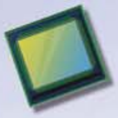


OV10640

application note

IMPLEMENTATION GUIDE

1/2.56 color CMOS 1.3 megapixel (1280 x 1080) high dynamic range (HDR)
high definition (HD) image sensor

Trademark Information

color CMOS 1.3 megapixel (1280x1080) high dynamic range (HDR) high definition image sensor

application note
IMPLEMENTATION GUIDE

version 1.1
april 2015

To learn more about OmniVision Technologies, visit www.ovt.com.
OmniVision Technologies is publicly traded on NASDAQ under the symbol OVTI.

table of contents

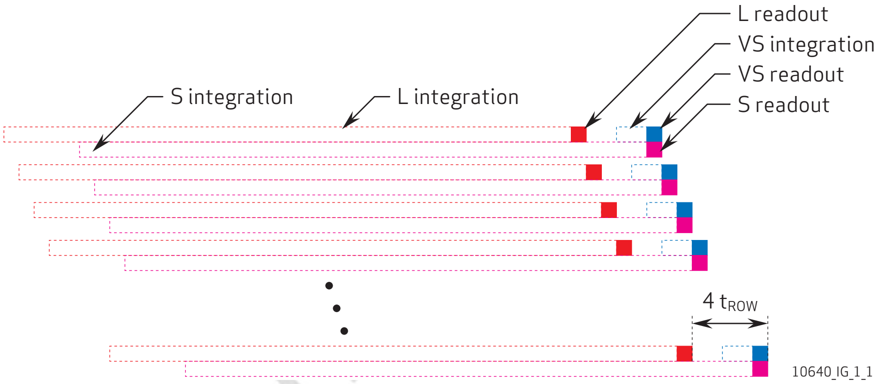
list of figures

list of tables

1 integration time

The OV10640 supports quasi-simultaneous captures of long (L) and short (S) integration time using two photodiodes per pixel, and an optional staggered capture of very short (VS) integration time, as shown in **figure 1-1**. The readout of L and S channels is always 4 row periods apart even when the VS capture is not used.

figure 1-1 L, S, and VS exposure and readout timing diagram



The integration time of L and S channels is in unit of row period, and the maximum is VTS -6 where VTS is the frame period in unit of row period. The precision of the integration time for VS capture is 1/32 row period. The maximum integration time of the VS channel is 3.95 row periods, the minimum integration time of the VS channel is 0.5 row periods. **table 1-1** lists the integration time control registers of each capture.

table 1-1 exposure time control registers

address	register name	default value	R/W	description
{0x30E6, 0x30E7}	EXPO_L	0x0040	RW	Long Channel (L) Exposure Time in Units of Row Periods
{0x30E8, 0x30E9}	EXPO_S	0x0040	RW	Short Channel (S) Exposure Time in Units of Row Periods
0x30EA	EXPO_VS	0x20	RW	Very Short Channel (VS) Exposure Time in Units of Row Periods

2 gain

The gain of L, S and VS can be applied independently. The OV10640 supports 1x, 2x, 4x and 8x analog gain, and fine digital gain. The conversion gain controlled by register 0x30EB[6] (for L) and 0x30EB[7] (for VS) is superior to the subsequent gain stages because it is applied at the very beginning of the signal path. It should be applied prior to other gain stages whenever possible. This gain stage is not available for S channel though. **table 2-1** summarizes the gain registers and the formula.

table 2-1 exposure time control registers

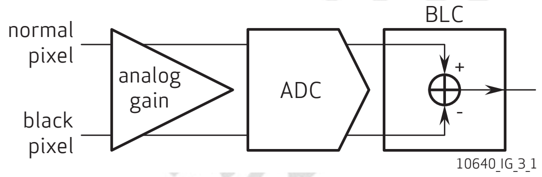
address	register name	default value	R/W	description
0x30EB	CG_AGAIN	0x00	RW	Bit[7]: VS channel conversion gain (CG) 0: Low CG 1: High CG Bit[6]: L channel conversion gain (CG) 0: Low CG 1: High CG Bit[5:4]: VS channel analog gain, equal to $2^{0x30EB[5:4]}$ Bit[3:2]: S channel analog gain, equal to $2^{0x30EB[3:2]}$ Bit[1:0]: L channel analog gain, equal to $2^{0x30EB[1:0]}$
{0x30EC, 0x30ED}	DGAIN_L	0x00100	RW	L Channel Digital Gain, equal to {0x30EC, 0x30ED} / 0x100
{0x30EE, 0x30EF}	DGAIN_S	0x00100	RW	S Channel Digital Gain, equal to {0x30EE, 0x30EF} / 0x100
{0x30F0, 0x30F1}	DGAIN_VS	0x00100	RW	VS Channel Digital Gain, equal to {0x30F0, 0x30F1} / 0x100

3 ADC range and minimum gain

The OV10640 implements a high precision low noise 12-bit ADC per column. The ADC range is maximized to leave room for dark current, and thus, is greater than pixel full well capacity (FWC), so a minimum gain must be applied to ensure that bright objects can always saturate and the pixel response is linear before saturation.

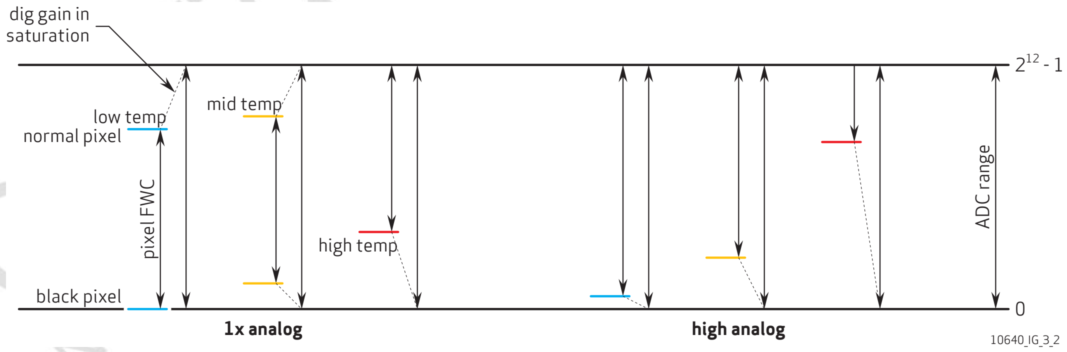
As shown in **figure 3-1**, the optical black pixel value is subtracted from the normal pixel in the black level calibration (BLC) block to keep black level consistent over temperature.

figure 3-1 ADC and BLC diagram



The black pixel's output value, y_Black, is not always 0 due to dark current. **figure 3-2** shows the black pixel value and the highest value of a normal pixel at different temperature and analog gain. A minimum (digital) gain must be applied after BLC to compensate the ADC range eaten by dark current.

figure 3-2 signal range, ADC range, and minimum gain



The minimum digital gain to stretch the remaining signal range to full ADC range is:

$$gain_{dig\_min} = \frac{4095 - C_{BLC}}{4095 - (y_{Black} + C_{ADC})}$$

where C_{BLC} is the black level pedestal, C_{ADC} is the ADC pedestal, y_{Black} is the black pixel output. y_{Black} is equal to dark current multiplied by analog gain.

$$y_{Black} = gain_{ana} \times y_{Dark}$$

where $gain_{ana}$ is analog gain and y_{Dark} is dark current. Dark current increases exponentially with temperature and linearly with exposure time. y_{Black} of L, S and VS channels can be read back from register {0x30D6, 0x30D7}, {0x30D8, 0x30D9} and {0x30DA, 0x30DB}, respectively.

$$y_{Black\_L} = register\ \{0x30D6, 0x30D7\}$$

$$y_{Black\_S} = register\ \{0x30D8, 0x30D9\}$$

$$y_{Black\_VS} = register\ \{0x30DA, 0x30DB\}$$

These registers are unsigned values. However, when y_{Black} is small, the calculation of y_{Black} may overflow and result in a big number (e.g., 0xFFE). In this case, the real y_{Black} can be calculated by subtracting 4096 from the register value. For example, 0xFFE means y_{Black} of -2. The dark current can be read back to distinguish real high y_{Black} from the overflowed value. When dark current is small, a big register value means y_{Black} is negative. The dark current of L, S and VS channels can be read back from registers {0x30D0, 0x30D1}, {0x30D2, 0x30D3}, and {0x30D4, 0x30D5}, respectively.

The ADC pedestal value, C_{ADC}, is also a function of analog gain. Please use following pedestal values to calculate the minimum digital gain (may change in the final setting). These values include the potential increase of y_{Black} before the next gain and exposure adjustment due to temperature change.

$$C_{ADC} = \begin{cases} 126, & when\ gain_{ana} = 1 \\ 312, & when\ gain_{ana} = 2 \\ 560, & when\ gain_{ana} = 4 \\ 1056, & when\ gain_{ana} = 8 \end{cases}$$

With minimum digital gain calculated following the above procedure, the max output of a normal pixel may be still less than 4095, so extra gain, preferably analog gain, is required. This gain can be calculated by the following formula:

$$gain_{extra\_min} = \frac{4095 - C_{BLC}}{y_{White} \times gain_{dig\_min}}$$

where y_{White} is the maximum value of normal pixel at 1x analog gain (i.e., the pixel FWC). For the OV10640, y_{White} is 3500 for L and VS, and 1600 for S. Please keep in mind that fractional analog gain is not supported and analog gain must be 1x, 2x, 4x or 8x.

