

Operation and performance of a color image sensor with layered photodiodes

David L. Gilblom^{*a}, Sang Keun Yoo^b, Peter Ventura^c

^aAlternative Vision Corporation, P.O. Box 4055, Los Altos, CA, USA 94024-1055;

^bHanVision Co., Ltd., KAIST-AVH, 373-1, Guseong-dong, Yuseong-gu, Daejeon, R.O. Korea

^cPeter Ventura, Foveon, Inc., 2820 San Tomas Expressway, Santa Clara, CA, USA 95051

ABSTRACT

A silicon image sensor has been developed and placed in production using standard 0.18 μ m CMOS process having three stacked photodiodes per pixel location to provide full-color imaging without external color filters. With a fill factor exceeding 50%, this image sensor achieves approximately 45% peak quantum efficiency in the mid range visible and provides usable response extending from the near-ultraviolet to the near-infrared. Initial results from a commercial digital still camera indicate that this device can produce excellent color reproduction with equivalent ISO film speeds from 100 to 400 and that it produces images free from color artifacts common in images made with sensors incorporating color filter arrays. Development is now underway on camera equipment designed to operate this image sensor in a variety of scan modes.

Keywords: Image sensor, CMOS, CCD, color, multispectral

1. INTRODUCTION

One of the barriers to efficient multispectral imaging is the complexity of acquiring and reconstructing images using traditional sensors. Various techniques have been employed during acquisition including use of a single sensor with sequential filters, multiple sensors on band-separation prisms, on-sensor filter matrices and sequential filtered illumination. Each of these techniques suffers from certain limitations – time separation with the sequential methods and reconstruction errors with filter matrix methods, for example.

Sensor designers have long sought to circumvent these limitations by developing techniques that provide independent detection of more than one spectral band at each photosite. Two-band techniques have been developed for infrared and workable implementations have been devised for use in visible CCDs. Neither of these, however, has proved suitable as a basis for widespread use in commercial color imaging.

Some ten years ago, a strategy for implementing multiple stacked photodiodes using commercial CMOS processes was proposed in which the natural wavelength-dependent absorption properties of silicon could provide inherent filtering in the visible band. Subsequent development has resulted in designs providing sufficient band separation to support high-quality color imaging that can be executed using common CMOS processing.

In this paper, we will present the technology behind a commercial color sensor incorporating layered photodiodes, the Foveon® X3™ Pro 10M, describe its configuration and performance and discuss the implications of its characteristics on common applications.

2. COLOR IMAGING

The fundamental requirement for production of color images, that is, images that can be presented in such a way that the eye perceives them as having color characteristics reasonably approximating the colors in the original, is that at least three samples of the original must be taken using detectors with spectral characteristics that can be combined to produce results close to the spectral characteristics of three normative curves known as the tristimulus (or color-matching) functions (Figure 1). It is significant that the spectral characteristics of the detectors need not match any actual spectral

* dave.gilblom@alt-vision.com; phone +1-650-625-0318; fax +1-650-240-4005; www.alt-vision.com

characteristic of receptors in the eye. Within these constraints a variety of schemes for providing three or more channels of spectrally-separated image data can be employed to produce reasonable color images. Of course, the accuracy of reproduction can vary widely depending on the quantity and the spectral shape of the individual detection channels.

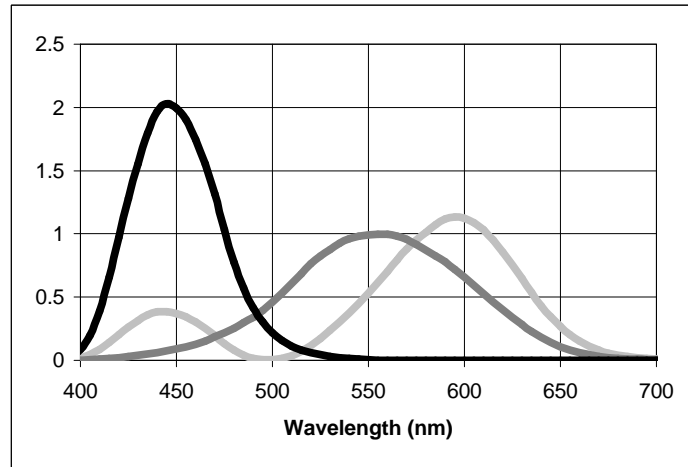


Figure 1 - Tristimulus (color matching) curves (CIE 10° 1964)¹

In addition, the frequency of occurrence of secondary effects such as metamerism, in which two different spectra from objects produce the same perceived color, varies with the characteristics of the channels. It should be noted that in imaging systems, metamerism is locked in at the time of detection. Once the detection channels produce the same signal set with two different stimuli, there is no way to distinguish these cases by subsequent analysis. This is a case similar to that found in another effect, aliasing. Once aliases, signals representing the difference between an incoming spatial frequency and the maximum sampling spatial frequency of the detector, are generated they cannot be removed. As a result, these effects and others must be considered before detection if they are to be minimized.

Often, the color images are not displayed; the goal is instead the accurate measurement of color. In such systems, the color separation requirement is unchanged because the accurate measurement of color requires that the detector channels produce signals that can be combined to produce a three signal set that is the same as the set that would be produced by detectors with the spectral characteristics of the tristimulus curves. To some extent, this is an easier task than color reproduction because the color gamut constraints of the display are removed. Fundamentally, though the requirement remains: can the spectral characteristics of the sensor channels be combined to approximate well enough the tristimulus curves?

