

The Effect of Single-Sensor CFA Captures on Images Intended for Motion Picture and TV Applications

Richard B. Wheeler, Nestor M. Rodriguez
Eastman Kodak Company

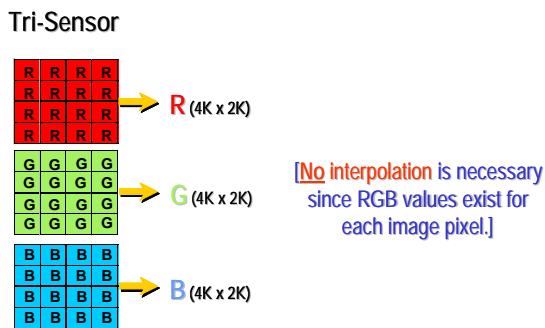
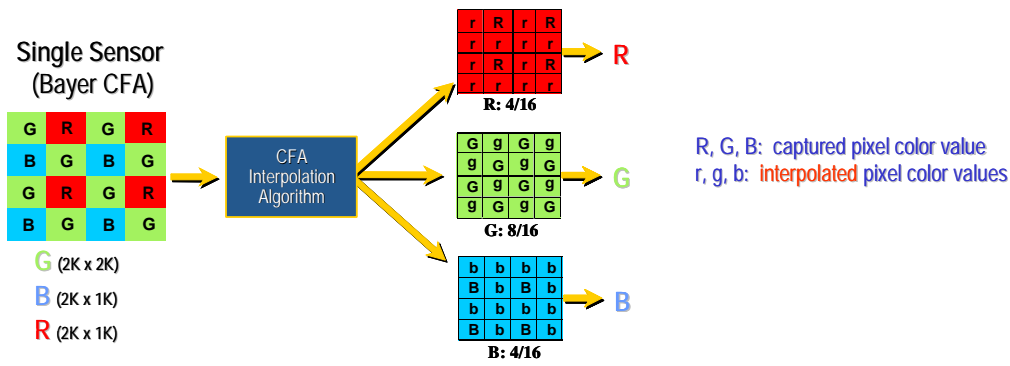
Abstract

Current digital cinema camera designs employ a variety of sensor architectures based on CCD and CMOS integrated circuit technologies, with either separate sensors for each color channel or a single sensor with a color filter array. Some of the sensor solutions that have proven robust for consumer still photography may fail to meet expectations at the higher system magnifications and greater viewing angles encountered in motion picture applications. In this paper, we compare the desirable attributes and unwanted artifacts that are generated by various single-sensor and multiple-sensor camera designs. To illustrate the key trade-offs in system design, we quantify the resolution, sharpness, aliasing, reconstruction errors, CFA artifacts, sensitivity (speed), and dynamic range that are associated with current designs. We also examine mitigation strategies that can improve overall image quality.

Introduction

Relative to film origination, digital cinematography offers an immediacy advantage when there is a need to have the captured image information in digital data form, such as in post-production operations for the Digital Intermediate (DI) process. Specifically, digital capture does not require the chemical processing and optical scanning required in film capture to convert the scene image content to digital data. However, digital camera recording bandwidth limitations (in the case of uncompressed image data), high-cost storage requirements, and the difficulty in obtaining suitable CCD/CMOS pixel-sensor size combinations that are compatible with film camera optics and capture/record full RGB color resolution with an exposure range and sensitivity comparable to film, have led electronic cinematography camera design to move from the traditional video camera tri-sensor/prism geometry to single-sensor designs with a color filter array (CFA). Single-sensor designs with a CFA are very prevalent in the still camera industry, and such designs include the Bayer pattern and striped arrays. Figure 1 provides a high level summary of the properties of tri-sensor and single-sensor Bayer CFA configurations.

This paper quantifies the image quality trade-offs that are associated with digital cinema cameras employing tri-sensor and single-sensor designs. When evaluating different systems, we assume that the color, tone, and noise characteristics are optimal and equal, and we employ the same high quality lenses, digital sampling, and digital projection components in each case. As a result, the quality differences reported herein arise solely from variations in sensor design, spatial interpolation method, and the effect of anti-aliasing filters.



NOTE:
 4K Tri-Sensor camera (versus a Single Sensor 4K type) would require considerable
 -- higher Bandwidth/Storage and
 -- larger size/weight, as well as being
 -- incompatible with Film Camera Lenses

Figure 1: R,G,B sampling with 4K single-sensor and 4K tri-sensor geometries.

Image Quality Assessment and Modelling: A Review

At Eastman Kodak Company, a consistent and integrated approach to image quality characterization and prediction has been developed and integrated into product development processes. A detailed description of the research enabling the creation of our predictive image quality models is beyond the scope of this paper, but it is summarized in Refs. [1-3], which cover the following topics: establishing image quality standards; performing psychophysical experiments; designing objective metrics correlating to perceptual attributes; predicting the overall quality of samples affected by multiple attributes based on knowledge of the impact of each attribute in isolation; constructing software to enable powerful system image quality modeling of capability and performance; and verifying the accuracy of predictions that are generated.

It is generally not feasible to perform factorial experiments that fully map out the dependence of quality on different aspects of a pictorial imaging system. An approach known as “multivariate formalism” allows this situation to be simplified, and it has proven to be very reliable. This approach greatly facilitates the construction of predictive models of image quality. In this approach, a set of seemingly distinct perceptual attributes (e.g., sharpness, noisiness, etc.) that span the quality variations of interest are identified. The dependence of overall quality on each attribute, varied one at a time, is then determined. The quality changes are expressed in terms of just noticeable differences (JNDs). Once the effect of each individual attribute is known, a variable-exponent Minkowski metric is applied to predict the impact on overall quality of all of the attributes in combination. The accuracy of the predictions from these models has been verified in numerous Kodak experiments and customer intercept studies.

The results that are discussed in this paper are based on attribute-specific objective metrics that were derived from quantitative perceptual experiments employing a still image

quality ruler using the method detailed in ISO 20462 [4, 5]. The attribute that is typically varied in a quality ruler is sharpness, because sharpness has several desirable properties: 1) it can be easily changed using simple image processing; 2) it has low scene and observer variability, and 3) it has a strong influence on image quality. Although it may seem surprising, a quality ruler that varies only in sharpness can be used to accurately assess other types of quality impairments such as compression artifacts or noise. This is because the task that is presented to an observer when assessing overall quality involves making a fairly simple judgment as to which image he would prefer to keep if the image represented a cherished memory. Image quality rulers based on sharpness variations have been successfully used to characterize 23 key image quality attributes other than sharpness (such as noise, color, tonescale, output dynamic range, aliasing, etc.) in studies involving thousands of individual observations [3]. A review of the techniques used to construct, calibrate, and use a motion quality ruler can be found in [6].

