Development Challenges of a New Image Capture Technology: Foveon X3 Image Sensors

Richard M. Turner and Rudolph J. Guttosch
Foveon, Inc.
Santa Clara, California, USA

1. Abstract

Relatively recently a new and different color capture technology, the Foveon X3 image sensor, has been developed. This technology competes with the Bayer CFA (Color Filter Array), the dominant color image capture technology in digital cameras of all types. The relative merits of the two color capture technologies are briefly discussed and a review of the pixel design, manufacturing, image processing, and system integration challenges involved in bringing this new image sensor technology to market is given.

2. Introduction

Image sensors are rapidly converging upon a relatively standard set of technologies for producing high quality and low cost devices for high volume markets. One rather different approach has been researched and commercialized at Foveon. This approach, called X3 Technology, is to leverage the natural color absorption properties of silicon and to stack pixels in a 3-dimensional array, resulting in a sensor design with fundamental advantages over the standard methods for making color image sensors.[1] Full color sensors have been investigated previously, although prior to X3 most methods have not been commercially or technologically viable.[2,3] The advantages of X3 Technology are many, and include the ability to make much smaller, and therefore less expensive, sensors for the demanding mobile device market, the elimination of artifact reducing blur filters, and the ability to provide fully sampled video through on-chip pixel aggregation rather than lower quality sub-sampled video that limits the video quality attainable with other sensors.

The advantages of the technology do, however, come at the price of some significant challenges. This paper describes the advantages of the technology as well as several of the challenges Foveon has faced in developing this technology. In particular, challenges in the areas of manufacturing, sensor design, image processing, and system integration will all be discussed. We explain that the challenges associated with X3 Technology are surmountable, and the benefits are very compelling.

Bayer Still Capture

The Bayer CFA (Color Filter Array) is made up of a repeating pattern of red, green, and blue filter material deposited on top of each pixel (Figure 1). These tiny filters enable what is normally a black-and-white sensor to create color images.

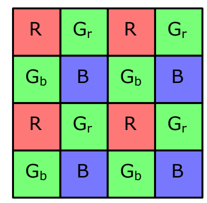

Figure 1: Typical Bayer CFA schematic showing the arrangement of red, green, and blue pixels.

By using two green-filtered pixels for every red and blue, the Bayer CFA is designed to maximize the perceived sharpness. However, since the green channel is under-sampled by 50%, and the red and blue channels are undersampled by 75%, the full luminance and chrominance detail delivered by the lens is not captured. Figure 2 shows a Bayer filter pattern decomposed into its constituent color components, illustrating the sparseness of the sampling. In order to create the required fully-populated color planes (i.e. three complete planes of red, green, and blue image information), the missing data in the Bayer CFA must be estimated.

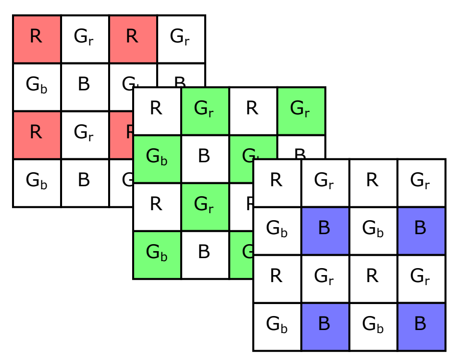

Figure 2: Decomposition of a Bayer CFA pattern into its components. Under-sampling in the image plane results in lower sharpness than could otherwise be achieved. Further, sampling gaps in the color planes lead to color Moiré artifacts.

In addition to the loss of detail, the relative phase shifts of the measured R, G, and B data in the spatial domain often result in color Moiré patterns. Subjects with repeating patterns (fabrics, building exteriors, and other fine, periodic detail) are known to create odd artifacts consisting of colors and frequencies not found in the original scene.

In order to overcome the system limitations imposed by the Bayer CFA, various schemes are employed. Optical Low Pass Filters (also known as blur filters) are used to reduce the effect of the relative phase shifts in the color sampling planes. Also, algorithms of increasing sophistication have been implemented to improve image sharpness and fidelity and to reduce color artifacts.[4]

Video Capture with Bayer CFA

Creating a video stream usually involves a combination of the use of on-chip binning registers, which aggregate pixels of the same color during readout, and decimation of sensor data to achieve the desired video stream resolution. Figures 3 and 4 illustrate two basic methods of binning and decimation. Depending on the readout speed of the sensor and the level of design complexity that can be tolerated, varying degrees of binning and decimation can be used to create a video stream of good quality. However, unless the sensor operates at full resolution with downstream digital binning, the CFA sensor will always sacrifice some performance due to the fundamental inability to simultaneously combine adjacent pixels in both dimensions and maintain accurate color information.

