

Canon

WHITE PAPER

EOS C300 / EOS C300 PL

RGB RESOLUTION CONSIDERATIONS IN A NEW CMOS SENSOR FOR CINE MOTION IMAGING

Created December 19, 2011



Written by **Larry Thorpe**
Professional Engineering & Solutions Division, Canon U.S.A., Inc.

For more info:
cinemaeos.usa.canon.com

© 2012 Canon USA, Inc. All rights reserved.

CINEMA EOS

RGB Resolution Considerations in a New CMOS Sensor for Cine Motion Imaging

© 2011 Canon USA, Inc. All rights reserved

Abstract

Large format single-sensor digital cine cameras have emerged as an important worldwide alternative to motion picture film origination. Most of the image sensors used in these cameras are based upon the established Bayer color filter array. Important advances in demosaicking algorithms have steadily improved the reconstruction of RGB video components that support the color correction and image manipulation required in postproduction. This paper will discuss an important new CMOS image sensor employing the classic Bayer color filter array that is specifically intended to deliver full equiband 4:4:4 HDTV RGB component video directly from the sensor — each at a 1920 (H) x 1080 (V) digital sampling structure. No demosaicking algorithm of any form is required. The three components created within the image sensor itself can be selected at picture capture rates of 60P, 50P, 30P, 25P, 24P, or 23.98P. The 1080-line 60i and 50i interlaced video formats are also created in the image sensor. Conversion to the 720-line @60P format can be made subsequently external to the CMOS chip. The paper will outline the structuring of the separate RGB digital video components and the associated MTF characteristics for both the progressive and the interlaced video formats.

1.0 Introduction

Development activities in large-format single-sensor motion imaging sensors began almost a decade ago. They encompass both CCD and CMOS technologies. Many use the classic Bayer color filter array while a few employ RGB color stripes. Global developments continue on increasingly sophisticated debayering algorithms to separate the RGB video components following the optoelectronic transformation and the photosite readout processes. Canon has developed a totally new CMOS image sensor with an active image size based upon the 3-Perf 35 mm motion picture film format — one that is specifically intended for digital cine motion imaging.

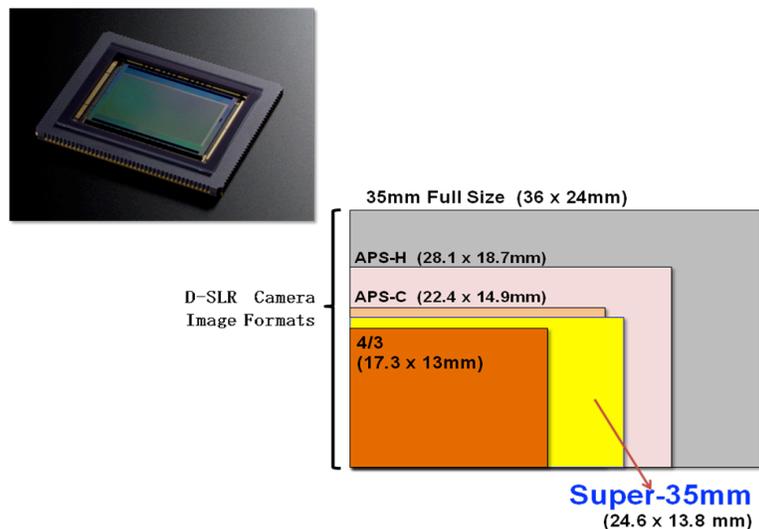


Figure 1 Canon CMOS image sensor compared to other well-known image format sizes

2.0 The Photosite Sampling Lattice

The new image sensor is a “4K” array in terms of the total number of photosites — that lattice being 4206 (H) x 2340 (V). The image sensor is, however, specifically designed to produce a standardized 1920 (H) x 1080 (V) high definition 4:4:4 RGB video component output set progressively scanned at an internal picture capture rate up to 60 frames per second. Accordingly, the number of active photosites is 3840 (H) x 2160 (V). The CMOS sensor array, by itself, is sensitive to the luminance of the projected image. The requisite color spectrum is sampled by a superimposed Color Filter Array (CFA) that is designed to spatially match the active CMOS photosite array. The CFA used in the new sensor is of the classic Bayer [1] pattern shown in Figure 2.