In each of the perceptual experiments pertaining to this analysis, the attribute or artifact of interest was added by digital image simulation to 12 representative scenes. The digitally altered scenes, now containing the artifact under study, were compared by 30 observers to images of the same scene having no artifacts, but varying in known JND increments of quality. The image quality ratings of the observers were then correlated with critical image chain engineering parameters to produce predictive equations that allow the impact of product design changes on image quality to be estimated without running additional perceptual tests. Because the results from this analysis were derived from still scenes, the quality loss from aliasing and CFA-related artifacts encountered in practical cinematography might be worse than reported here, particularly when scenes with fine detail are captured with subject motion and camera panning [7].

JND Units: A Scale for Assessing Image Quality

All of the results in this analysis are expressed in the perceptual units of image quality known as 50% JNDs. In this section, the statistical and perceptual significance of the JND is illustrated using the paired comparison image quality assessment method. In the paired comparison method, an observer is asked to select one of two stimuli that are simultaneously presented. Typically, the observer is instructed to choose the stimulus that has greater or lesser image quality (or a specific attribute of quality, such as sharpness). This method is also referred to as forced choice paired comparison (FCPC).

The paired comparison method leads naturally to the concept of a JND, which corresponds to a stimulus difference that produces a 75:25 proportion of responses [8]. That is, if one stimulus is selected 75% of the time in a paired comparison test, the two stimuli differ by one JND. This is sometimes referred to as one 50% JND, because a 75:25 split in responses means that the difference between the two stimuli is actually detected 50% of the time, and the observers are guessing in the remaining cases, half of which will be correct by chance ($50\% + \frac{1}{2}(50\%) = 75\%$). The JNDs for other response proportions can be computed using:

$$\text{JNDs} = \frac{12}{p} \cdot \sin^{-1}(\sqrt{p}) - 3, \quad (1)$$

where p is the probability of a given stimulus being selected [4].

One 50% JND can be thought of as the smallest meaningful unit on a continuous scale of quality. As the quality difference between two images increases, so does the number of JNDs separating them. For example, it takes about 6 JNDs to move from one subjective quality category to another; e.g., from “good” to “fair”. A quality difference of one JND is generally not thought of as “advertisable”, because at that level, only half of the observers detect the difference. A difference of 2 JNDs (visible to 87% of observers) is typically considered to be the smallest advertisable quality difference. One of the important characteristics of the JND scale is perceptual uniformity. This means that two images

separated by a known number of JND units will cause observers to experience the same sensation of image quality difference, regardless of where the images are positioned on the scale.

Key Spatial Quality Attributes for Camera Sensor Evaluation

In this paper, we focus on those aspects of image quality that are most strongly affected by incomplete sampling of all three color records at each pixel position, which is brought on by the use of color filter arrays.

Many video and still cameras employ a single image sensor covered with a CFA to capture real-time color images. Each pixel only records one of the three colors (e.g. red-green-blue or cyan-magenta-yellow) that are required to produce a full three-channel color image. The missing pixels in each channel are obtained by interpolation from the neighboring existing pixels to reconstruct the full three-color image. Examples of commercially used CFA patterns and interpolation schemes can be found in Refs. [9-11]. In addition to some degradation of sharpness, CFA sensors and the associated interpolation algorithms can also incur aliasing, which results in several visually different phenomena.

Image sensors, like CCD arrays, record image information at regularly spaced, discrete locations, i.e., the image is spatially sampled. The Whittaker-Shannon sampling theorem implies that only those scenes that are band-limited to less than half of the sampling frequency can be fully reconstructed from the sampled image. Spatial frequency information beyond half of the sampling frequency is aliased to lower spatial frequencies. Aliasing manifests itself mainly as a distortion in patterns and changes in textures.

If the image sensor is covered with a CFA, each of the color channels is sampled at a lower rate than that of the full sensor. The sampling rate can vary between the channels, e.g., the green channel in the Kodak Bayer CFA pattern is sampled at twice the rate of the red and blue channels. Also, the regular array of sampling points in a color channel is always shifted by a known number of pixels relative to another channel. This property of CFAs gives rise to the appearance of colored textures and colored fringes at edges. Chromatic aliasing has a distinctly different visual appearance than achromatic aliasing.

Besides aliasing, additional artifacts are introduced when images are displayed using non-ideal reconstruction. Reconstruction artifacts can appear as jagged edges and/or ringing. A non-published Kodak study showed that aliasing rarely occurs in isolation from reconstruction artifacts. The impact of these artifacts on image quality is quantified in a separate reconstruction error metric.

CFA interpolation is a superb example of the need for the application of multivariate theory in image quality analysis [2]. The four artifacts mentioned above (loss of sharpness, achromatic aliasing, chromatic aliasing, and reconstruction error) are perceptually distinct. Psychophysical transforms between an objective metric and JNDs of image quality degradation were previously established for each individual artifact, and the results can be combined by multivariate theory to produce the overall JNDs of quality loss associated with CFA interpolation.

Because resolution and the related attribute of sharpness are often cited when camera designs are compared, the JND values for each camera design will be expressed in terms of sharpness only and also as a multivariate quality sum that includes the combined effects of loss of sharpness, achromatic and chromatic aliasing, and reconstruction error.

System Components Used in Simulations

The component specifications were selected to cover the design space in which most current digital cinema cameras reside. None of the systems under study is intended to replicate a specific product that is available in the marketplace. When evaluating different systems, we assumed that the color, tone, and noise characteristics are optimal and equal, and our simulations always employed the same high quality lenses, digital sampling, and digital projection components. This section summarizes the system components that were

used in our baseline simulations to quantify the perceived sharpness and the multivariate quality of images captured at 2K, 3K, 4K and 6K resolution with single-sensor and tri-sensor cameras, and subsequently displayed with 2K and 4K digital projectors. In our artifact minimization examples, we vary the pixel fill factor and the CFA interpolation method, and we also illustrate the benefit provided by birefringent anti-aliasing filters.

Components common to all systems:

- High-quality capture optics
- Camera sensor with 8.0 micron pixel pitch and 0.7 fill factor (70% active area)
- Optimized prefilters prior to downsampling (e.g., 4K capture → 2K projection)
- Output file resizing via cubic convolution
- No digital boosting (sharpening) in post production
- High-quality digital projector optics
- 2K and 4K digital projection devices with 13.7 micron pixel pitch
- Viewing at two picture heights (2 PH)

Components unique to specific systems:

- All single-sensor CFA cameras include a birefringent anti-aliasing filter
- Striped CFA cameras use linear interpolation to populate the color planes
- Bayer CFA cameras use adaptive interpolation to populate the color planes

Camera and projector resolutions (and format dimensions):

- 2K Camera: 2048 x 1080 (16.4 x 8.6 mm)
- 3K Camera: 3072 x 1620 (24.6 x 13.0 mm)
- 4K Camera: 4096 x 2160 (32.8 x 17.3 mm)
- 6K Camera: 6144 x 3240 (49.2 x 25.9 mm)

- 2K Projector: 2048 x 1080 (28.0 x 14.8 mm)
- 4K Projector: 4096 x 2160 (56.0 x 29.6 mm)

Quality Predictions for Tri-Sensor and Single-Sensor Cameras

The results of the analysis are summarized on plots with JNDs as a function of camera capture resolution. The “0 JND” level represents an artifact-free image having ideal color and tone reproduction, with a sharpness level above which no visually perceptible improvements are possible. The larger the negative JND value, the lower the perceived quality of the image. As discussed previously, one JND can be thought of as the smallest meaningful unit on a continuous scale of quality. As the quality difference between two images increases, so does the number of JNDs separating them.

