A 4M pixel full-PDAF CMOS image sensor with 1.58μm 2×1 On-Chip Micro-Split-Lens technology

Sozo Yokogawa† [1], Isao Hirota[1], Isao Ohdaira[1], Masao Matsumura[1], Atsushi Morimitsu[1], Hiroaki Takahashi[1], Toshio Yamazaki[2], Hideki Oyaizu[2], Yalcin Incesu [3], Muhammad Atif [3], Yoshikazu Nitta[1]*

[1] Sony Corporation, Atsugi Tec. 4-14-1 Asahi-cho, Atsugi, Kanagawa, 243-0014, Japan
[2] Sony Corporation, Sony City Osaki 2-10-1 Osaki, Shinagawa-ku, Tokyo, 141-8610, Japan
[3] Sony Deutschland GmbH, Stuttgart Tec. Center, Hedelfinger Str. 61, 70327, Stuttgart, Germany
†Corresponding author: Sozo.Yokogawa@jp.sony.com, Tel: +81-50-3141-0982

ABSTRACT

We introduce a 1.58μm back illuminated CMOS image-sensor (BI CIS) which contains 4 Mega 2×1 shared on-chip micro-lenses (OCL). With this architecture, we realize a full phase-detection auto focus (PDAF) CIS, where all photo detectors (PDs) are L/R paired with a common OCL, the same color on-chip-color-filter (OCCF) and without metal grid between the L/R pixel boundary. The device has a good separation ratio of 2.59 and is more than twice sensitive under normal incident 550nm plane wave compared to current state-of-the-art PDAF pixels with half metal-shielded aperture. Because the device has no obstacle structure on its optical path, it hardly suffers from reflection and diffraction of the incident light. This feature is adequate to realize PDAF function especially for the CISs with smaller pixels. In addition, the OCCF configuration of the device is a 45º rotated to Bayer CF pattern, and hence, every 2 lines contain only Green pixels. This CF configuration makes it suitable to detect the phase difference between L and R sub-images. Furthermore, this architecture is in particular suitable for HDR imaging by varying the exposure times for each L and R sub-images because it is possible to capture two set of sub-images with almost identical optical characteristics to the on-focus object. In this paper, we report on the device performance and propose versatile applications including high-sensitive AF and HDR for the full-PDAF CIS.

In the past decade, the rapid progress of CMOS image sensor technology has drastically improved their image quality along with the trend of shrinking pixel size, increasing pixel number and smart signal processing. Recently, Back illuminated CMOS image sensor with a stacked logic structure were reported, which realize various on-chip functions including PDAF and HDR functions[1].

We have developed a novel prototype device with 1.58μm◇, 2×1 shared OCL technology, so-called micro-split lens (μSL) pixel, as shown in the pixel schematic of Figure 1. Each pixel is equipped with L and R PDs. There is no metal ridge at the L/R pixel boundary which minimizes the reflection and diffraction caused by such obstacle structures. Therefore, the device achieves a two times higher sensitivity than conventional architectures with half-shielded PDAF pixels and provides a better L/R discrimination capability[2].

Figure 2(a) shows a block diagram of the chip. A chip is equipped with a 4 Mega 2×1 shared OCL pixel array, diagonally 4.60mm (Type 1/4) in size, load transistors, H/V scanners and column parallel ADCs in the peripheral. The inset (b) is a micrograph of the chip including a close-up view of the OCL of the pixel.

We designed the pixel structure to achieve good L/R discrimination capability by using FDTD simulation. Figure 3(a), (b) show schematics of the pixel and (c), (d) are the snapshots of magnetic field |H| distributions cut along the sensor plane obtained with the simulations. The simulations were performed with the incident plane wave at λ = 550nm with the incident angles of θ = +15º (a), (c) and θ = 0 º (b), (d), respectively. Here the angle is defined to be θ = 0 º for the normal incident light and is changing along with the plane including the horizontal axis of the pixels. As shown in (c), (d), the incident light is mainly guided into the left pixel at the incident angle of θ = +15º and is equally distributed to both the L and R pixels at θ =0º.