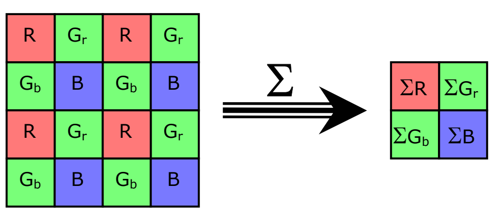

Figure 3: Bayer CFA pixel summing scheme taking 16 (4x4) pixels and binning to 4 (2x2) pixels suitable for video applications. Such a scheme is more suitable for Bayer CFA CMOS sensors.

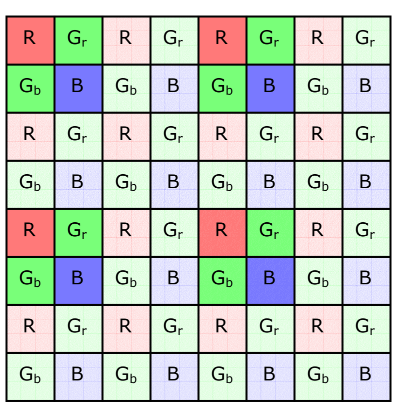

Figure 4: Alternate method for creating a video stream with a Bayer CFA sensor which employs skipping every other RGGB group of pixels

Bayer System Infrastructure

A wide variety of Bayer CFA sensors in a wide variety of sizes are available from a large number of manufacturers. There exists a plethora of image processing ICs for producing fully-processed RGB images from these sensors: off-the-shelf ASICs and DSPs are available from commercial sources, and a large number of proprietary image processing ASICs exist. Many low-cost CMOS image sensors include the image processing directly on the sensor chip. The existence of these image processing solutions ensures that a total digital image capture and processing solution for a Bayer image sensor can be created with relative ease.

Foveon X3 Technology

An alternative method for obtaining color images from a single chip monolithic imaging array is the Foveon X3 image sensor.[5,6] This unique sensor captures red, green, and blue light at each point in an image during a single exposure. To achieve this, Foveon sensors take advantage of the natural light absorbing characteristics of silicon. Light of different wavelengths is absorbed at different depths in the silicon -- blue photons are absorbed near the surface, green photons in the middle, and red photons are absorbed deeper in the material.[7] This mechanism results in spectral sensitivity curves (see Fig. 7) for the red, green, and blue channels that are broader than those found in typical CFA sensors, and as we shall see later, the combination of broad spectral curves and full color sampling at every point in the image plane also has advantages in the design of the sensor topology.

A schematic representation of the vertical arrangement of X3 pixels is shown in Figure 5. Figure 6 shows the color planes that result directly during image capture, without interpolation.

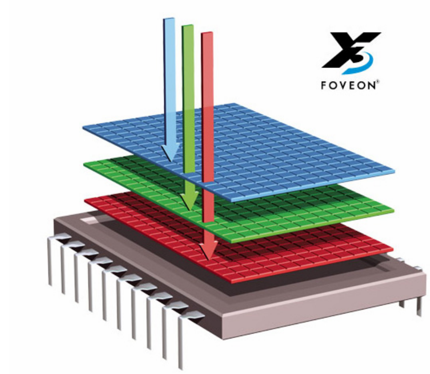

Figure 5: Schematic depiction of a Foveon X3 image sensor showing stacks of pixels, which record color channels depth-wise in silicon.

Note also that each pixel in the array is surrounded by pixels of the same color, a feature that greatly simplifies video readout and enables new sensor architectures not possible with Bayer CFAs.

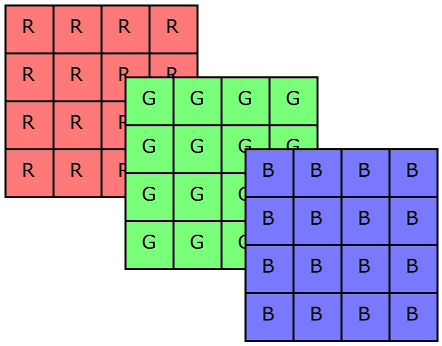

Figure 6: Fully populated image sampling found in film scanners, color-separation prism cameras, and also in Foveon X3 image sensors.