Because resolution and the related attribute of sharpness are often cited when camera designs are compared, JND values for each camera design will first be determined for sharpness only, and then as a multivariate quality sum that includes the combined effects of loss of sharpness, chromatic and achromatic aliasing, and reconstruction error.

Sharpness with 2K and 4K digital projection

Figures 2 and 3 show the sharpness levels for the tri-sensor and single-sensor systems at three capture resolutions and at 2K and 4K projector resolutions, respectively. It can be seen that the tri-sensor and the single-sensor Bayer CFA with adaptive interpolation deliver similar sharpness, but the striped CFA with linear interpolation has significantly lower sharpness in all cases. Not surprisingly, it can also be seen that increasing the capture resolution from 2K to 4K provides minimal benefit with 2K projection, but it has a significant benefit with 4K projection.

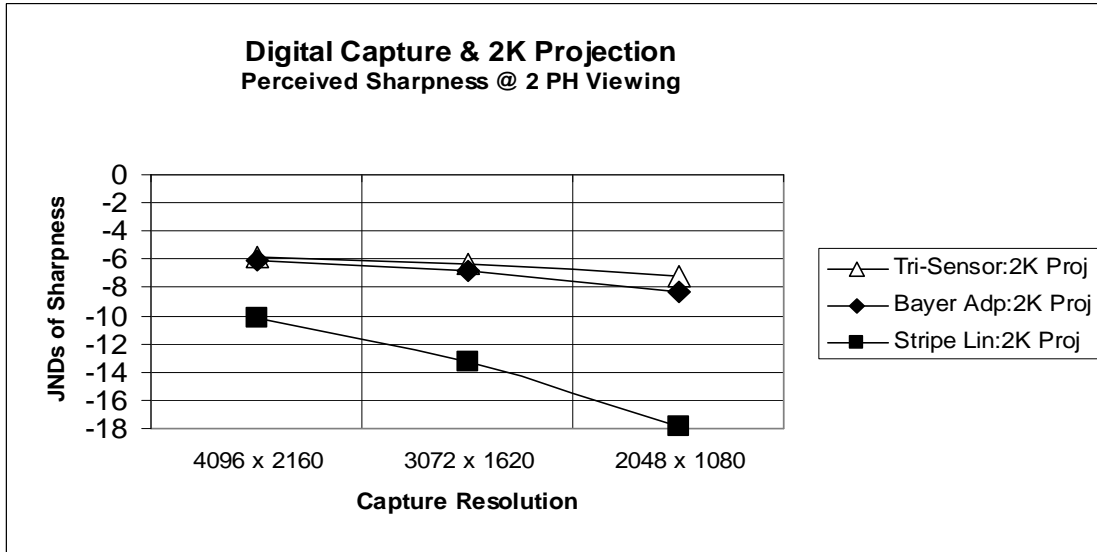


Figure 2: JNDs of sharpness for each of the camera sensor designs and capture resolutions when projected at 2K.

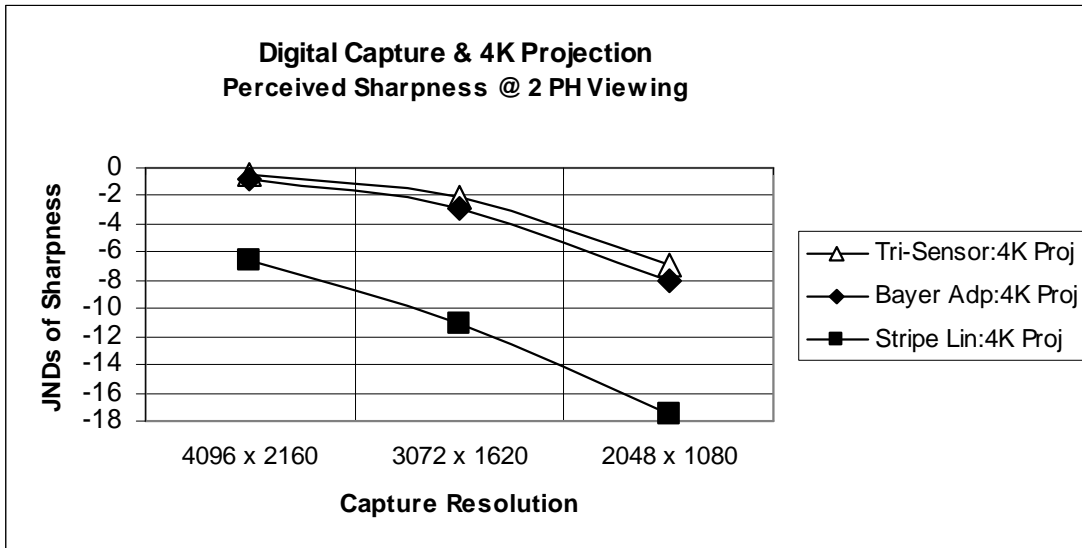


Figure 3: JNDs of sharpness for each of the camera sensor designs and capture resolutions when projected at 4K.

Multivariate quality with 2K and 4K digital projection

Figures 4 and 5 show the multivariate quality losses for the tri-sensor and single-sensor systems at three capture resolutions and at 2K and 4K projector resolutions, respectively. The tri-sensor system delivers superior, artifact-free quality in all cases. While the Bayer CFA had similar sharpness as the tri-sensor (Figs. 2 and 3), the artifacts that arise from the subsampled colors and corresponding interpolation cause a 4-6 JND quality drop, which is about one quality category. The artifacts with a striped CFA produce an even larger quality drop of 10-14 JNDs, which is two quality categories. It can also be seen that increasing the capture resolution from 2K to 4K provides only modest benefits with CFA sensors because degradations from the artifacts dominate the overall quality that is achievable.