Figure 4 shows the incident angle dependency of a green pixel sensitivity of the device with the simulation (c) and measurements (d).The data are taken along the horizontal plane denoted as the magenta lines on (a),(b). The incident light is set to λ = 550nm both in the simulation and measurement and the angle sweep ranges from θ = +30º to θ = -30º in steps of 2º. Solid lines represent the data of this μSL pixel (1.58μm◇), and the dotted lines are for those of the half shielded PDAF pixel (1.12μm□), showing good consistency between the simulation and the measurements. The μSL pixel at θ = 0º achieves a 2.27 times higher sensitivity compared to the PDAF pixel and almost equal sensitivity to the normal green pixel, because the area of each L/R pixel of the 1.58μm◇ PDs is same as that of 1.12μm□ pixel. The μSL pixel has a separation ratio as high as 2.59 whereas that of the PDAF is about 1.65. The definition of separation ratio is:

$$Separation\ Ratio = Ave\left(\int_0^{+10} \frac{Sens\ L_{\mathrm{pixel}}}{Sens\ R_{\mathrm{pixel}}} d\theta, \int_{-10}^{0} \frac{Sens\ R_{\mathrm{pixel}}}{Sens\ L_{\mathrm{pixel}}} d\theta\right)$$

To evaluate and demonstrate the image quality of the device, we prototyped the AF lens module with F2.4, focal length of 2.87mm and 4 piece plastic lenses. As the CF configuration of the device is different from the Bayer format, a specifically tailored remosaicing block is required for the system. In a prototype, we implemented this function on the FPGA board and delivered the output data in Bayer format. This data is subsequently processed with a general demosaicing block to get the final RGB image. Figure 5(a) shows a sample image of the star-chart[3] taken with the device after full image reconstruction pipeline. Figure 5 (b) is the close-up view of the area which is denoted as yellow line in (a).

One of the advantage of this device is the density of PDAF pixels and 100% pixels are paired PDs. State-of-the-art devices that have embedded PDAF pixels allow only few percent for the PDAF function pixel in order to maintain the image quality. Furthermore, the 45º rotated Bayer CF

pattern provides in every second line Green pixel only as shown in Figure 1. This CF configuration is suitable to calculate a disparity between L and R images. To demonstrate these advantages, we developed the PDAF function by using the ZSAD[4] block matching algorithm for the disparity estimation (DE) between L and R images.

To demonstrate PDAF function, we prepared software-based PDAF demo system. We prepared a LUT for the module, which is a correspondence table describing the relationship between an amount of disparity and an AF driver position, by taking the random dot chart images at a number of distances. By using the LUT and the estimated disparity of the captured image, we implemented the PDAF functionality and evaluated its capability under various illumination condition. The test conditions are summarized as follows:

- Target: B/W random dot chart at the distance of 30cm
- Illumination levels are set to make the sensor signal levels 30% of Qs (Qs = Saturation signal level)
- Exposure times is fixed to 33mS
- ROI is set to the center 400×300 pixels. This is corresponding to H=1/8, V=1/8 of the full image size
- Gain is variable from 0dB to 42dB

The evaluation steps are as follows. At first, we set the AF driver position to infinity (Count = 250) for every set of the trial. Then we calculated the amount of disparity with the DE algorithm within the ROI of the chart. From the comparison between the derived disparity value and the LUT value which we prepared beforehand, we received the target AF driver position of the first AF trial, and then send command to the AF driver for focusing to the target. To check the convergence , we continued this procedure 10 times. This is the 1 set of the AF evaluations. We executed the evaluation 30 times to confirm repeatability and determine standard deviations of the PDAF results under all gain conditions. The results is documented in Table 1.