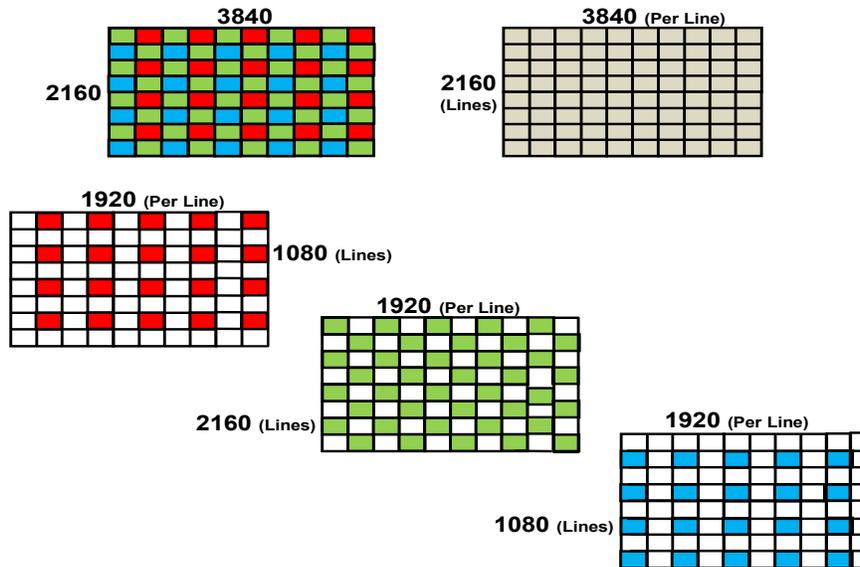


Figure 2 *Illustrating the separate CFA array and the CMOS imager while also showing the CFA separated into its component color filters to better expose the structure of their respective sampling lattices*

3.0 Unique Avoidance of a Demosaicking Interpolation Process

The usual Bayer mosaic requires a subsequent demosaicking process in order to reconstruct the required separate R, G, and B video components. This can be implemented in real-time within a camera, or later in non real-time within a postproduction process. The reconstruction process requires “filling in” the missing pixels shown in the three sparse-sampled color components of figure 2 [2]. Interpolation strategies from neighboring pixels are utilized to do this. These processes are always less than perfect, and are accompanied by various artifacts that can include impaired sharpness, aliasing (both monochromatic and chromatic), and reconstruction errors [3] [4]. Certain scene content, and motion capture activities (camera pans and tilts, and subjects in motion within a scene) can exacerbate the visibility of these artifacts. Over the years, many highly sophisticated demosaicking interpolation schemes continue to be developed that are intended to both increase the accuracy of the reconstruction and to minimize the artifacts. Some of the more sophisticated are quite excellent in terms of subjective quality.

The new Canon image sensor was designed to avoid the demosaicking process entirely. It instead relies on innovations in pixel addressing and associated readout mechanisms to separately extract the three RGB video components — as described in a separate Canon white paper [5].

4.0 Duality at the Green Sampling Lattice Level

The image sensor readout strategy radically departs from the customary “De-Bayer” deployment of quincunx sampling of the green photosites to maximize the green video resolution (and hence the matrixed Luma resolution). In the traditional de-Bayering process used in various “4K” single sensor cameras the resolution quest is typically in the neighborhood of 2500-3000 TVL/ph horizontal Luma limiting resolution — but this is, of necessity, accompanied by a significant shortfall in red and blue resolution [6].

The design strategy of this new sensor is not to seek any form of “4K” resolution — but rather to specifically confine the reconstruction of each of the R, G, and B video components to a full digital sampling structure of 1920 (H) x 1080 (V) — according to the SMPTE 274M HDTV Production Standard. Thus, a radically different strategy in dealing with the Green quincunx photosite sampling is undertaken.

The Green photosite array is essentially composed of two separate lattices of 1920 (H) x 1080 (V) as shown in Figure 3. These two sampling lattices are separated in time from each other by a sampling period — the specific time duration being a function of the pixel sampling processes. In the horizontal domain the two pixels are separated by a horizontal sampling period and in the vertical direction by a one-line period. No interpolation whatever is involved. A number of unique performance advantages are inherent in this process — as discussed in ref [5].