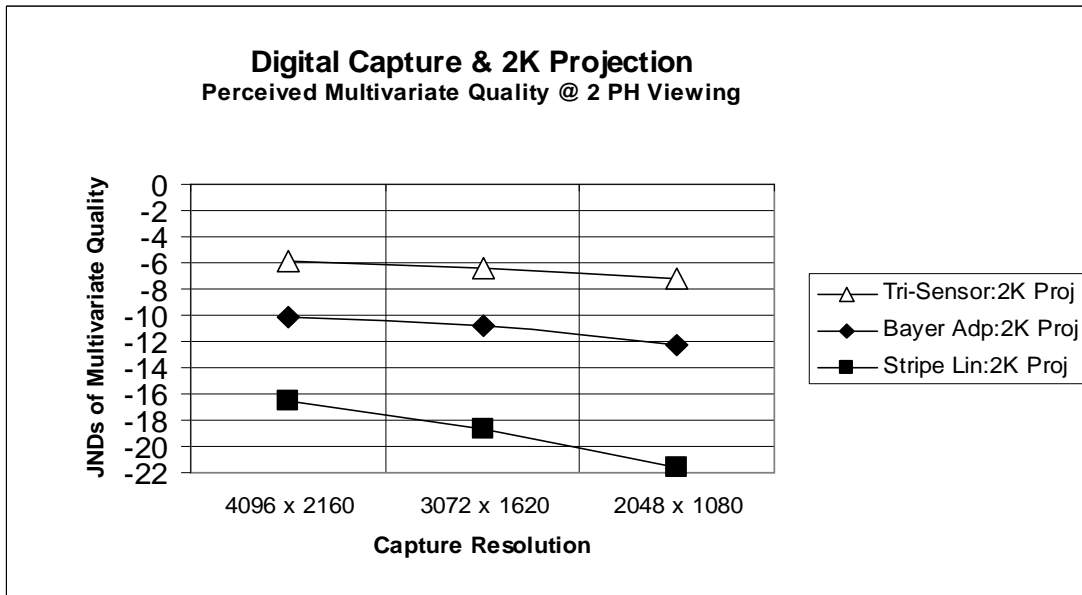


Figure 4: JNDs of multivariate quality for each of the camera sensor designs and capture resolutions when projected at 2K.

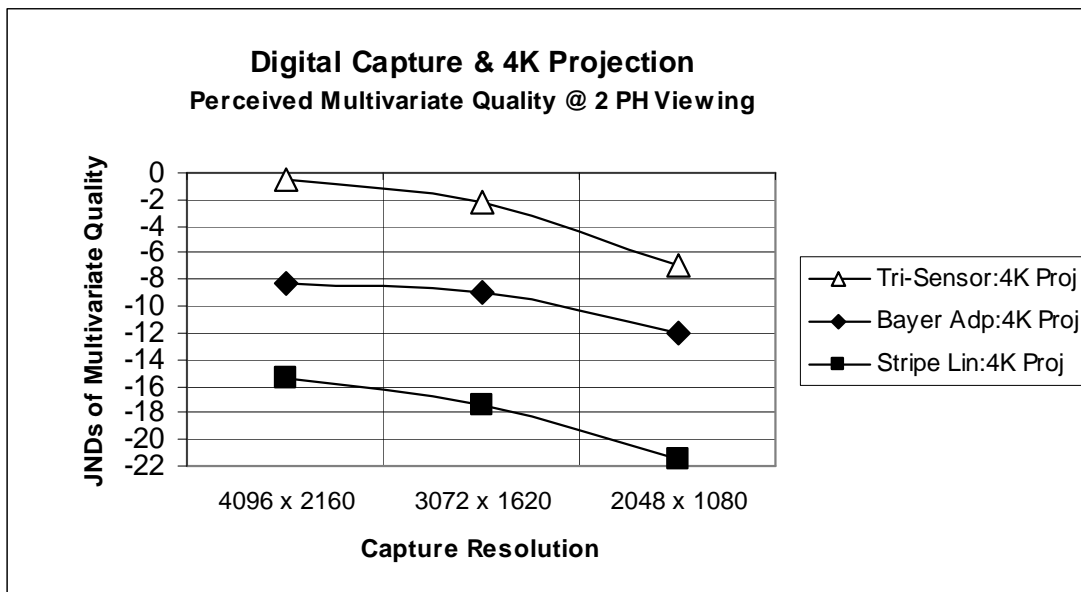


Figure 5: JNDs of multivariate quality for each of the camera sensor designs and capture resolutions when projected at 4K.

Artifact Minimization Examples

The results shown in Figs. 2-5 reflect what can be achieved with digital cameras that have components tailored to produce the best results from each sensor configuration. For cost and/or compatibility reasons, it may be necessary to use off-the-shelf components with different characteristics than we selected. In the following artifact minimization examples for a Bayer CFA system, we explore pixel fill factor, anti-aliasing filters, CFA interpolation method, and oversampling with optical blurring.

Pixel fill factor (active area) refers to the portion of the pixel that records scene information. For example, when the fill factor is 0.5, only half of the area of each pixel records light. Because the light collection (sampling) window is 50% smaller, it is capable of recording higher frequency information than a pixel having 100% active area. However, a 0.5 fill factor pixel has 50% inactive portion (dead area) separating it from neighboring pixels, which makes it more prone to aliasing.

An anti-aliasing filter limits the spatial frequencies that reach the sensor. The optical blur filter that is examined in this section for anti-aliasing is a birefringent type with a spot separation equal to one-half the pixel pitch.

For interpolating full resolution color from a Bayer CFA, an adaptive interpolation method that is similar to the one employed in this analysis is described in Ref. [10]. The accuracy of color plane reconstruction is improved with adaptive interpolation by locating areas of high spatial activity (e.g., edges) and subsequently interpolating parallel to the edge to minimize sharpness losses and artifacts.

Oversampling (intentionally capturing a higher resolution image than will be viewed) with integral optical blurring has been proposed as a means of simulating the benefits of tri-sensor cameras, which capture R,G,B at every pixel location, in single-sensor CFA cameras. This technique involves spreading a portion of the light that was formerly striking a single color (e.g. G) pixel over the other color neighboring pixels (e.g. R, B) to provide a better approximation of a fully populated tri-color image plane.

Impact of pixel active area (fill factor)

Figures 6 and 7 show the sharpness and multivariate quality losses for a Bayer CFA with pixel fill factors of 0.5 and 0.9 at three capture resolutions and at 2K and 4K projector resolutions, respectively. In all cases, adaptive interpolation is used, along with an anti-aliasing filter. It can be seen that the smaller fill factor (0.5) provides a slight sharpness benefit because of its improved response at high frequencies. However, the larger fill factor (0.9) is less prone to aliasing, so it has a slight advantage in terms of the overall multivariate quality. These trade-offs between sharpness and artifacts also lead to the result that the 0.5 fill factor is slightly better for 2K capture, while the 0.9 fill factor is better for 4K capture, in terms of overall quality.

Impact of birefringent anti-aliasing filter (BF)

Figures 8 and 9 show the sharpness and multivariate quality losses for a Bayer CFA with and without an anti-aliasing filter at three capture resolutions and at 2K and 4K projector resolutions, respectively. In all cases, adaptive interpolation is used with a pixel fill factor of 0.7. It can be seen that the anti-aliasing filter causes a slight loss in sharpness for 2K capture, but it also leads to a slight reduction in artifacts as well. In contrast, the 4K captures with the Bayer CFA and anti-aliasing filter achieve significant artifact reduction (3 - 4 JNDs) with only a minimal loss in sharpness.

Impact of CFA interpolation methods

Figures 10 and 11 show the sharpness and multivariate quality losses for a Bayer CFA using either linear or adaptive interpolation at three capture resolutions and at 2K and 4K projector resolutions, respectively. In all cases, an anti-aliasing filter is used with a pixel fill factor of 0.7. It can be seen that adaptive interpolation leads to improved sharpness and reduced artifacts. The benefits of adaptive interpolation are greatest (about 6 JNDs) for the lower capture resolutions. However, it is worthwhile to note that even with adaptive interpolation, the multivariate quality that is produced by single-sensor Bayer CFA systems is significantly lower than the image quality produced by a tri-sensor system (Figures 4 and 5).