Figure 6(a) shows that the result with 0dB condition. First AF trials are distributed around the counts of 520-528 and from second to tenth, they are converged around 512. At higher gain conditions, the results are plotted in Figure 6(b-d), showing small scattering even at 42 dB condition with 8 lux illuminance (b). The scattering becomes larger at 2 lux illuminance for (c) and 1lux illuminance for (d). The standard deviations of the repeatability are summarized in Table 1. Although there is a trend that the deviation becomes larger with the higher gain, the number is quite constant between 0dB to 42 dB, starts to increase from 1-2 lux, and finally breaks down at 0.5 lux. These results suggest that the PDAF sensitivity is almost comparable to the sensor sensitivity, because the device has only a few percent of Qs signal level at 0.5 lux condition with an exposure time of 33 mS. Figure 6(e) shows the results of monitoring the contrast index taken with a Bayer sensor whose pixel size is 1.12 μm, 8 Mega pixel and with the same lens module as a reference. The results show that the distinct contrast peaks at low gain conditions from 0 dB to 30 dB, but suffer from large scattering at 42 dB conditions. At a gain of 42 dB and half illumination or darker conditions, it is expected to be difficult to find the contrast peak position accurately.

A high dynamic range (HDR) function is an important feature for most of the image sensor applications because the intrinsic CIS dynamic range is narrower than that of human eyes. Numerous approaches have been developed to improve the performance, such as multi-exposure HDR during a single frame. The 2×1 shared OCL device is very suitable to realize a HDR function in a simple yet robust way. Because each OCLs is equipped with two PDs, one promising approach to get HDR function is to make simply shorten the exposure time for one pixel (Left images) while expanding the exposure time for the other (Right images) as shown in Figure 7. Because the L and R images are almost optically identical toward the on-focused objects, it is hardly suffering from color artifacts or degradation of spatial resolution which often arise in Bayer sensors during the HDR synthesis. However motion blur artifacts still remain.

Gain		Exp. Time	Luminosity*	AF driver STDEV	Signal Level
Analog	Digital	[mSec]	[lux]	[1σ]	Relative
0	0	33	800 (D)	—	1.0
18	0		130 (H)	0.32	1.0
	6		70 (H)	1.61	1.0
	12		35 (H)	1.94	1.0
	18		20 (H)	2.30	1.0
	24		9 (H)	1.28	1.0
			4 (H)	2.11	0.5
			2 (H)	3.56	0.25
			1 (H)	4.95	0.125
			0.5 (H)	17.4	0.0625

Table 1: Results of the PDAF repeatability test under the gain conditions from 0 dB to 42 dB with the fixed exposure time of 33 mS. Luminosities are selected to set the average G-signal levels of the device to ~30% of Qs. D and H of the luminosity column stands for D65, Halogen lamp were used during the test.

Here, we simulated HDR images by using two set of images taken by the device, one is with short exposure (40 mS) and the other is with long exposure (160 mS). By selecting the short exposure data for Left image and the long exposure data for Right image, we synthesized HDR data. Figure 8(a) is the simulated HDR image and (d) is the normal image, comparable spatial resolution as shown in both the panels (b), (e), whereas big differences in HDR can be seen in the panels (c),(f).

In summary, we developed a 1.58μm 4M pixel CIS with 2×1 shared OCL technology. The device has good separation ratio of 2.59 and higher sensitivity compared to commonly implemented half metal-shielded PDAF pixels. This results in good PDAF capability at high-gain and low-light conditions. The feature of this device that can take 2 set of images, are highly compatible to the HDR function.

References:

[1] SUKEGAWA, Shunichi, et al. A 1/4-inch 8mpixel back-illuminated stacked cmos image sensor. In: *Solid-State Circ uits Conference Digest of Technical Papers (ISSCC), 2013 I EEE International*. IEEE, 2013. p. 484-485.

[2] HIROTA, Isao. *Solid-state image pickup device and cam era system.*US 2014/0071244 A1

[3] LOEBICH, Christian, et al. Digital camera resolution measurement using sinusoidal Siemens stars. In: *Electronic Imaging 2007*. International Society for Optics and Photonics, 2007. p. 65020N-65020N-11.

[4] HANNAH, Marsha Jo. Computer matching of areas in stereo images. STANFORD UNIV CALIF DEPT OF COMPUTER SCIENCE, 1974.