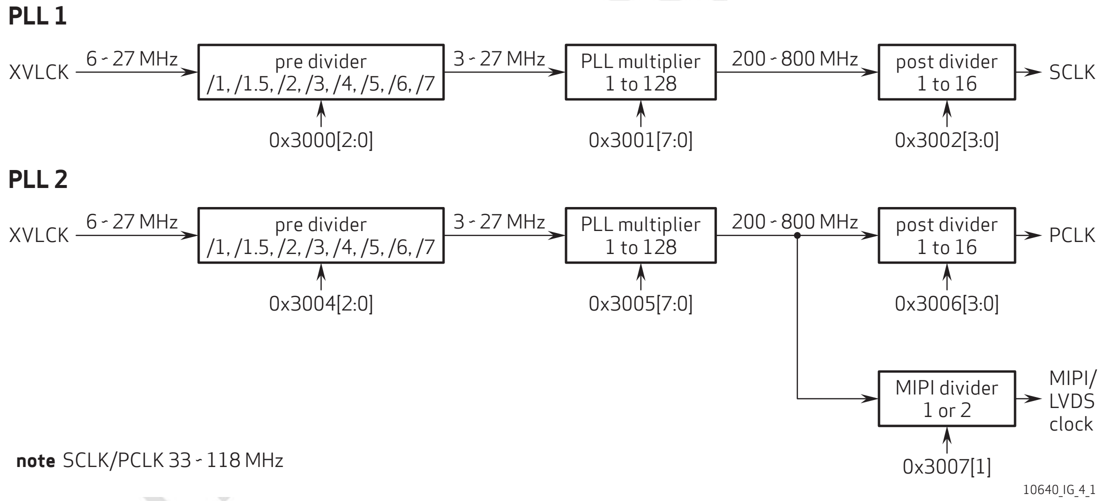
In extremely strong illumination (e.g., the Sun or reflection from car body), the pixel reset level will drop quickly before the correlated double sampling (CDS) can result in a low readout value. This is known as black sun issue. The OV10640 has a built-in anti-Black Sun function. A minimum 1.35x digital gain is required for this circuit to operate properly. If the minimal digital gain calculated earlier is lower than 1.35x, the minimum gain must be set to 1.35x. Very short channel has the same minimum gain as long channel and short channel gain is 3x.

4 PLL and frame timing

The OV10640 implements two phase lock loops (PLLs) to generate clocks. PLL1 generates a system clock for all internal blocks. PLL2 is used mainly to generate a clock for the output interface. EMC issue is usually related to the interface. Decoupling the interface clock PLL from the system clock PLL allows adjustment of the interface clock speed while keeping the frame rate unchanged.

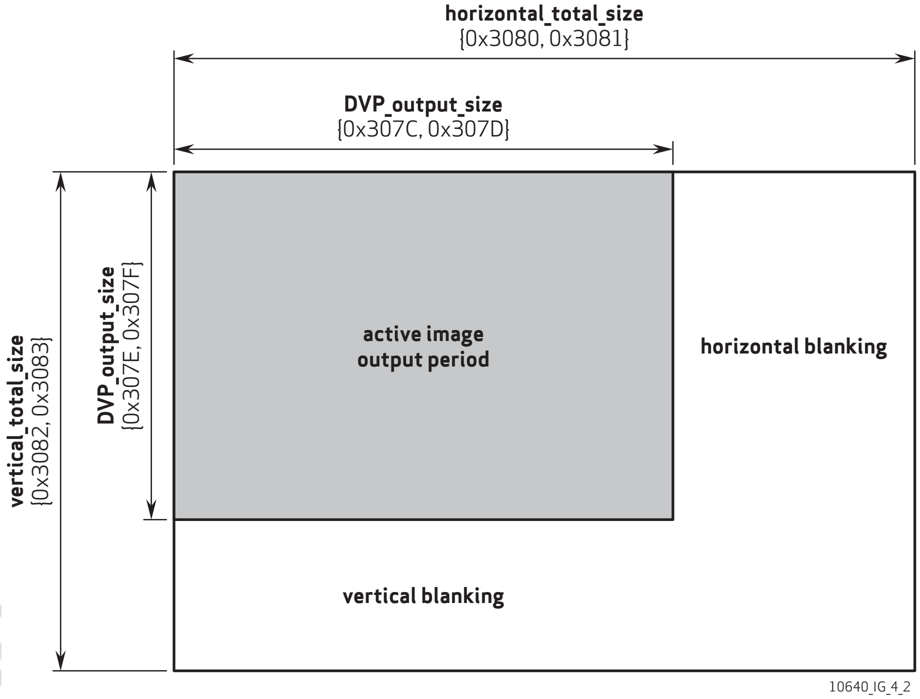
The frequency range of the PLL reference clock is 3~27MHz. The input clock frequency range is 6~27MHz. Though it is possible to slow down XVCLK to 3MHz with a pre-divider of 1, the maximum SCCB speed decreases with XVCLK, so it is

figure 4-1 PLL block diagram



Frame timing is generated from the system clock. The row length (or horizontal total size) is set by registers {0x3080, 0x3081} in unit of system clock period. The frame length (or vertical total size) is set by registers {0x3082, 0x3083} in unit of row period. The image width and height are set by registers {0x307C, 0x307D} and {0x307E, 0x307F}, respectively. **figure 4-2** shows a diagram for frame timing.

figure 4-2 frame timing



The frame rate can be calculated from the system clock frequency, row length and frame length.

$$n_{frame\_rate\_fps} = \frac{f_{sys\_clock\_Hz}}{n_{row\_length\_CLK} \times n_{frame\_length\_row}}$$

The maximum frame rate is 60 fps at full resolution.

The minimum horizontal blanking period is limited by the read out circuit speed and is dependent on the system clock period. During vertical blanking period, the sensor needs to finish black level calibration (BLC) etc., so a minimum number of blanking rows is also required.

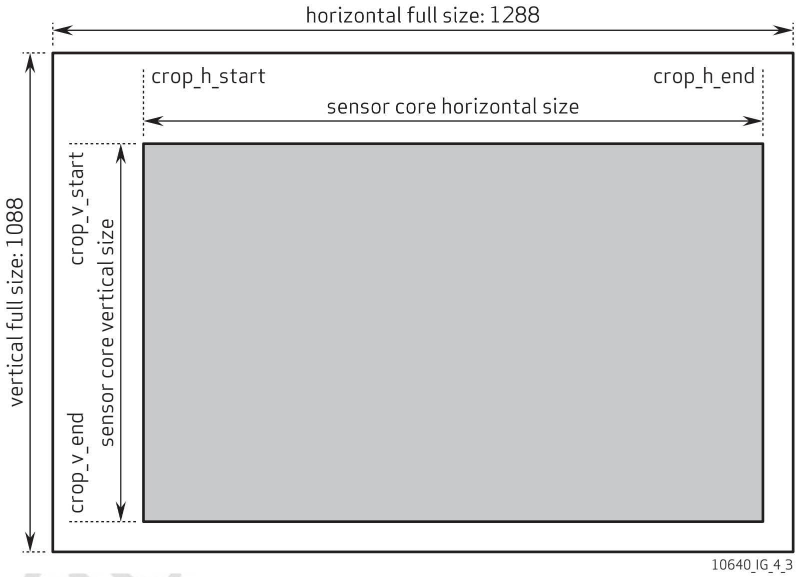
To minimize rolling shutter effect, it is recommended to keep the row period at minimum value. The reference setting provided usually sets the horizontal blanking period to minimum already.

The OV10640 has a pixel array of 1288 columns by 1088 rows. The output window size is programmable from 256 to 1288 columns in steps of 4 (8 in sub-sampling mode), and 20 to 1088 rows in steps of 2 (4 in sub-sampling mode). The window must start at an even row and column and end at odd row and column. The starting column is set by crop_h_start (registers {0x3074, 0x3075}), and the starting row is set by crop_v_start (registers {0x3076, 0x3077}). The number of columns read out is set by crop_h_end - crop_h_start+1 and the number of rows read out is set by crop_v_end -

crop_v_start +1. The readout window size is usually slightly bigger than the display image size. The extra border pixels are used by the Image Signal Processor (ISP) and trimmed out before it is output to display.

Refer to the *OV10640 Preliminary Specification* for additional readout control (e.g., mirror and flip, etc.).

figure 4-3 readout area and output image size



5 image format and output

5.1 image format

The OV10640 can output Bayer raw data of one exposure channel only through the MIPI interface or pixel interleaved fashion through the DVP interface, or the raw data combining three exposure channels, or the raw data combining two exposure channels. The OV10640 does not have line buffers to support line interleaved output (with different virtual channels. It is always pixel interleaved (MIPI, LVDS, and DVP). The single channel output is not useful for real application and is only used for some debug purposes.