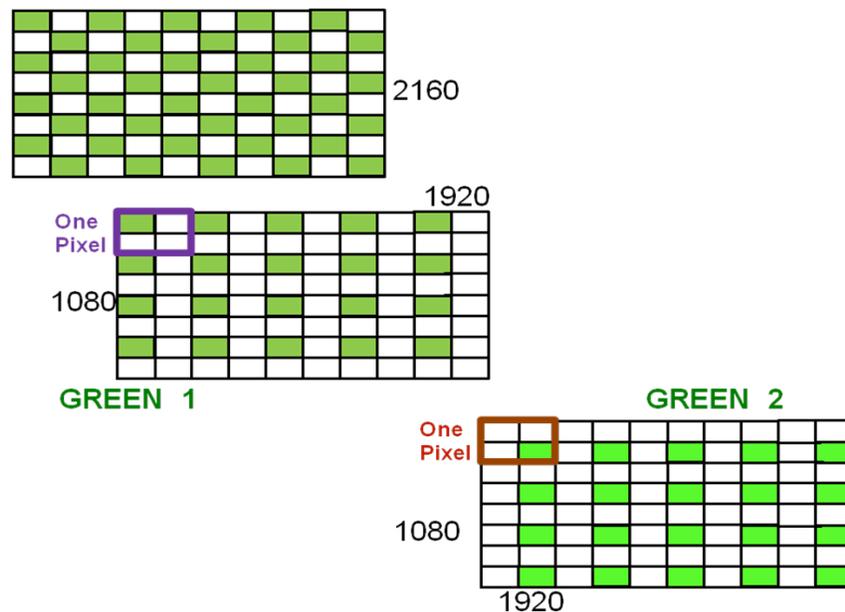


Figure 3 The green photosite array within the Bayer CFA is separated into two distinct Green sampling lattices during readout — with each having a 1920 (H) x 1080 (V) structure

5.0 Modulation Transfer Function of the Image Sensor Video Output

For digital cine motion imaging an important performance parameter of the image sensor is the separate horizontal and vertical MTF produced in the individual RGB video component outputs. This ultimately determines the picture sharpness that can be created by a lens-camera system deploying that image sensor. This MTF is, of course, largely determined by the sampling lattice of the sensor array — that is, the number of samples per horizontal line and the number of vertical line samples. The MTF characteristic of the image sensor (horizontal and vertical) follows the classic $\text{Sin } x / x$ curve (where x is the product of the pixel horizontal dimension multiplied by the variable of spatial sampling frequency) — with a modification to this that is dependent upon the photosite pitch. The MTF of the image sensor video output is, of course, further modified by that of the lens that will ultimately be projecting optical imagery on to it. The final compromise between a desirable MTF for the RGB video components will be determined by the design of the necessary optical low-pass filter (integral to any camera in which the sensor might be deployed) that curtails the sideband spectra that can create aliasing artifacts.

For HD origination having a 16:9 aspect ratio, the desired digital video sampling structure for each RGB video component is 1920 (H) x 1080 (V). The associated horizontal spatial sampling frequency is given by $1920 \times 2 \times 9 / 16 = 2160 \text{ TVL/ph}$, and accordingly, the Nyquist frequency is $2160 / 2 = 1080 \text{ TVL/ph}$.

6.0 Horizontal MTF Considerations within a Digital Cine Camera

To examine the potential resolution characteristics offered by the new image sensor it is necessary to consider its implementation in the new Canon EOS C300 digital cine camera. Specifically, the management of the component video MTF and the aliasing associated with its sampling mechanism must be determined.

The all-important green video component will be discussed first. In Figure 4 a cosine-shaped optical low-pass filter is chosen with its zero located above the effective carrier frequency of 2160 TVL/ph — and it convolutes with the basic sensor aperture to produce the effective MTF response for each of the two separate green video outputs as shown.