2.1.1 Traditional color separation techniques

The techniques to approximate the tristimulus curves fall in two broad categories, those which try to do it with exactly three detector channels and those that use more. Those that use many more channels are essentially hyperspectral imagers, using ten or 20 or 50 narrowband channels to provide accurate spectral distribution curves for the objects being viewed. With these systems, matching the tristimulus curves becomes a purely mathematical operation in which the spectral sensitivities of the numerous channels are linearly weighted and combined. Increasing the number of channels can bring these approximations arbitrarily close to the normative tristimulus curves. Of course, this technique carries substantial burdens of cost, complexity and speed.

Color separation can, of course, be achieved with just three channels. In the trivial case, the three channels have spectral characteristics that match the tristimulus curves and no further adjustments are required. This was the common approach for many years in instruments that measure color at a single point simply because the cost of developing very accurate shaping filters was less than the cost of additional detector channels. Now that silicon sensors are so inexpensive, even

the simplest color measuring instruments often use six or seven channels, improving overall accuracy while relaxing the requirements for accurate spectral characteristics in the individual channels.

In imaging, cost is still very much the issue so techniques using one or three image sensors account for virtually all commercial implementations. These have been of three basic types: one-sensor sequential, one-sensor simultaneous and three-sensor.

2.1.1.1 One-sensor sequential

Sequential color imaging is simply an extension of the usual method of monochrome frame scanning. For color acquisition, three successive frames are taken, each with a color filter inserted in the optical path. Several varieties of filters can be used. Historically, these have been three Kodak gelatin filters – Wratten types 25 (red), 58 (green) and 47 (blue) are a widely used triplet – that separate the spectrum into three relatively non-overlapping sections. Other dye and interference filters sets with this spectral behavior have also been developed. Alternatively, the filter set may use complementary colors – cyan, magenta and yellow – each of which takes in two of the RGB color bands. The CMY scheme is more light-efficient than the RGB but requires calculating strong differences between channels to produce the tristimulus approximations. Which technique is better depends on a variety of image sensor characteristics and the relative importance of various image quality measures.

2.1.1.2 One-sensor simultaneous

The most common method of color separation uses an image sensor provided with a color filter array (CFA), with one color filter per pixel location. Various CFA configurations are possible but the most common is the Bayer arrangement² in which a four-pixel group is provided with two green filters and one each of red and blue in a diagonal configuration (Figure 2). A complementary version of this design is also used.



Figure 2 - Bayer filters: direct and complementary

The Bayer arrangement is particularly useful because it allows the use of relatively simple calculations to estimate missing luminance data and to smooth the color information. Luminance resolution in cameras with color filter arrays is approximately 70%³ of what would be produced by a monochrome sensor with the same pixel count. This results partly from the not having direct luminance samples from each pixel and partly from the anti-aliasing filters that are needed in CFA cameras to control color artifact generation. Color resolution is lower. As a result, aliasing from undersampling begins at different frequencies for the various color components and the direction of strong components varies by color. This nearly always results in visible color artifacts when significant energy at spatial frequencies above the various sampling limits is allowed to reach the sensor. Typically, the missing color information is filled in with estimates made from close neighbors. This can generate color errors due to mismatches in the estimates from pixel to pixel⁴. In some implementations, no attempt is made to fill in the missing information; each color signal is simply treated as complete but of lower resolution. This procedure avoids generation of reconstruction artifacts but reduces perceived resolution and can cause ambiguities in the positions of color edges.

2.1.1.3 Three-sensor

In order to combine full resolution and simultaneous color imaging, three sensors can be used. Most commonly, these sensors are fitted on a prism assembly incorporating internal color filters to direct the R, G and B bands to the three sensors. In these imagers, the monochrome and color resolution are identical except for any errors, generally small, in the alignment of the sensors to each other. Separation prisms of this type must be carefully designed to provide identical optical path lengths for the three bands. Often, special optics are required for low color aberration because the lens design must take into account the thick glass region between the lens and the imager⁵. Because dichroic multilayer filters are almost universally used in prism-based imagers, overlap in the color bands is minimal and the CMY model is not applicable.

2.1.2 Color spaces

A color space is a numerical model with a specified mapping, usually involving three or four values, which provides convenient methods of describing or transmitting color information in specific applications or environments. A common example is the sRGB color space⁶ used in almost all computers that maps tristimulus values to the colors capable of reproduction by the phosphors in CRT displays. The numerical representation for sRGB is the familiar 0-255 scale for each of the three RGB channels. Generally, one color space may be converted to another through the use of a set of simultaneous equations. This process may, however, produce negative values, indicating that a color that can be represented in one color space is out of range for another.

2.1.2.1 Color matrix

Sets of simultaneous equations for color space conversion are often expressed in matrix form. Where the conversions are linear, as in the conversion of tristimulus values to sRGB, the matrix is of the simple 3 x 3 element type (Figure 3).

$$\begin{bmatrix} R_{,sRGB} \\ G_{,sRGB} \\ B_{,sRGB} \end{bmatrix} = \begin{bmatrix} 3.2410 & -1.5374 & -0.4986 \\ -0.9692 & 1.8760 & 0.0416 \\ 0.0556 & -0.2040 & 1.0570 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Figure 3 - Tristimulus to sRGB conversion matrix

Clearly, this can produce both negative values and values greater than 1 for the sRGB coordinates. In practice, those values are discarded because they cannot be reproduced by typical displays.