*: Heidelberg Collaboratory for Image Processing, Universität Heidelberg, Speyerer Straße 6, 69115 Heidelberg, Germany

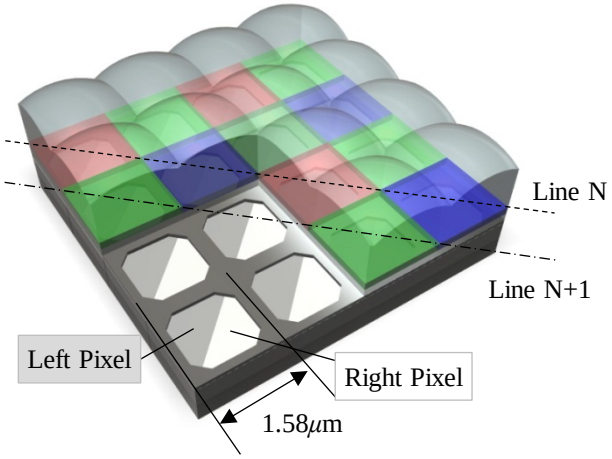

Fig.1: A schematic image of the pixel design.

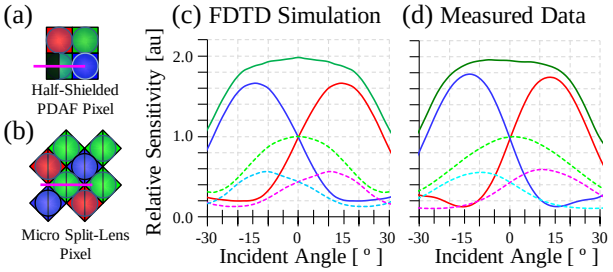

Fig.4: Angular response of the sensor sensitivity for the half-shielded PDAF pixel (a) and the μSL pixel (b). The solid lines are for μSL. Blue, Red, Dark Green curves are corresponding to Left, Right, and L+R pixels, respectively. The dotted lines are for the PDAF pixels. Light Green, Cyan, Magenta are corresponding to the normal green pixel, Left and Right PDAF pixels.

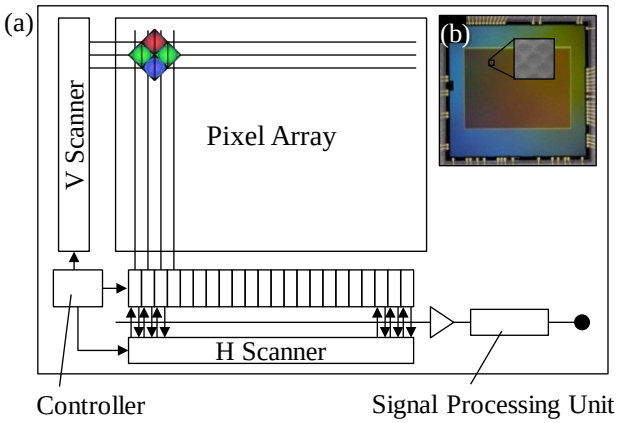

Fig.2: (a) A block diagram of the chip. The inset (b) is a micrograph of the chip including a close-up view of the pixels.

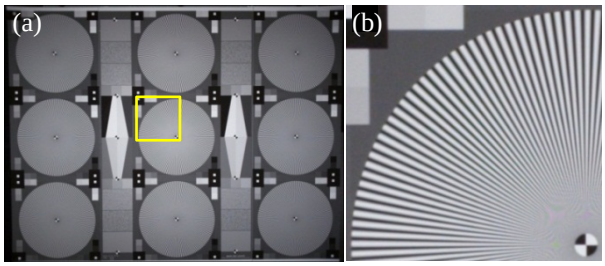

Fig.5(a): A sample image of the Siemens 9 star chart after full image reconstruction. (b): A close-up view which is denoted as yellow line in (a).

