X3 Technology –

3 Color Sensors per Pixel

Allen Rush, Foveon, Inc.

Overview

Conventional image sensors incorporate a single layer of photodetectors covered by a color filter arrangement, for example the RGB Bayer pattern or the CMYG complementary color pattern. The result is that conventional image sensors report only one color value per pixel. Foveon has developed the first commercial image sensor that reports three distinct colors within every pixel. This novel approach to image capture, based on X3™ technology, eliminates many of the tradeoffs and limitations that have existed for single-layer image sensors.

This paper presents the basic principles of X3 technology, along with initial data that demonstrates the design effectiveness for image capture applications of interest – digital cameras, digital video cameras, etc. In addition, the paper introduces new features and functions that are enabled by the X3 pixel architecture.

In an image acquisition system, the image sensor provides spatially based light information to the back-end image processing section of the system. X3 technology is designed to overcome the key limitations of single-layer sensors, which provide only one color value for each spatial location. Rendered images in standard representations (including TIFF, sRGB, and others) incorporate three or more color values per spatial location. Accordingly, raw image data reported by the conventional single-layer sensor needs to be converted or rendered in to the final form through a series of calculations that typically uses information from neighboring spatial locations of different colors to arrive at a final full-color output data set.

The enabling physical characteristic for X3 sensor operation is the wavelength absorption response of visible light (and IR) in silicon. Photocurrent is generated at various depths in silicon as a function of the wavelength of the incident light. By placing photodetectors at corresponding depths in the silicon, it is possible to distinguish color bands within the pixel structure and directly report three distinct color values for each spatial location.

Motivation

Many factors are involved with determining image quality and system capabilities for image capture devices. These include noise, pixel size, fill factor, system computation, quantum efficiency, etc. In a color filter array, the colors are distributed spatially on the array. A conventional single-layer image sensor samples only 25% of the red, 25% of the blue, and 50% of the green colors. Since a completed image represents all color components (usually in RGB form, although other color spaces are possible), the image acquisition system attempts to calculate the missing color components from the spatially sampled values. There are many approaches to estimating missing colors. Some have been exhaustively studied for accuracy and analysis, such as methods presented by Ramanath et al (1). Most demosaicing algorithms use a combination of filtering and neighboring color heuristics to determine a best approximation for the missing colors. More computationally intensive approaches incorporate adaptive, gradient, and other image-information-based data-driven algorithms. Errors that occur can be noticed; the extent of the error can be correlated by measuring the ΔE^*_{ab} of the resulting missing color calculation. In a test of five different algorithms and eight different test images, Ramanath reported a ΔE^*_{ab} range of 0 (no noticeable difference) to a worst case of over 65 (a ΔE^*_{ab} of 2-3 is usually considered to create a just noticeable difference, JND). Moreover, all of the approaches failed to adequately estimate images exhibiting high spatial frequencies.

Errors created by approximating missing color have various appearances, depending on the type of error. Moreover, some errors, such as moiré, are minimized by optically low-pass filtering the image, thereby removing the potential for aliasing. In this case, the resulting capture of the image has correspondingly less resolution than the theoretical Nyquist limit. Further attempts to restore the appearance of a sharper image are accomplished by sharpening filters, such as unsharp mask. This processing adds further errors to the result, although in many cases the sharpening has a beneficial perceptual effect. The net result is a series of cause – effect – remedy that adds considerable complexity and cost to the capture system, while the best quality remains unachievable.