2.1.2.2 Color gamut

Each color device has a range of color over which it can operate. This range is termed the gamut. Printers, displays and most other output devices have color gamuts that cover perhaps the central 30% of the tristimulus chart; they are not capable of reproducing saturated colors due to the fairly broad spectral characteristics of the inks or phosphors used. Because no two output devices have exactly the same gamut, an additional process, color profiling, is often applied. Color profiles are developed for each device. These are typically non-linear lookup tables that map the standard color space into the actual color gamut of the specific device so that every value presented to the device has some output. Attempting to get the colors in these actual outputs to resemble one another is the science of color management. Color profiles are generally defined according to a format developed by the International Color Consortium (ICC).

2.1.2.3 Detector “gamut”

Strictly speaking, a detector alone does not have a gamut because the definition of gamut requires presentation to the eye. However, detectors do have ranges over which they can produce accurate data representing colors presented to them. Except in rare cases, detectors used for imaging do not have spectral response curves that are designed to match the tristimulus curves directly. Instead, the signals from detectors may have a conversion matrix applied of the same form as those used for output devices and may also have color profiles applied. In fact, the entire ICC color effort is focused on supporting the concept that all input and output devices should communicate through a common color space. The color space of choice is, of course, the tristimulus space because it incorporates all colors discernable to the eye. Thus, the capability for accurate color reproduction for detectors rests on their ability to produce signals that can be accurately transformed (preferably by linear simultaneous equations) into tristimulus values.

2.1.3 Color channel response

Mapping into tristimulus space can be done directly or indirectly. In the case noted above in which many narrowband filters are used, the mapping is indirect, occurring mathematically only after a spectral estimate is made. Extensive testing has demonstrated that the spectral estimation process can be reasonably accurate with as few as six channels⁷. For the most common arrangement, three channels, a 3 x 3 conversion matrix is applied to the camera signals to produce tristimulus data directly. This may subsequently be transformed by another matrix into data suitable for display. In some cases, the two matrices may be combined to provide direct camera-to-display conversion.

2.1.3.1 Color channel characteristics

The filter sets used for color separation generally fall into one of two categories – absorption or dichroic. The Bayer filters are almost all of the absorption type, consisting of arrays of filter elements composed of transparent plastic containing organic dyes. In the typical implementation, these filters have moderate overlap. Dichroic filters, typically used in 3CCD cameras have, in comparison, relatively small overlap (Figure 4). These curves are generic, not representative of a specific sensor.

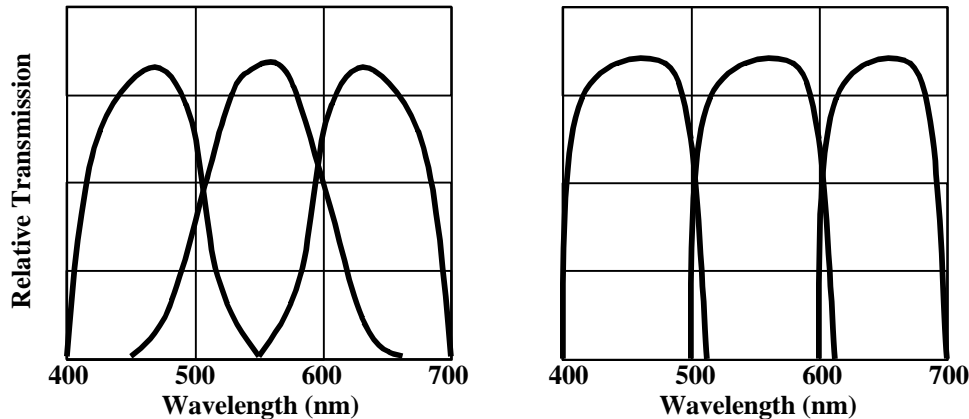


Figure 4 - Dye and dichroic filter response curves

Each of these has advantages. The overlapping curves tend to be able to detect wavelength shifts in narrowband sources better than the separated curves while the separated curves are better able to discriminate between narrowband and wideband sources with the same central wavelength. In addition, the separated curves often have difficulty accurately detecting changes in wavelength near the transitions. In typical implementations, the overlapping curves have high light losses as well because the light not used is absorbed. In the 3CCD designs, the dichroic filters are placed on a color separation prism so minimal light is lost.

Both, however, share one problem. Outside the transmission peaks of the red and blue filters, there is little response from the other channels. As a result, neither type can accurately represent the wavelength of narrowband sources at the spectral extremes. In the red region, this problem has minimal impact because the apparent color to the eye is not much dependent on wavelength. In the blue, though, the eye continues to register color change to the end of its sensitivity range moving from blue to indigo to violet (note the secondary blue peak in the red tristimulus curve in figure 1). In both of these filter schemes, much of the shift can be lost, the wavelength change being recognized only as blue of fading intensity.

3. LAYERED PHOTODIODE SENSING

Semiconductor materials have characteristic photon absorption lengths that vary with the energy of the photons absorbed. Silicon, fortunately, has such a curve that varies by nearly two orders of magnitude over the visible range (Figure 4)⁸. This variation provides sufficient space to stack multiple diode junctions at depths that are both capable of separating photons of various wavelengths and amenable to fabrication using standard CMOS manufacturing processes.

