in Proc. 2011 International Image Sensor Workshop,
Hokkaido, Japan, June 8-11, 2011.

Investigation of Two-Layer Photodetectors for YSNR10 Improvement in Submicron Pixels

Eric R. Fossum

Samsung Electronics Semiconductor R&D Center, Giheung-Gu, Yongin, South Korea
and Thayer School of Engineering at Dartmouth, Hanover, New Hampshire USA

Abstract—Two-layer pixels and CFA kernels are investigated for improving luminance signal-to-noise ratio (YSNR) in color image sensors. It is found that some improvement is possible compared to conventional single-layer pixels for both YSNR and color resolution at the possible expense of color fidelity and fabrication complexity.

I. INTRODUCTION

An important figure of merit for color CMOS image sensors is the low-light signal-to-noise ratio (SNR) for the individual colors and the luminance. For example, a luminance-SNR (YSNR) of 10 is desired at 100 lux scene illumination by camera-phone manufacturers [1].

Conventional single-layer pixels have been shrunk in size to remarkably small dimensions. Pixel pitches are routinely 2200 nm and as small as 1100 nm. Next generation devices will shrink to possibly to 900 nm. For a given scene illumination and optical F-number, the number of photons per pixel drops as the pixel pitch shrinks. Maintaining a YSNR of 10 at 100 lux in sub-micrometer pixel pitches is difficult.

For most modern small-pixel devices, the noise is dominated by the photon shot noise, not read noise, so SNR can only be improved by increasing the number of photoelectrons from the detector. Photons lost by reflection or absorption prior to entering the pixel result in reduced SNR. For example, much effort has been placed on improving microlens performance, and enabling commercialization of backside illumination (BSI) to increase quantum efficiency [2-4].

The conventional red-green-blue (RGB) color filter array (CFA), by its nature, admits about 30% of the incident photons and the rest are absorbed in the CFA and lost. It would be advantageous to utilize a higher proportion of these photons.

One possible strategy is to use multi-layer photodetection so that more photons can contribute to the signal level and improve SNR. In a triple-layer (3L) photodetector, the CFA is omitted and the photodetector provides the color selectivity [*e.g.*, 5-9]. Two-layer (2L) photodetection has been previously investigated using yellow-cyan and green-magenta checkerboard CFA yielding promising results despite performance problems due to the detector implementation [10, 11].

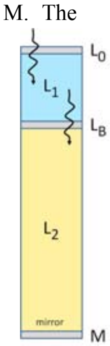
Application of multi-layer photodetection to small pixels is a challenge for several reasons. First, making contact to buried photodetection layers reduces fill-factor. Second, readout of the additional photodetector layers usually introduces additional noise (e.g. kTC noise) and often lag, and the corresponding readout circuits also reduce pixel area. Third, fabrication of a complex vertical structure requires both significant development expense and can add to total wafer fabrication cost in a very cost-sensitive market. Fourth, removal of spectral overlap in the digital color processing can reduce SNR of the color signals. For these reasons, multi-layer photodetection devices are presently non-existent in almost all digital camera applications including cell phone cameras and web cameras.

This paper reports on the exploratory theoretical investigation of 2L photodetectors for improving the YSNR in next generation sub-micron pixels. With the advent of commercial backside-illumination fabrication processes, the constraint on using pixel area for vias or readout circuits is reduced. The use of a 2L structure has less fabrication complexity compared to implementing a three-layer structure. 2L devices can yield improved color resolution compared to single-layer (1L) pixels for same pixel pitch. Finally, by moving kTC read noise to less critical components of the YSNR, the impact of higher read noise can be minimized. 2L photodetection can also improve YSNR in larger pixels as well.

In this paper, the generation of RGB color signals from CFA and layer combinations is investigated. The color and luminance SNR is estimated for several possible combinations. This modeling helps to illuminate future paths for image sensor pixel device development.

II. STRUCTURE AND RESPONSE

In this work a general 2L structure is modeled. As shown in Fig. 1 it consists of two main absorber layers L_1 and L_2 and three ancillary layers called L_0 (dead), L_B (barrier) and M. The mirror layer M has been used in BSI infrared detectors to increase the effective absorption path of photons [12]. Above L_0, two filter layers are assumed (but not shown): a NIR/UV filter and the conventional color filter layer. The complete model, from illuminant, carrier generation and collection thru color optimization for each case was specially developed for this work and is quite detailed and complex. However, this paper focuses only on the results.



Some of the structures considered by the author include (1) the structure of Findlater, (2) BSI structures with direct readout of L_1 and intra-pixel charge-transfer readout of L_2, and (3) a novel BSI vertical-charge-transfer (VCT) device in which all kTC noise is suppressed. The results here apply to all these candidate implementations.