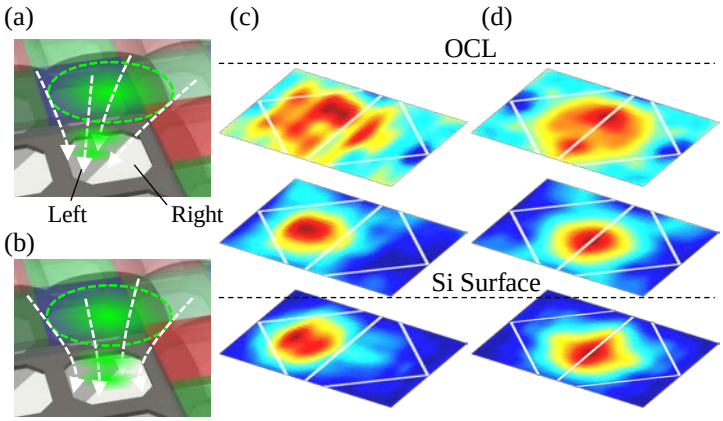

Fig.3: (a), (b) are schematic images of the pixel at the incident light of $\lambda = 550nm$, and the incident angle θ = +15º and θ = 0 º, respectively. (c), (d) are the corresponding |H| distributions cut along the plane under the OCL, above the Si surface, and under the Si surface.

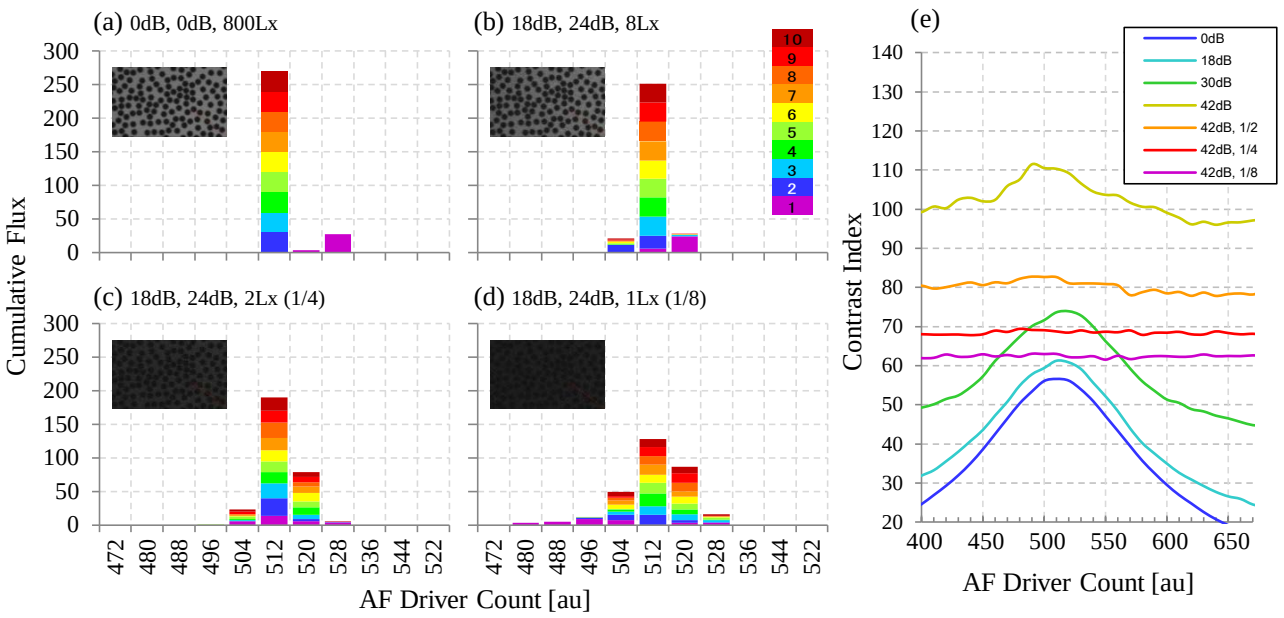

Fig.6: AF driver count repeatability test at 0dB condition (a) and at 42dB conditions (b-d). Each color is corresponding to the results of N-th trial. Inset images are the close-up view of the random dot chart raw images at each lighting conditions. (e): Contrast index vs AD driver count taken by the Bayer sensor for reference.