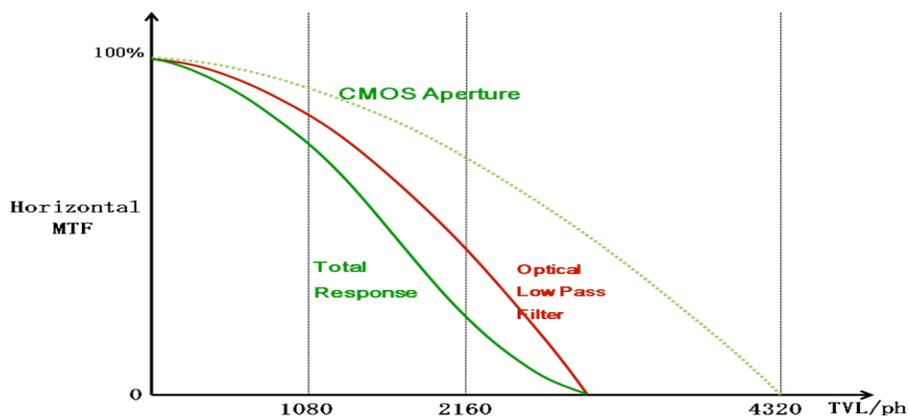


Figure 4 Showing a possible choice for a Cosine-shaped optical low pass (OLP) filter that is placed in front of the CMOS image sensor and operates equally on the separate MTF characteristic of the two green and the red and blue video components

A very important electronic modification to the MTF of the final green video is introduced by the summation of the two separate green components referred to earlier. As shown in Figure 5 the individual photo sites of each green lattice is offset with respect to each other by a half pixel — both horizontally and vertically.

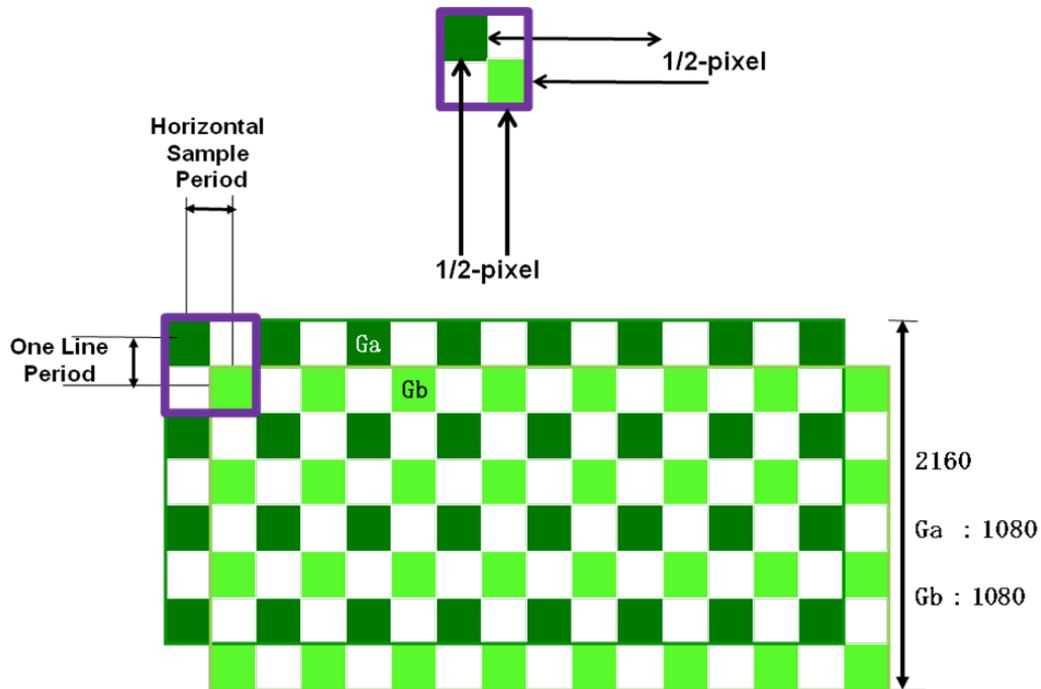


Figure 5 Showing the two separate green 1920 (H) x 1080 (V) photosite lattices and the horizontal and vertical timing offsets between each of the two “diagonal” pixels that are summed during the readout process

This half-pixel offset between two sampling lattices introduces a powerful anti-aliasing mechanism into the final summed green video component — both horizontally and vertically. That pixel offset shifts the phase of the first order sideband of Green Gb by 180 degrees with respect to that of Green Ga. Thus, when the two components are summed to form the final green video signal output these sideband spectra cancel each other. The cancellation of these sidebands effectively defeats the Nyquist limitation and allows the useful green MTF to extend beyond the Nyquist frequency of 1080 TVL/ph as shown below in Figure 6.