The nominal layer thicknesses are shown in Table I. Variations in these thicknesses were also examined but the main conclusions can be drawn from using the values in the Table. The effect of thickness on optical crosstalk was not examined in this work.

Table I. Parameter values.

Parameter	Value
L_0 thickness	20 nm
L_1 thickness	1000 nm
L_B thickness (optical)	0 nm
L_2 thickness (effective)	6000 nm
Scene illuminant	CIE A (2856K)
Scene reflectivity	18%
Lens F-number	2.8
Total trans. from lens to Si	55%
Pixel pitch	900 nm
BSI eff. fill factor	90%
Carrier collection efficiency	100%
Integration time	67 msec (15 fps)
Ideal red center/halfwidth	600/50 nm
Ideal green center/halfwidth	555/66 nm
Ideal blue center/halfwidth	450/33 nm

The number of photons absorbed between depth z_1 and depth z_2 at wavelength λ, $N_{abs}(z_1,z_2,\lambda)$, is simply:

$$N_{abs}(z_1,z_2,\lambda) = N_o(\lambda)\ (\ exp(-\alpha(\lambda)z_1) - exp(-\alpha(\lambda)z_2)\ ) \qquad (1)$$

where $N_o(\lambda)$ is the number of photons entering the silicon surface and $\alpha(\lambda)$ is the characteristic absorption length. Layers L_1 and L_2 each have their own characteristic response.

The on-chip filter layers are central to the operating principle of two-layer structures. The filter cut-on and cut-off characteristics depend on how they are made and can be somewhat tailored. Generic filter characteristics are used in this model. It is noted that improved sensor performance might be obtained by optimization of the filter characteristics. It is also noted that many non-idealities exist in actual filter characteristics but these non-idealities do not have a significant effect on the modeling results. Reflection and interference effects are also present in real structures but ignored in this model.

The product of the color filter response with the L_1 or L_2 response gives the photodetector layer spectral response. The 3 normal filters red (R), green (G), and blue (B), the 3 complementary filters, Cyan (C), Yellow (Y) and Magenta (M), and a clear filter which is referred to as white (W) can be considered. These 7 filters along with the L_1, L_2, and L_1+L_2 responses yield 21 different photodetector relative response curves. The signals generated in L_1 and L_2 depend on scene, camera and sensor parameters. These parameters are shown in Table I.

This paper does not address the detailed implementation of a 2L pixel but many possibilities exist. In the implementation of a 2L pixel, the readout of L_1 and L_2 may be done separately with each layer having a different read noise contribution. It is noted that the L_1 readout will mostly affect the blue signal and hence have a relatively small contribution to luminance noise (see Eq. 6). In a BSI structure, readout of L_1 – closest to the backside and furthest from readout circuitry – may thus be permitted to incur kTC noise.

III. COLOR PROCESSING

A. Nomenclature

Color image sensors generally consist of an array of pixel *kernels* where each kernel consists of four pixels. In the conventional Bayer-type CFA, the kernel consists of 2 X 2 (four) 1L pixels that are covered respectively by R, G, G, and B filters, referred to as a RGBG kernel. The red, green and blue signals for each pixel, *R*, *G*, and *B*, are determined by a digital color processing algorithm that typically consists of interpolation, color correction and white balance [13]. It has been proposed to replace one of the green pixels by a white (clear) pixel to form a RGBW kernel, or to expand the kernel by adding more white pixels - a RGBnW configuration [e.g., 14 and references therein].

This work is concerned with 2L pixels. For nomenclature, the pixels in the kernel are labeled using the first letter of the filter (R, G, B, C, Y, M, or W) and the number of layers used. For example, G2M2 would be a two-pixel kernel with one 2L pixel covered by a green filter, and one 2L pixel covered by a magenta filter. This kernel would have four output signals, $G\text{-}L_1$, $G\text{-}L_2$, $M\text{-}L_1$ and $M\text{-}L_2$.

It is found that even for 2L pixels, some useful kernels include pixels where the responses of the two layers are added together. For example, a G1M2 kernel would be the same as the above example except the $G\text{-}L_1$ and $G\text{-}L_2$ signals are summed as *G*.

In the simplest conventional RGBG kernel color interpolation, *R, G, B* signals are generated for each of the four pixels by interpolating the missing signals from adjacent pixels. Similarly in a kernel which may contain 2L pixels, interpolation must be performed to generate a full set of spectral signals for each pixel.

