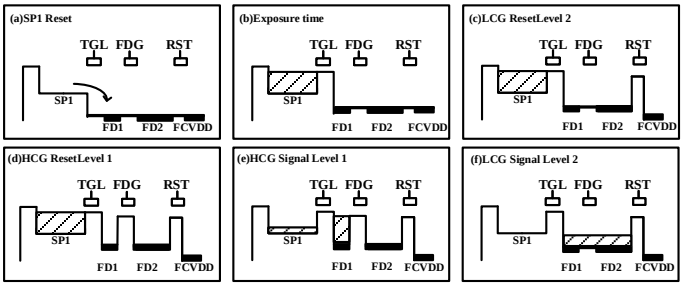


Fig. 7. SP1 potential diagram (A-B in Fig.2)

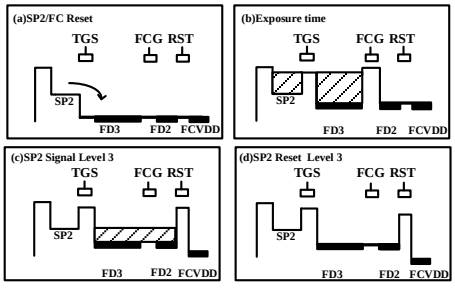


Fig. 8. SP2 potential diagram (C-B in Fig.2)

III. SENSOR CHARACTERISTICS

Fig. 9 shows the photo responses curve for SP1H, SP1L, and SP2L. The linear FWCs of SP1 and SP2 are 10000e- and 78500e-, respectively. By multiplying SP2L and a gain of 10, its linear FWC becomes equivalent to 785000e-, thus a dynamic range of 121 dB is achieved. Fig. 10 shows the signal-to-noise ratio (SNR) curve of the synthesized signal comprising three types of signals, i.e., SP1H, SP1L, and SP2L used in low-light, medium-light, and high-light scenes, respectively. SNR ≥ 20 dB can be maintained even when connecting SP1L to SP2L at 60 °C.

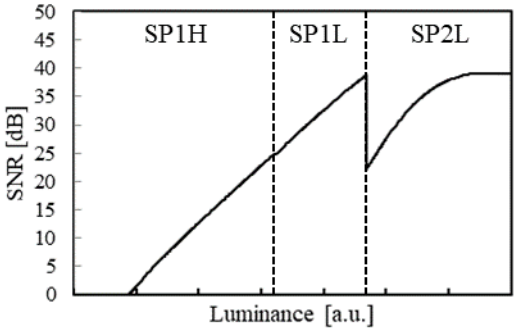


Fig. 9. SNR curve of synthesized signal

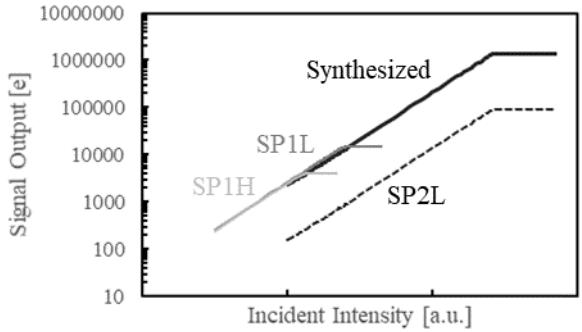


Fig. 10. Photo response curve of each signal

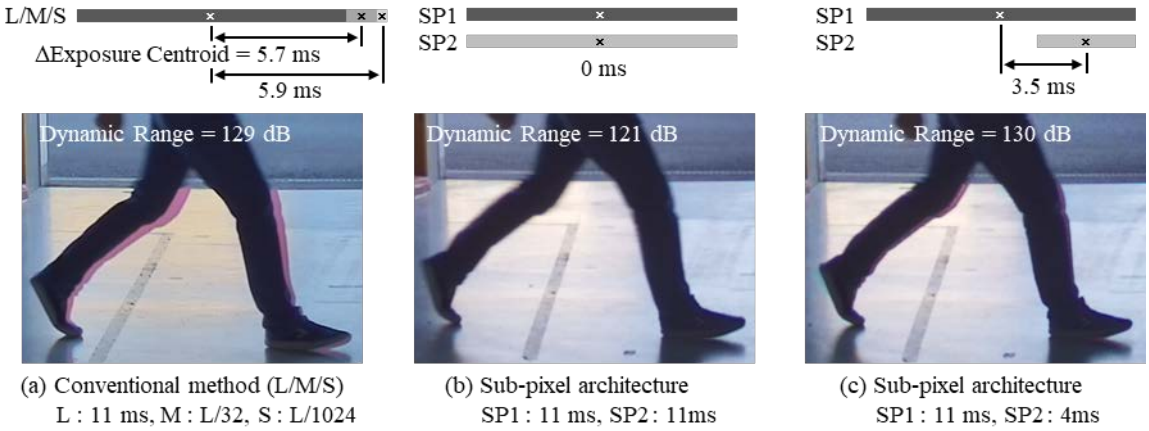
Fig. 11 shows the synthesized image of an HDR scene. As shown in Fig. 11-c, the inside of the tunnel and the outside landscape are accurately captured by synthesizing SP1H, SP1L, and SP2L.

