Development of a 70mm, 25 megapixel Electronic Cinematography Camera with Integrated Flash Recorder

John Galt
Panavision, john.galt@panavision.com

Branko Petljanski
Panavision, branko.petljanski@panavision.com

Stacey Campbell
Panavision, stacey.campbell@panavision.com

Written for presentation at the SMPTE 2012 Fall Technical Conference

Abstract. This paper will describe the system design of the world’s first 70mm, 25 megapixel, electronic-cinematography camera with an integrated flash memory recorder.

Although this camera shares many of the attributes of NHK’s Ultra High Definition Television and benefits from NHK’s pioneering research in the area of very high-resolution imaging, what we are about to describe is an electronic cinematography camera and recorder – not a television camera. While the unique requirements of the cinematographic process allow us to take advantage of certain processes that would not be practical in a television environment, we must also make provision for requirements that are unique to movie making.

The single CMOS sensor in the new Panavision camera has the same four times HDTV horizontal resolution of 7680 photo-sites as that proposed by NHK for UHDTV. Our sensor’s vertical resolution is only three times the HDTV vertical for a resolution of 3240 photo-sites and an aspect ratio of 2.37:1 rather than UHDTV’s 16 x 9, 1.78:1 aspect ratio.

Bucking the recent trend to take a lower resolution imager and interpolate to a higher resolution, the Panavision camera utilizes an oversampling technique to output a full bandwidth 19 megapixel RGB image from the available 25 million photo-sites.

The recorded frame size is 3840 x 1620 x 3 (RGB) or 18,662,400 pixels.
Keywords. 4K, UHDTV, Super-70mm, oversampling, electronic cinematography
Introduction

In the motion picture business, it is now quite common for camera manufacturers to make references to resolutions that describe their cameras as 2K, 4K, and 6K. Describing area array imaging systems using this metric is misleading. This metric can only correctly be applied to line array scanners and cameras.

When camera manufacturers compare a "Bayer pattern" sensor camera to the 4K resolution scans that are common in high-end feature film Digital Intermediates, they are confusing an up-converted interpolation of photo-sites with a native, full-bandwidth, RGB image capture.

The most obvious difference between a “4K” camera (whose claim for "4K resolution" is achieved by adding all the colored pixels in a line) and a true 4K film scan is this: When the film is scanned by a 4K line array camera, the output consists of 4K of red pixels plus 4K of green pixels and 4K of blue pixels, resulting in a total of 12K color pixels per line.

Scanning a 35 mm Academy aperture negative would require a camera with an area array sensor of 36 million photo-sites (a Super-35 scan would be only 28 million). These scans are usually 10bit log. per-color, uncompressed, resulting in an image data size of 45 megabytes per frame.

In this paper we will use the term “photo-site” when referring to a sensor photosensitive element and “pixel” when we are referring to an interpolated single color output, as the term “pixel” has now become synonymous with the pixel guessing that would enable an 8.3-million photo-site imager to become a 25 million-pixel camera.

The “K” confusion

Despite all the technical challenges associated with creating the first “True 4K” large format digital cinema camera, one of the greatest challenges we are faced with is how to end the confusion surrounding the term ‘4K’.

We believe one consequence of this confusion is the mismatch between acquisition and display in digital cinema. With the introduction by Sony, Barco, NEC and Christie of real “4K” projection systems with resolutions of 4096 x 2160 x 3 or 26.5 million projected sensor elements, for the first time in history our cinema display systems now exceed the resolution of most available camera systems.

Our industries have an endless fascination with the number of photo-sites in a solid-state imager, assuming that this number equates with resolution. However, this assumption is only valid if the optics employed can resolve the sensor elements with sufficient contrast. So the question is: just how much contrast is sufficient?

System MTF as a resolution metric

The Modulation Transfer Function (MTF) of all components in an imaging system will define the image performance of the system [1]. The most important aspect of MTF is that the final MTF of the system will always be less than the lowest-performing component. Therefore an ideal
system will have matching components. For example: photo-site count need not exceed the lens performance or vice versa.