5.1.1 3x12b linear raw

In this mode, the OV10640 outputs 12-bit L, S and VS channel Bayer raw data. The backend processor can collect the statistics of each exposure channel and use it to control exposure, gain, white balance, lens shading correction, defect pixel correction, HDR combination, tone mapping, de-mosaic, color correction, gamma correction, sharpness and de-noise, etc. Optionally, the sensor can perform defect pixel correction, lens shading correction, and apply white balance (WB) gain.

table 5-1 shows the settings to set up this format.

table 5-1 3x12b key settings

register	DVP	MIPI
0x3119	0x40	0x50
0x3127[0]	1’b0	1’b0

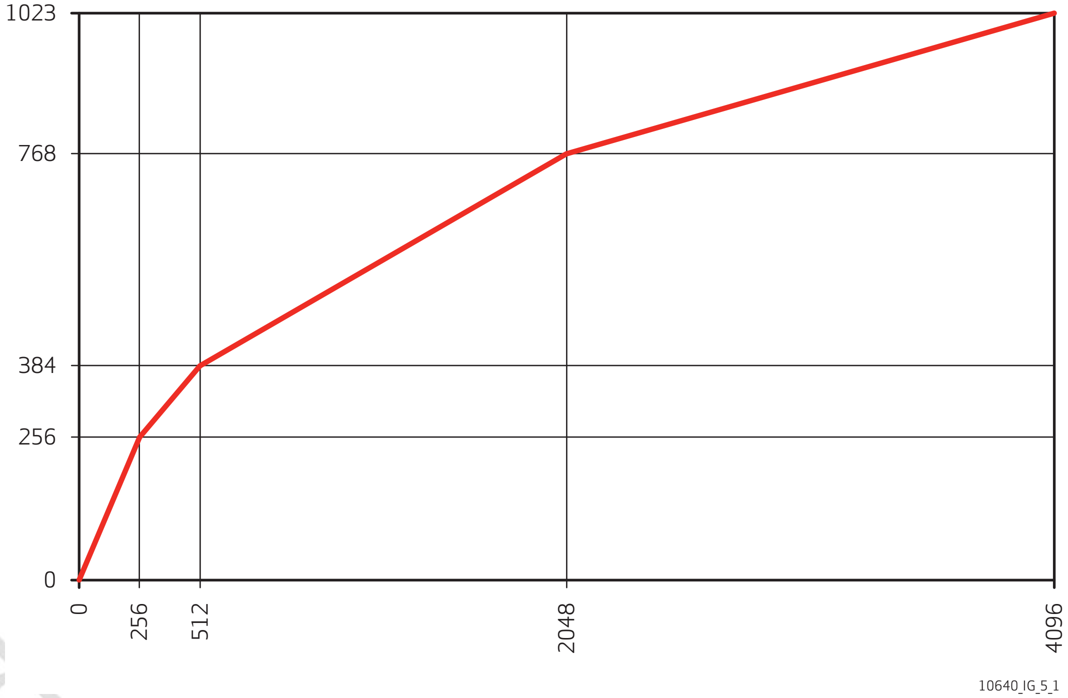
5.1.2 2x11b PWL raw

The ratio from L and VS is usually large. For any pixel, either the L channel data is saturated, or the VS channel data is nearly black, so there is no need to send out both L and VS channel data. 2x11b format is defined with this in mind to save output data bandwidth. S is between L and VS, so S is always output together with either L or VS.

In this mode, the L, S and VS channel firstly compressed to 10 bit from 12 bit by a 4-piece piece-wise linear (PWL) curve defined by following formula and illustrated in **figure 5-1**.

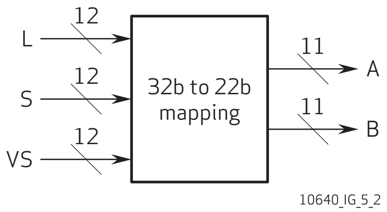
$$y_{out\_10b} = \begin{cases} y_{in\_12b}, & y_{in\_12b} < 256 \\ \dfrac{y_{in\_12b}}{2} + 128, & 256 \le y_{in\_12b} < 512 \\ \dfrac{y_{in\_12b}}{4} + 256, & 512 \le y_{in\_12b} < 2048 \\ \dfrac{y_{in\_12b}}{8} + 512, & y_{in\_12b} \ge 2048 \end{cases}$$

figure 5-1 RAW12 to RAW10 compression curve



To identify whether the data is L or VS, an extra bit is output with the 10-bit data. The mapping of 3x12b to 2x11b is illustrated in **figure 5-2**.

figure 5-2 36b to 22b mapping



$$A[9:0] = \begin{cases} PWL_{12\_10}(L[11:0]), & |A[10] = 0 \\ PWL_{12\_10}(VS[11:0]), & |A[10] = 1 \end{cases}$$

$$B[9:0] = PWL_{12\_10}(S[11:0])$$

$$B[10]: \text{reserved}$$

The backend processor can recover L and VS using following formula:

$$y_{out\_12b} = \begin{cases} y_{in\_10b}, & y_{in\_10b} < 256 \\ 2 \times (y_{in\_10b} - 128), & 256 \le y_{in\_10b} < 384 \\ 4 \times (y_{in\_10b} - 256), & 384 \le y_{in\_10b} < 768 \\ 8 \times (y_{in\_10b} - 512), & y_{in\_10b} \ge 768 \end{cases}$$

table 5-2 shows the settings to set up this format. When register bit 0x3119[3] is set to 0, the sensor outputs L channel data (PWL12_10(L[11:0])) instead of L/VS (A[10:0]).

table 5-2 2x11b key settings

register	DVP	MIPI
0x3119	0x49	0x59
0x3127[0]	1’b0	1’b0

5.1.3 16b log domain combined raw

The OV10640 can combine L, S and VS channels together and output combined raw. The combination takes place either between L and S, or between S and VS. The weighting is based on a 2D lookup table defined by registers 0x3136~0x3145. The threshold values at which the weight is defined are programmable by registers 0x3133~0x3155. There are two options:

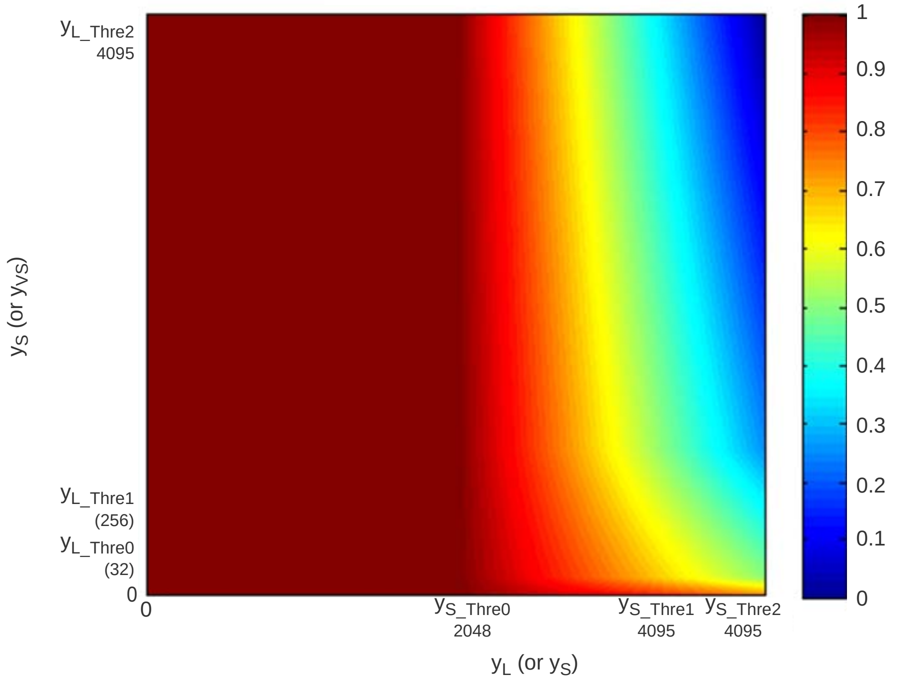
- when register bit 0x31BE[0] is 1’b1, the combination occurs on the original 12 bits of data and max threshold should be 2^12
- when register bit 0x31BE[0] is 1’b0, the combination occurs on the original 10 bits of data (2 LSBs are lost) and max threshold should be 2^10

The following is the formula for the threshold and weight:

$$
\begin{aligned}
y_{L\_Thre0} &= 2^{reg\ 0x3133[7:4]+2} \\
y_{L\_Thre1} &= 2^{reg\ 0x3134[7:4]+2} \\
y_{L\_Thre2} &= 2^{reg\ 0x3135[7:4]+2} \\
y_{L\_Thre0} &= 2^{reg\ 0x3133[3:0]+2} \\
y_{L\_Thre1} &= 2^{reg\ 0x3134[3:0]+2} \\
y_{L\_Thre2} &= 2^{reg\ 0x3135[3:0]+2} \\
w_L\,(0,\ 0) &= reg\ 0x3136/0x80 \\
w_L\,(y_{L\_Thre0},\ 0) &= reg\ 0x3137/0x80 \\
w_L\,(y_{L\_Thre1},\ 0) &= reg\ 0x3138/0x80 \\
w_L\,(y_{L\_Thre2},\ 0) &= reg\ 0x3139/0x80 \\
w_L\,(0,\ y_{S\_Thre0}) &= reg\ 0x3139/0x80 \\
w_L\,(y_{L\_Thre0},\ y_{S\_Thre0}) &= reg\ 0x313B/0x80 \\
w_L\,(y_{L\_Thre1},\ y_{S\_Thre0}) &= reg\ 0x313C/0x80 \\
w_L\,(y_{L\_Thre2},\ y_{S\_Thre0}) &= reg\ 0x313D/0x80 \\
w_L\,(0,\ y_{S\_Thre1}) &= reg\ 0x313E/0x80 \\
w_L\,(y_{L\_Thre0},\ y_{S\_Thre1}) &= reg\ 0x313F/0x80 \\
w_L\,(y_{L\_Thre1},\ y_{S\_Thre1}) &= reg\ 0x3140/0x80 \\
w_L\,(y_{L\_Thre2},\ y_{S\_Thre1}) &= reg\ 0x3141/0x80 \\
w_L\,(0,\ y_{S\_Thre2}) &= reg\ 0x3142/0x80 \\
w_L\,(y_{L\_Thre0},\ y_{S\_Thre2}) &= reg\ 0x3143/0x80 \\
w_L\,(y_{L\_Thre1},\ y_{S\_Thre2}) &= reg\ 0x3144/0x80 \\
w_L\,(y_{L\_Thre2},\ y_{S\_Thre2}) &= reg\ 0x3145/0x80 \\
w_S\,(0,\ 0) &= 1 - w_L
\end{aligned}
$$

figure 5-3 illustrates the default combination weight. The same lookup table is used for L/S combination and S/VS combination. The horizontal and vertical axis of the drawing is the maximum color component of the long and short exposure channels of each pixel.