A major difference between most Bayer CFAs and Foveon X3 technology lies in the spectral sensitivity curves for the three color channels. As can be seen in Figure 7, the X3 spectral sensitivity curves are broader than those of typical Bayer CFAs. The broad character of the curves does not affect the ability to produce color–accurate images. The curves provide a set of basis functions which are quite suitable for producing images with high color accuracy and which reduce the likelihood of metamerism. Broad curves do require a color matrix with larger off-diagonal coefficients which can increase noise in the final image. Figure 8 gives sample matrices for X3 and Bayer sensors.

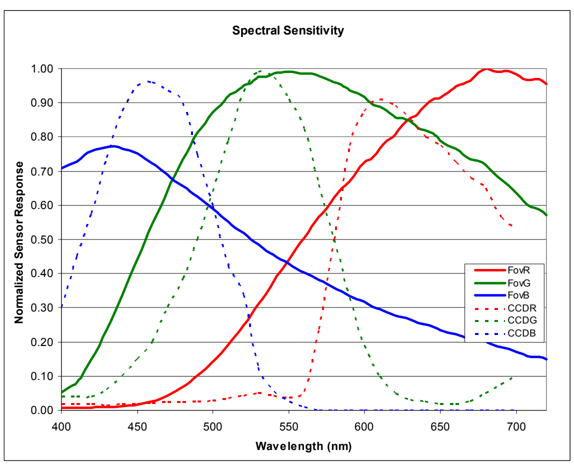

Figure 7: Spectral curves for a Foveon X3 sensor (solid lines) and a typical Bayer CFA (dsahed lines) used in consumer digital still cameras

$$\begin{bmatrix} 4.9 & -5.4 & 1.5 \\ -3.2 & 8.1 & -3.9 \\ 2.0 & -8.1 & 7.1 \end{bmatrix} \quad \begin{bmatrix} 1.4 & -0.9 & 0.7 \\ -0.9 & 2.3 & -0.2 \\ -0.5 & -1.5 & 3.3 \end{bmatrix}$$

Figure 8: Example Color matrices for an X3 sensor (left) and a Bayer CFA sensor (right) to transform images from camera color space to sRGB color space

Advantages of Stacked Pixels

Vertical stacking increases pixel density, thereby increasing the sharpness per unit silicon area in the resulting image. The stacks of red, green and blue pixels also eliminate the relative phase differences among the samples in color planes of a Bayer CFA, thus eliminating false color patterns. The combination of increased pixel density and the elimination of the need for a blur filter results in sharpness improvements in both luminance and chrominance that can be measured using industry-standard techniques as well as newer methods.[8, 9]

Further, since the image data is captured as three complete planes, the Bayer-to-fully-populated interpolation step can be skipped.

The color filter layers used in Bayer CFAs increase the thickness of the optical stack above the pixel. By eliminating the need for these layers, X3 image sensors can be made with a thinner optical stack, simplifying microlens design, an important feature as pixel pitches shrink.

The X3 architecture also enables a fundamental advantage in the way that video is generated. Since each pixel in an X3 pixel stack has neighbors of the same color, pixels can be directly combined on-chip to create images at video resolutions. For example, a Foveon X3 sensor with a pixel layout of 1280 x 960 x 3 pixels can be binned on-chip in 2x2 blocks to a VGA video stream of 640 x 480 x 3 pixels. During readout, each of these 2x2 blocks is read out of the sensor as a single value, thereby increasing the frame rate to video speeds, while maintaining the image quality advantages of a fully-sampled image plane.

In addition to enabling high quality video, the fact that each pixel is surrounded horizontally by pixels of the same color also creates some unique advantages in readout architecture. For example, the planes do not all have to be read out at the same resolution to achieve a desired image quality level. This topic is discussed in greater detail in the following section.

3. Foveon X3 Technology Challenges

Manufacturing

X3 sensors make use of many of the same manufacturing techniques that traditional sensors use, however the key differences relate to the implementation of color sensitivity. Traditional image sensors start as monochrome devices and are then converted to color sensing devices through a series of masking, deposition, and etching steps that involve the use of some type of organic or plastic color filter material. The exact number of steps varies from manufacturer to manufacturer, but in general a minimum of 3-5 steps will be required to achieve a 3-color device.

