

# A 1280x1080 4.2 $\mu$ m Split-diode Pixel HDR Sensor in 110nm BSI CMOS Process

Trygve Willassen<sup>1</sup>, Johannes Solhusvik<sup>1</sup>, Robert Johansson<sup>1</sup>,  
Sohrab Yaghmai<sup>1</sup>, Howard Rhodes<sup>2</sup>, Sohei Manabe, Duli  
Mao<sup>2</sup>, Zhiqiang Lin<sup>2</sup>, Dajiang Yang<sup>2</sup>, Orkun Cellek<sup>2</sup>, Eric  
Webster<sup>2</sup>, Siguang Ma<sup>2</sup>, and Bowei Zhang<sup>2</sup>  
<sup>1</sup>OmniVision Technologies, Oslo, Norway,  
<sup>2</sup>OmniVision Technologies, Santa Clara, USA  
<sup>1</sup>Gaustadalléen 21, 0349 Oslo, Norway  
Phone: +47 45660481, email: trygve.willassen@ovt.com

A triple exposure, high dynamic range (HDR), CMOS image sensor with an active array size of 1280 x 1080, and sub 1e- noise floor is presented. This sensor is the first sensor that utilizes OmniVision's split-pixel technology ported to the BSI process technology. The pixel size is 4.2 $\mu$ m, and the pixel incorporates programmable conversion gain (CG). There are three exposure values: the long exposure channel (L), captured by the large photo diode (LPD); the short exposure (S), captured by the small photo diode (SPD); and the very short exposure (VS), captured by the LPD. The three exposure values are A/D converted and processed digitally to generate HDR pixel values with a 20-bits linear range. The sensor is able to output full resolution at 60fps, with both serial MIPI and parallel DVP output being supported.

Fig. 1 shows the pixel architecture. The floating diffusion (FD) nodes for the SPD (FDS) and LPD (FDL) are separated by the DFD transistor. Low CG (LCG) is obtained by asserting the DFD, which increases the equivalent capacitance at the gate of the pixel source follower transistor (SF). High CG (HCG) is obtained by de-asserting the DFD. The LPD can be read using both LCG and HCG, whereas the SPD can only be read in LCG mode. This is reflected in fig. 2, which illustrates the readout timing diagram for both the SPD and the LPD.

The sensitivity ratio between the LPD and the SPD is 6.5:1, which is equivalent to 16.3dB dynamic range (DR) extension. Furthermore, the exposure time is controlled independently for L, S, and VS in order to obtain the target overall exposure ratios and DR. Fig. 3 shows the operation of the pre-charge; *i.e.*, reset, and readout pointers in the pixel array. The split-diode pixel architecture enables overlapping exposure of L and S. This is beneficial with regard to artefacts arising from motion or from pulsed light sources in the

scene [1]. The VS exposure is captured on the LPD, and since S and VS have the same readout address, the VS exposure will partly overlap the S exposure.

Fig. 4 shows the chip architecture. In order to output HDR processed images at 60fps, the A/D conversion rate needs to be equivalent to 180fps. In order to conserve power dissipation, a column parallel successive approximation (SA) A/D converter (ADC) architecture was implemented. The correlated double sampling was implemented in the digital domain (DCDS) to suppress ADC offsets, and to minimize vertical FPN (VFPN). A column level arithmetic logic unit (ALU) was implemented to perform the DCDS operation. The ALU also supports multi-sampling and averaging (MSA) to further suppress the readout noise; in addition, the digital output values from the ALU are written to a column memory. The black level correction (BLC) was implemented in the digital domain with individual correction for the L, S, and VS exposure channels. In order to align L with the S and VS pixel values, the L data is written to a stack of line buffers. The lens correction (LENC) and the defect pixel correction (DPC) processing steps are both performed before the HDR combination. Before data is transmitted on MIPI or DVP, the pixel values can be tone mapped and compressed to 12-bits.