Various technologies have been applied to this idea over the past 30 years. Eastman Kodak built CCDs with two stacked photosensitive layers in the seventies, several organizations have fabricated two or three layer devices with amorphous silicon layers and other have fabricated devices using a variety of BiCMOS and CMOS processes. Infrared devices have also been fabricated using several different compound semiconductors and, most recently, QWIP technologies.

All of these implementations unfortunately failed to produce devices usable for true color sensing with performance and cost suitable for large-volume imaging applications. The key has been developing fabrication technologies capable of placing the three junctions at just the right depths in silicon.

3.1 Junction depths

In practice, junctions at depths around 0.2 μm , 0.8 μm and 3.0 μm provide workable spectral separation for true color imaging. With appropriate junction construction, the top diode can collect charge nearly to the surface of the silicon to extend response into the near ultraviolet and the bottom diode can deplete well into the substrate to provide extended near-infrared response.

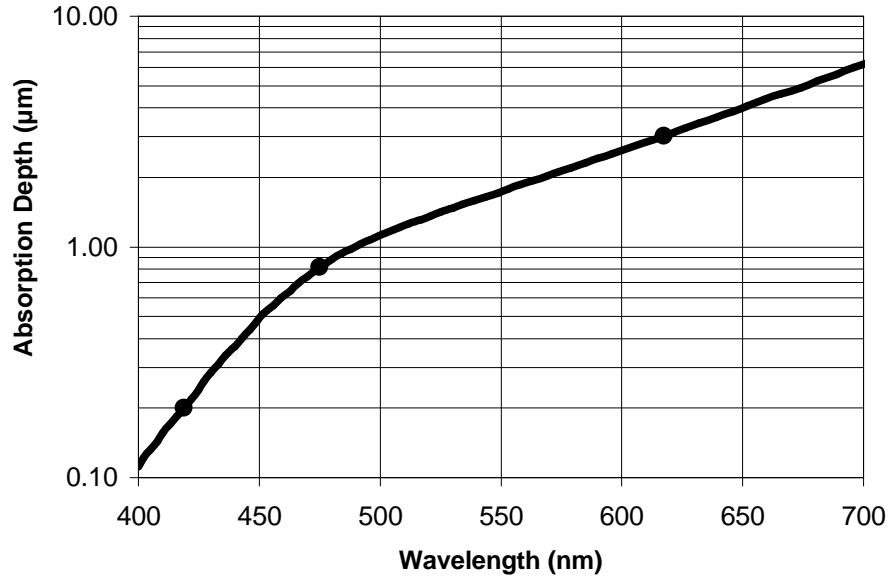


Figure 5 - Silicon photon absorption depth with practical junction depths indicated

3.2 Spectral characteristics

Even with the large changes in absorption depth with wavelength, the response curves of devices using the semiconductor material overlap considerably (Figure 6)⁹. The steep slope in the silicon curve in the 400-475 nm range provides substantial separation of the blue signal from the red and green below, but the relatively shallow slope above 475 nm results in a significant contribution of longer wavelength illumination to the top two signals. Fortunately, the relatively thin absorption regions of the top two diodes minimize this. In addition, some of the short-wavelength photons will make their way into the middle diode. It is this overlap that makes possible the discrimination of wavelength below 450 nm that is so difficult using color filters. The extended response at both ends of the visible spectrum also makes incorporation of a sharp-cut visible filter essential. The curves in figure 6 include the effects of a filter with cutoffs at 400 and 660 nm.

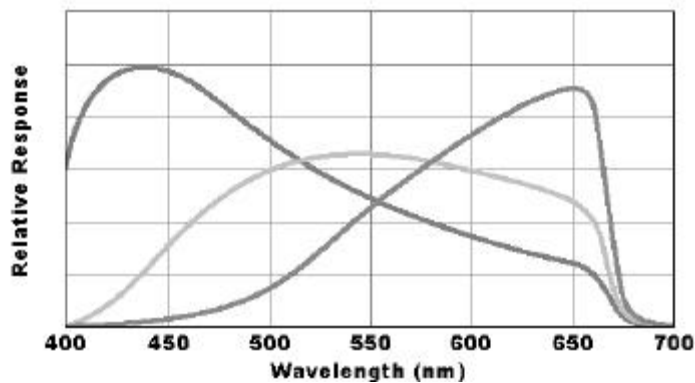


Figure 6 - Relative spectral response of three stacked junctions

3.3 Color matrix effects

The overlap in the spectral response of the three channels, while providing robust information on the wavelength of narrowband sources, leads to relatively large off-diagonal terms in the color transformation matrix needed to produce tristimulus values. As an example, it can be easily seen that the blue channel in the tristimulus function is much narrower than the blue photodiode response. To narrow this channel, a large component of the green must be subtracted. The green is reduced by the blue and the red to narrow the peak and the red channel is augmented with the green to shift the peak to a slightly lower wavelength. Figure 7 shows a typical matrix¹⁰ that incorporates those features.