For X3, the color capability is built-in during the initial formation of the stacked pixels. The stacked pixels are created through a series of epitaxial growth and masking/ion implantation steps. The processes required to perform these steps are well developed, standard, semiconductor manufacturing techniques. The three photodiode layers are formed at precise depths and with precise doping concentrations that are designed to create the most optimal color response (which is directly proportional to detector depth) combined with the highest light collection and conversion efficiency.

One of the challenges is deciding the proper tradeoff between light collection, color separation, and red-channel MTF. As with all image sensors, the carriers absorbed most deeply in the semiconductor will have the greatest opportunity to diffuse to other pixels, and therefore the red channel has the potential to have lower MTF if the collection depth is not correctly controlled. Related to this, there is a direct tradeoff between color accuracy and light collection – collection of a large amount of red signal will increase the red SNR, but will also require that the color processing apply more aggressive correction to the raw sensor data to achieve accurate or desired color, resulting in a higher noise gain due to the transfer from sensor raw color space to an appropriate rendered image color space.

A second challenge associated with epitaxial deposition and photodiode construction is one of the unique aspects of X3. Unlike most other epi growth processes, the stacked pixel construction requires that new epitaxial layers be grown on top of layers that have been patterned. Maintaining alignment of high quality epi layers requires specialized techniques that have been developed for this task.

The color response and dark leakage currents are also sensitive to the epi quality and uniformity/thickness. Through careful control and epi reactor operating techniques, the thickness is maintained to within a few percent across a wafer, and the purity is kept to a level that results in no distinguishable leakage current difference between the grown epi layers and highly pure bulk silicon.

Once the fundamental photodiode depth and color responses are optimized, the next major manufacturing challenge is establishing a connection from the surface, where the active pixel circuitry resides, to the buried photodiode layers. Foveon employs a plug structure that relies upon highly conductive silicon connections to the buried photodiodes. The plugs require multiple high energy n-type implants, which are derived from the retrograde N-wells used in CMOS. Thick photoresist is capable of blocking the implants and still forming the openings used for the plugs. The deepest implants are those which form the plugs between the red and green photodiodes, and fortunately due to the pixel design geometries employed, the spacing requirements between the deep implants are less demanding.

Sensor Design and Architecture

In order to maximize the benefits of the X3 architecture, Foveon has explored and employed a number of distinct pixel designs. All of the pixel designs are targeted toward addressing one of the major issues with the stacked pixel technology: due to the parallel nature of the light collection and photo generation, as well as the increased density of data that an X3 image sensor provides, there is a requirement for more readout circuitry per-unit silicon area than for a Bayer CFA traditional sensor. Traditional sensors generate one data value per pixel location because there is only a single pixel sensor in that location, whereas an X3 image sensor generates 3 distinct data values per pixel location due to the stack of 3 pixel sensors.

The initial pixel design employed by Foveon was a 4T per color design based on Foveon's initial electronic shutter sensor design.[10] The drawbacks with this approach related to image lag and threshold voltage variations and nonlinearities in the electronic shutter switch. More modern 4T designs which employ a CCD-like depleted photodiode and readout structure can eliminate these drawbacks, but at that time the X3 process was not merged with a depleted photodiode design.

Foveon's two prior SLR image sensors, the F7 and F7N, were both 10.2 megapixel (2304x1536x3) sensors that were based on a standard 3T readout structure. Due to the extra circuitry required to readout the 3 pixels in parallel, the design made use of column lines as both an output channel as well as a reset input channel for each color. The advantages of this design are high fill factor at more than 60% without a micro lens as well as efficient use of metal layers, which is critical to achieve high yield for large sized image sensors.

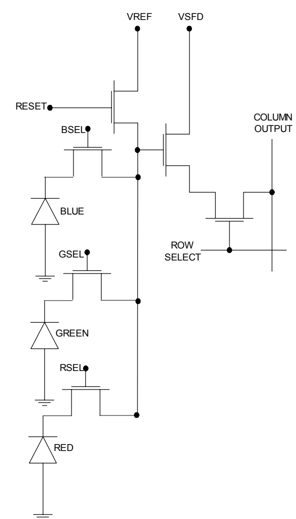

Figure 9: Serial pixel schematic. Separate pixels are read out sequentially, sharing amplifier and row select transistors.

