A comparison of high dynamic range CIS technologies for automotive applications

Johannes Solhusvik, Jiangtao Kuang, Zhiqiang Lin, Sohei Manabe, Jeong-Ho Lyu, Howard Rhodes, OmniVision Technologies

Johannes.Solhusvik@ovt.com, Gaustadalleen 21, 0349 Oslo, Norway

1 Introduction

A comparison of four high dynamic range (HDR) CMOS image sensor (CIS) technologies is presented. Pixel size is 6um, technology is FSI, and target application is automotive where low-light performance, dynamic range, and fast motion is of particular concern.

2 HDR methods

The sensors considered in this work combine two captures with high and low sensitivity into one HDR output image.

Skimming HDR (fig.1) obtains its response knee-point by asserting after a time (T1) a mid-level voltage pulse on the transfer gate (TG) whilst photon charge is being integrated [1-2]. The falling edge of this pulse defines the start of the 2nd (short) exposure (T2). The ratio between the long and short integration times, i.e. (T1+T2)/T2, defines the DR extension. This assumes that the CDS time (T3) illustrated in fig.1 is negligible compared to T1 and T2. The mid-level voltage defines a potential barrier under TG in such a way that any PD charge above this barrier level is skimmed (drained) to the supply (AVDD) via the floating diffusion node (FD) and the reset transistor (RST). This effectively splits the PD full well capacity (FWC) in two parts. A 50-50 split is assumed in this comparison.

Staggered multi-capture HDR (fig.2) combines one long and one short integration time capture [3-6]. The rolling shutter readout is staggered (row interleaved) so that the short integration (Tshort in red) starts immediately after sampling of the long integration (Tlong in green). Tlong pixel values are stored in line buffers until Tshort values for the same pixel row become available to perform HDR combination in digital domain.

Down-sampling HDR (fig.3) trades off pixel resolution for increased DR by combining neighboring pixels with different integration times [7-8]. One of OmniVision's alternating line approaches is illustrated in the figure below, where T1/T2 represents long/short integration times, respectively. Note this particular sensor also uses RGBC CFA for improved light sensitivity. RGBC is equally applicable to the other HDR schemes and is therefore not included in the below light sensitivity comparison.

The forth technology is called split-diode HDR [9] and is developed for automotive applications (fig.4). Each pixel has one large photodiode (LPD) and one small photodiode (SPD). The diodes are exposed simultaneously, which makes the sensor inherently immune to ghost artifacts caused by scene motion (ref below). Sensitivity ratio and DR is further enhanced by use of dual CG readout (CGC). Ratio of Hi/Lo CG is 3x, and the same CG ratio is applicable to all HDR schemes to make the comparison valid.

3 Dynamic range

We define dynamic range (DR) as Smax/Smin, where Smax=FWC (full well capacity) and Smin=Nfloor (read noise floor). Two captures combined, having high/low sensitivity ratio R, yield a DR as follows (in bits): DR = $\log_2$(FWC x R / Nfloor).

Nfloor is set to 1.2e- rms for all HDR schemes (same CG, BC-SF and CDS circuitry). The max value of R is constrained by the minimum allowable S/N ratio at the knee-point. We set minimum SNR to 25dB to ensure high image quality across the entire signal range. Resulting DR is provided in the table below.

HDR method	Pixel size (um)	Nfloor (e- rms)	FWC (e-)	Ratio, R	DR (bits)
Skimming	6	1.2	15000	47.4	19.2
Staggered	6	1.2	30000	94.9	21.2
Down-sampling	6	1.2	30000	94.9	21.2
Split-diode	6	1.2	30000	94.9	19.6

Best DR performance (21.2bits) is achieved with staggered HDR and down-sampling HDR. Skimming HDR is limited to 19.2bits. This is because of the potential well which is split into two parts, thus reducing FWC by 50%. Split-diode HDR is limited to 19.6bits due to reduced FWC of SPD (10k vs 30k) as illustrated in fig.5.

4 Low-light performance

Low-light performance depends on the optical lens, pixel responsivity, noise floor, and integration time (Tframe). From these parameters we calculate scene illumination equivalent to SNR=1 (E_{SNR1}) assuming a lambertian white object:

$$E_{SNR1} = \frac{4F^2 \times N_{floor}}{T_{lens} \times S \times T_{frame}}$$

, where F is lens number, Tlens is lens transmission, and S is responsivity (e-/lux-s). Using Tlens=0.92 the calculated result is shown below.