Impact of oversampling with optical blur

Figures 12 and 13 show the sharpness and multivariate quality for a 6K Bayer CFA using adaptive interpolation, and for a 6K striped CFA using linear interpolation, when the images are intentionally subject to controlled optical blurring at the time of image capture and then downsized to 2K resolution for projection. In both cases the 8 micron pixel pitch and 0.7 fill factor of the previous examples were retained in order to hold constant the dynamic range, sensitivity, and noise characteristics. However, because the pixel count was increased, the sensor dimensions grew to 49.2 x 25.9 mm, necessitating the use of specially designed lenses. If compatibility with existing 35mm lenses and a 6K pixel count are requirements, the pixel pitch would decrease to 5.3 microns, leading to reductions in dynamic range and/or increasing noise, as illustrated in Figure 15.

Figure 14 provides a schematic diagram of a single sensor geometry that simulates the R,G,B sampling resolution of a tri-sensor camera by implementing a CFA consisting of a striped pattern that repeats every three horizontal pixels. However, unlike the tri-sensor configuration where R,G,B information is available at every point in the image, the R,G,B pixels of the striped CFA are horizontally offset from one another. Since the R,G,B sampling of the image is not co-sited in the striped geometry, the signals from three adjacent R,G,B pixels are typically “mixed” by optically blurring and/or by signal processing to produce an approximation of the actual R,G,B signal level at each photosite. After this blurring takes place, the effective horizontal sampling interval of this geometry is the combined R,G,B pixel triad width, as opposed to the width of a single pixel as is the case for a tri-sensor design.

Figure 12 shows the effect of optical blurring using isotropic birefringent blur filters with varying spot separations. Figure 13 shows the effect of optical blurring by intentionally choosing a lens with lower MTF response. In both cases, the introduction of optical blur degrades sharpness, but leads to an improvement in overall quality due to CFA artifact reduction. This strategy leads to an improvement in overall quality, because the sensor has a native resolution that exceeds the display requirements. Employing a similar technique with a sensor having a resolution of 2K would lead to unacceptable losses in sharpness. The birefringent blur filter with optimized spot separation is preferred, because it minimizes CFA artifacts with less sharpness loss than is incurred by lowering the MTF of the capture lens. The degraded capture lens causes more sharpness degradation because it suffers modulation loss at all frequencies, while the birefringent filter preferentially reduces high frequency information.

Referring to Figure 12, the 6K single-sensor with Bayer CFA, adaptive interpolation and a birefringent blur filter with a spot separation of 0.9 pixels compares favorably to the 2K tri-sensor for overall quality (refer to Figure 4), when both are projected at 2K. The 6K single-sensor with striped CFA and optical blurring shows a significant improvement over the 2K single-sensor striped CFA, but falls short of the 2K tri-sensor for overall quality when projected at 2K.

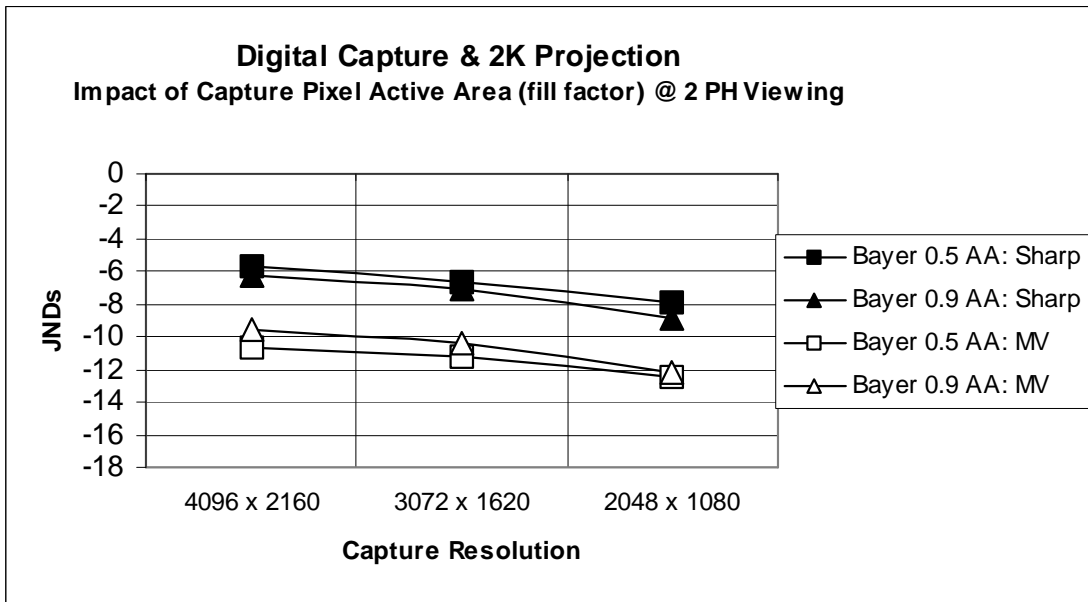


Figure 6: JNDs of sharpness (Sharp) and multivariate quality (MV) for single-sensor Bayer CFA cameras with 0.5 and 0.9 pixel fill factors when projected at 2K.

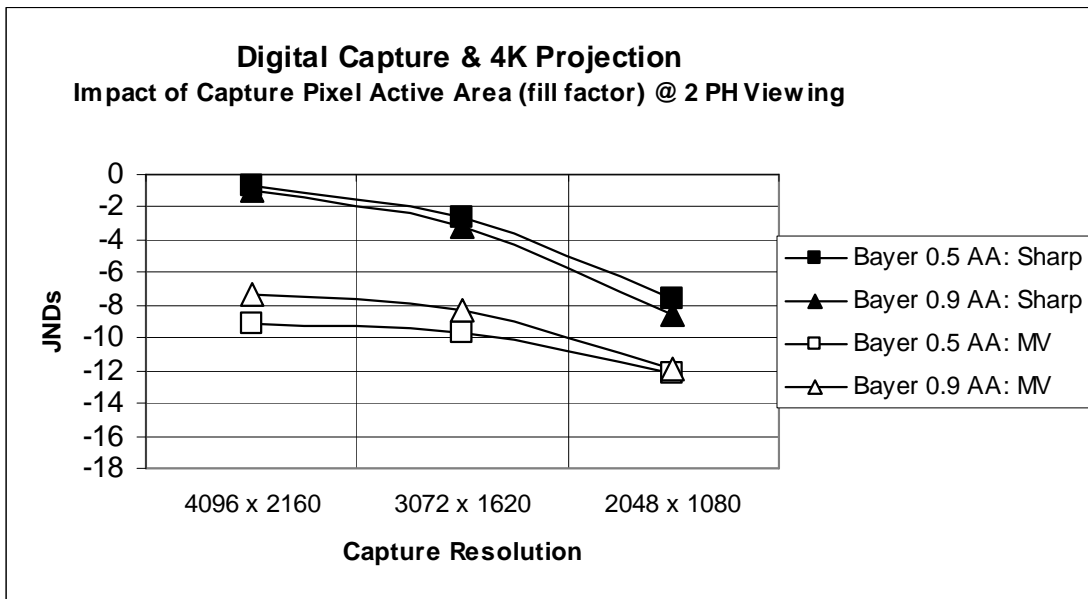


Figure 7: JNDs of sharpness (Sharp) and multivariate quality (MV) for single-sensor Bayer CFA cameras with 0.5 and 0.9 pixel fill factors when projected at 4K.