Sensor Inputs			
	R	G	B
X	0.381	0.597	-0.03
Y	-1.18	3.472	-1.3
Z	0.887	-3.91	4.112

Figure 7 - Sample Color Matrix

While large off-diagonal terms have a negative effect on signal-to-noise ratio, this is generally not more significant than the effect of the absorption losses in color filter arrays. For color filter arrays using yellow-magenta-cyan or other overlapping bands to increase light throughput, large off-diagonal color matrix terms are also needed, negating some of the sensitivity improvement. Those arrangements still absorb at least one-third of the light.

3.4 Photosensor crosstalk

In CCD sensors, the relatively shallow depth of the wells allows lateral diffusion of charge that increases quickly with increases in illumination wavelength. This creates crosstalk between pixels that is wavelength-dependent. Most importantly, it allows charge generated by red light to be collected by the green and blue diodes. In CMOS sensors, these effects are reduced somewhat by the complex structures that surround the photodiodes but there remains even in these direct paths among the diodes receiving all three colors.

4. AN EXAMPLE DEVICE

To produce a practical layered color imager, Foveon developed in a 0.18 micron, 3.3V CMOS process, a triple-diode structure with junction depths of 0.2, 0.8 and 3.2 μm joined with active pixel signal elements and peripheral timing and readout functions. This image sensor, designated the F7 - now known commercially as the X3™ Pro 10M – (Figure 8), has the characteristics shown in Table 1.

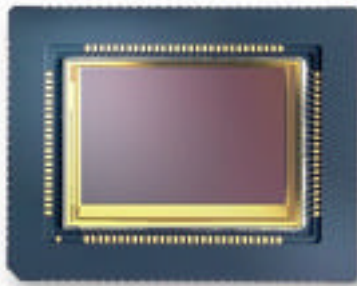


Figure 8 - F7 CMOS color image sensor

Parameter	Value
Pixel pitch	9.12 μ m x 9.12 μ m
Pixel locations	2304 x 1536 (total); 2268 x 1512 (active)
Total photosensors	10.2 million
Active area	20.7mm x 13.8mm
Fill factor	~54%
Output	3 analog
Package	100-pin ceramic leaded chip carrier (CLCC)
Window	Glass with 400-660nm multilayer visible pass filter

Table 1 – F7 image sensor characteristics

In addition, this sensor includes precautions against internal reflection: the inside of the window is coated with a visible range antireflectance layer and the non-active areas between the photosensors are covered with a black mask material. There is no anti-aliasing filter included.

4.1 Operating functions

The use of standard CMOS process affords the opportunity to include extensive control functionality in imagers. In the F7, several useful functions have been implemented.

4.1.1 Power management

Low voltage CMOS and fully-static clocking in the F7 keep the power consumption to 80mW during readout when all counters and signal paths are active. To support battery operation, the F7 also includes a standby mode with 10mW consumption and a power-down mode that maintains register contents while consuming 100 μ W.

4.1.2 Scan control

Scanning in the F7 is controlled by two sets of counters – one each for horizontal and vertical control. The settings of these counters are loaded through a dedicated serial port as a single data stream. Loading takes less than 50 μ s. The counters determine which row is to be activated and which pixel location in that row is to be read out. Once the counters are loaded, the counter operation is controlled by external clocks. Each counter has three registers to set the start count, the increment and the stop count separately in the horizontal and vertical directions. These can be set in any combination to define a rectangular region of interest from one pixel location to the entire array. The increment controls permit sparse and reverse scanning.

To facilitate more rapid scanning of any selected area, the F7 supports a function similar to binning in CCDs in which adjacent pixel locations can be read out in groups designated Variable Pixel Size (VPSTM). VPS uses two controls in each of the horizontal and vertical directions – one to set the number of pixel locations, in powers of 2, to be grouped and the increment control to match the spacing to the VPS settings. Grouping in CMOS devices is not like binning in CCDs. In CCD binning the charges are added from the binned areas and read out in one packet. This serves to increase the signal but also runs the risk of overloading the readout registers in bright areas. In the F7, the grouping connects the voltages generated by the grouped pixels to an amplifier node. This reports an analog combination of the applied voltages rather than a sum and reduces the fixed pattern and random noise rather than increasing the signal.

4.1.3 Exposure control

The F7 includes a global reset control that allows the charge stored in all photodiodes to be dumped and a line reset control that dumps only a selected line. The global reset is used primarily for still shot (SS) imaging of the type common in digital still cameras. In this mode, which uses an external mechanical shutter, the exposure cycle is very simple. To start the cycle, the sensor is globally reset, then, scanning is stopped, the external shutter is opened, an external flash may be fired and the external shutter is closed. After completion of the exposure, scanning is initiated. The external shutter may be a blade or electro-optic type, in which the entire sensor is exposed simultaneously, or a curtain shutter, in which a slot is drawn across the sensor, as is typical in 35mm film cameras. An equivalent cycle may be used where the shutter

is replaced by a pulsed light source, but, in this situation, care must be taken to assure that the sensor is kept in the dark during scanning.

For flexibility, the F7 includes a separate counter to enable lines for reset, which be used together with the vertical readout counter to set a wide range of integration times between reset and readout of each line. This rolling shutter (RS) scan mode is an electronic equivalent of a curtain shutter. The counters can be configured to wrap around at the last line so that the integration time is constant for all lines. In the RS mode, the global reset is never activated.