Most definitions of image performance cite limiting resolution. This particular criterion is most useful in determining the smallest detail that can be extracted from a still image such as in those found in aerial surveillance. Motion picture systems, though, have a different requirement.

Otto Shade, an imaging scientist working at RCA in the late 1940's, recognized that limiting resolution did not adequately define image performance of a moving image [2]. He proposed that a more valid measure of image performance would be to square the MTF values and use the area under the squared value as a quality metric. This approach tends to emphasize lower frequencies at the expense of high frequency detail. In a simplification of Otto Shade’s concept, we aim to maximize MTF at half our Nyquist frequency in our optical designs.

**Optics**

When Sony and Panavision developed Genesis, the first Super-35 format electronic cinematography camera, the camera resolution requirement was dictated by the performance of the existing Primo lenses for the 35mm cine format. Figure 1 shows the MTF measurements of Panavision Primo 100mm lenses with EK5277, a medium speed color negative film manufactured by Kodak [3]. The same figure shows the system MTF, which is a result of multiplying lens MTF with film MTF. Primo lenses historically had MTFs in the range of 85 to 90% at 20 line pairs per millimeter (lp/mm), which was a good complement to the 35mm film emulsions with which they were designed to image. This dictated a Nyquist frequency for the sensor of 40 lp/mm.

![Figure 1: The solid black line indicates MTF of Kodak Vision 320T color negative film (EK5277). The dashed line represents MTF of Panavision Primo 100mm lenses and the lightly shaded curve shows system MTF.](image)

Despite the inevitable aberrations caused by optical low pass filters, I.R. cutoff filters and other filters being placed between the lens and the sensor, the performance was quite acceptable. Figure 2 illustrates the MTF degradation of Primo 35 mm lenses when the planar optical elements representing OLPF, IR filter, and sensor cover glass are inserted between the cine lens and the image plane.
Figure 2: MTF comparison of a Primo 35mm lens with or without the filter pack between the lens and the image plane. These Primo lenses were designed for film cameras that assume free air between the last lens element and the film emulsion. (Solid line is tangential. Dotted line is sagittal)

**Sensor Design Considerations**

Consider the imager and optical requirements of a true 4K camera, i.e. four times the HD resolution of 6.2 million photo-sites. A true 4K image at four times this resolution would require 24.8 million photo-sites. Squeezing this number of photo-sites into a Super-35 image size would result in a 3.6 µm photo-site-pitch. To avoid the image artifact of a rolling shutter, the smallest global shutter photo-site-pitch that our foundry could guarantee was 5 µm. We also felt that a 3.6 µm photo-site-pitch would have an unacceptably low dynamic range.

The most important consideration, however, was that the Nyquist frequency for a 3.6 µm photo-site is just over 138 lp/mm. Lenses would then have to be designed and built for a half Nyquist frequency response of 70 lp/mm. Since we also needed to have better than 85% MTF at this frequency, our lens designers felt this would be almost impossible over the series of lens focal lengths that our customers would require.

The final choice of image size and photo-site size was mostly based on technical considerations, but there was also a historical component. In 1959, Panavision introduced the Super Panavision 70 and Ultra Panavision 70 formats (camera pictured in Figure 3). Although fewer than 40 features were shot using these formats they are among the most lauded in cinema history.

Figure 3: Ultra/Super Panavision 70 camera from 1959.
Figure 4: Theatrical release poster for the 1962 film “Lawrence of Arabia”, filmed in the Super Panavision 70 format.

Anyone who has seen a 70mm print of a film originated on a 65mm negative would recognize that large image formats have a unique image quality. This was amply demonstrated at the recent screening of the digitally restored Lawrence of Arabia (see Figure 4) by the Motion Picture Academy on a Sony “4K” projector.

Despite the film emulsion and optics technology of the early 1960s, the image quality easily matched that of modern 35mm technology. Two times oversampling for the final 4K RGB output was done at 8192 x 3584 pixels, creating 10 bit log digital picture file exchange format (DPX) files.