HDR method	Pixel size (um)	Nfloor (e- rms)	Responsivity (e-/lux-s)	Lux@SNR1, 60fps, F1.4
Skimming	6	1.2	113650	0.0054
Staggered	6	1.2	113650	0.0054
Down-sampling	6	1.2	113650	0.0054
Split-diode	6	1.2	113650	0.0054

This shows that 5.4mlux minimum illumination is achieved for each HDR capture scheme. Even though the split-diode concept uses a more complex pixel light sensitivity is maintained thanks to large pixel size which enables close to 100% fill-factor of LPD+SPD with dual micro-lens and other process optimization steps.

5 Motion performance

Combining high and low sensitivity values into one HDR output value is trivial when there is zero motion involved. For instance, one straightforward approach is to add the two signals together and apply a gain (R+1) after the knee-point, thus forming a linear (straight line) output response from zero to max light level. However, in most HDR schemes the high and low sensitivity captures do not take place simultaneously in time. And since scene motion causes time variant pixel illumination the captured pixel values become non-linear and the combined HDR output suffers from image artifacts such as ghosting, color imbalance, color rings, loss of details in the ghost images, etc. The degree of such side-effects depends on the specific HDR capture method. Given that motion (distance travelled) is proportional to the time difference between the high and low sensitivity captures, we define a simple motion FOM as follows

$$FOM_{motion} = 1 - \frac{T\_COG_{hi} - T\_COG_{lo}}{T_{frame}} = 1 - \frac{\Delta T}{T_{frame}}$$

, T_COGhi and T_COGlo represent the mid-points (centers of gravity) of the high and low sensitivity captures, respectively. Tframe is the frame time (=1/frame-rate).

Note this FOM also measures vulnerability to light flicker or blinking LED illumination which creates similar artifacts for similar reasons (time varying pixel illumination).

For a given light sensitivity ratio R, and assuming Tlong and Tshort are always maximized (constrained by Tframe) for light sensitivity, the term ΔT can be calculated for each HDR method:

Skimming HDR: ΔT = Tframe/2-Tframe/R+Tframe/R/2 => FOM_motion=1-(1/2-1/R+1/2R)
Staggered HDR: ΔT = Tframe/2 => FOM_motion=1/2
Down-sampling HDR: ΔT = Tframe/2-Tframe/R+Tframe/R/2 => FOM_motion=1-(1/2-1/R+1/2R)
Split-diode HDR: ΔT = 0 => FOM_motion=1

FOM_motion is plotted as a function of R in fig.6. For skimming and down-sampling HDR the FOM_motion goes asymptotically towards 0.5 as R increases. This is because the exposure overlap becomes smaller and smaller relative to the total integration time (Tframe).

The capture process of an ideal HDR sensor with unlimited DR has been simulated (fig.7). A dark object moving laterally 64 length units from left to right in front of a white background was used. The observation was 640 length units in the horizontal direction. The background brightness is about 6500 light units (aka DNs) and the dark object brightness is about 200 light units. The object width is 180 length units. A regular 10bit CIS would clip the background at 1023 light units which would give a relatively narrow transition from background to foreground of the dark object. Thus, the edge would look relatively sharp simply because of poor DR as in non-HDR CIS cameras. For an ideal HDR camera, however, with unlimited dynamic range, information content is higher but the edge transitions are wider and more visible in a linear response plot (dotted line in fig.7a and fig.7c).

In skimming HDR, staggered HDR, and down-sampling HDR there is a time lapse between high and low sensitivity captures. When combined with scene motion, this can result in ghost effects as illustrated in fig.8a (simulated), and fig.8b-c. The high sensitivity capture is saturated and the low sensitivity capture has very low signal level. Thus, after HDR combination the output value is a grey tone with very poor S/N ratio.

In the case of split-diode (fig.9) it is possible to keep both high sensitivity (LPD) and low sensitivity (SPD) captures 100% overlapping during photon integration. This way motion performance is similar to an ideal HDR sensor without ghosting. This is illustrated in fig.9 in a 1-D simulation plot and in the image capture examples using split-diode HDR sensor with 8x and 24x exposure ratios.

6 Conclusions

Overall conclusions from the comparison are listed in the below table.