X3 Goals

The solution to the motivation for considering an alternative to single color filter pixels needed to satisfy certain technical, practical, and business objectives. Above all, a solution needs to achieve a minimum technical performance that allows it to be considered as a viable alternative to existing CFA based solutions. In digital still camera applications, this most often implies CCD based image sensors. The basic goals of X3 are:

1. Capture all of the available visible light in a single pixel, or spatial sampling site;
2. Capture three different color wavelength bands;
3. Sufficiently separate or distinguish the color bands to achieve high color accuracy;
4. Capture image information with low noise in each detector;
5. Achieve minimum cross talk, both vertically (intrapixel) and horizontally (interpixel);
6. Design to take advantage of intrinsic CMOS image sensor capabilities: random access, region of interest, crop, pan, etc.
7. Design for standard CMOS processing equipment and methods;
8. Design for low manufacturing cost;
9. Design for scalability – in pixel size and array configuration

As prototypes and working systems evolved, it also became clear that other goals would emerge, mainly relating to taking advantage of flexibility in pixel readout, ability to combine pixel values, etc. These derived goals served to expand the potential application space, suggesting methods for utilizing the capture system in ways previously not possible with traditional system design and technology.

X3 Principles of Operation

An X3 pixel contains 3 photodetectors, located at different depths in the silicon. Shorter wavelengths (blue) are absorbed near the surface, medium wavelengths are absorbed further down, and long wavelengths (red and IR) are absorbed deep in the silicon. In general, the X3 detectors operate in a similar way to single color conventional CMOS image sensors – a photo diode with a co-located pixel amplifier, classically referred to as APS, or active pixel sensor. The major difference is in the fact that three values are reported, or read, at each pixel location. Figure 1 shows the basic structure of an X3 pixel, in abstract form, and Figure 2 shows an example, for comparison, of a CMOS pixel with a color filter array (CFA).

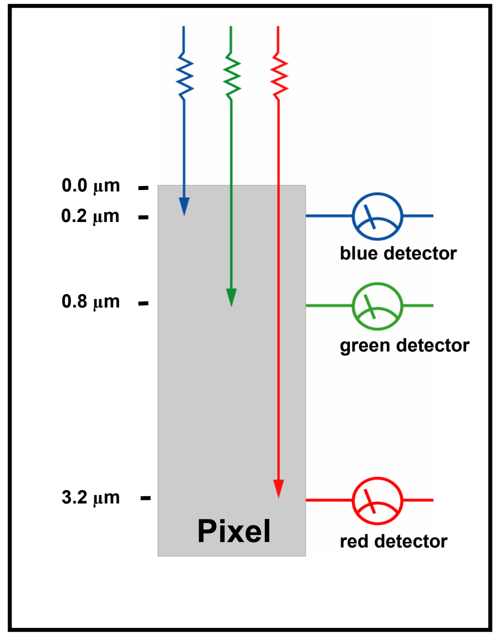

Figure 1 X3 Basic Structure

The X3 technology is compatible with 0.18um CMOS processing methods, and is fabricated using the National Semiconductor fab in Portland, Maine. All of the aspects of the standard 0.18um technology are incorporated to take advantage of yield, process stability, design rules, etc.

A major goal for the pixel performance is to determine the color response for each channel. Since there are three photodiodes, all of the available light in terms of wavelength is available for conversion. The QE, quantum efficiency, of the pixel is best approximated by the sum of the QE's for each color channel. A typical QE curve for an X3 sensor is shown in Figure 3, and a typical interline CCD QE curve is shown in Figure 4. Aside from the difference in relative gain from one channel to another, the most notable difference between the two responses is in the bandwidth of the color filters. For CFA's the bandwidth is narrow, more distinctly defining each of the color bands. For X3, the response is broader. When calculating color matching functions, this will result in a color correction matrix that has high off-axis correction terms. A typical color correction matrix is shown in Figure 5.