B. Color Correction Matrix (CCM) and White Balance

The purpose of the color correction matrix (CCM) is to (typically) produce *R,G,B* signals from the pixels that match some ideal RGB response as closely as possible. Generally, if there are *n* spectral signals S_1, $S_2 \ldots S_n$, represented by the matrix *[S]*, then the color correction matrix is a 3x*n* matrix such that the corrected signal colors *[RGB]* is given by:

$$[RGB] = [CCM] \times [S] \tag{2}$$

For a 1L pixel kernel with conventional RGB color filters, the response of the pixels is fairly close to the ideal RGB response and the main issues are in color overlap. The off-diagonal CCM values are usually small leading to both excellent color reproduction and SNR.

In 2L photodetectors, there can be large spectral overlap between the signals leading to significant negative-value CCM elements. The signals subtract but the combined noise grows, reducing SNR. It is this same effect that limits the utility of complementary color filters in 1L image sensors. By their nature, complementary color filters admit two bands of color and in the CCM, one of these must be subtracted out. The color comes out fine, but the noise can be large so that the SNR is reduced. Some tradeoff between color quality and SNR can be achieved by judicious choice of CCM values during CCM optimization by error minimization.

The white balanced RGB signals are given by:

$$R' = r_b R,\ G' = g_b G,\ and\ B' = b_b B \tag{3}$$

where r_b, g_b, and b_b are the white balance gain values.

C. Luminance

The luminance from R, G, B is determined by a linearly weighted sum of the R, G and B signals. The luminance is brightness of the signal as seen by the human eye as described by the photopic curve or luminosity function $y(\lambda)$ [15]. The luminance signal is given by $Y = r_y R' + g_y G' + b_y B'$, where r_y, g_y and b_y measure the weighting of the color response by the photopic response and are calculated as:

$$r_y = (1/y_o) \int R'(\lambda)\, y(\lambda)\, d\lambda, \tag{4a}$$
$$g_y = (1/y_o) \int G'(\lambda)\, y(\lambda)\, d\lambda, \tag{4b}$$
$$b_y = (1/y_o) \int B'(\lambda)\, y(\lambda)\, d\lambda \tag{4c}$$
$$\text{and } y_o = \int [R'(\lambda)+G'(\lambda)+B'(\lambda)]\, y(\lambda)\, d\lambda \tag{4d}$$

The luminance coefficients depend on the choice of ideal RGB response curve used for CCM optimization. Usually one fixes the R, G, B response for the luminance calculation in terms of ideal response curves. The ideal response is not a universal standard and color quality is often in the eyes of the beholder. In this work, the ideal RGB response is taken from the human eye spectral response with Gaussian fits shown in Table I. For these ideal responses and in this work, the luminance is determined by:

$$Y = 0.370\, R + 0.577\, G + 0.053\, B \tag{5}$$

Thus noise in the green signal can have great effect on the luminance SNR whereas the same noise in the blue signal has ten-fold less effect on the luminance SNR.

To achieve the coefficients used in sRGB standards (0.2126, 0.7152, and 0.0722 respectively), the red ideal response should be shifted from 600 nm to 640 nm in Table I. It was found that shifting the red ideal response curve to 640 nm and using the sRGB coefficients vs. those in Eq. 5 had small effect on the relative results of this study.

IV. SNR AND KERNEL COMBINATIONS

A. Noise.

In this work only two sources of noise were considered. These are input-referred read noise and photon shot noise. The input-referred read noise n_r for kTC noise-suppressed readout is well under 5 e- r.m.s. in state-of-the-art devices. Feedback-noise-cancelling schemes claim a read noise of about 7-10 e- rms. Without kTC noise suppression, noise can be estimated at 23 e- r.m.s. for a conversion gain of 50 uV/e-. The shot noise in signal S_i is just $\sqrt{S_i}$.

For the model, photodetector signal S_i has noise n_i given by:

$$n_i = \sqrt{(S_i + n_{ri}^2)} \quad (6)$$

B. YSNR Calculation

After CCM and white balance, the luminance signal Y is given by:

$$Y = \sum_i (r_y r_b r_i + g_y g_b g_i + b_y b_b b_i)\, S_i \quad (7)$$

Where r_i, g_i and b_i are the CCM column elements corresponding to the signal S_i. The noise in the luminance n_y is given by:

$$n_y^2 = \sum_i [(r_y r_b r_i)^2 + (g_y g_b g_i)^2 + (b_y b_b b_i)^2]\, n_i^2 \quad (8)$$

and the YSNR is just given by:

$$\text{YSNR} = Y / n_y \quad (9)$$

The term "YSNR10" refers to the lux level at which the YSNR has the value of 10 and it is desired that this lux level be less than 100 lux if possible in each succeeding generation of pixel.