Following the color parallel approach employed in the SLR sensors, Foveon next adopted a color serial design which is better suited to the smaller pixels required for compact format image sensors. In this design, which was produced for the compact DSC market and is shown in Figure 9, the three color pixel sensors all share the same readout amplifier and column circuitry. The benefit to this is the reduction in transistor count per array location; however the penalty comes in the form of a relatively high capacitance of the sense node combined with a slower frame readout that delivers pixels in a line-sequential color format. The data format is problematic because most downstream processing requires RGB pixel data. To address the speed issue, the sensor was designed to achieve a higher than usual pixel rate for a sensor of its class, and a high speed pipeline A/D converter was employed.

Although the color serial approach accommodates pixel sizes down to 3.3 microns (or smaller), Foveon has also investigated a different readout mode which can accommodate even smaller pixels without sacrificing performance. This readout mode is intertwined with the image processing architecture, and is termed Y-RGB readout (see Figure 10). This readout leverages one of the aspects of X3 that is usually considered to be a hindrance: the broad spectral responses of the color channels. Because each channel in the sensor has a broad spectral response,

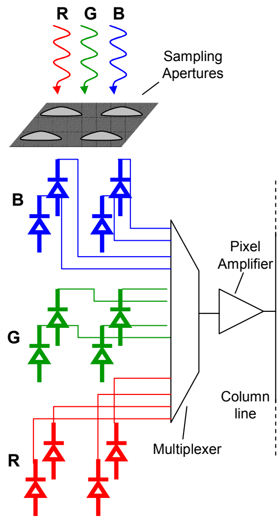

Figure 10: Conceptual schematic representation of a pixel design suitable for the Foveon Y-RGB image processing pipeline.

it is possible to extract complete sharpness or "luminance" data from the green channel. Using this natural property of X3, the Y-RGB readout reduces the quantity of data that must be handled by combining the output signals from two of the three color channels (R,B) in the multiplexer shown in Figure 10, or alternatively downstream in the digital domain. The G pixel layer is then used as both a sharpness channel and a color channel. In this way, the resulting dataset can effectively be reduced to a single high resolution sharpness channel and three low resolution chroma channels. Depending on whether or not the chroma downsampling is done in the analog or digital domains, there are potential power and readout bandwidth benefits, signal-to-noise advantages, as well as the possibility for increased fill factor in the pixel design through sharing of readout and amplifier resources between the channels. Because the chroma signal aggregation/downsampling occurs after the optical sampling and the green channel covers the visible spectrum, there is no color aliasing and no visible loss in resolution, as shown in Figure 11. Although this readout scheme requires a different approach to the image processing (discussed in the following section), it turns out to be a very natural way to operate the sensor, and is consistent with the way most digital video systems operate on data. In essence, the nearly universally required steps of changing to a luminance-chrominance color space and downsampling the color data are performed early in the processing, as opposed to downstream after most of the computationally complex processing has been done, as is usually the case.

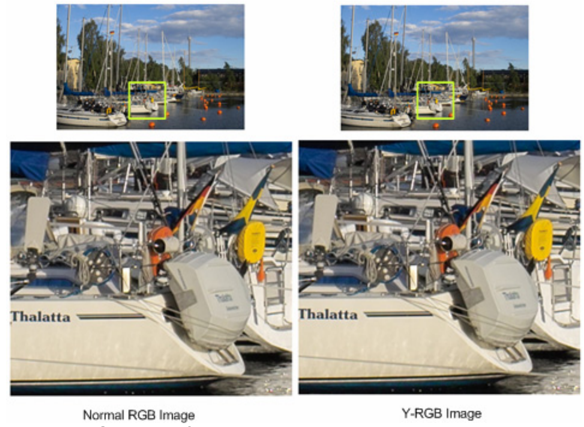

Figure 11: Samples of images processed using standard pipeline and Foveon Y-RGB. The images show equivalent sharpness.

Image Processing

The development of the pixel architectures creates an opportunity to build image processing pipelines specifically tailored to the sensor output. Two processing pipelines have been created, one for operating on color-parallel and color-serial data, and the other for operating on Y-RGB data.[11] As far as the specific steps involved, the pipelines are similar, the main difference being that the Y-RGB data uses a split pipe, one for the low resolution chroma data, and one for the high resolution sharpness data (see Figure 12). The split pipeline offers the advantage of handling the more processing-intensive color data at ¼ the resolution of the sharpness data, cutting processing resources proportionally.