HDR method	Min pixel size	Pros	Cons
Skimming	Any size	Small die	Motion, blinking LEDs, split FWC, FPN/PRNU around knee-point (Vth variation, temp dependence)
Staggered	Any size	Best DR	Motion, blinking LEDs, die size (memory)
Down sampling	Any size	Best DR, small die	Motion, blinking LEDs, aliasing/Moiré, pixel resolution
Split diode pixel	≥ ~3um	Best motion+LED performance, small die	Complex pixel, difficult to scale below ~3um

Table-1: Comparison Summary

7 References

[1] ST patent application US20110221944
[2] Toshiba patent US7586523
[3] Yadid-Pecht et al, “Wide Intrascene Dynamic Range CMOS APS Using Dual Sampling”, Workshop on CCDs and AIS, Bruges, 1999
[4] Egawa et al., “A 1/2.5 inch 5.2Mpixel, 96dB Dynamic Range CMOS Image Sensor with Fixed Pattern Noise Free, Double Exposure Time Read-Out Operation”, ASSCC, 2006
[5] N. Ide, et al, “A Wide DR and Linear Response CMOS Image Sensor With Three Photocurrent Integrations in Photodiodes, Lateral Overflow Capacitors, and Column Capacitors”, IEEE JSSC, 2008
[6] Solhusvik et al, “A 1280x960 3.75um pixel CMOS imager with Triple Exposure HDR”, IISW, Bergen, 2009
[7] Rhodes, “BSI CMOS image sensors with RGBC technology”, IS2013
[8] Cho et al, "Alternating line high dynamic range imaging," 17th Int. Conf. on DSP, 2011
[9] OmniVision patent pending, US App 13-434,124

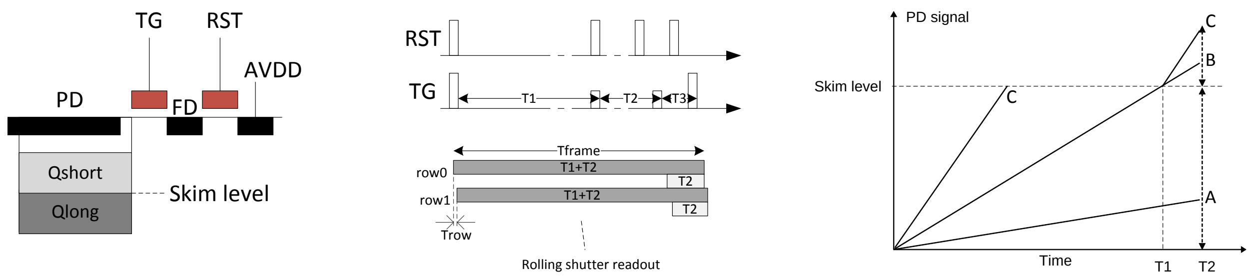

Fig. 1 Illustration of Skimming HDR (Lateral-Overflow) approach

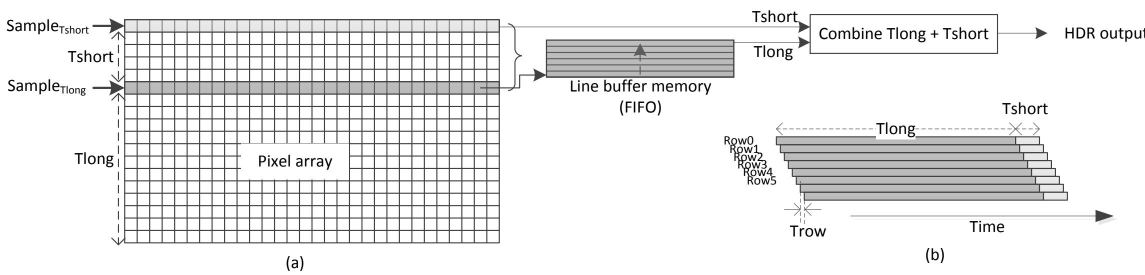

Fig. 2 Illustration of (a) staggered HDR concept and (b) timing diagram of each rows exposure periods