figure 5-3 pseudo color plot of combination weight



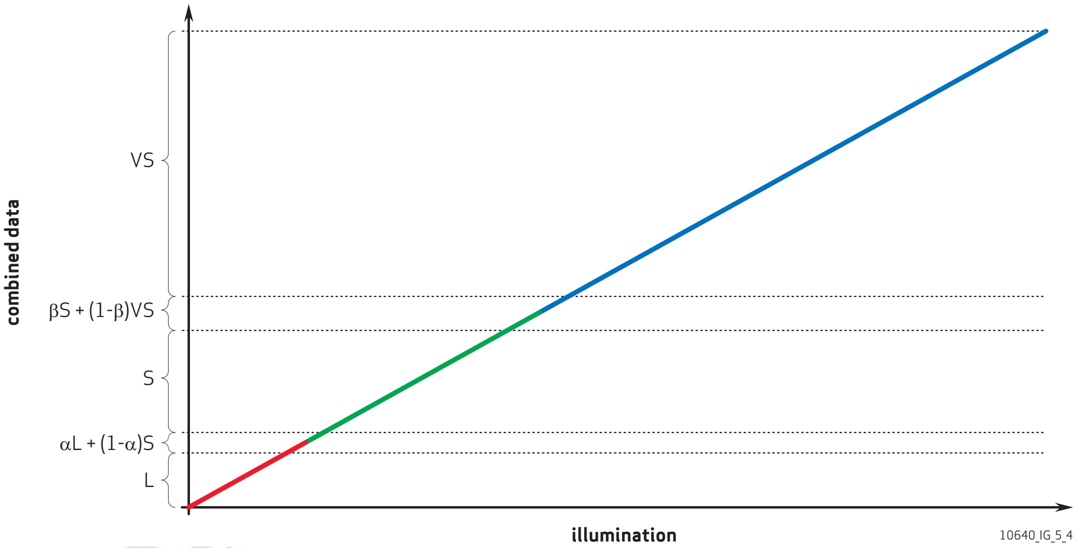
The image combination is done in log domain and the sensor can output the 16-bit combined raw data. The formula to convert the data back to linear domain is:

$$y_{Linear\_20b} = 2^{\frac{y_{Log\_16b}}{2048}} + 1$$

This 16-bit data is not normalized, meaning the maximum value is dependent on the exposure ratio between L and VS. The exposure ratio has to be limited in order to limit the bit depth of linear data converted from this combined data.

The statistics of L, S and VS channels, which is useful for AEC/AGC/AWB, cannot be simply taken from the log raw data. After converting it back to linear data, it is possible to collect the statistics of L, S and VS channels between the combination knee points since the backend processor knows the exposure ratio (and thus, the knee points), as shown in **figure 5-4**. Because the combination weight is based on the maximum color component of each pixel, the data range of each exposure channel shown in this diagram is not true for every pixel, so the statistics are not very accurate, but should be reasonably good for AEC/AGC/AWB control after excluding the values around the knee points.

figure 5-4 linear combined data



The combined 16-bit log raw contains all the information and requires a lower bandwidth than 3x12b and 2x11b. However, it requires a pixel-wise calculation to convert it back to linear data.

Any processing applied on an individual exposure channel (e.g., the compensation on S channel to match the spectral response with L and VS) can be done only before the combination in sensor only. For example, the sensor can apply a pre-WB gain and a pre-color correction matrix (CCM) on S channel before the combination to compensate the spectral difference between S and L/VS channels. Applying a pre-WB gain on S channel is done because its blue channel response is relatively stronger than its red channel response as compared to L and VS channels. The pre-CCM is done to further match the S channel color to L and VS channels. Because pre-CCM must be applied on white balanced data, it is not possible to apply it when the white balance gain is applied in the backend processor. The pre-CCM is usually close to the unit matrix. It is recommended for the customer to evaluate whether the color without this correction is acceptable.

table 5-3 16b log raw key settings

register	DVP	MIPI
0x3119	0x43	0x53
0x3127[0]	1’b1	1’b1

5.1.4 20b linear combined raw

The sensor can also convert the 16-bit log raw to 20-bit linear raw. This requires a higher bandwidth, but eliminates the need for conversion in the backend processor. The formula of the conversion is as shown below:

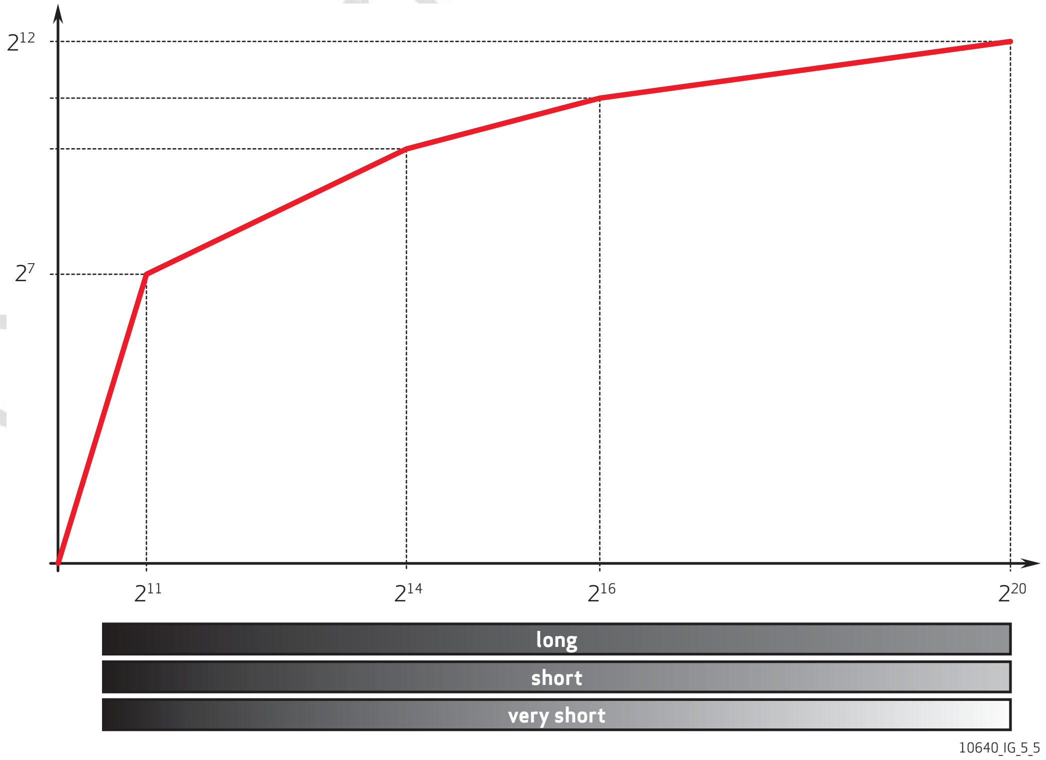
$$y_{Linear\_20b} = 2^{\frac{y_{Log\_16b}}{2048}} + 1$$

All other limitations of the 16b log raw apply to 20b linear raw as well. When converting data from 16 bits log to 20 bits linear, please pay attention to the normalization feature. After combination, there is an option which can normalize the data or not. If enabled, normalization (register bit 0x328A[1] = 1’b0), 20b linear raw data is MSB aligned; else, it is LSB aligned. If normalization is off, there is data overflow when the total ratio is larger than 1024x (register bit 0x31BE[0] is 1’b0) or 256x (register bit 0x31BE[0] is 1’b1).

5.1.4.1 normalization ON

Suppose L/S/VS ratio is 1, L=S=VS. So, valid combination data only occupies the lower 12 bits in 20b linear raw range. After PWL, L, S, and VS saturation levels do not change and are always 4095.

figure 5-5 L/S/VS when normalization is ON



There is a difference between how the OV10640 (rev 1D) sensor and OV10640 (rev 1E) sensor handles normalization. For the OV10640 (rev 1D) sensor, the user should manually set the max log value (registers {0x31BF, 0x31C0}, such as a value of 20480 + (L/VS-1) * 2048. If the user does not want to set the max log value manually, the user can set it as 0x9FFF and the sensor will always output 20 bits of data which will be LSB aligned. For the OV10640 (rev 1E) sensor, the sensor can automatically update the max log value if the auto calculation feature is set (register bit 0x328A[4] = 1’b1 and register bit 0x328A[0] = 1’b0).