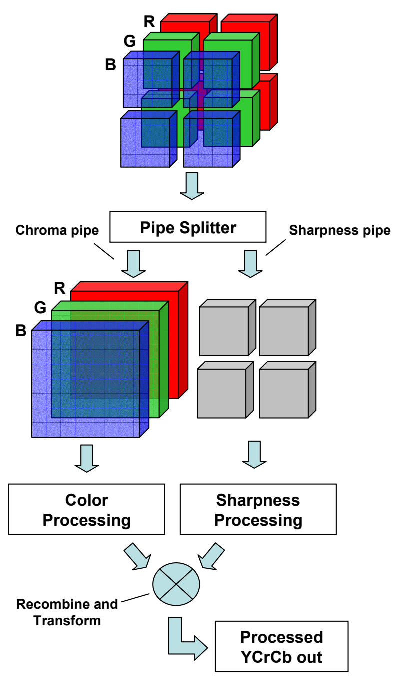

Figure 12: Split image pipeline data concept. Chroma data used for color processing is ¼ resolution compared to the panchromatic sharpness data.

As noted earlier, a wide variety of image processing options exist for Bayer sensors. The degree to which these ICs can be used for an X3 application has a lot to do with the flexibility and programmability of the individual processing blocks. First among these is simply the ability to read out an X3 sensor with the pixels being assembled in the correct order. Other limitations encountered in the image processing section of Bayer CFA support ICs include hardwired demosaicking blocks with no opportunity to bypass, limits on the size of the color matrix coefficients, and 10-bit internal pipelines (12-bits are typically used with X3 for best results).

With these HW-based limitations in place, initial image processing development at Foveon focused on implementing an X3-specific pipeline in software. In particular, the work concentrated on 3 areas: sensor linearization, noise reduction, and highlight neutralization. The first area, linearization, uses look-up tables or polynomials to ensure that the sensor response is optimized to account for saturation effects and to increase the dynamic range of the device. The linearization LUTs can also be used to improve low ISO performance by ensuring that saturated pixels are rendered correctly.

Advanced noise reduction algorithms were also developed in order to reduce the effects of the color matrix noise gain. Noise reduction is generally performed in camera color space, or similar wide gamut space on 12-bit data. Most of the noise reduction functions are non-linear neighborhood operators and are suitable for use in either image processing pipeline. In addition to the neighborhood operators, there is also a function for eliminating single pixel outliers. The development of these algorithms, which included coring, blurring, and localized color saturation adjustment functions, continues to pay dividends as pixels shrink in size.

Another image characteristic that must be dealt with for both Bayer and Foveon X3 sensors is the rendering of saturated highlights. Without extra processing, specular highlights can show unnatural color. In the X3 case, this effect is magnified by the color matrix. To control this effect, highlight and highlight-neighbor neutralization algorithms were also developed, which use the computed scene white balance to neutralize affected pixels.

As noted above, software-based processing was initially developed, but hardware-based solutions are now coming on-line.

System Integration

Among the system integration issues that a camera designer takes into account when planning an X3 project are sensor readout, memory requirements and architecture, and specification of the lens and IR filter.

The majority of image processing ICs only support Bayer pixel output modes, i.e., they require an (RGRG…, GBGB…)-type pixel stream coming into the IC. While the Foveon X3 sensor can support this output mode, [12] it sacrifices one of the key benefits of X3 and is therefore not an option when the best image quality is required. Image data is delivered either as one color line at a time (RRR…, GGG…, BBB…,), or as RGB triplets, which requires some buffering and a slightly different memory management scheme before being passed to the image processing section of the IC.

As far as optical issues are concerned, as noted earlier there is an advantage in the optical stack since there is no color filter layer in X3 sensors. The lens should be chosen such that the MTF is optimized for the pixel size and an IR filter must be placed in the optical path. In addition to blocking IR, the filter used for X3 sensors must also block UV because the X3 sensor has a significant UV response. In addition, the cut filter also helps shape the red and blue color responses for optimizing the tradeoff between color matrix noise gains and desirable/accurate color reproduction. There is usually some optimization work that is done to match the microlens design to the requirements of the image forming lens and this is especially relevant in the case of high chief ray angles often associated with compact optical systems. The corrections applied with X3 sensors are not significantly different than those used with Bayer CFA sensors.