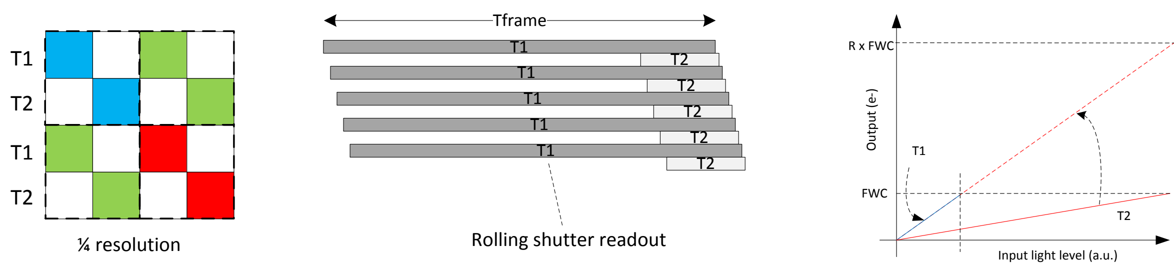

Fig. 3 Down-sampling HDR concept with RGBC CFA pattern for increased sensitivity

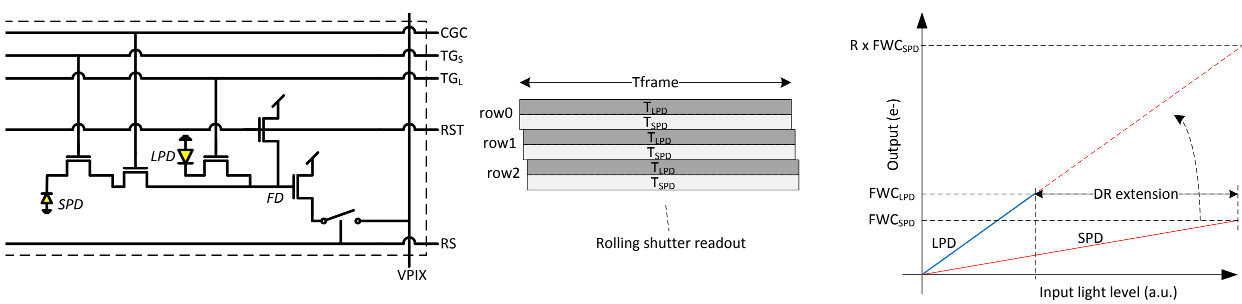

Fig. 4 Example of split-diode pixel architecture and timing

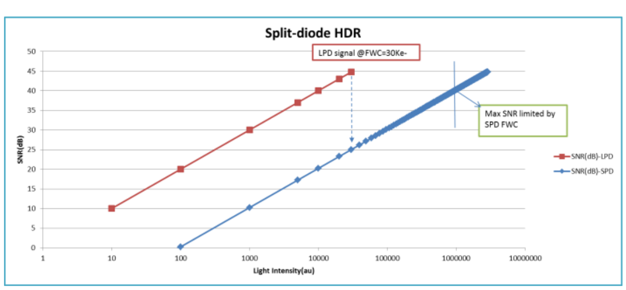

Fig. 5 Split-diode SNR vs Light intensity

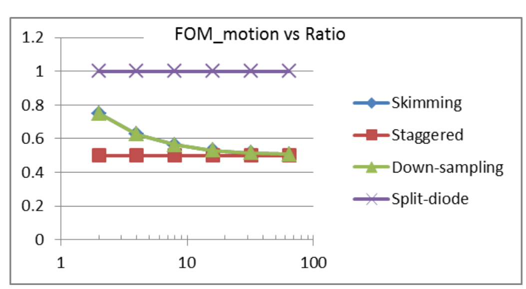

Fig. 6 FOM_motion vs Hi/Lo sensitivity ratio