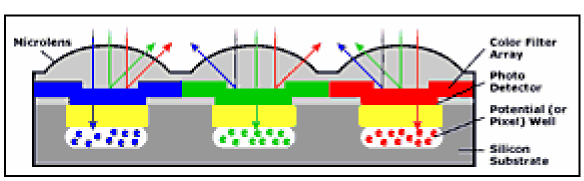

Figure 2 CFA Structure

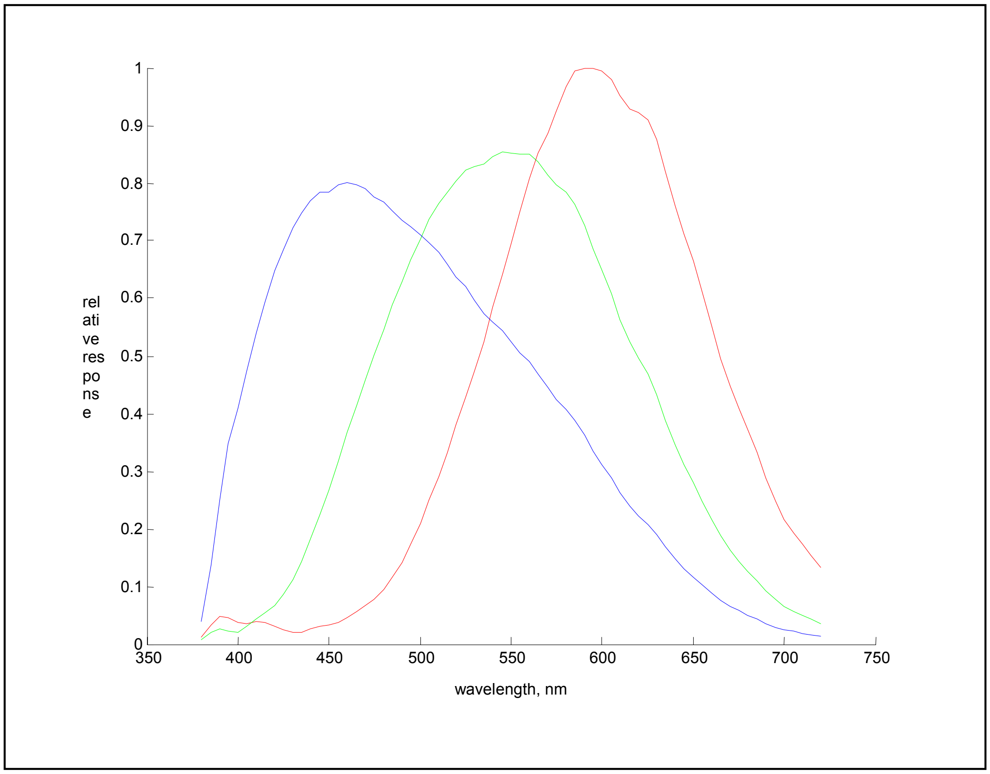

Figure 3 Typical QE Response for X3 Pixel

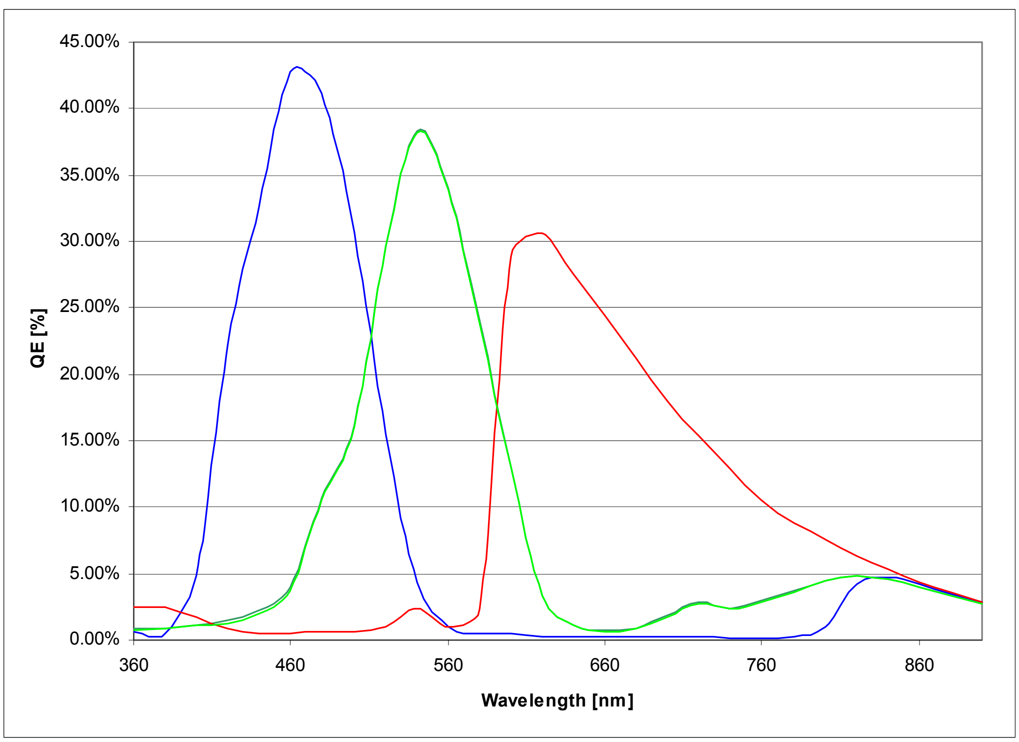

Figure 4 Typical CFA QE
(Source: Baer, IEEE CCD and AIS Workshop, 1999)

0.381	0.597	-0.03
-1.18	3.472	-1.3
0.887	-3.91	4.112

Figure 5 Sample X3 Color Matrix

One of the goals of X3 was to sufficiently separate or distinguish the color bands to achieve high color accuracy. While the color separation contains a relatively high level of overlap from one band to another, it maps to the CIE Standard Observer Curves XYZ with a high degree of accuracy. Furthermore, the color spectral response meets the Luther-Ives condition which requires that the matrix solution is a nonsingular transformation of the CIE color matching functions (3). Figure 6 shows a response of the X3 colors when mapped to a CIE Standard Observer.