There is a second anti-aliasing benefit to the summation of the two offset green video components. When the two are summed as part of the readout mechanism, a horizontal Finite Impulse Response (FIR) filtering action is invoked because of the time separation between the two horizontal pixels that are summed, and this introduces the electronic prefilter that is also shown in Figure 6.

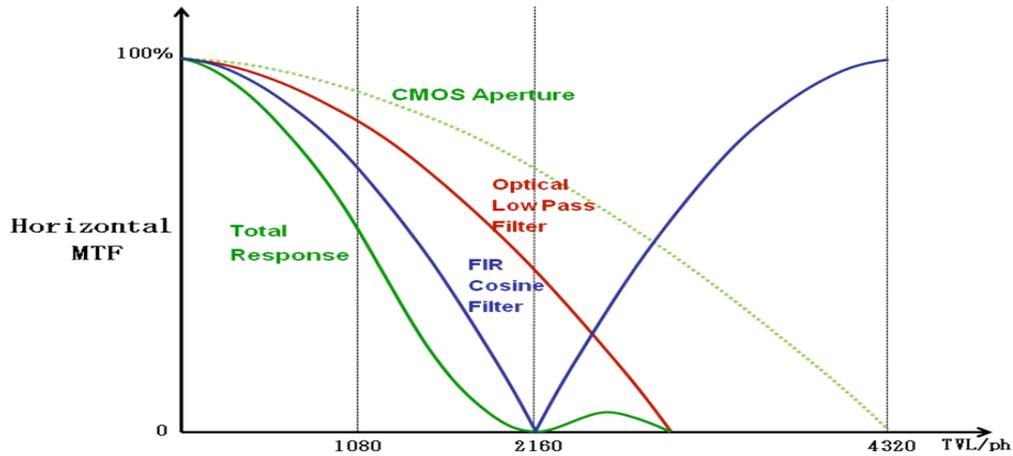


Figure 6 Showing the final response of the green video component following convolution of the horizontal FIR filter (created by the summation of the two original green sampling lattices during readout) and the optical low pass filter

The final green video MTF characteristic is produced by the convolution of the FIR filter, the optical pre-filter, and the $\text{Sin } x / x$ characteristic of the image sensor photosite sampling process and is shown as “Total Response” in Figure 6. The final green video spatial frequency characteristic has a high MTF across the HDTV passband of 0 to 1080 TVL/ph and a quite well controlled falloff above the Nyquist frequency as shown in Figure 7.

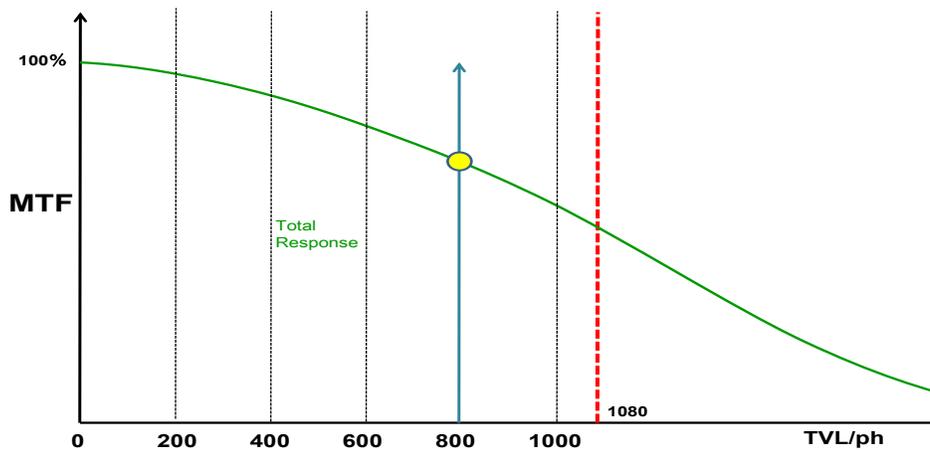


Figure 7 Showing the final horizontal MTF characteristic of the final green video

In high-end tri-sensor 2/3-inch studio cameras it is traditional to specify a depth of modulation at the spot frequency of 800 TVL/ph — which is close to the high-end of the HDTV passband (comfortably below the Nyquist frequency of 1080 TVL/ph and the prescribed SMPTE digital HD band-limiting filter having a turnover frequency of 872 TVL/ph). For this new CMOS image sensor, this particular MTF measurement is about 70% -- which is very high (this will lower to approximately 60-65% depending upon the lens employed).