Fig. 5 shows the SA, ADC architecture and the timing diagram. The feedback to the comparator is applied from a 12-bit D/A converter, which was based on a scaled capacitor array. In order to reduce the die size, a split capacitor array was implemented to reduce the capacitor ratio. The comparator was based on a low noise differential amplifier followed by a regenerative latch. The quantization noise can be reduced by programming a lower ADC reference voltage (VREF). The overall sampled noise in the analog readout channel can be suppressed by MSA, and fig. 5 shows the timing diagram for MSA2, meaning two samples have been averaged. The sensor supports programmable timing control, and MSA of different orders can be used for all exposure channels. Typically MSA2 or MSA4 is only used for the readout of L, whereas regular CDS (MSA1) is used for S and VS. By using MSA4, a readout noise of 0.94e-rms is achieved for HCG.

A scene with a pulsed LED source was used to compare the split-pixel HDR technology with the more common staggered HDR technology, which uses only one photo diode per pixel [2, 3]. The exposure control and frame operations are illustrated in fig. 6. Another OmniVision sensor was used to capture staggered HDR images. The two sensors were

operated at 30fps, and the LED source was pulsed at 100Hz, with a pulse duration of 0.1ms. In the case of staggered HDR, the short LED pulse can result in similar pixel output values for the different exposure channels. This will lead to distortion after HDR combination, since gain equivalent to the exposure ratio is applied to generate the combined output image. The distortion of the pulsed LED source is clearly seen in fig. 7, which shows the image captured with staggered HDR sensor. The image captured with the split-pixel-pixel sensor, presented in this paper, has significantly less distortion as shown in fig. 8. There is still some distortion resulting from the VS capture as shown in fig. 8 (brighter horizontal line inside dotted line). The distortion is avoided in cases where the scene dynamic range matches the L and S channels and VS becomes redundant.

Fig. 9 shows the quantum efficiency (QE) characteristics for the LPD, with the peak QE of 55% at wavelength 530nm.

The SNR was characterized vs. scene illumination level. The measurement was done at F/8 aperture with a 530 nm wavelength illumination source, in line with the EMVA 1288 standard [4]. The L and S channels were measured with an integration time of 43ms, and the VS channel was measured with an integration time of 43μs. SNR=1 for the L channel was reached at 3.5photons/pixel, and from that a dynamic range of >120 dB was achieved before the VS became saturated, all the while maintaining SNR>1, as shown in fig. 10.

The sensor was fabricated in our oldest CMOS BSI technology at 110nm, 1P4M. Future HDR sensors will be further improved with our 65nm BSI2 technology.