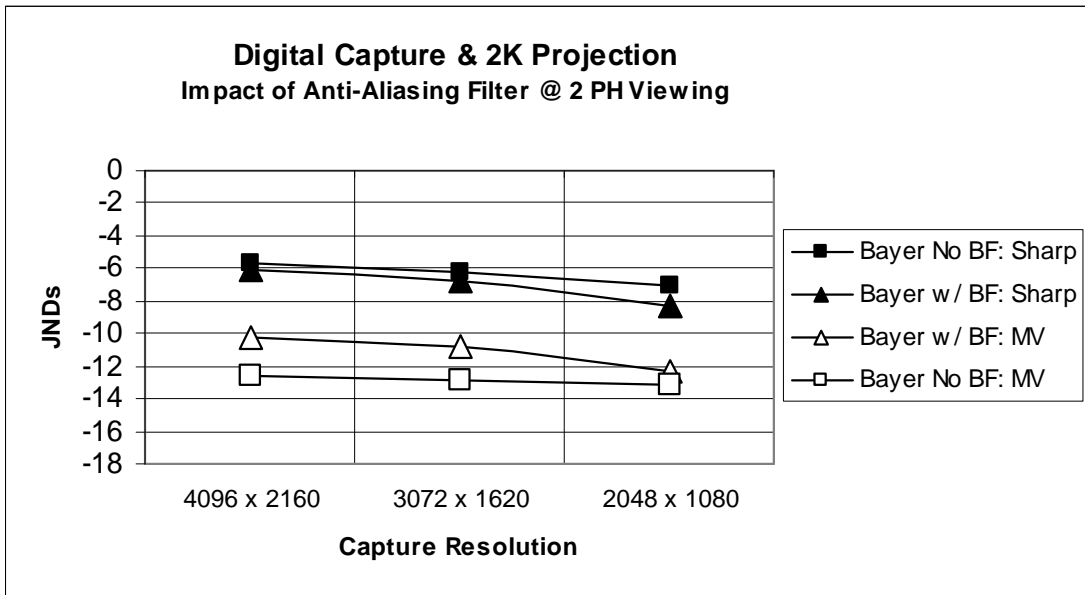


Figure 8: JNDs of sharpness (Sharp) and multivariate quality (MV) for single-sensor Bayer CFA cameras with and without a birefringent anti-aliasing filter when projected at 2K.

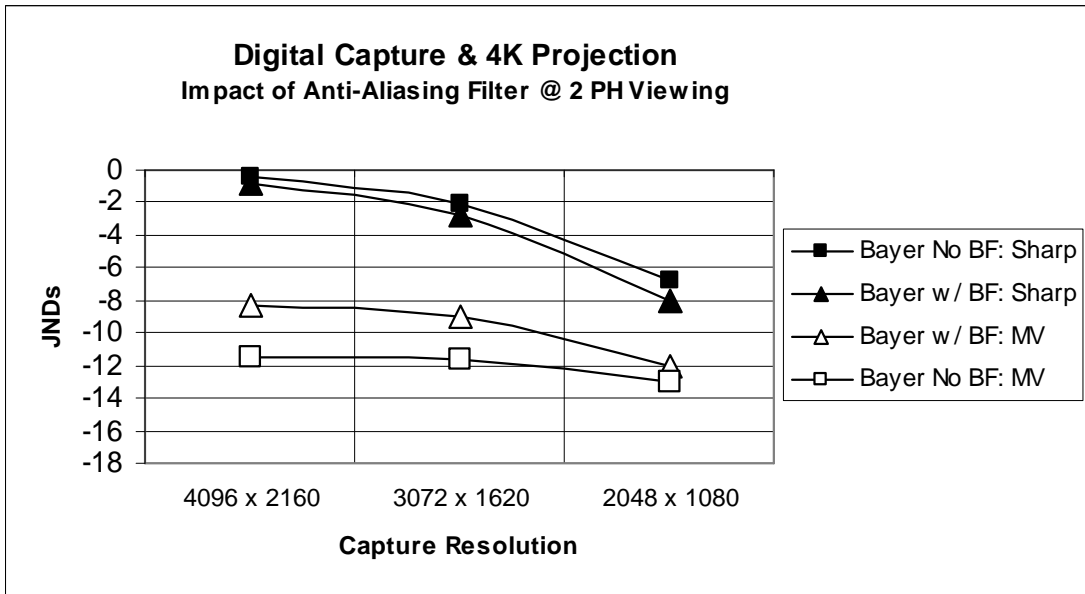


Figure 9: JNDs of sharpness (Sharp) and multivariate quality (MV) for single-sensor Bayer CFA cameras with and without a birefringent anti-aliasing filter when projected at 4K.

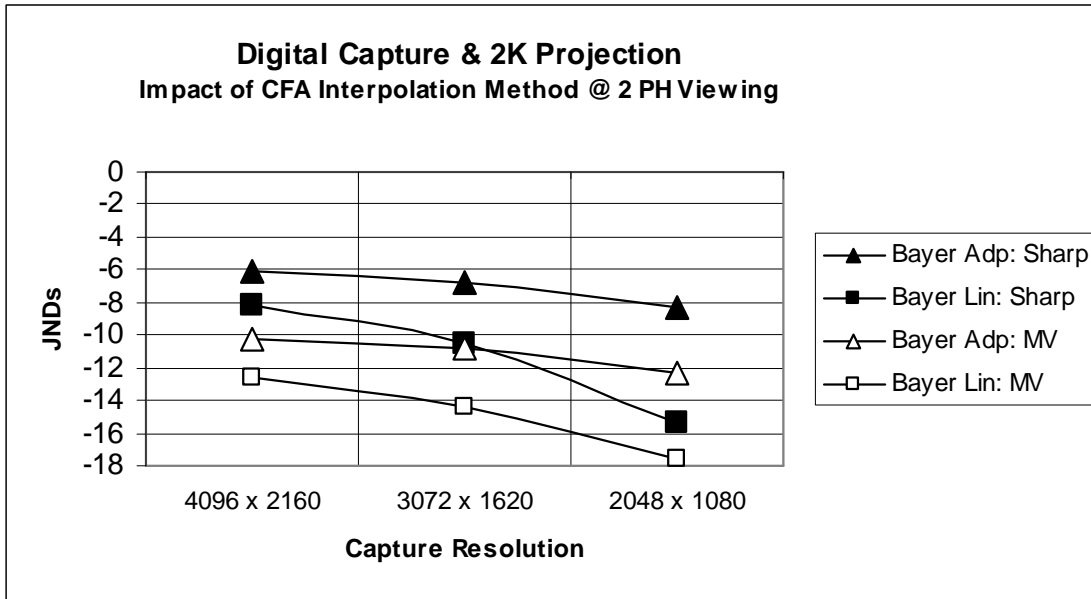


Figure 10: JNDs of sharpness (Sharp) and multivariate quality (MV) for single -sensor Bayer CFA cameras with linear and adaptive interpolation when projected at 2K.