7.0 Vertical MTF of the Progressive Scan Green Video Frame

As earlier outlined, the final Green video component is produced by summing the two separately originated Green 1 and Green 2 video signals. Because they are separated vertically with respect to each other by one horizontal line period the actual summation of the two is schematically represented in Figure 8.

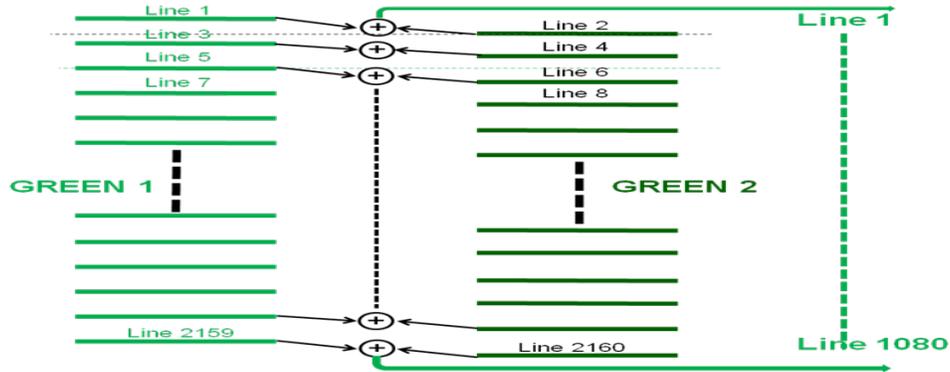


Figure 8 Illustrating the principle entailed in the formulation of the final progressive (24P, 25P, or 30P) 1080-line green video output from a summation of the two separate progressive green video signals which are separated vertically from each other by one horizontal line period.

The summation of the two time-displaced green signals creates an effective vertical FIR filter having a cosine characteristic and a zero at 2160 TVL/ph as shown below in Figure 9. This produces a high MTF in the desired 0 to 1080 TVL/ph vertical passband, but there remains a potentially high alias component beyond the 2160 TVL/ph carrier frequency. Again, the vertical half-pixel offset between the two green video signals cancels out the two reverse-phase vertical sideband spectra to produce an essentially alias-free vertical progressive MTF. The optical low-pass filter was designed with a zero chosen above the same spatial frequency as its horizontal counterpart and further protects against second order sidebands. Figure 9 depicts the total green vertical MTF response in the progressive scanning mode — it is identical to the horizontal response.

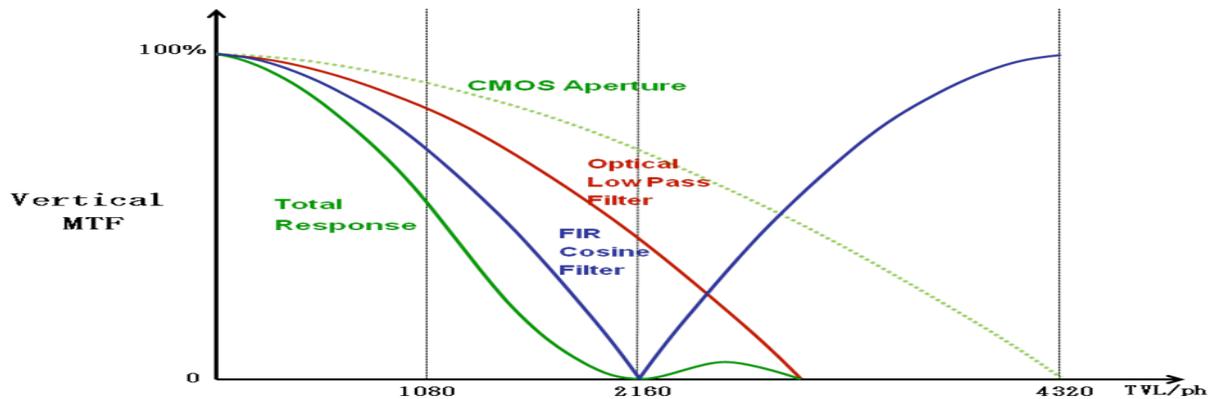


Figure 9 Showing the MTF of vertical progressive scan green video being shaped by the convolution of the electronic FIR filter (formed by summing the two vertically displaced Green video signals) and the optical low pass filter.