The final image size for the new Panavision digital camera is 48mm x 20.25mm – exactly three times the area of a Super-35 film frame. The number of photo-sites is 7680 (1920 x 4) by 3240 (1080 x 3) for a total of 24,883,200 photo-sites, with a photo-site size of 6.25µm (see Figure 5).

Figure 5: Imager size comparison between Super-35 and Super-70 digital image formats.

Although the 1962 optics do not perform at the level of modern optics, they were imaging to a film format with twice the horizontal dimension of Super-35 film or digital image formats, and the image area was also three times that of Super-35. There is one major reason why this 50-year-old technology should challenge the performance of today’s best optical and imager technology: by projecting the same amount of information on a larger area, the lens-resolving power and MTF requirement are proportionally (with area) less demanding [2].

Given the absence of a reflex mirror, the latest 70 mm optics (Primo-70) are able to take advantage of a new short working distance design concept [4]. They also utilize multiple
aspherical elements and new optical glass types that have resulted in dramatic advances over the original Super-70 lens performance. Figure 6 illustrates the improvement in MTF performance between Super Panavision 70 from 1962 and the latest designs.

![Figure 6: MTF comparison of Super Panavision 70 lens from 1959 and a new Primo-70 lens.](image)

The most important feature, which significantly reduces the burden on the lens MTF requirements, is that the Primo-70 lenses are designed to cover a 70mm image format with a 6.25 µm photo-site.

Digital cameras are often required to include several planar optical elements between the last lens element and the sensor surface. Typically, this optical stack includes sensor cover glass, optical low-pass filter (OLPF), IR/UV blocking filter, and neutral density filter. Historically, cine lenses that were intended for film-based cameras were designed without consideration for parallel glass elements between the lens and the film.

Since the new digital camera requires several parallel surfaces between the lens and sensor, one of the requirements was to optimize the lens design for an optical stack that is the equivalent of a 7.0 mm thick parallel plate made of BK7 glass (see Figure 7).

![Figure 7: shows the performance of a 50mm focal length lens for the new Super-70 format imaging in free air, compared to the performance through the equivalent of 7.0 mm of BK7 glass](image)
Figure 8 shows MTFs of the Primo-70 lens (focal length 100mm) for several different image heights at 10, 20, 40 and 80 lp/mm [4].

One additional design requirement that differentiates the new Primo-70 mm lenses from previous Panavision cine lenses is the ability to provide information to the camera (metadata) such as lens identification, focus distance, zoom position, aperture, and temperature. Captured metadata is transmitted to the camera and recorder for use in post-processing. In addition to having traditional geared manual controls, these lenses have internal motorized control of focus, zoom, and aperture (see Figure 9). Switching between internal or manual mode of operation does not require clutches to disengage.

Figure 9: Contact pins on the Primo-70 lens for metadata.
Sampling lattice and Nyquist limits

This imager size and photo-site count results in a 6.25 µm photo-site. A 6.25 µm photo-site has a Nyquist frequency of 80 lp/mm and therefore a half-Nyquist of 40 lp/mm. As mentioned earlier, 40 lp/mm was deemed by our optical designers to be an achievable frequency response with high MTF. The new Primo-70 optics provide a 90% MTF at 40 lp/mm.

![Diagram](image)

Figure 10: The green sampling lattice is shown on the left. The horizontal (vertical) distance between centers of the green samples is 6.25 µm. The horizontal (vertical) distance between centers of red/blue samples is 12.5 µm. On the right are shown Nyquist limits for green samples (light gray) and red/blue samples (dark gray).

Green samples are located on a diagonal lattice, while red/blue samples are located on a rectangular lattice. Green samples have a horizontal and vertical Nyquist frequency of 80 lp/mm, while red/blue samples have a horizontal and vertical Nyquist frequency of 40 lp/mm. Observing diagonally, the green samples have a Nyquist limit at 56.74 lp/mm, which is the same as the diagonal Nyquist limit for red/blue samples (see Figure 10).