5.1.4.2 normalization OFF

Suppose L/S/VS ratio is 1, L=S=VS. So, valid combined data only occupies the lower 12 bits in 20b linear raw range. After PWL, L, S, and VS saturation levels change from 4095 to 640.

figure 5-6 L/S/VS when normalization is OFF

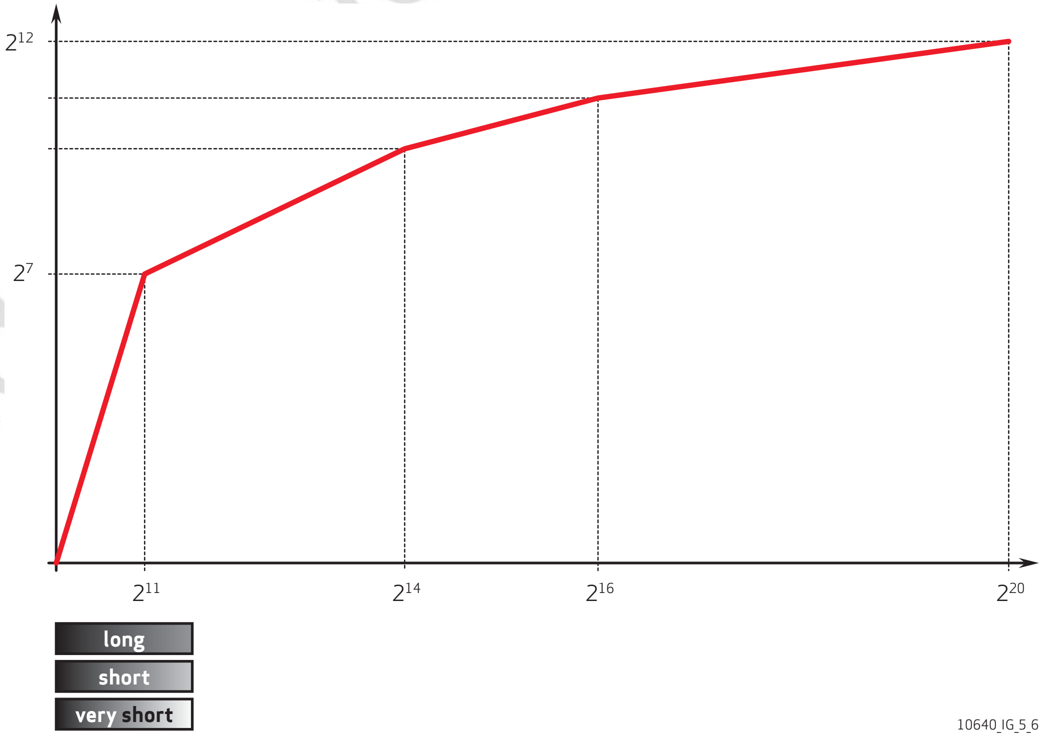


table 5-4 20b linear combined raw key settings

register	DVP	MIPI
0x3119	0x42	0x52
0x3127[0]	1’b1	1’b1

5.1.5 12b PWL combined raw

To save bandwidth, the 20-bit linear raw can be compressed to 12-bit by a 4-piece PWL curve as shown below:

$$y_{out\_12b} = \begin{cases} \dfrac{y_{in\_20b}}{4}, & y_{in\_12b} < 2048 \\ \dfrac{y_{in\_20b}}{16} + 384, & 2048 \le y_{in\_12b} < 16384 \\ \dfrac{y_{in\_20b}}{64} + 1152, & 16384 \le y_{in\_12b} < 65536 \\ \dfrac{y_{in\_20b}}{512} + 2048, & y_{in\_12b} \ge 65536 \end{cases}$$

The 12-bit raw can be converted back to 20-bit raw using following formula in the backend processor.

$$y_{out\_12b} = \begin{cases} 4 \times y_{in\_12b}, & y_{in\_12b} < 512 \\ 16 \times (y_{in\_12b} - 384), & 512 \le y_{in\_12b} < 1408 \\ 64 \times (y_{in\_12b} - 1152), & 1408 \le y_{in\_12b} < 2176 \\ 512 \times (y_{in\_12b} - 2048), & y_{in\_12b} \ge 2176 \end{cases}$$

table 5-5 shows the settings to set up this format.

table 5-5 12b PWL combined raw key settings

register	DVP	MIPI
0x3119	0x44	0x54
0x3127[0]	1’b1	1’b1

5.2 interface and maximum frame rates

The OV10640 can output all image formats via the digital video port (DVP) and the mobile industry processor interface (MIPI).

5.2.1 DVP interface

The OV10640 DVP bus is 12-bit wide and the pixel clock can run up to 118 MHz. This bandwidth limits the maximum frame rate of each format. **table 5-6** lists the maximum frame rate that can be achieved via DVP interface.

table 5-6 maximum frame rates supported via MIPI interface

format	resolution	PCLK	max frame rate
3x12b linear raw	1288 x 1088	100 Mbps[a]	22 fps
3x12b linear raw	1288 x 968	118 Mbps[b]	30 fps
2x11b PWL raw	1288 x 1088	100 Mbps	30 fps
16b log combined raw	1288 x 1088	100 Mbps	30 fps
20b linear combined raw	1288 x 1088	100 Mbps	30 fps
12 PWL combined raw	1288 x 1088	100 Mbps	30 fps

a. 5 pF maximum loading

b. 20 pF maximum loading

In DVP mode, the video stream outputs in pixel-interleaving fashion. **figure 5-7** ~ **figure 5-13** show the output timing of 3x12b, 2x11b, 16b, 20b and 12b, respectively.

For 16b data format, timing option 1 enables transmitting 16-bit data through the eight most significant data lines. With timing option 2, the 12 most significant bits are transmitted in the first clock cycle and the rest of the bits are transferred in the second clock cycle, which requires specific data packing in the backend processor. Similarly, two options are provided for 20b data format. These options are selected by register bit 0x3123[3], where a value of 0 means option 1 and a value of 1 means option 2.