C. Kernel Size

The minimum kernel size for 1L Bayer RGB pixels is 3 pixels although 4 is commonly used. For 2L pixels, the minimum kernel size is 2 pixels since 4 signals are sufficient to generate R,G and B and the CCM is a 3x4 matrix. For a 1.5 L kernel a 3x3 CCM is used. Color resolution can be improved with 2L photodetection since all colors can be determined in half the area for the same pixel spatial pitch.

It is important to compare different kernels on an equal basis. At the same pixel pitch, a 2-pixel kernel receives half the total number of photons as a 4-pixel kernel and would be at a disadvantage for total SNR if one did not correct for this. To do this, the 2-pixel kernel is assumed to have a 2-pixel mirror (resulting in a checkerboard pattern). If the signals from the two 2-pixel kernels are added, the luminance signal will double and the noise will increase by √2. Thus, the SNR will increase by √2. This is included when reporting the results of this work.

D. Effect of White Pixels

As discussed above, white pixels are sometimes added to RGB pixels to form a RGBW pixel. In this case the color correction is performed for the RGB pixels. The computed R, G, and B values from the RGB pixels give an estimate of how the white pixel signal should be partitioned into R, G and B components. The partitioning is assumed to be a noiseless process – that is, even though the ratio of R G B is determined using noisy signals, it is assumed that the noise introduced by the partitioning is 2^{nd} order and can be ignored. Of course the partitioned signals still contain shot noise according to the number of electrons assigned to a particular color.

Such white pixel partitioning can be done with 2L kernels as well (using the L_1+L_2 white pixel signal). One needs to determine the ratio of R G B using the other appropriate signals from the kernel, and then partition the white pixel signal(s) in the kernel accordingly.

V. RESULTS

Many different kernel combinations have been explored in this work and the results are summarized in Table II.

For each CFA kernel explored, first color deviation was minimized and the YSNR computed for 100 lux scene illumination. Color deviation is measured unconventionally as sum-of-squares deviation from ideal post CCM spectral response. Next, the color deviation constraint was relaxed so that the color deviation did not exceed that of the CYMW kernel and YSNR recomputed. This color quality relaxation generally improves YSNR. These two conditions on color were explored for two different levels of L1 signal read noise. The first case was with no CDS and read noise of 23 e- rms. L2 and L1+L2 signals were modeled with 2 e- rms read noise. In the second case, L1 was modeled also with 2 e- rms read noise.

In Table II, YSNR is expressed as a percentage of the theoretical maximum YSNR one could obtain from 4 white pixels that contributed completely to the luminance, which corresponds to 135 e- or SNR=11.6 in this model. This is done for easy comparison. The first 7 entries in the table are for 1L conventional kernels. The other kernels use 2L photodetectors, except for the last which was done to model a 3L device named X3.

While most of the 2L kernels that were explored have lower YSNR performance than conventional 1L structures, the G1W1W1M2 kernel shows superior performance in YSNR when the color quality is relaxed compared to all 1L structures. The two-pixel 2L kernel G1M2 also shows promise for improved color resolution with good YSNR and color quality.

The effect of layer thickness was also explored. In Table III the L1 and L2 layer thicknesses are varied for the 1L RGBW and 2L G1W1W1M2 kernels. For each thickness, the CCM was optimized for maximum YSNR for an allowed max. color deviation. For most cases, the 2L structure outperforms the 1L structure.

It is noted that changing relative pixel sizes within the kernel, esp. enlarging the white pixel ~2x, can improve YSNR without much color quality penalty.

VI. CONCLUSION

The use of 2L photodetectors can improve the YSNR and color resolution of an image sensor when compared to conventional 1L RGB and RGBnW image sensors. YSNR improvement can be traded with loss in RGB color fidelity. However, implementation of 2L devices may add undesirable manufacturing complexity and cost.

ACKNOWLEDGMENTS

The author appreciates useful discussions with his colleagues at Samsung Electronics, especially Dr. Ilia Osiannikov and Dr. Jungchak Ahn. The support of Dr. YD Park is also appreciated.

REFERENCES

[1] J. Alakarhu, "Image sensors and image quality in mobile phones," in Proc. 2007 International Image Sensor Workshop, Ogunquit Maine, USA June 7-10, 2007.

[2] J-C Ahn, et al ,"Advanced image sensor technology for pixel scaling down toward 1.0 um" in Proc. 2008 IEEE International Electron Devices Meeting, pp. 1-4, San Francisco, CA, Dec. 2008

[3] For example, see Proc. Of International Image Sensor Workshop Symposium on Backside Illumination of Solid-State Image Sensors, Bergen, Norway, June 25, 2009.