Our output full-bandwidth RGB image is 3840 x 1620 pixels, resulting in maximum output frequency (for the Super-70 optical format) of 40 lp/mm.

Optical Low Pass Filtering

In order to minimize aliasing of frequencies above the Nyquist limits, most contemporary still and video cameras employ an optical low-pass filter (OLPF) as a prefilter. The majority of OLPFs are based on optically birefringent materials whose thickness and optical crystal orientation determine the distance between exiting rays of the filter and, consequently, the cutoff frequency of the filter. Due to the nature of birefringent filters, it is not possible to achieve a steep transition between pass-band and stop-band. This forces camera designers to select between less aggressive filters to maintain image resolution and allowing some aliasing, or deploying more aggressive filters that result in less aliasing, but compromised image MTF.

In order to improve the sharpness of the resulting image, our approach to this problem is to double oversample horizontally and vertically for each output pixel. In this way, the design...
constraints for the OLPF may be relaxed and a reasonably flat MTF can be achieved up to the maximum output frequency (see Figure 11).

![MTF plot](image)

Figure 11: MTF of two 4-point separation filters is described as a cosine function. The position of the first zero crossing is determined by the distance between points. Shaded area illustrates pass-band MTF improvement by using an OLPF with half the distance between points.

In the case of the widely used OLPF with a square four-point separation configuration, designers often select the horizontal distance between green samples as a design parameter. Such a filter has a cut-off frequency equal to the horizontal Nyquist frequency of the green samples. Due to the inherent slow transition characteristic of the OLPF filters, MTF of the pass-band signal is compromised. When the image is spatially digitized, a digital low-pass filter with better transition characteristics than the OLPF may be used to produce desired output. Such an approach results in camera outputs with higher MTF values for pass-band signals.

**Processing**

**Image Processing Overview**

The main objectives of our image processing pipeline are to correct the image artifacts that are part of image acquisition and to generate an output image that does not require extensive post-processing such as demosaicing, color correction, etc.

Features and requirements of the image-processing pipeline are:

- Pipeline input is a 7680x3240 Bayer pattern stream of 12 bits per photo-site at up to 60 fps.
- Image processing pipeline output is 3840x1620 10 bit perceptually encoded RGB triplets (30 bits per pixel).
- All photo-site values undergo digital flat-field correction.
- Configurable demosaic kernels and down sampling kernel.
- User configurable color correction matrix, color balance, and output pixel lookup table (LUT).
- SMPTE ST 2042-1:2012 HQ (Dirac Pro as developed by British Broadcasting Corporation R&D) [5] compression to meet the solid-state recorder maximum bandwidth capability.
**Flat Field Pixel Correction**

The camera platform provides sufficient data bandwidth, multipliers, and adders to perform full-frame per-pixel correction of pixel dark current (DSNU), pixel photo response (PRNU), and—in sensor array locations where pixel data is uncorrectable—pixel replacement.

**Demosaic**

The goal of a demosaic filter is to produce as accurate an RGB interpolation of the image feature at each sensor photo-site as possible, maximizing MTF and minimizing aliasing and noise. In addition to utilizing the spatial correlation between samples to calculate the missing values, the demosaic algorithm also takes into account the correlation among red, green and blue values. Configurable kernel sets can be matched to the highest frequency response of the optical system, I.R. filter selection, crosstalk characteristics of the sensor, and the MTF and SNR requirements of the user.

**Color Correction and Color Balance**

A linear 3x3 matrix multiply converts camera-native RGB to a known color space.

\[
\begin{bmatrix}
R_{out} \\
G_{out} \\
B_{out}
\end{bmatrix} =
\begin{bmatrix}
M_{rr} & M_{rg} & M_{rb} \\
M_{gr} & M_{gg} & M_{gb} \\
M_{br} & M_{bg} & M_{bb}
\end{bmatrix}
\begin{bmatrix}
R_{in} \\
G_{in} \\
B_{in}
\end{bmatrix}
\]

User configurable digital gain is then applied on a per-channel basis.