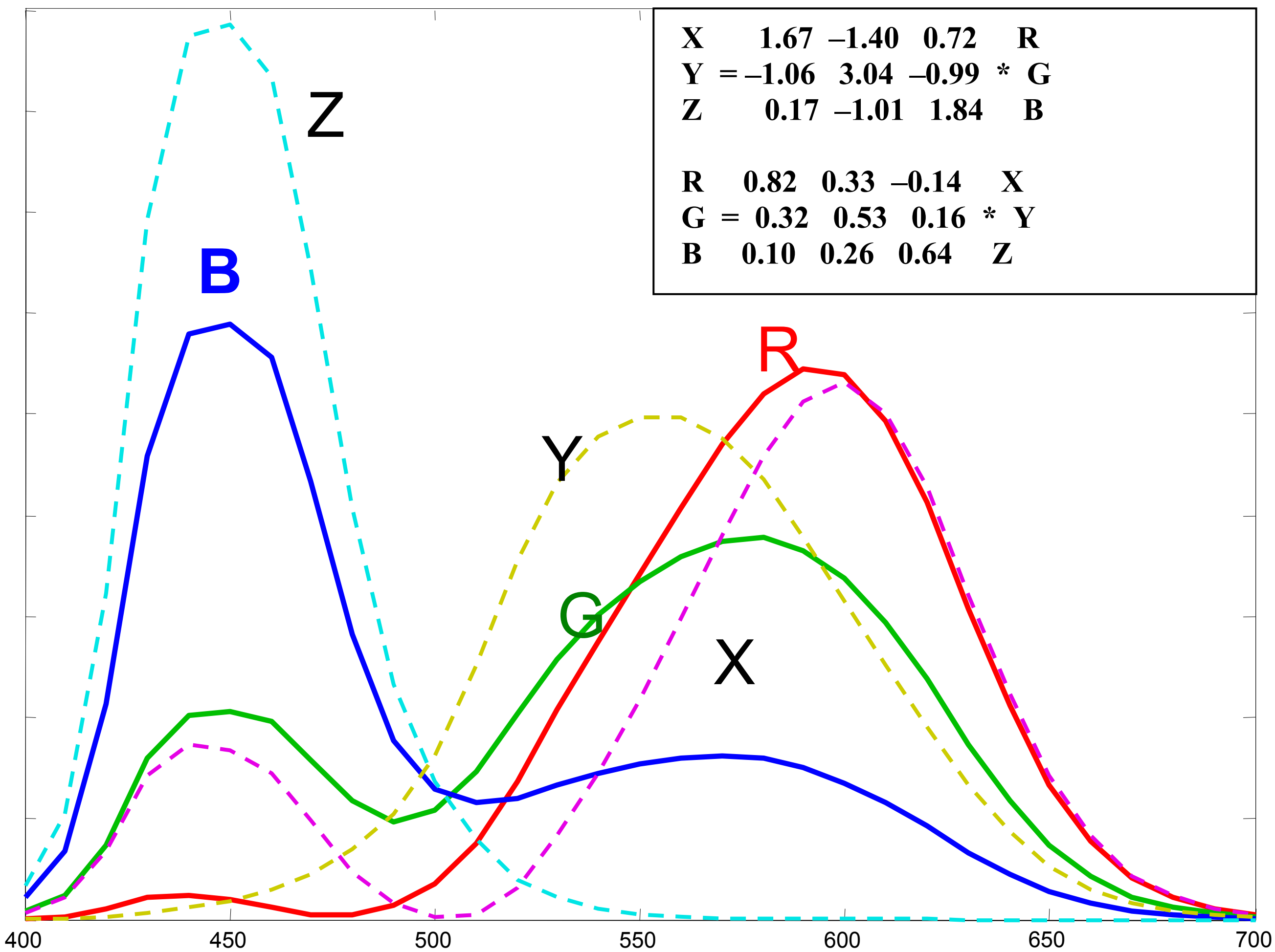

Figure 6 RGB mapping for CIE Standard Observer

Each of the three photodiodes detects and collects light and acts as an independent photodetector. After a reset, the photodiodes accumulate electrons during an exposure. The initial X3 devices are frame readout types; there is no electronic shutter mechanism. During readout each pixel reports three values, roughly corresponding to red, green, and blue channels. The readout is done in parallel, such that the three values are reported as separate and simultaneous outputs at the device pin out boundary. In devices that contain an integrated A/D converter, the analog outputs may be connected to either one or three A/D converters.

The readout speed for a basic X3 pixel is similar to other image sensors – typically 12-24 Mhz. The tradeoffs for readout speed are essentially the same as they are of other devices – faster readout usually generates more noise.

The pixel performance is defined in isolation by its response to wavelengths of light and corresponding noise that is generated by various sources within the pixel. It can be characterized by many of the same methods as other pixel use: pixel pitch, fill factor, well capacity, noise floor, etc. One key difference is seen by the total number of electrons collected – three times as much light converted per pixel area. The fill factor is kept to a reasonable level, typically in excess of 50%, by taking advantage of 0.18um CMOS design rules.

At the array level, surrounding pixels and readout circuitry also influence the performance. Traditional measures, such as photo-response non-uniformity (PRNU) need to be expanded to include additional measures of color channel cross talk and color dependent noise influence. For example, a CFA array will experience cross talk from the red channel to neighboring pixels, as red wavelengths are converted deep in the silicon. The result is color cross talk, or color noise, shown in Figure 7. In an X3 array, the red detectors are deep, so that cross talk is minimized. The cross talk that does occur will mostly be collected in neighboring red detectors, resulting in a loss of red MTF.