8.0 Vertical MTF of the Red and Blue 60i Interlaced Frame

The image sensor's readout system constructs a 1080-line progressive digital video format for each of the Red, Blue, Green Ga, and Green Gb video component signals as earlier described. When a 1080-line interlaced frame is required for the video output component signals, the readout mechanism is switched to perform a summing process of adjacent lines. The Red and Blue digital formats are processed in the traditional way — where a video line summing process constructs a 540-line vertical field every 1/60 of a second in the manner shown in Figure 10 for the Red component. The two sequential fields then structure the interlaced 30-frame digital video format.

The creation of an interlaced 30-frame video format introduces an additional unwanted sideband component over and above the fixed sideband that accompanies the progressive frame format. The latter sideband is centered at 2160 TVL/ph, whereas the new interlaced sideband is centered at 1080 TVL/ph and reverses its phase each field — thus producing the highly undesirable 30 Hz “flickering alias” that is characteristic of interlaced scanning.

Note however, the crucial overlapping of lines that takes place in the summation process between the two interlaced fields. This introduces an electronic vertical FIR filter that has a zero at the vertical Nyquist frequency of 1080 TVL/ph — as shown in Figure 11 — which considerably reduces the magnitude of the interlace-related “flickering” alias .

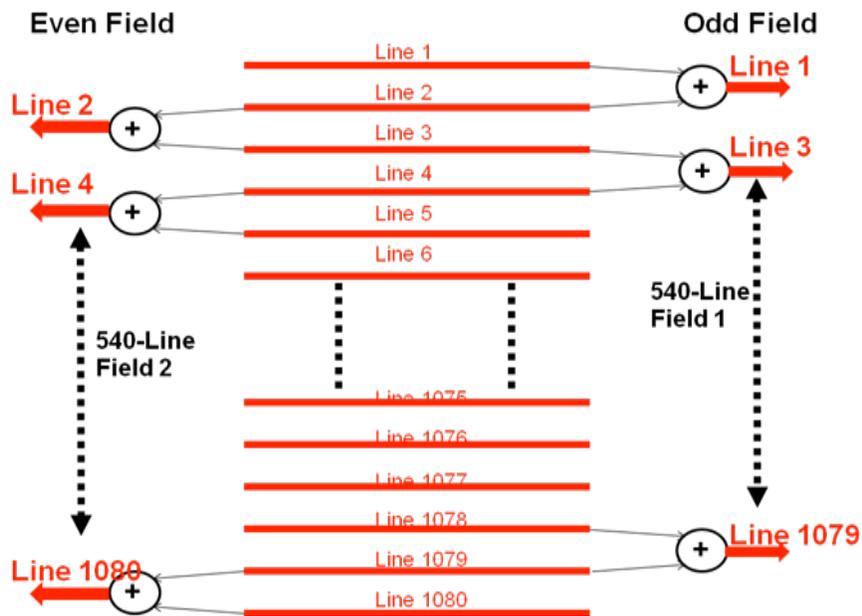


Figure 10 Illustrates the television line-summing process that structures two sequential 540-line fields each at 1/60 sec that make up the final 1/30 sec interlaced Red video frame.