[4] S-H. Cho, et al., "Optoelectronic Investigation for High Perormance 1.4 um pixel CMOS image sensors," in Proc. 2007 International Image Sensor Workshop, Ogunquit Maine, USA June 7-10, 2007.

[5] C.F. Gay and R.D. Wieting, "Vertically Integrated Solid State Color Imager," US patent 4,581,625, 1986.

[6] M. Ben Chouikha, "Photodetector based on buried junctions and a corresponding method of manufacture," US Patent 5,883,421, 1999.R. B. Merrill, "Color Separation in an Active Pixel Cell Imaging Array Using a Triple-Well-Structure," US patent 5,965,875, 1999.

[7] R. F. Lyon and P.M. Hubel, "Eyeing the Camera: into the Next Century," IS&T/TSID 10th Color Imaging Conference Proceedings, Scottsdale, AZ USA, 2002

[8] S. Takada, M. Hayashi, T. Mitsui, Y. Maehara, M. Ihama, "CMOS Color Image Sensor Overlaid with Organic Photoelectric Conversion Layers: Depression of Dark Current," in Proc. 2007 International Image Sensor Workshop, Ogunquit Maine, USA June 7-10, 2007.

[9] K. M. Findlater, et al , "Buried double junction pixel using green and magenta filters," 1999 IEEE Workshop on CCDs and Advanced Image Sensors, June 10-12 1999 Karuizawa Japan.

[10] K. M. Findlater, et al., "A CMOS Image Sensor Employing a Double Junction Photodiode," 2001 IEEE Workshop on CCDs and Advanced Image Sensors, June 7-9 Lake Tahoe, Nevada, USA, 2001.

[11] W. F. Kosonocky, F.V. Shallcross, T. Vallani, and J. Groppe, "160 x 244 element Pt-Si Schottky-barrier IR-CCD image sensor," IEEE Trans. Electron Devices, vol. 32(8) pp. 1564-1573 (1985).

[12] J. Nakamura, ed., Image Sensors and Signal Processing for Digital Still Cameras, CRC Press, Taylor and Francis Group LLC, 2006.

[13] H. Honda, Y. Iida, Y. Egawa, and H. Seki, "A Color CMOS Imager With 4×4 White-RGB Color Filter Array for Increased Low-Illumination Signal-to-Noise Ratio," IEEE Trans. Electron Devices vol. 56(11), pp. 2398-2402 (2009).

[14] C. I. E. 1931 Standard Colorimetric Observer data.

Table III

EFFECT OF THICKNESS				
L1 (um)	L2 (um)	L2 (EFF)	YSNR 1L	YSNR 2L
1	2	4	85%	93%
1	1.75	3.5	85%	90%
1	1.5	3	84%	86%
1	1.25	2.5	83%	80%
0.5	1.75	3.5	84%	84%
1	1.75	3.5	85%	90%
1.5	1.75	3.5	85%	91%

Table II.

		L1 NOISE = 23 e- rms L2 NOISE = 2 e- rms			L1 NOISE = 2 e- rms L2 NOISE = 2 e- rms		
				COLOR=1			COLOR=1
CFA	NAME	YSNR	COLOR	YSNR	YSNR	COLOR	YSNR
	RGBG				55%	0.27	67%
	RBCY				56%	0.47	52%
	RGBW				74%	0.27	85%
	RCGW				64%	1.12	
	CYMW				71%	1.02	
	RWGW				75%	2.06	
	RWBW				51%	2.20	
	C2M2	14%	0.85	28%	35%	0.85	46%
	C2Y2	7%	0.62	16%	28%	0.62	43%
	G2M2	12%	0.40	32%	46%	0.40	62%
	G1M2	19%	0.77	33%	52%	0.77	58%
	G2W2	8%	1.42		35%	1.42	
	G1W2	32%	4.07		58%	4.07	
	M2Y2	7%	0.53	18%	34%	0.53	53%
	R1C2	23%	0.84	33%	49%	0.84	58%
	R1W2	20%	3.74		52%	3.74	
	B1Y2	7%	0.68	10%	31%	0.68	39%
	B1W2	39%	4.96		63%	4.96	
	W1M2	16%	3.29		51%	3.29	
	W1W2	30%	6.78		46%	6.78	
	G1W1 W1M2	34%	0.77	55%	81%	0.77	90%
	G1W1 W1W2	52%	4.07		86%	4.07	
	X3	10%	5.52		54%	5.52	