Since the reset and readout functions are separately controlled, it is also possible to read the same line repeatedly without resetting it. This non-destructive reading permits monitoring of signal buildup during extended integration times to assure maximum dynamic range without overload. With careful signal management, high-speed binned monitoring can be followed by full-resolution readout without damage to the image data.

4.1.4 Signal control

Because the F7 is essentially three image sensors stacked one upon another, three independent sets of setup voltages may be supplied. External access is provided to these three sets of analog amplifiers and references wherever an adjustment might produce useful results. In signals where channel tracking is most important, the inputs are internally tied. A full discussion of the uses and effects of the analog voltage adjustments is beyond the scope of this paper, however, these do affect well capacity, linearity, antiblooming, voltage output levels, frequency response and reset levels.

4.2 Signal characteristics

The output from the F7 is three trains of negative-going pulses representing the voltage values read from the three layers of photodiodes. The clock rate is nominally 12 MHz, although versions of the device operating to 24 MHz are available. Each line time consists of a transfer interval, during which internal signal stabilization takes place followed by a clocked readout. For 12-bit accuracy, the transfer interval is 49 μ s. This duration is fixed regardless of the number of pixel locations read out per line. There is no requirement for waiting at the end of the frame; however, scanning can be stopped here to increase the integration time in the RS mode.

The response of the photodiodes to incoming light is sublinear, primarily as a result of the transfer characteristics of reverse-biased diodes used in a voltage-readout mode. Additional suppression of the response curve is intentionally added using an anti-blooming control voltage to prevent accumulation of charge in the photodiodes in excess of that to be used in the final signal.

Analog processing required off-chip is minimal, consisting typically of a buffer amplifier and then a 12-bit digitizer per channel. Output impedance is 200 ohms. Since each column (or the whole frame) is reset simultaneously, access to individual pixel reset signals at the output is not available. The conversion efficiency is 7.14 μ V/electron and the maximum signal output level is 550 mV.

4.3 Performance

The total quantum efficiency of the F7 at 625nm is approximately 49% including the effects of fill factor. Total quantum efficiency is over 45% from about 530nm to beyond 660nm. Testing is underway to establish the limits of wavelength response. The F7 is expected to have useful sensitivity extending from below 300nm to 1000nm or higher. Well capacity is approximately 77,000 electrons per photodiode but the usual operating point (for restricted non-linearity) corresponds to about 45,000 electrons. Photo response non-uniformity (PRNU) is less than $\pm 1\%$.

Several fixed-pattern and random noise reduction techniques have been incorporated into the F7 design to realize very good noise performance for the CMOS technology. The total fixed pattern noise from all sources is less than $\pm 1\%$. The primary contributor to dark noise is kTC noise from diode reset. This noise is approximately 70 electrons. It is possible to reduce this to about 40 electrons by implementing a reset-read-expose-read cycle for the frame and then subtract the first frame from the second. Lag is zero.

The typical dynamic range of the F7 is 61db. This can be increased by the use of a larger portion of the non-linear portion of the transfer curve as long as care is taken to carefully correct the non-linearity. The signal-to-noise ratio is shot-noise limited with an exposure approximately 10% of the nominal maximum signal point.

Dark current is approximately 1.0 nA/cm^2 at 25C, allowing exposures up to several seconds without cooling. Noise contribution from dark current is very small and dark current uniformity is better than $\pm 1\%$.

5. IMAGE CHARACTERISTICS

The images produced by the F7 have been described by numerous photographers using the Sigma SD-9, the first production camera incorporating this sensor, as “accurate”, “clean” and “clear”^{9, 10, 11}. This impression results largely from the lack of color aliasing and reconstruction noise that affect all color filter array cameras and from the ability of the overlapping spectral bands to provide reasonably accurate tristimulus data for objects with both narrow and broad spectral characteristics across the visible band.

5.1 Optical considerations

Although the stacked photodiode structure eliminates artifacts generated by the offsets of the color receptors in color filter array sensors, some subtle effects remain that should be considered. First, there is a potential with stacked photodiodes to have the sensitivity vary with f-number because lower-angle rays might escape detection by escaping from the side of the diode. However, the diode geometry is designed to minimize this and the effect has not been demonstrated. Next, the presence of low-angle rays might shift the color response due to the variation in ray path length through the silicon. This is possible, but because of the high index of refraction of silicon, incoming rays are bent strongly toward the normal. For most practical optics, the path length variation would be 20% at most. This effect could probably be demonstrated but has not yet been rigorously examined. Finally, the variation in absorption depth in silicon with temperature could result in a significant color shift in cryogenically-cooled cameras. This effect will need to be studied in detail and may require application of a set of temperature-dependent color matrix values. Similar effects are seen in two color single photodiodes.

Although the sensor optical effects are generally minor, the quality of the optics used with the stacked-photodiode sensors can seriously impact image quality. Most significant is chromatic aberration, which is quite visible in monochrome images taken with the F7. Generally, color filter arrays mask chromatic aberration because geometrically the effects are on the order of one pixel. However, F7 images clearly show radially-symmetric color fringing resulting from lenses that suffer from an excess of this problem. Similarly, unsharpness in focus, astigmatism, and other lens defects can have clearly-observable negative effects on images made with stacked-photodiode sensors. Users of the SD-9 camera have often noted how important it is to use the best optics in obtaining the highest image quality from these sensors.