(a) SP1L 0.3 ms (b) SP1L 11 ms (c) SP1H+SP1L+SP2L 11 ms

Fig. 11. Synthesized image (high dynamic range scene)

Fig. 12-a shows a synthesized image for moving object by using a conventional multiple exposure method. An exposure time ratio of middle/long is 32 times, that of short/long is 1024 times. A large motion artifact occurs due to synthesizing a large different centroid of exposure time. Fig. 12-b shows a synthesized image of SP1 and SP2 which exposure time are set to 11 ms. In this case, a motion artifact free with an HDR of 121dB can be achieved. Fig. 12-c shows a synthesized image of SP1 and SP2 which exposure time are set to 11 ms and 4 ms, respectively. In this case, an HDR of 130dB can be achieved thanks to

set a different exposure time ratio between SP1 and SP2. With different sensitivity photodiodes and an FC with large FWC, a centroid difference of exposure time can be minimized compare to multiple exposure method and it can decrease motion artifact and can extend a dynamic range.



(a) Conventional method (L/M/S) L : 11 ms, M : L/32, S : L/1024

(b) Sub-pixel architecture SP1 : 11 ms, SP2 : 11ms

(c) Sub-pixel architecture SP1 : 11 ms, SP2 : 4ms

Fig. 12. Motion artifact comparison for dynamic range priority mode

Table 1 shows the performance of the developed sensor, and Table 2 shows a comparison of the characteristics reported previously [4]–[6]. Well-balanced characteristics are achieved using the CMOS image sensor with the sub-pixel architecture.

Table 1. Sensor performance

Parameter	Unit	Value
Power Supply	V	2.9
Process Technology	–	90 nm 4Cu1AL
Pixel Array	pixels	1920 × 1200
Pixel Pitch	μm	3
SP1L Linear Full-well Capacity	e-	10000
SP2 Linear Full-well Capacity	e-	78500
Sensitivity (Green, 3200K with IR cut filter)	e-/lx-sec	36000
Sensitivity Ratio (SP1 : SP2)	–	10 : 1
Random Noise	e-rms	0.68
Dynamic Range	dB	120

Table 2. Comparison of sensor characteristics

Parameter	Unit	This Work	VLSI 2018 [2]	ISSCC 2016 [3]	IISW 2017 [4]
Pixel Pitch	μm	3.0	3.0	6.0	2.8
Random Noise@RT	e-rms	0.68	6.2	5.4	1
Full-well Capacity	e-	78500	489K	600K	50000
Dynamic Range	dB	121	121	123.8	94

IV. CONCLUSION

We developed a new CMOS image sensor using a sub-pixel architecture, which consisted of a single large photodiode, a single small photodiode, and an in-pixel floating capacitor with a pixel pitch of 3μm. This sensor achieves an HDR of 121 dB with the same exposure time of SP1 and SP2. Furthermore, the dynamic range can be expanded by reducing the exposure time of SP2 with less exposure centroid difference than the conventional multiple exposure method. This technique is suitable for less motion artifact characteristics.

REFERENCES

[1] Trygve Willassen et al., "A 1280x1080 4.2μm Split-diode Pixel HDR Sensor in 110 nm BSI CMOS Process," in IISW 2015.
[2] Sergey Velichko et al., "140 dB Dynamic Range Sub-electron Noise Floor Image Sensor," in IISW 2017, pp. 294-297.
[3] Chris Silsby et al., " A 1.2MP 1/3" CMOS Image Sensor with Light Flicker Mitigation," in IISW 2015.
[4] M. Takase et al., "An Over 120 dB Wide-Dynamic-Range 3.0 μm Pixel Image Sensor with In-Pixel Capacitor of 41.7 fF/μm2 and High Reliability Enabled by BEOL 3D Capacitor Process," in Symp. VLSI 2018, pp. 71-72.
[5] Kazuko Nishimura et al., "An Over 120dB Simultaneous-Capture Wide-Dynamic-Range 1.6e- Ultra-Low-Reset-Noise Organic-Photoconductive-Film CMOS Image Sensor," in ISSCC 2016, pp. 110-112.
[6] Johannes Solhusvik et al., "A 1392x976 2.8μm 120dB CIS with Per-Pixel Controlled Conversion Gain," in IISW 2017, pp. 298-30.