One item that can be removed from the optical path is the blur filter, reducing cost. As discussed earlier, blur filters are not typically necessary for X3 sensors, since X3 sensors have an inherently higher sampling rate for all three color channels in the image plane. X3 sensors will, of course, produce some spatial aliasing artifacts if presented with frequencies above the Nyquist frequency corresponding to the pixel pitch, however the aliasing will not contain false color, and is therefore not objectionable for most natural subjects.

4. Conclusion

Foveon's unique approach to image sensors has created a new way to capture digital images. The advantages of capturing all three colors with three pixels at each location on the sensor array come in the form of superior image quality and sharpness, as well as in the form of a cost savings due to the ability to use less silicon area to achieve the required performance for a given imaging application. The inherent advantages have compelled others to continue to investigate the path to multi-layer sensors.[13]

The manufacturing, design, image processing, and integration challenges that have been discussed have been addressed through a combination of image sensor design and a system solutions approach, resulting in products that offer high performance at a competitive cost.

While many of the challenges associated with bringing this new technology from the R&D lab to consumer products have been identified and addressed, there remain challenges. Pixel geometries are shrinking, noise and leakage requirements have become more aggressive, and the image sensor market has exploded, generating stiff price and performance competition. In order to continue to offer a compelling solution, X3 technology must continue to evolve with the market and the technologies that support the digital imaging industry. The natural behavior of silicon that Foveon has leveraged to create the stacked pixel image sensor scales to meet the new demands, and therefore X3 technology will continue to advance to meet those demands.

5. Author Biographies

Richard Turner received BS and MS degrees in Optics from the University of Rochester, and a PhD in Electrical Engineering from the University of Colorado. Prior to joining Foveon, Dr. Turner designed CCD image sensors at EG&G Reticon, and CMOS image sensors at National Semiconductor.

Rudy Guttosch directs the marketing applications effort Foveon, Inc. Prior to his current position, he held positions in sales and product management at Swedish camera maker, Hasselblad. He received BS and MS degrees in Imaging Science from the Rochester Institute of Technology.

6. References

[1] Merrill, R. B., "Vertical Color Filter Detector Group and Array", US Patent 6,727,521, 2004

[2] Carr, W. N., "Multi-spectrum photodiode devices," US Patent 4,238,760, 1980.

[3] Chouikha, M., "Buried Triple p-n Junction Structure in a BiCMOS Technology for Color Detection," IEEE BCTM 6.4, pp. 108-111, 1997.

[4] Ramanath, et al., Journal of Electronic Imaging 11(3), 306-315 (July 2002)

[5] R. B. Merrill, "Color Separation in an Active Pixel Cell Imaging Array Using a Triple-Well-Structure," US Patent 5,965,875, 1999.

[6] R. F. Lyon and P. M. Hubel, "Eyeing the Camera: into the Next Century", Tenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications; Scottsdale, Arizona; November, 2002.

[7] A. J. P. Theuwissen, Solid-State Imaging with Charged-Coupled Devices, Kluwer Academic Press, Dordrecht, 1995.

[8] P. M. Hubel, J. Liu, and R. J. Guttosch, Spatial Frequency Response of Color Image Sensors: Bayer Color Filters and Foveon X3; Proc. SPIE Vol. 5301, p. 402-407, 2004.

[9] P. M. Hubel, M. Bautsch, "Resolution for Color Photography," Electronic Imaging: Digital Photography II, SPIE proc. 6069, 2006.

[10] Merrill, R. B., Turner, R. M.; Mead, C. A.; Lyon, R. F.; "Intra-Pixel Frame Storage Element, Array, and Electronic Shutter Method Suitable for Electronic Still Camera Applications"; US Patent 6,369,853, 2002.

[11] P. M. Hubel, "Foveon Technology and the Changing Landscape of Digital Camera Technology," keynote paper at the 13th IS&T Color Imaging Conference, Scottsdale, AZ, pp. 314-317, 2005.

[12] Merrill, R. B., Turner, R. M.; Guttosch, R. G.; Lyon, R. F.; "Vertical Color Filter Detector Group and Array"; US Patent 6,864,557, 2005.

[13] Steibig, H., Street, R. A., Knipp, D., Krause, M., Ho, J.; "Vertically integrated thin-film color sensor arrays for advanced sensing applications"; Appl. Phys. Lett., 88 (2006).