5.2 Processing requirements

In order to optimize dynamic range, noise and uniformity, certain modifications can be introduced into the signal as it is generated on the sensor. These modifications then need to be accurately reversed by the post-processing to provide maximum image quality. In the F7, these modifications are non-linear so maintaining the proper order in reversal is important.

5.2.1 Linearization

The non-linearity in the transfer characteristic must be accurately reversed to prevent shifts in color with brightness. Without correct linearization, the color matrix calculations will produce values that depend on the position of each of the color components on the non-linear curve. If gamma correction is to be applied to the image, this must be done after all linear color calculations are completed.

5.2.2 Dark field subtraction

Several improvements can be obtained by subtracting an image acquired without illumination from the scene image. If the integration time is long enough to produce a dark current signal that is significant, then dark field subtraction can reverse the baseline shift. Shifts in the baseline will affect color matrix calculations. In the still image mode, the dark

field should always be subtracted because the dark current will vary linearly from the top of the image to the bottom. This can cause visible shifts in color vertically in the final image. Dark subtraction can also effectively reduce stationary fixed pattern noise.

In the best case, the dark field should be calculated from a series of frames so that the reset noise can be averaged out. In practice, a single frame must often be used. Still, the increase in noise is usually more than offset by the reduction in non-uniformities. The dark field must undergo the same column filtering and linearization as the scene image to assure accurate subtraction.

5.2.3 Optional steps

Depending on the application, some additional processing steps might be included. For viewed images, compensating for sensor blemishes might be useful. The algorithms for these are the same for the F7 sensor as for monochrome sensors. No special processing is required to accommodate the geometry of color filter arrays. Finally, although the F7 has relatively low non-uniformity, shading correction might be applied, especially if significant optical rolloff exists.

5.2.4 Color space conversion

The actual color space to be used will be determined by the application. If the image is to be directly displayed then a single sensor to sRGB space with gamma correction might be sufficient. If the image is to be used in a color management system, then sensor to XYZ conversion might be appropriate. Understanding the nature of various color spaces and the effect of white point selection is very important to producing images with the desired color fidelity. In many industrial and scientific applications, no consideration is given to color space issues because the narrow filters and the narrow display primaries restrict the range of accurate color representation sufficiently that simple white balance of the output signal is sufficient to realize the full potential of these systems. With the simplest color requirements, selecting the best color matrix for the F7 may not be essential but better color performance is available when it is needed.

Certain types of processing may require two color conversions. For example, if it is desired to sharpen the luminance of the signal without affecting color saturation, a first conversion should be made to XYZ space so the green channel can be sharpened and then a second conversion to sRGB space for display. White point correction can be done with a simple gain matrix in which only the diagonal elements are non-zero. These matrices are standardized and widely available for a variety of illuminants.

5.2.5 Common processing

Once the signal is in a display color space, any of the standard processing steps can be applied, as they would be with other color signals. These might include noise reduction by recursive or median filtering, aperture correction or other sharpening operations, gamma correction or low-bit encoding.

5.3 Monochrome imaging

Because devices with stacked photodiodes have co-located sensors for the entire range of detected wavelengths, the signals from the diodes (after proper first-stage corrections) can be summed to produce a monochrome image. The effect is that of having a very good broadband photodiode.

5.4 Image samples

Many of the effects mentioned in this paper can be seen in actual images. These do not reproduce well in halftoned images and especially do not survive conversion to black and white. Samples and links to additional images can be found online at <http://www.alt-vision.com/r/5074.htm>.

6. IMPLEMENTATION REQUIREMENTS

Building a camera to accommodate the F7 sensor is complex because of the flexibility in both the selection of analog operating points and the variety of potential scan modes. Tools are available to assist with this process. In addition, two measurements must be made for each sensor to assure highest image quality. These measurements correspond to the

sensor-based corrections – linearization and color conversion. These parameters can vary, although over fairly narrow ranges, from sensor to sensor. Both linearization tables and color matrices should be determined for each sensor to assure the most accurate color results. These procedures are straightforward but the details are beyond the scope of this paper. Other data that might be measured for correction as required by the application include blemishes, shading and fixed pattern noise. It should be noted that all of these parameters can be influenced to some degree by the sensor operating conditions.

7. APPLICATIONS

The original application for the F7 sensor anticipated by Foveon was for straightforward use in color digital still cameras intended for professional and consumer photography. However, the stacked photodiode architecture provides advantages in other potential uses. Some of these are closely related to photography and, to some extent, already involve the use of off-the shelf photographic-type cameras. Microscopy is an example. Stacked photodiodes should contribute to microscopy the same benefits of image clarity and color accuracy that are demonstrated in standard photography. Cooling should produce much the same benefits in these devices as it does in CCDs although there will remain a noise difference. Astronomy should see similar benefits. In fact, any use in which there is a benefit to using three image sensors for color work rather than one should ultimately realize the advantages of a single-sensor implementation using a stacked-photodiode device. These might include remote sensing, broadcast, endoscopy, cinema special effects and ophthalmology.