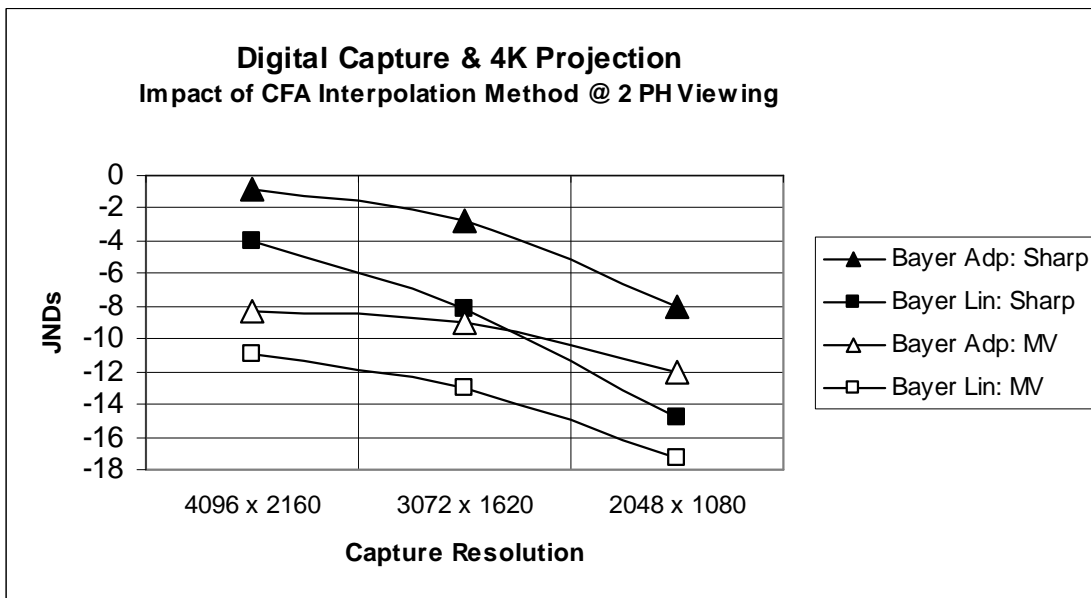


Figure 11: JNDs of sharpness (Sharp) and multivariate quality (MV) for single -sensor Bayer CFA cameras with linear and adaptive interpolation when projected at 4K.

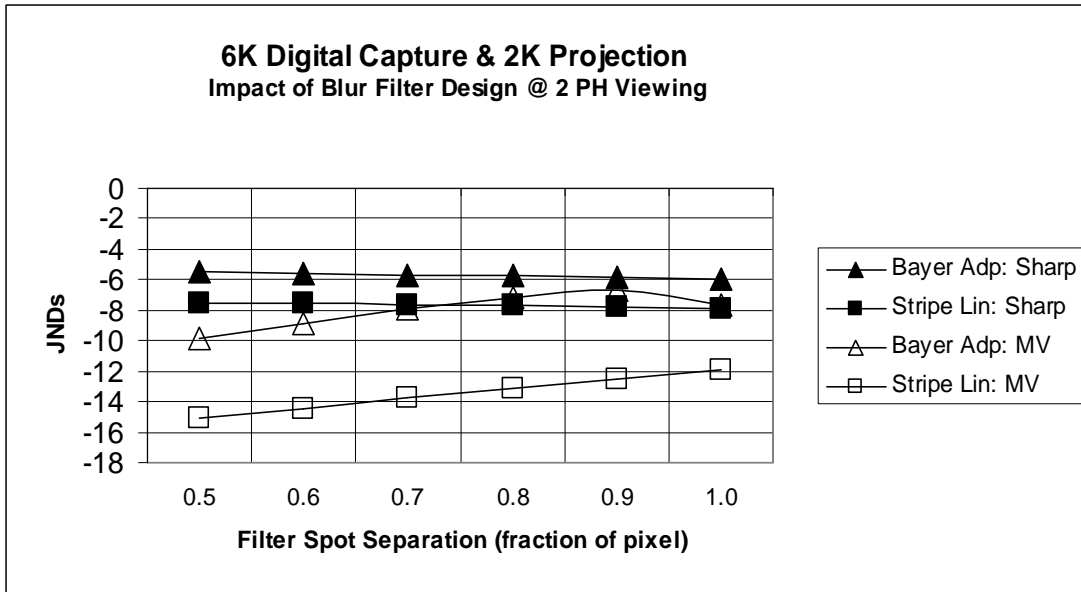


Figure 12: JNDs of sharpness (Sharp) and multivariate quality (MV) for 6K single-sensor Bayer CFA and striped CFA cameras after blurring with birefringent filter and downsizing for 2K projection.

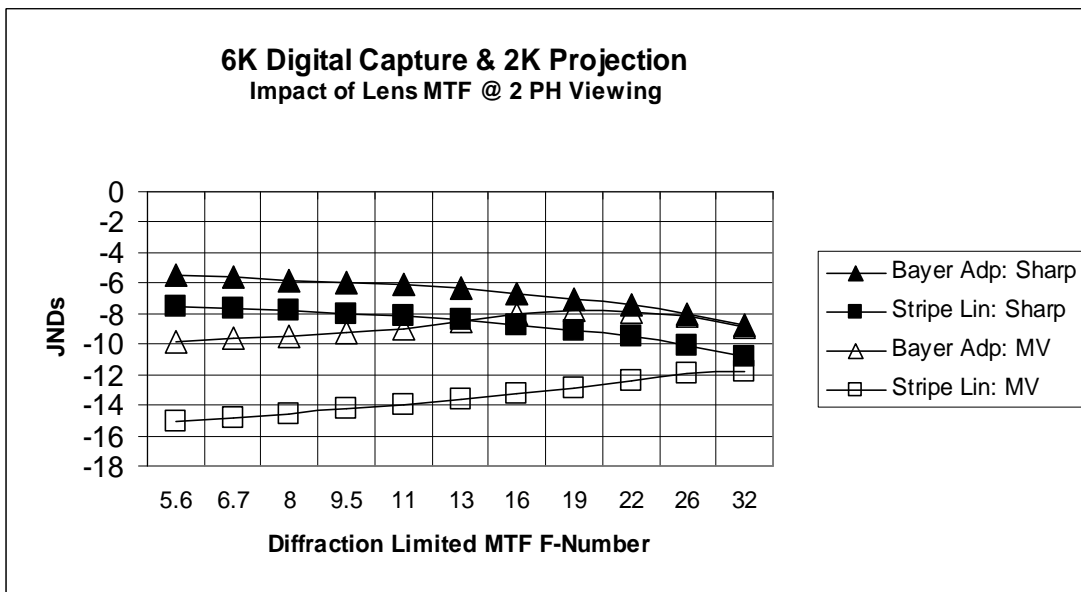


Figure 13: JNDs of sharpness (Sharp) and multivariate quality (MV) for 6K single-sensor Bayer CFA and striped CFA cameras after blurring with taking lens and downsizing for 2K projection.