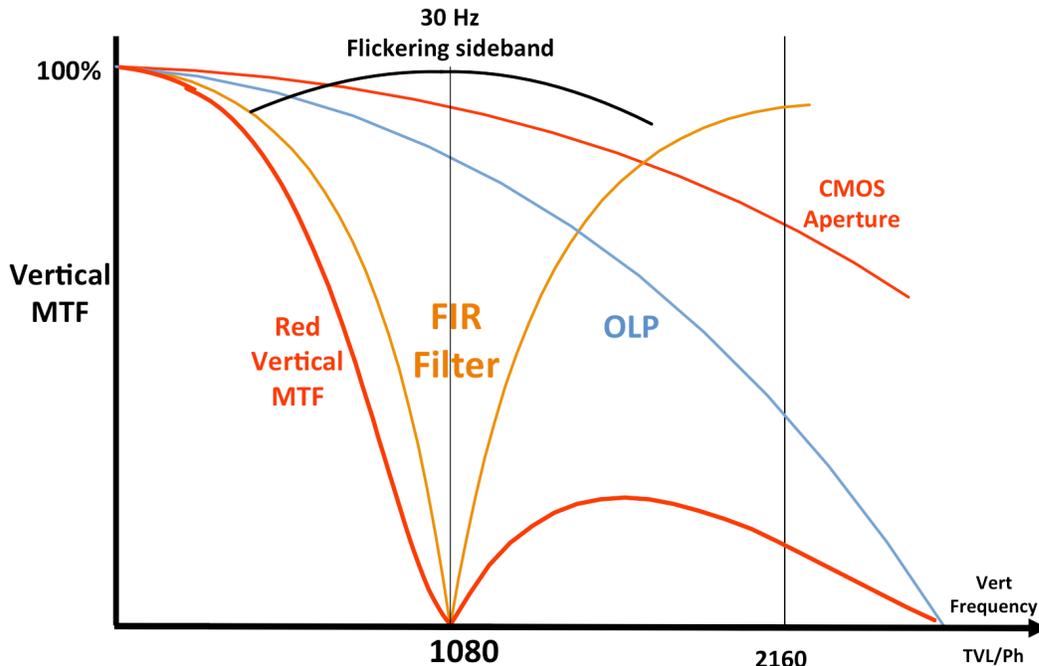


Figure 11 Constructing the Red vertical MTF characteristic in interlaced scanning mode (Blue video is identical).

9.0 Summary

A new CMOS image sensor has been described. It represents a definitive decision by Canon to enter the global field of digital cinematic motion imaging. It is anticipated that there will be many progressive advances in the years ahead. Accordingly, a priority was assigned to taking a first step into this important field of imaging by placing an initial focus on originating a very high quality RGB video component set specifically intended for high-performance High definition video production. Today, HDTV plays a significant global role in high-end television program origination, television commercial production, and to an increasing degree in high-end theatrical movie production, as well as numerous corporate, government, and educational applications. Large-format single-sensor digital cameras play an important role in extending creative flexibilities in each of these production arenas — because of the deployment of 35mm lenses and the attendant desires for short depth of field and the long-established craftsmanship of the world's film cinematographers.

The new image sensor capitalizes on the ingenuity of the classic Bayer CFA sampling lattice to accomplish the optoelectronic transformation of the 35mm optical image. But, it harnesses innovative pixel readout strategies to totally obviate the need for demosaicing interpolation techniques. As a consequence, full 4:4:4 video RGB digital components — each having the 1920 (H) x 1080 (V) representation — are delivered at the output of the image sensor system.

Coupled with the excellent noise characteristics of the photosites and the wide dynamic range (described in a companion white paper), it is anticipated that this image sensor will offer an overall image quality that will meet the aspirations of the world's cinematographers.

REFERENCES

- [1] B.E. Bayer “*Color imaging Array*” U.S. Patent No. 3,971,06
- [2] B.K. Gunturk et al, “*Demosaicking: Color Filter Array Interpolation in Single-Chip Digital Cameras*”
IEEE Special Issue on Color Image Processing
- [3] R.B. Wheeler and N. Rodriguez “*The Effect of Single-Sensor CFA Captures on Images Intended for Motion Picture and TV Applications*”
SMPTE J., 2007
- [4] R. Ramanath et al, “*Demosaicking Methods for Bayer Color Arrays*”
Journal of Electronic imaging 11 (3), p: 306-315, July 2002
- [5] L.Thorpe “*New 35mm CMOS Image Sensor for Digital Cine Motion Imaging*”
- [6] Adam Wilt, “*An Unfair Comparison of Three Different Cameras*”
Camera Log, ProVideoCoalition website, March 2008
- [7] Otto H. Schade, Sr. “*Image Quality: A Comparison of Photographic and Television Systems*”
J.SMPTE, June 1987 p: 567 — 595