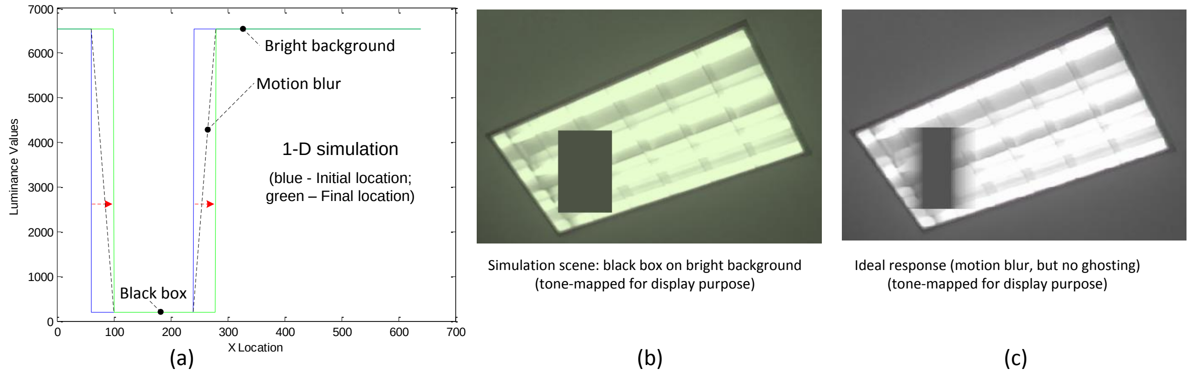

Fig. 7 Simulation of moving object during capture: (a) 1-D model, (b) 2-D static scene, (c) simulated ideal HDR response

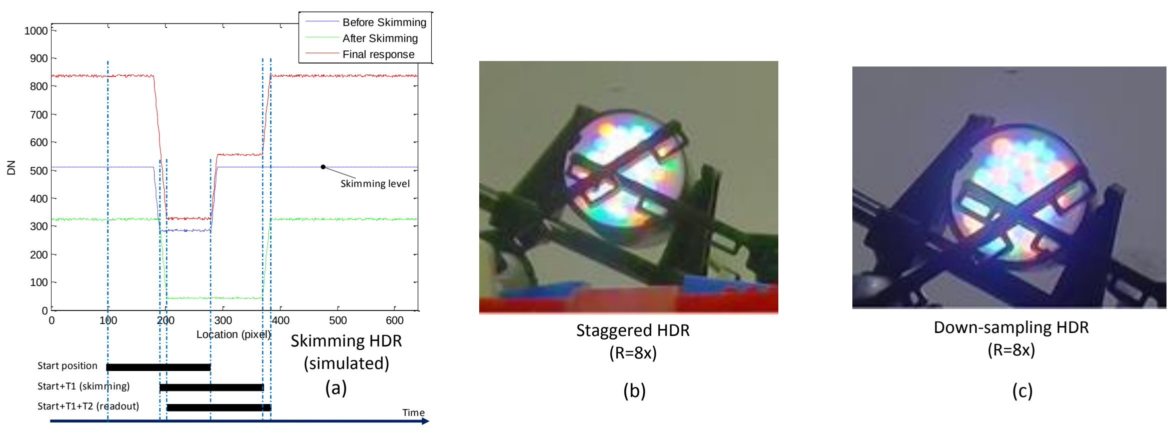

Fig. 8 Ghost effect in (a) skimming HDR (simulated), (b) staggered HDR (R=8x), and (c) down-sampling HDR (R=8x)

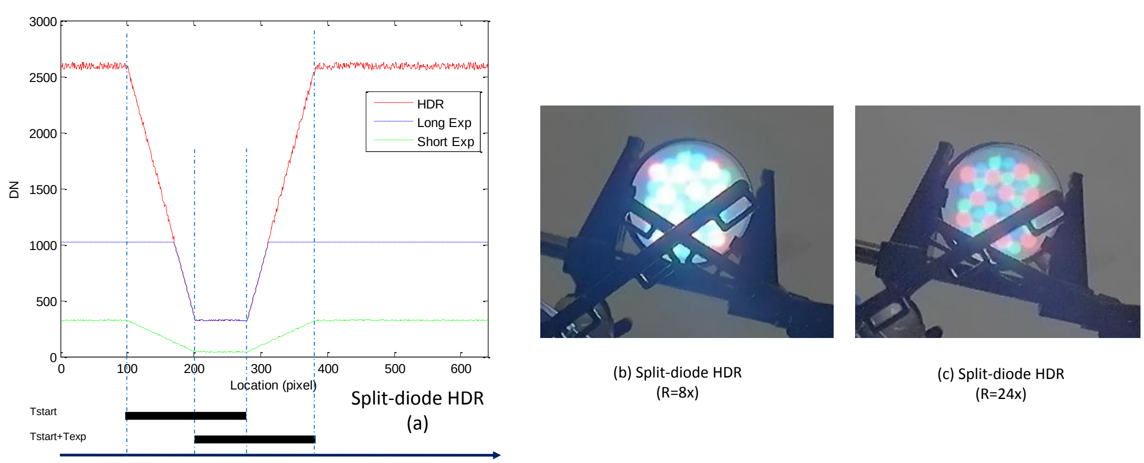

Fig. 9 Split-diode HDR simulation and captures with 8x and 24x exposure ratios