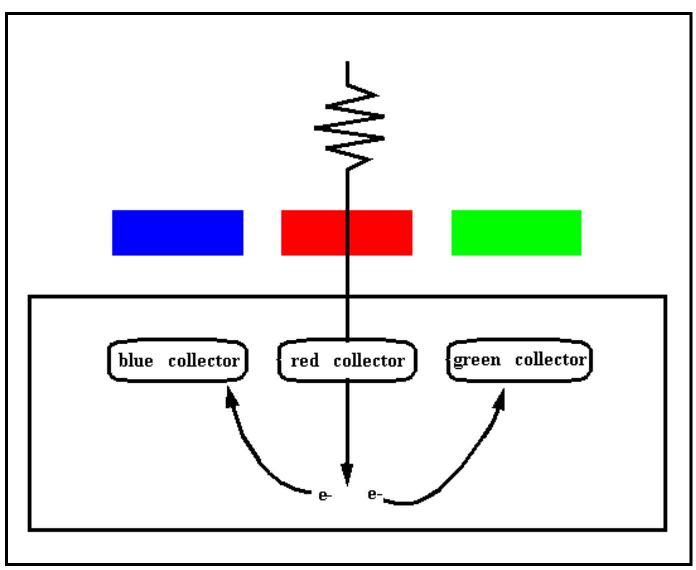

Figure 7 Red Cross Talk in CFA Sensor

Image Quality Results

Significant testing has been done to compare the results from X3 and color filter array type sensors. In all of the tests, attempts were made to keep the tests fair, including pixel sizes, resolution, and lighting conditions. An example of the difference between an image captured by a CFA sensor and one captured by an X3 sensor is shown in Figure 8.

Figure 8 Image Quality Difference Between CFA and X3 – Color Accuracy

As seen in the examples, errors in calculating the missing pixel values show up as color aliasing in high contrast, high frequency image content.

X3 Readout Options

The X3 pixel and array design can provide additional function and benefit when different readout methods are used to extract data from the sensor. These techniques include VPS (Variable Pixel Size), crop, pan, and randomly defined regions of interest.

VPS is defined as changing the effective, or reported, pixel resolution of the array by altering the virtual size of the pixel. In certain CCD's a limited version of this is achieved through binning, or combining of charges from neighboring pixels. With CFA's this has limitations due to the inability to combine charges from pixels of different colors. Alternately, complex system computation is used to decimate the sampled data to reduce resolution.

One early example of multiresolution capability was reported by JPL (2). It used banks of capacitors to store samples of adjacent pixels. By combining and averaging, the resolution can be reduced by 2x2, 3x3, 4x4, etc. The monochrome image results clearly showed the benefits of averaging neighboring pixels and more correctly low pass sample the image.

VPS provides for pixels to be combined, color separated, resulting in virtual pixel sizes ranging from 2x2 to any combination of mxn. This combination is accomplished in the analog domain, and results in an effective noise reduction of approximately sqrt(mxn). Also, the frame readout speed is increased by approximately the factor of resolution reduction.

VPS control is managed through the digital control block of the device. The setup and control for different VPS configurations can be changed on a frame boundary. If the primary mode of operation is low resolution video, then a still frame can be configured on a frame boundary, and a full resolution image can be read out instead of a normal video frame. A full frame readout normally takes longer- sometimes several frame times, so a few video frames may be dropped in exchange for the one full still image readout. A missing frame estimation algorithm can be used to fill in the missing frame data. Figure 9 shows a possible video frame and still capture scenario.