figure 5-7 3x12b linear raw output timing

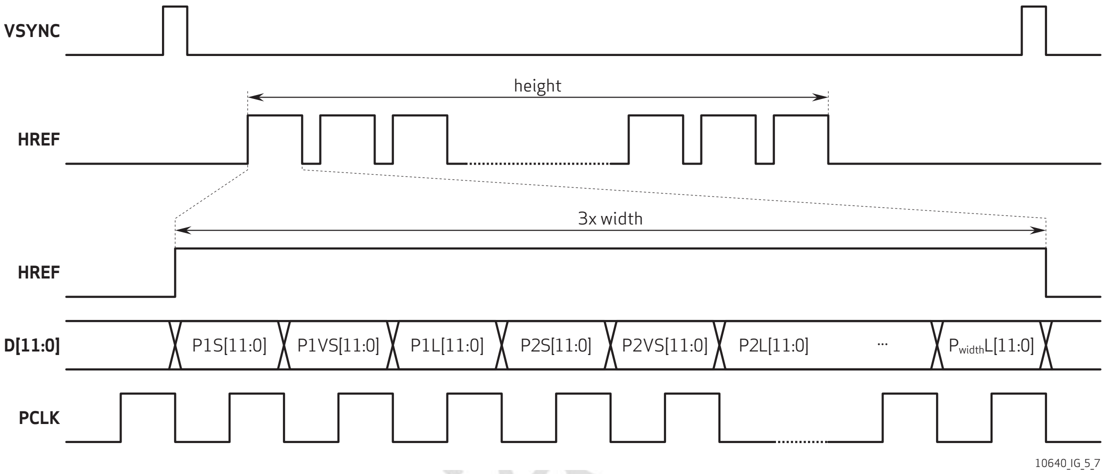


figure 5-8 2x11b linear raw output timing

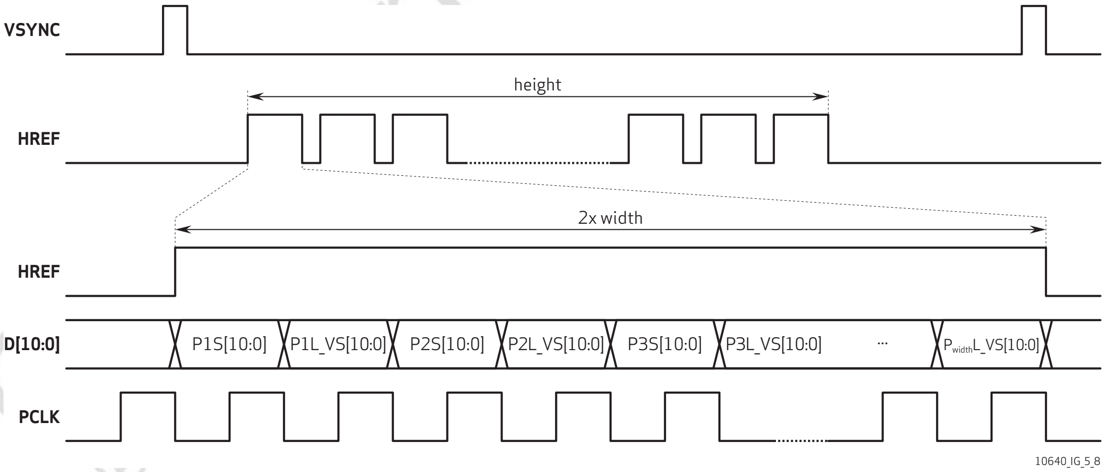


figure 5-9 16b log raw output timing option 1

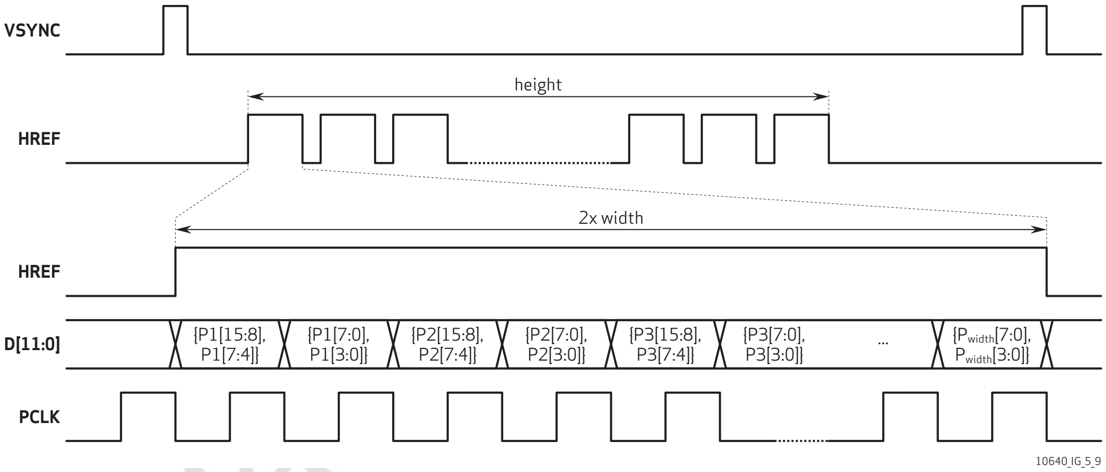


figure 5-10 16b log raw output timing option 2

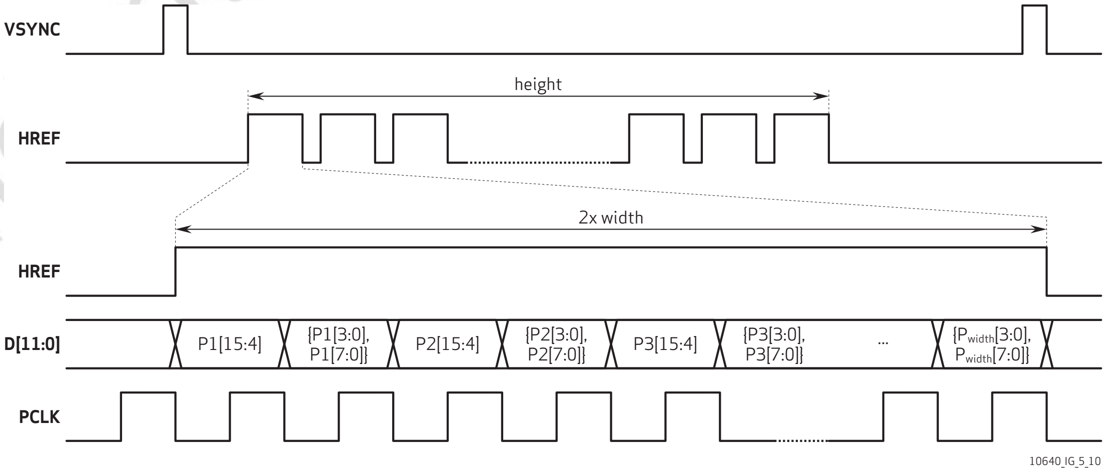


figure 5-11 20b linear combined raw output timing option 1

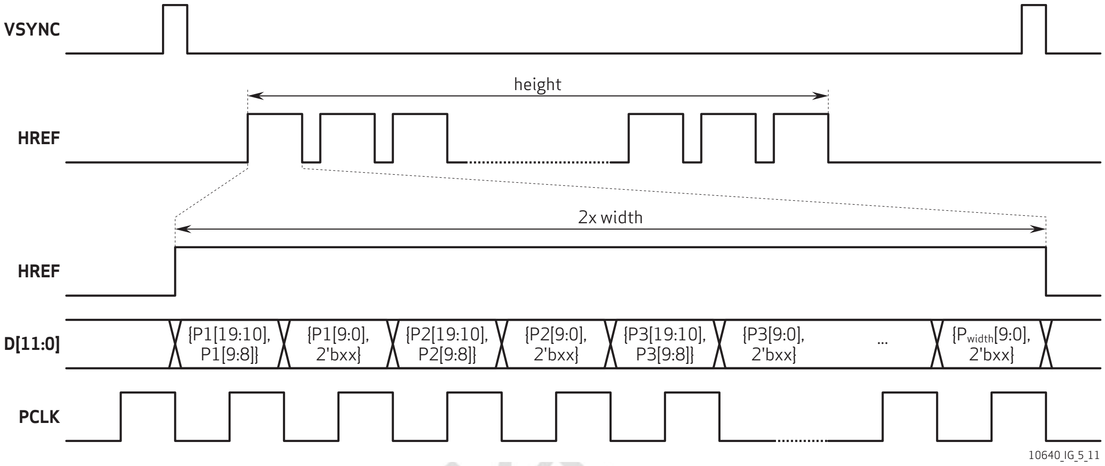


figure 5-12 20b linear combined raw output timing option 2

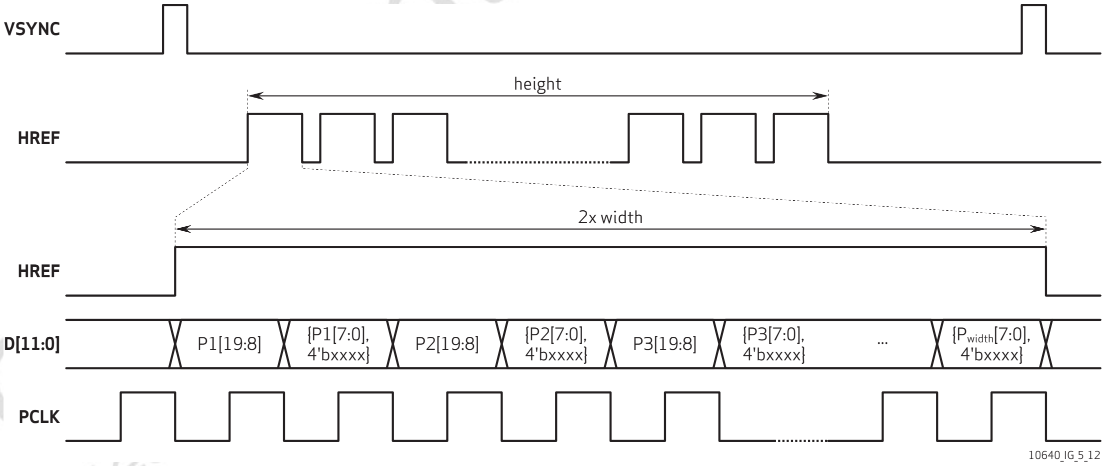
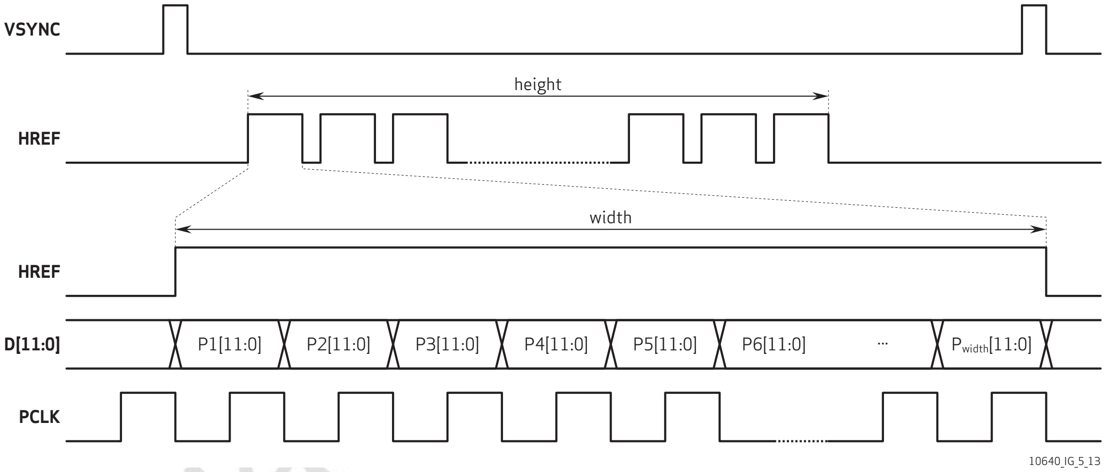


figure 5-13 12b PWL combined raw output timing



5.2.2 MIPI interface

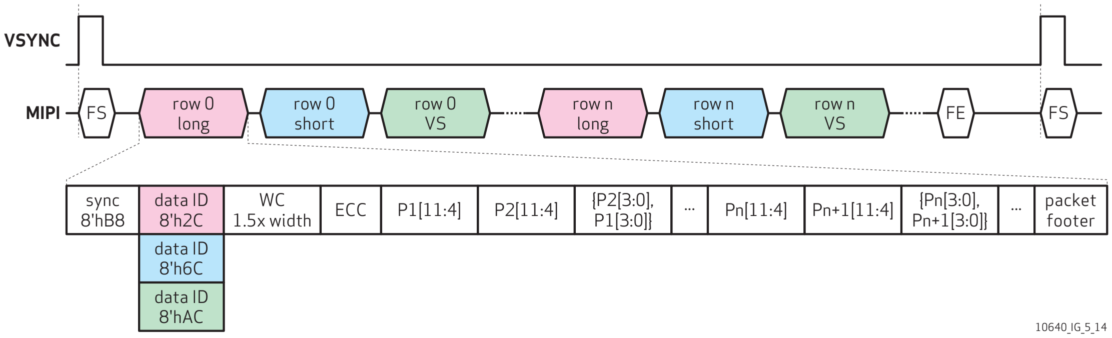
The OV10640 supports up to four data lanes. The maximum data rate is 800 Mbps per lane and the total bandwidth is about 3.2Gbps, which is enough to output full resolution at 60 fps for any data format.