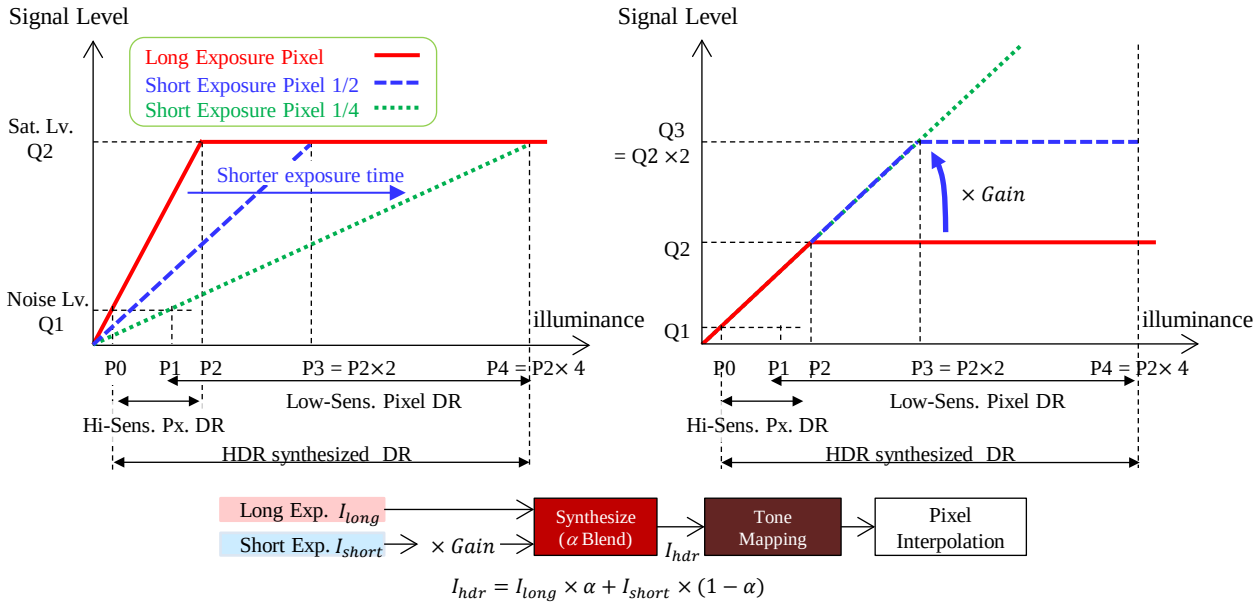

Fig.7: An example of the HDR synthesis and the block diagram.

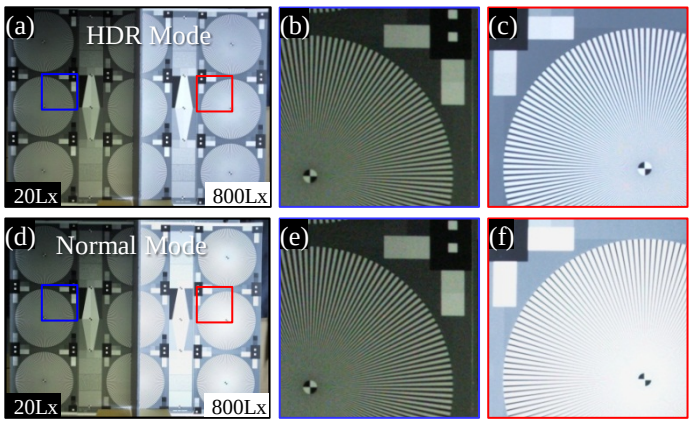

Fig.8: Comparison between the simulated HDR mode (a-c) and the normal mode (d-f). Close-up area is denoted as Blue and Red lines in (a), (d). The ratio of the exposure times is assumed to be 4:1.