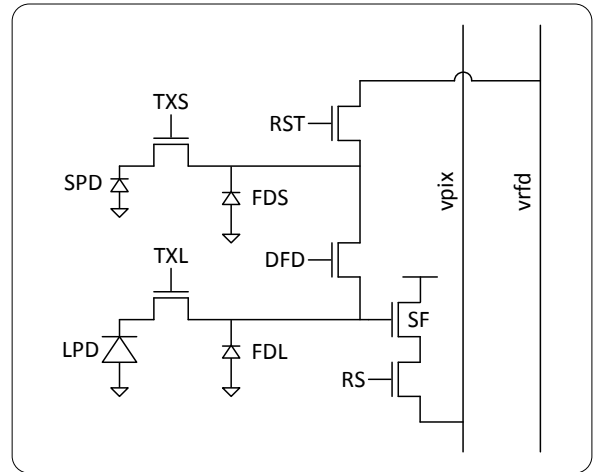


Figure 1

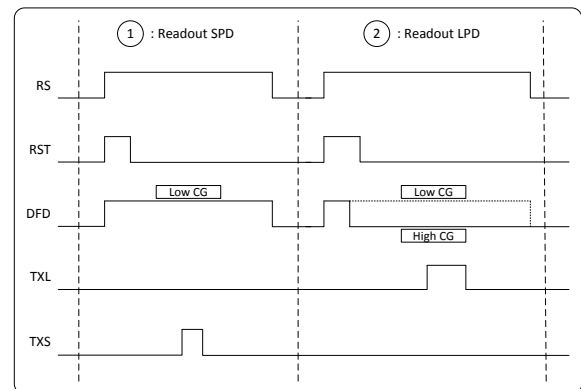


Figure 2

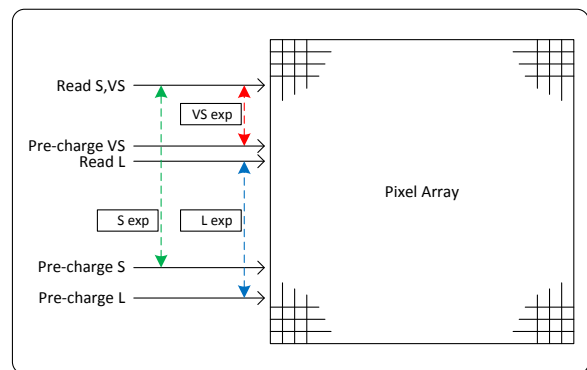


Figure 3

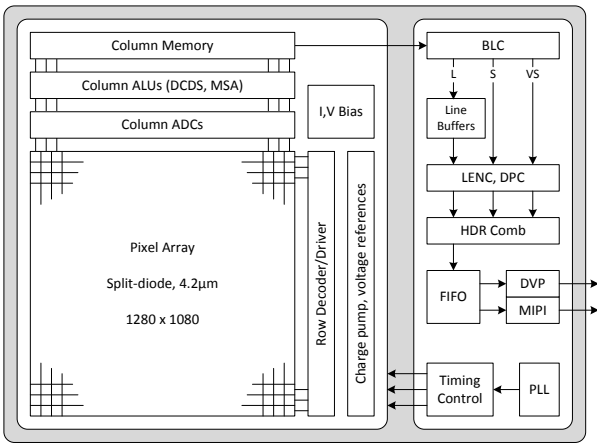


Figure 4

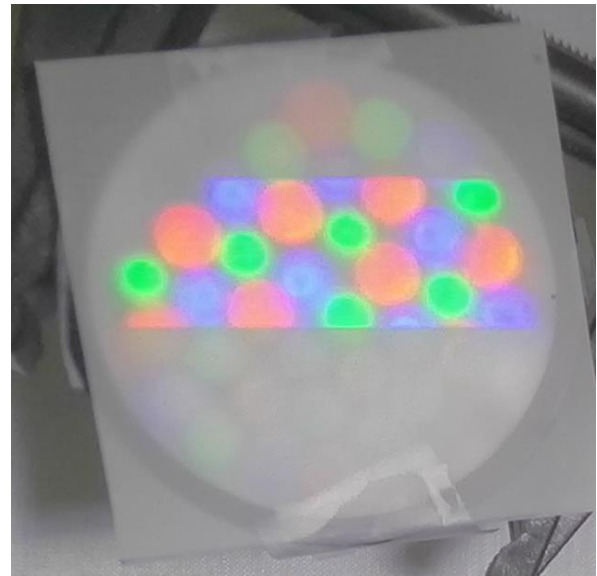


Figure 7

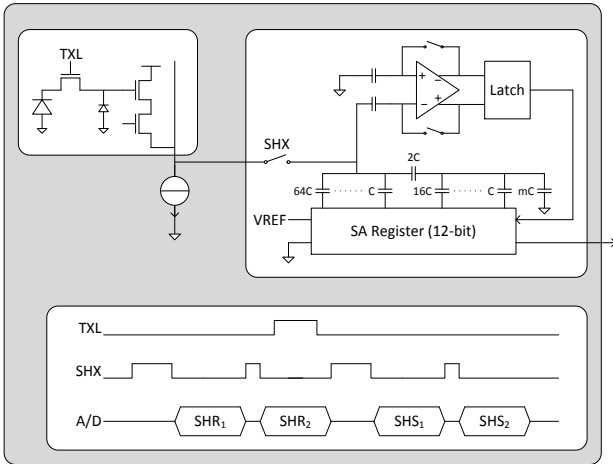


Figure 5

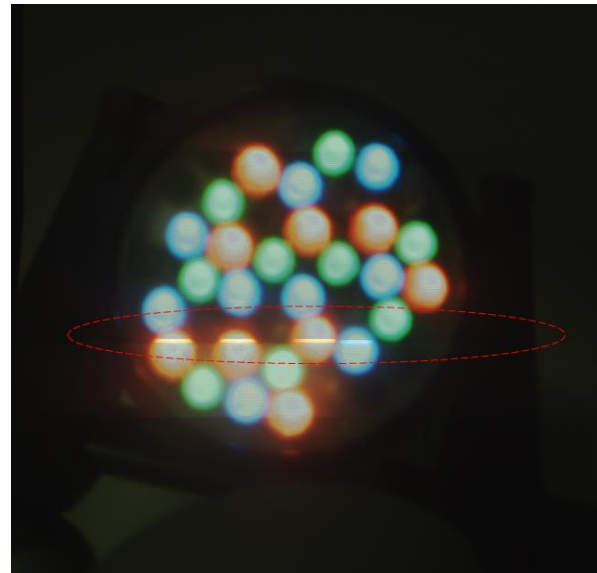


Figure 8

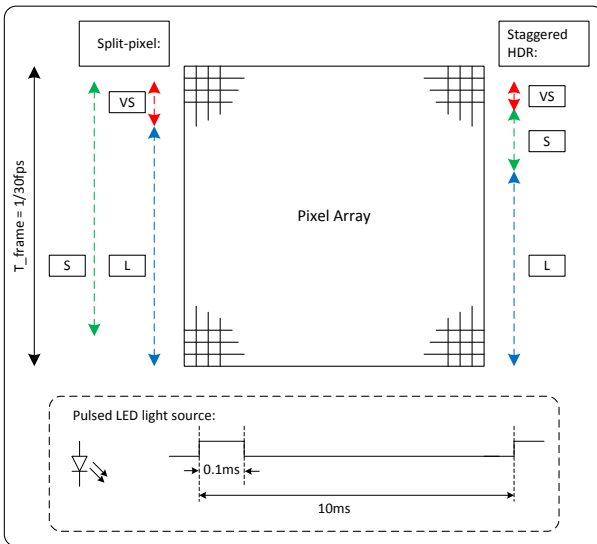


Figure 6

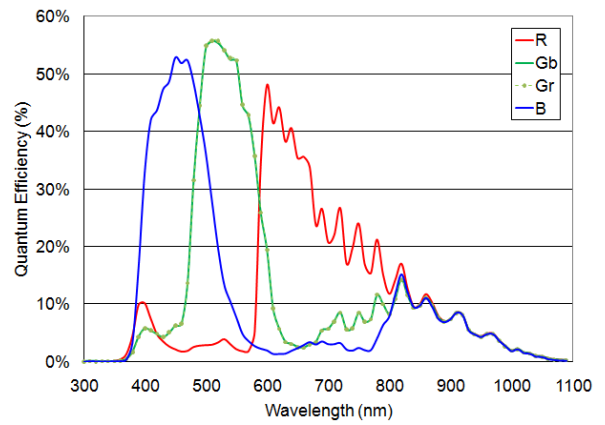


Figure 9

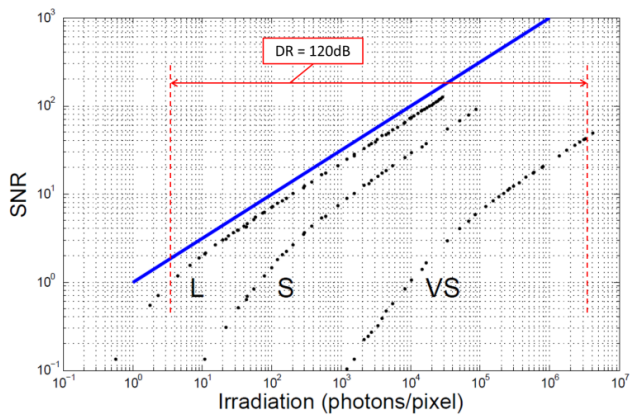


Figure 10

### References:

- [1] J. Solhusvik et al., "A comparison of high dynamic range CIS technologies for automotive applications", IISW 2013
- [2] Yadid-Pecht and E.R. Fossum, "Wide Intra-scene Dynamic Range CMOS Image Sensor with Fixed Pattern Noise Free, Double Exposure Time Read-Out Operation", ASSCC 2006
- [3] J. Solhusvik et al., "A 1280x960 3.75um pixel CMOS imager with Triple Exposure HDR", IISW 2009
- [4] European Machine Vision Association ([www.emva.org](http://www.emva.org)), "EMVA Standard 1288, Standard for Characterization of Image Sensors and Cameras", Release 3.0, Nov 29, 2010