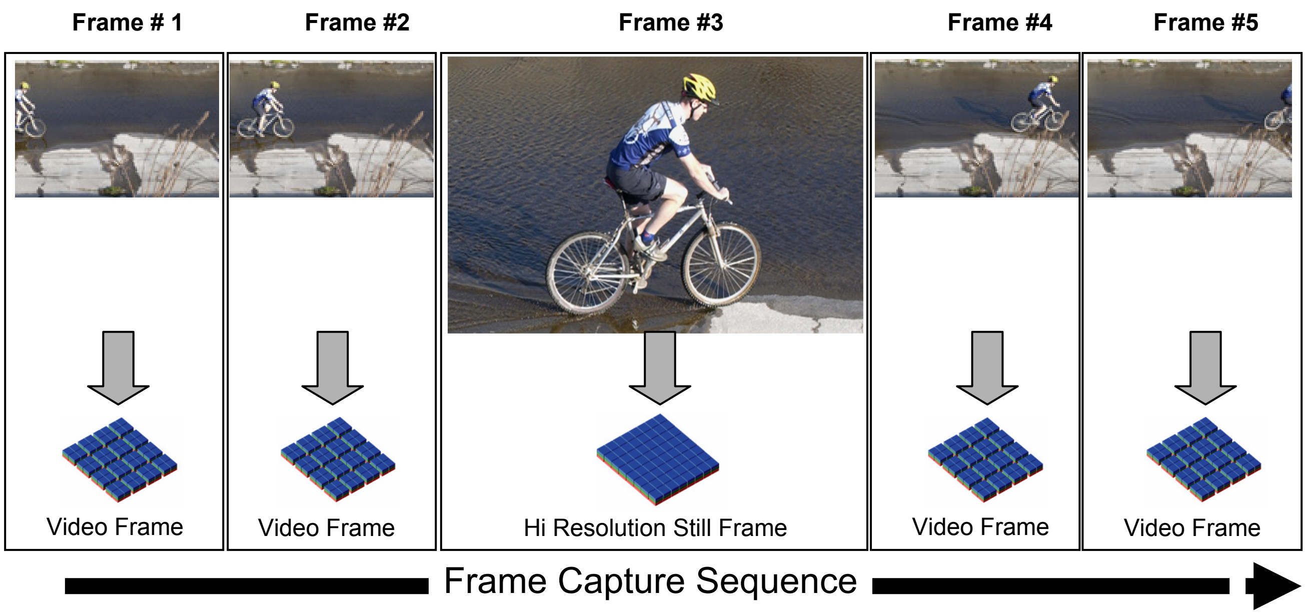

Figure 9 Video Frame Sequence with Full Resolution Still Capture

Optical Considerations

X3 pixel arrays have been tested with and without microlens. For small pixels with limited fill factors, the addition of the microlens boosts the effective fill factor and sensitivity. The gain of the microlens depends on the starting point or initial fill factor. The results that have been measured show a similar gain among the three channels, and a similar roll-off as a function of angle of incidence.

For angle of incidence, the X3 response is similar to a frame transfer CCD or interline transfer CCD in the vertical direction. In the horizontal direction, the interline transfer CCD falls off rapidly due to the asymmetric construction of the IT pixel. (5). The response is very acceptable out to angles between 15 and 20 degrees. This will support optical designs beyond f/2.8, which requires a minimum angle response of 10 degrees.

One unique characteristic of the X3 sensor is its immunity to chromatic optical aberration. In the case of lateral color shift, the sensor can easily detect an axial shift of one or more pixels, and the corresponding correction can be applied. In high-end lens design this is not a noticeable problem, but lower cost optics may have this type of aberration, and the appropriate corrections can be embedded into the image processing chain.

Conclusion

X3 is a new technology for capturing scene image data with three separate color bands and associated detectors. A successful color separation and filtering is achieved by taking advantage of the depth dependent photon current generation of different wavelengths of light. The motivation for this type of sensor is a combination of reduced cost and complexity as well as a higher level of image signal fidelity during capture. Test arrays and first production versions of the X3 technology have been designed and tested. The results show a predictable improvement in overall sharpness as well as color accuracy.

A very flexible readout method is enabled with the X3 technology, called Variable Pixel Size, or VPS. The virtual size of a pixel can be created by combining neighboring pixels of each color separately, resulting in larger pixels and lower array resolution. This technique is useful for creating multiple functions within the same array – high quality still image capture as well as high sensitivity video capture.

The X3 technology is highly scalable, and is fabricated in a standard 0.18um CMOS facility.

Bibliography

1. Ramanath et al, "*Demosaicing Methods for Color Filter Arrays*", Journal for Electronic Imaging, July 2001.

2. Kemeny et al, "Multiresolution Image Sensor", in *IEEE Transactions on Circuits and Systems for Video Technology*, Aug 1997, pp 575-581

3. Sharma et al, "*Color Imaging for Multimedia*", Proceedings of the IEEE, June 1998

4. Vora, P.L. & Trussel, H.J., "*Mathematical Methods for the Analysis of Color Scanning Filters*", IEEE Transactions on Image Processing, Feb. 1997.

5. Baer, R. L., "*CCD Requirements for Digital Photography*", IS&T 2000 PICS Conference

Backup Material