In MIPI mode, L, S and VS are sent out via different virtual channels, so they are output in line-interleaving fashion.

5.2.2.1 3x12b linear raw format

L, S and VS are sent out via virtual channel 0, 1 and 2, respectively. For each exposure channel, every 2 pixels are packed into 3 bytes. **figure 5-14** illustrates the MIPI output packet.

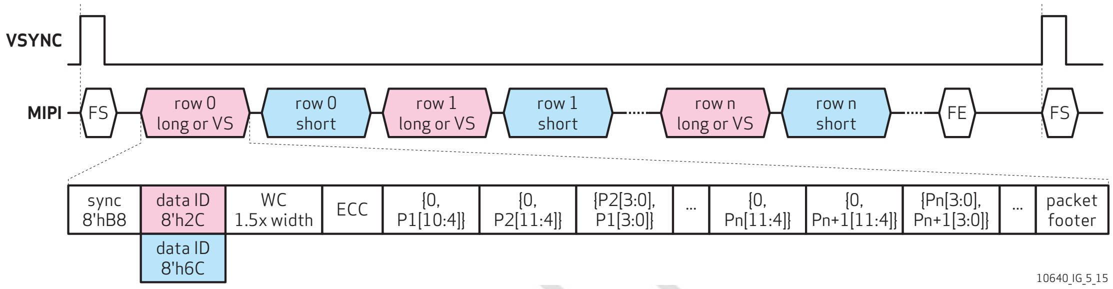
figure 5-14 3x12b MIPI output packet diagram



5.2.2.2 2x11b PWL raw format

The 11b data is extended to 12b with MSB set to 0 and then sent out following CSI-2 raw12 format.

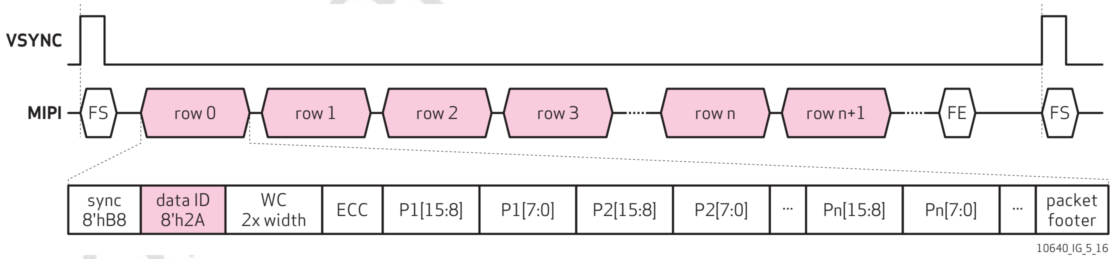
figure 5-15 2x11b MIPI output packet diagram



5.2.2.3 16b log combined raw format

16b log combined raw is sent out following CSI-2 raw8 format. Every pixel is packed into 2 bytes. **figure 5-16** illustrates the MIPI output packet.

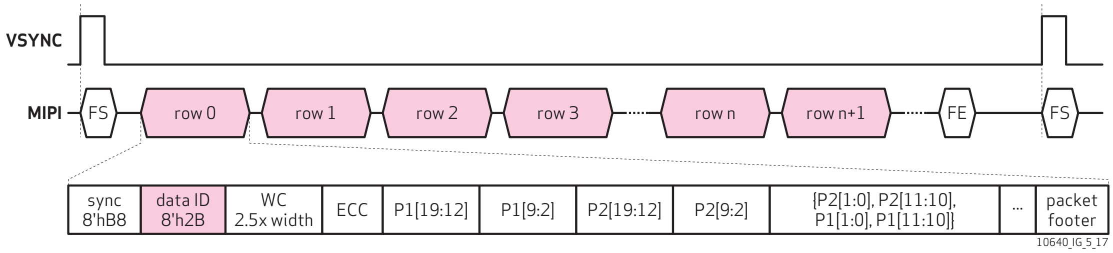
figure 5-16 16b MIPI output packet diagram



5.2.2.4 20b linear combined raw format

20b linear combined raw is sent out following CSI-2 raw 10 format. Every 2 pixels are packed into 5 bytes. **figure 5-17** illustrates the MIPI output packet.

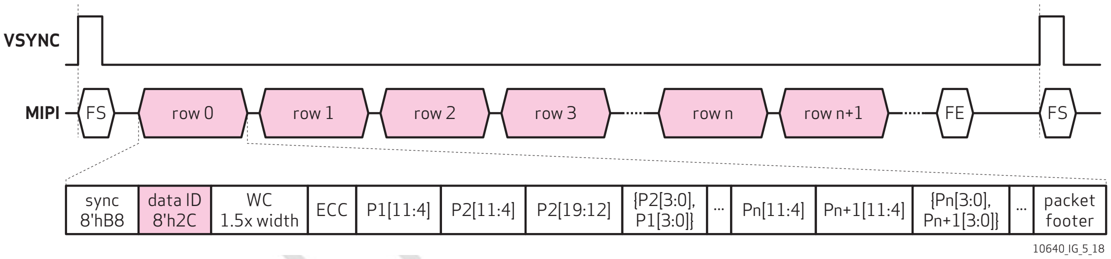
figure 5-17 20b MIPI output packet diagram



5.2.2.5 12b PWL combined raw format

20b linear combined raw is sent out following CSI-2 raw 10 format. Every 2 pixels are packed into 5 bytes. **figure 5-18** illustrates the MIPI output packet.

figure 5-18 20b MIPI output packet diagram



6 embedded row

Embedded row is used to send register values and internal statistics to the backend processor. It is strongly recommended for the backend processor to get the information from the sensor through an embedded row instead of the SCCB. When the OV10640 interfaces with the backend processor via a pair of SerDes and LVDS links, the SCCB is supported by the LVDS link back channel, so it is slower to read registers through the SCCB. The other benefit is that the information is synchronous with the video frame, which is very critical for any control loop (e.g., AEC/AGC).

The OV10640 can output four embedded rows - two top rows before video frame and two bottom rows after video frame. The top embedded rows are enabled by register bit 0x3091[2], and the bottom rows are enabled by register bit 0x3091[3]. The top embedded rows are usually used to transfer information known before the video frame (e.g., the exposure, gain, black row average, image size, etc.). The bottom embedded rows are usually used to transfer statistic information which is available only after the video frame finishes.

The register list of each embedded row is programmable by specifying the start and end address of the registers. The OV10640 register address is continuous to allow easy programming of embedded rows. The following is an example.

```
; Top embedded rows register address range

     60 3030 30 ; start address is 0x3000

     60 3031 00

     60 3032 35 ; end register address is 0x3500

     60 3033 00

; Bottom embedded rows register address range

     60 3034 40 ; start address is 0x4000

     60 3035 00

     60 3036 42 ; end register address is 0x4200

     60 3037 00
```

Each register value is tagged by 0xDA. The register value and tag are MSB aligned in 10-bit (used for in MIPI interface for 20b format) or 12-bit data format. In 2x11b and 3x12b image format, the register value with its tag is repeated two and three times, respectively, as if each register is a pixel.

7 backend processor ISP consideration

7.1 AEC/AGC and dynamic range control

The OV10640 extends the dynamic range by the ratio of exposure between captures, so exposure control of each capture is the key for dynamic range. It is highly recommended to adjust the exposure of each capture based on the histogram of three captures together. Simple average-based AEC/AGC will not cover the scene dynamic range very well in some cases.

7.1.1 collect statistics

3x12b is the format without any information loss so the backend processor can collect the statistics data it requires to control the exposure and gain. The drawback is that the bandwidth requirement is high, so the MIPI interface is recommended. For the DVP interface, the maximum frame rate of 3x12b is 30fps at a resolution of 1280x960.

If interface bandwidth is the bottleneck of the system, the second recommended format is 2x11b format. The bandwidth requirement of 2x11b is 1/3 less than 3x12b format. Though some information is lost in 2x11b compression, the loss is below shot noise so the image quality is very close to 3x12b format. The backend processor can recover all three captures for each pixel and get the statistics.

For 2x11b format, the backend processor should use L or VS throughout its pipeline for each pixel. It should not derive VS from L when L becomes saturated somewhere in the pipeline after amplified by any digital gain. The reason is that VS and L are not correlated to each other for a moving object or an object with varying intensity, so the VS value derived from L will be different to the VS read out from the pixel. In a smooth area, if some pixels come from the VS capture and some other pixels are derived from L capture, the difference between derived VS and the real VS capture will make noticeable artifacts.

As mentioned, the interface bandwidth is greatly reduced for combined raw formats and the pixel wise computation requirement is also much less. However, it is difficult to collect the statistics for each exposure channel accurately. The statistics of each exposure channel can be extracted from their data range accordingly, but the result is not as accurate as 3x12b or 2x11b format.

7.1.2 calculate exposure and gain

The pixel response is proportional to the product of exposure time and gain. The first step is to calculate the product of the new exposure and gain for L, S and VS channels from their current image level, the target image level, and their current exposure and gain.