\[
\begin{bmatrix}
R_{out} \\
G_{out} \\
B_{out}
\end{bmatrix} =
\begin{bmatrix}
R_{gain} & 0 & 0 \\
0 & G_{gain} & 0 \\
0 & 0 & B_{gain}
\end{bmatrix}
\begin{bmatrix}
R_{in} \\
G_{in} \\
B_{in}
\end{bmatrix}
\]
**Down sampling**

Down sampling is performed by pre-filtering each internal 7680x3240 15-bit RGB channel with a kernel implementing a low-pass filter, then subsampling the filter output. The low-pass filter serves a dual function: it spatially pre-filters the full bandwidth image to limit its bandwidth before down-sampling to prevent aliasing. In addition to the spatial pre-filtering, it aligns all three color planes so the resulting RGB triplets are spatially aligned (see Figure 12). Fully configurable filter coefficients can be utilized, and coefficients will be determined by a set of factors including system MTF and SNR requirements and the active demosaic filter set.

Down-sampling rejects common demosaic artifacts such as zippers, chromatic aliasing, and magenta fringing at the edge of clipped regions.

![Figure 12: Red crosses on the left indicate the sampling output lattice. Colored squares represent original sensor photo-sites. The horizontal distance between the centers of the input (prior to down sampling) samples is 6.25 µm. The horizontal distance between the centers of the output samples is 12.5 µm. On the right (light gray) are shown Nyquist limits for the full bandwidth samples. Dark gray represents the outline of the frequency response of the low-pass pre-filter.](image)

**Output Look-Up-Table (LUT)**

A user configurable LUT maps each linear 12 bit R, G, and B value to a 10 bit transform value to decrease bandwidth. The transfer function used to convert from 12 bit to 10 bit values is typically based on perceptual coding.

**SMPTE 2042 VC-2 HQ Compression**

At LUT output the maximum pipeline bandwidth is 3840x1620 10 bit RGB (4:4:4) at 60 fps. A requirement of the camera system is that it contain a removable solid state recording system. The maximum recorder input bandwidth is half the maximum pipeline output bandwidth.

To match LUT output bandwidth with recorder input bandwidth a study was undertaken to evaluate constant bit rate compression codecs that produce a visually lossless decompressed output at a compression ratio of 2:1 or less. A transport implementation of the HQ profile of SMPTE ST 2042-1:2012 was selected and implemented.
Recorder

The record medium is flash memory arranged in a RAID-5 configuration to provide an uncompressed equivalent of 2 terabytes of data with a maximum record time of 40 minutes at 24 fps or 17 minutes at 60 fps. In addition to the primary low compression ratio recording, we also record a proxy channel in Avid’s DNxHD (SMPTE 2019-1-2008 VC-3 Picture Compression and Data Stream Format) for immediate dailies and editorial use.

Mechanics

Although the camera contains a Super-70 format sensor we have tried to keep it small and light enough to be both hand-held or used on Steadicam™ or similar devices.

Figure 13: Ultra Panavision 70 from 1962 and the new 70mm, 25-megapixel electronic cinematography camera with integrated flash recorder.
Conclusion

In this paper we present design concepts that are behind a new 70mm, 25-megapixel electronic cinematography camera with integrated flash recorder. We use quality metrics—beyond the simplistic photo-site count—as a guide in the design of the first true 4K camera. As a result, we believe that this camera qualifies as a true 4K design.

Our goal to create a true 4K camera lead us to analyze the limitations of current optics and image formats. We undertook a path that resulted in a significantly larger optical format than contemporary digital cinema cameras.

Combining optics specifically designed with digital imagers in mind, with a large format sensor and minimally compressed data stored on removable media, will allow cinematographers to create a visual experience rarely seen today.

Acknowledgements

The authors wish to acknowledge Tak Miyagishima of Panavision. Tak was an early and enthusiastic supporter of this project. We also wish to acknowledge the contributions of James Pearman, David Macintosh, Barry Rynk and other employees of Panavision who were instrumental in this project. Special thanks go to Brian Caldwell for initially proposing the short working distance optics concept.

References