Other applications are more esoteric:

- Digital radiography, in which two phosphors sensitive to different x-ray energies are simultaneously detected to produce a dual-energy image in one exposure.
- Imaging colorimetry, in which a sensor and a switchable filter (or two sensors, filtered) produce six-channel images capable of very accurate color measurements using spectral estimation.
- Spectroscopy, in which the pixels are scanned and summed in strips to cover the band from the near UV to the near IR at high speed.
- Web inspection, in which the sensor can be operated in a line mode for production imaging and frame mode for system alignment and test.
- Tracking, in which a small raster on the sensor is moved rapidly to follow a moving object, supported by full-sensor binned images for fast location of new targets.
- Low-energy x-ray multiband imaging, in which the sensor directly detects x-rays in the 500-2000 eV range using the wavelength-dependent differential absorption properties of silicon to separate the energy ranges.
- Optics quality control, in which the wavelength-dependent aberrations of optical elements are evaluated with true color information.
- Multispectral imaging, in which a prism is used but where only one port is needed for color while the others can image near-infrared bands.

8. ADDITIONAL DEVELOPMENTS

As of this review, only one three-layer, stacked-photodiode color detector is in production, the Foveon F7 described above. Other devices have been built in small quantities and additional devices intended for commercialization are under development. Foveon has already announced that it will be introducing a device 1088 x 1440 array with 5 μ m pixels and microlenses and scan control modes that will allow 640 x 480 operation at 30 frames per second. In contact with a variety of potential users, the authors have received requests for many other configurations:

- Line scan imagers, to solve the current problems inherent in multi-sensor prism assemblies and trilinear color imagers.
- Much larger arrays, to bring the stacked photodiode technology to the highest end of professional photography – large-format film backs.
- Much smaller arrays, to provide better images for cellphones and all the other portable devices that absolutely, positively need cameras embedded.

- Quieter imagers, to really compete with CCDs in fluorescence microscopy and astronomy - eliminating some of the requirements for separate filters but preserving monochrome capability for very narrowband imaging.
- Different diode depth arrangements, to provide optimum color bands for applications where true color is not the object but where off-diagonal matrix elements are minimized.
- Four, five or six diodes stacked to provide a UV band and a couple of infrared bands and to drive the semiconductor process people completely nuts.
- Rad-hard imagers.

9. CONCLUSION

CMOS imager development is now in about 1976 in CCD terms. A few devices are out and a few organizations are starting to learn how to use them. There is a lot to learn but already the advantages are showing. The journey is going to be very interesting and all are welcome to come along.

10. REFERENCES

1. Color and Vision Research Laboratories, <http://cvrl.ucl.ac.uk/ciexyz64.txt>
2. Bruce E. Bayer, "Color imaging array," *U.S. Patent 3,971,065*, 1976.
3. Signa SD-9 Review, <http://www.dpreview.com/reviews/sigma-sd9/page23.asp>
4. Maya R. Gupta and Ting Chen, "Vector Color Filter Array Demosaicing", Proc. SPIE Vol. 4306, p. 374-382, 2001
5. Douglas R. Dykaar and Graham Luckhurst, "Chromatic Aberrations and Color Balancing Issues with Common Optical Axis CCD Cameras", Proc. SPIE Vol. 3965, p. 205-213, May 2000
6. Michael Stokes, Matthew Anderson, Srinivasan Chandrasekar, Ricardo Motta, "A Standard Default Color Space for the Internet - sRGB", Version 1.10, November 5, 1996, from <http://www.w3.org/Graphics/Color/sRGB.html>
7. Shoji Tominaga, Spectral imaging by a multichannel camera, *Journal of Electronic Imaging*, October 1999, Volume 8, Issue 4, pp. 332-341
8. Michael H. Jones, Stephen H. Jones, "Optical Properties of Silicon", Virginia Semiconductor, Inc., August 2002
9. Foveon, Inc., "Foveon X3 Pro 10M CMOS Image Sensor", January 2003
10. Allen Rush, "X3 - Three Color Sensors per Pixel", Internal Report, Foveon, Inc., 2002
11. Bob Shell, "A Detailed Look at Last", PEI, p. 24-26, Nov/Dec 2002
12. COBA 8th Meeting, October 9, 2002, notes at http://www.tow.com/photogallery/20021009_coba/
13. Uwe Steinmueller, http://www.outbackphoto.com/reviews/equipment/sigma_sd9/sigma_sd9_Diary.html

11. FURTHER INFORMATION

The Web abounds with information on imaging and color. Here is a very abbreviated list of interesting starting points for further study. Many of these sites have extensive reference lists of their own.

Principles of Color

1. With an emphasis on application to photography - <http://www.photo.net/photo/edscott/spectsel.htm#01>
2. Charles Poynton's Color FAQ - <http://www.poynton.com/ColorFAQ.html>

Color Spaces

3. Conversion applets in Java - http://www.cs.rit.edu/~ncs/color/a_spaces.html
4. International Color Consortium, about color profiles - <http://www.color.org/>

Image Sensors

5. Univ. of Edinburgh Vision Systems Bibliography Database: Solid State Imaging Entries - <http://oldeee.see.ed.ac.uk/~vision/biblio/solid/>

Foveon X3 Image Samples

6. DPReview.com, 2k by 2k prototype samples- <http://www.dpreview.com/news/0202/02021103foveonx3preview.asp>
7. Stephen Johnson, 2k by 2k samples - <http://www.sjphoto.com/web-special/index.htm>
8. Pbase, Sigma SD-9 user galleries - http://www.pbase.com/sigmasd9/user_home
9. Mark Fritz, Extreme no-color-aliasing example - <http://www.pbase.com/image/15293766>