Calculating exposure and gain from their product is not trivial. A certain amount of digital gain and possibly a minimum analog gain must be applied. Refer to **section 3,** ADC range and minimum gain and the following formula:

$$gain_{dig\_min} = \frac{4095 - C_{BLC}}{4095 - (y_{Black} + C_{ADC})}$$

$$gain_{extra\_min} = \frac{4095 - C_{BLC}}{y_{White} \times gain_{dig\_min}}$$

$$y_{Black} = gain_{ana} \times y_{Dark}$$

Please keep in mind that y_{Black} is a function of exposure time and analog gain and the new minimum gains must be based on the new exposure and analog gain. The normalized dark current (per row period at 1x analog gain) can be calculated from current y_{Black} divided by current analog gain and current exposure time. Then, multiply this normalized dark current by the new exposure time, new analog gain to get the new y_{Black}. If the digital gain is less than 1.35x, it has to be set to 1.35x for the anti-Black Sun function to work properly.

Because dark current increases exponentially with temperature, high analog gain at high temperature may lead to too high y_{Black} and only leave a very small ADC range for photon current. In this case, the signal-to-noise ratio (SNR) is poor due to a small signal. It is recommended to keep y_{Black} below 1/3 of the full ADC range whenever possible. This can be done by applying the proper analog and digital gain. y_{Black} is a function of analog gain but does not change with digital gain. At cold temperature, dark current is low so y_{Black} is still low even with high analog gain. When temperature increases, analog gain must be controlled to not make y_{Black} too high. In summary, analog gain is preferred at cold or nominal temperature, but not at high temperature.

7.2 AWB

7.2.1 WB gain of L, S and VS channels

Unlike exposure and gain, the WB gain of three exposure channels is very similar. In fact, the L and VS are from the same photo diode, so their gain should be the same. The S channel is from a different photo diode with slightly different spectral response, so its WB gain is slightly different to L and VS channel. Based on experimental data, it is recommended to calculate S channel WB gain from L/VS channel by a fixed ratio as below.

$$R_{Gain\_S} = R_{Gain\_L\&VS} \times C_R$$
$$B_{Gain\_S} = B_{Gain\_L\&VS} \times C_B$$

where C_R is recommended to be 1.xx, and C_B is recommended to be 0.yy.

The OV10640 will apply AWB gain on the image data with BLC offset so the user must change AWB offset at the same time AWB gain is changed.

$$AWB_{offset} = (AWB_{Gain} - 1) \times BLC_{target} \times 256$$

7.2.2 3x12b or 2x11b format

For these two formats, it is recommended to get the WB gain of L and VS channel from L channel statistics according to the AWB algorithm, and then using the above formula to calculate WB gain for S channel.

The WB gain can be applied in the backend processor, or in the sensor depending on the image processing pipeline. The sensor WB gain registers of L, S and VS are listed below. Keep these gains at 1x if the WB gain is applied in the backend processor.

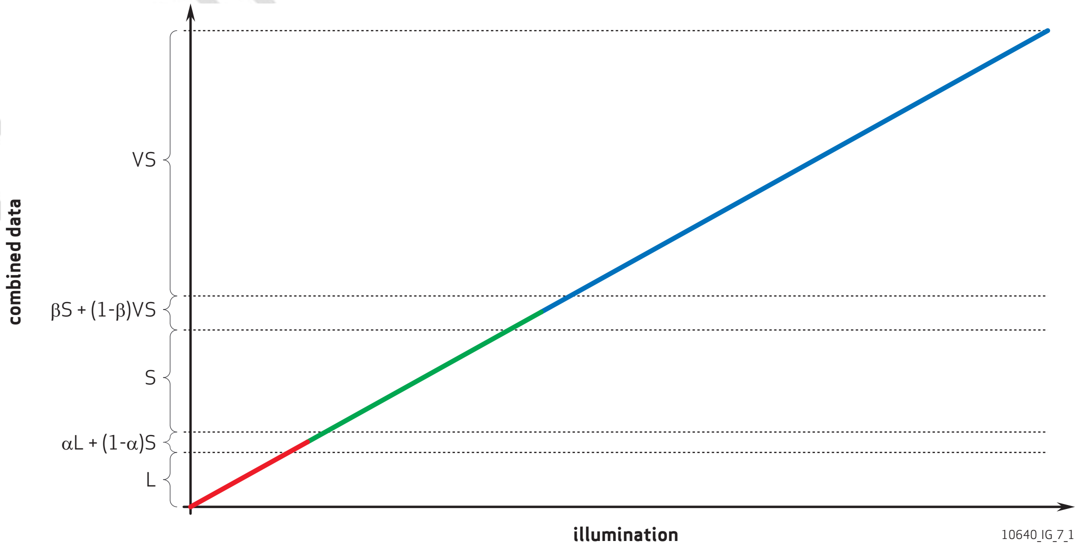
7.2.3 combined raw format

If the AWB algorithm allows applying WB gain in the sensor (i.e., it can calculate the new WB gain from the combined raw data with WB gain applied), it is very straightforward to apply the WB gain in the sensor. Refer to the sensor WB gain registers listed in the last section.

If the WB gain must be applied in backend processor, it is recommended to apply above C_R and C_B gain on S channel in the sensor, so the backend processor can apply the WB gain on the combined data. To keep the gain no less than 1 (otherwise, it causes a saturation issue and color artifact), the blue channel gain is set to 1, and blue and green channel gain is multiplied by $1/C_B$. This is equivalent to a global digital gain of $1/C_B$ on S channel. It is recommended to consider this into the dynamic range control.

Referring to **figure 7-1** showing a graph of the combined data versus illumination, it is recommended to collect the WB statistics data from the data range of L channel or L and S channels. The VS data is usually too bright and the number of pixels is usually small, so it is better to ignore them for AWB. Also, try to exclude the combination value of L and S, because the L of some pixels may be saturated and the linearity is not very good.

figure 7-1 linear combined data



7.3 defect pixel correction

The OV10640 supports dynamic defect pixel correction. This algorithm, however, is optimized for a white balanced input image. If the WB gain is applied by the backend processor and defect pixel correction is done by the OV10640, the defect correction will causes color artifacts at high contrast edges (e.g., thin black line on a bright background, or a bright line on a black background). It is recommended to do defect pixel correction with the backend processor if the WB gain is applied by the backend processor.

revision history

version 1.0 09.03.2014

- initial release

version 1.1 04.22.2015

- at the bottom of page 3-2, changed sentence " For the OV10640, y_{White} is 3500 for L and VS, and 2048 for VS (may change for final production version)." to " For the OV10640, y_{White} is 3500 for L and VS, and 1600 for S."
- in chapter 3, changed the last two sentences from "A minimum 1.4x digital gain is required for this circuit to operate properly. If the minimal digital gain calculated earlier is lower than 1.4x, the minimum gain must be set to 1.4x." to "A minimum 1.35x digital gain is required for this circuit to operate properly. If the minimal digital gain calculated earlier is lower than 1.35x, the minimum gain must be set to 1.35x." and added "Very short channel has the same minimum gain as long channel and short channel gain is 3x."
- in chapter 4, changed second sentence after figure 4-2 to "The maximum frame rate is 60 fps at full resolution." and removed next two sentences
- in section 5.1, deleted ", or all three channels in line interleaved fashion" from first sentence and added "The OV10640 does not have line buffers to support line interleaved output (with different virtual channels. It is always pixel interleaved (MIPI, LVDS, and DVP)."
- in section 5.1.3, added "There are two options: when register bit 0x31BE[0] is 1’b1, the combination occurs on the original 12 bits of data and max threshold should be 2^12 and when register bit 0x31BE[0] is 1’b0, the combination occurs on the original 10 bits of data (2 LSBs are lost) and max threshold should be 2^10"
- in section 5.1.4, added "When converting data from 16 bits log to 20 bits linear, please pay attention to the normalization feature. After combination, there is an option which can normalize the data or not. If enabled, normalization (register bit 0x328A[1] = 1’b0), 20b linear raw data is MSB aligned; else, it is LSB aligned. If normalization is off, there is data overflow when the total ratio is larger than 1024x (register bit 0x31BE[0] is 1’b0) or 256x (register bit 0x31BE[0] is 1’b1). to second paragraph
- in section 5.1.4, added subsections 5.1.4.1 and 5.1.4.2 including figures 5-5 and 5-6
- in section 7.1.2, changed last sentence of third paragraph to "If the digital gain is less than 1.35x, it has to be set to 1.35x for the anti-Black Sun function to work properly."

defining the future of digital imaging™

OmniVision Technologies, Inc.

UNITED STATES

4275 Burton Drive
Santa Clara, CA 95054

tel: +1 408 567 3000
fax: +1 408 567 3001
email: sales@ovt.com

UNITED KINGDOM

Hook, Hampshire +44 1256 744 610

GERMANY

Munich +49 89 63 81 99 88

INDIA

Bangalore +91 80 4112 8966

CHINA

Beijing + 86 10 6580 1690
Shanghai + 86 21 6175 9888
+86 21 5774 9288
Shenzhen + 86 755 8384 9733

JAPAN

Yokohama +81 45 478 7977
Osaka +81 6 4964 2606

KOREA

Seoul + 82 2 3472 3769

SINGAPORE +65 6933 1933

TAIWAN

Taipei +886 2 2657 9800
Hsin-chu +886 3 5656688

website: **www.ovt.com**