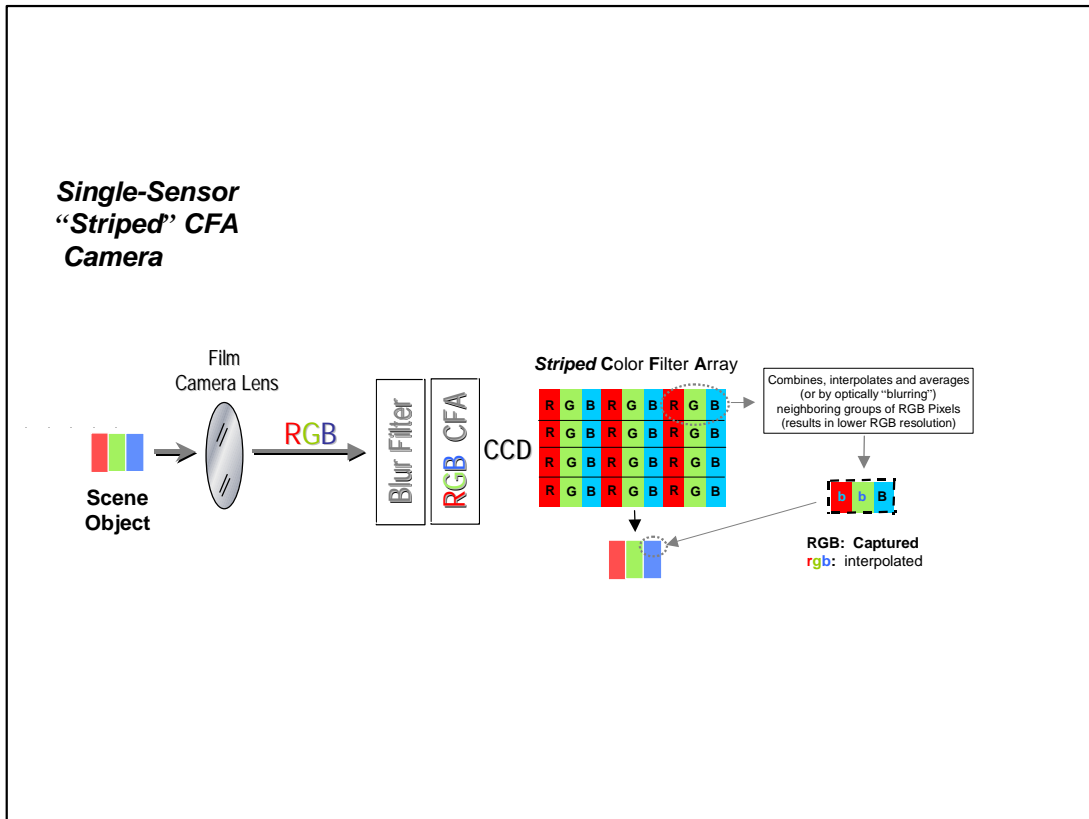


Figure 14: Schematic diagram of a single sensor CFA geometry that simulates the sampling of a tri-sensor camera by optically blurring and/or by digitally processing signals from three adjacent R,G,B pixels.

Impact of pixel density for constant sensor size

Another means that is often considered when attempting to reduce sharpness losses and artifacts resulting from the unequal RGB sampling of CFA sensors is to increase the pixel density (i.e., resolution) of the sensor while maintaining the same sensor size. However, this will have a negative impact on other image capture attributes, such as a reduction in dynamic (exposure) range and sensitivity (ISO Speed), assuming comparable dark current and noise. For a constant sensor size, increasing the pixel density to achieve higher image resolution results in smaller pixels, and the smaller pixels will have a lower full well capacity (maximum signal count) and a smaller light capture area. Figure 15 depicts these trade-offs between resolution, sensitivity, and dynamic range as a function of pixel density, for a constant sensor size.

Sensor Dynamic Range & Resolution Trade-off

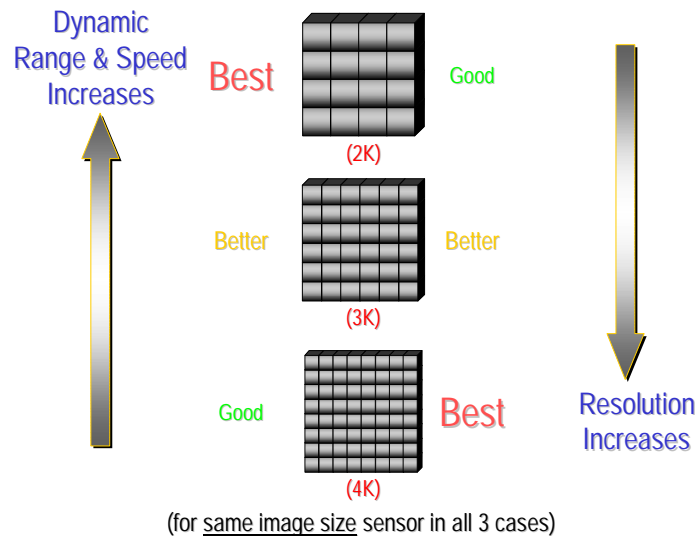


Figure 15: Impact of pixel density (resolution) change at constant sensor size.

The sensor size could be made larger to accommodate a larger pixel, but increasing the sensor size beyond the 35 mm frame size would require a different set of camera optics (e.g., nonstandard film camera lenses) to maintain similar depth-of-field for an equivalent field-of-view. One possible approach to achieving improved dynamic range and sensitivity without increasing the pixel/sensor size is to produce CCD/CMOS sensors with lower noise levels. The noise level is particularly important for these devices because they exhibit a linear response to light exposure, unlike film which has a nonlinear (logarithmic) response to exposure. Because of the nonlinear response, film can achieve a greater exposure range for an equivalent signal-to-noise ratio (SNR) as compared to linear detectors.

As discussed, CFA single-sensor designs that approximate the 35 mm film size will have a higher dynamic (exposure) range than an equivalent resolution conventional 2/3-inch video sensor, which exhibits a smaller pixel size. In addition, there is also an MTF advantage from the camera lens performance because of the increased pixel pitch (samples per unit length) of the larger size sensor format as shown in Figure 17.

Figure 16 demonstrates why a sensor exhibiting a “native” logarithmic (or power function) type response to light produces a wider exposure range for a given pixel “full well” (or dynamic range) than a sensor exhibiting a linear response. Consequently, at a constant full well capacity, the linear type sensor requires a significantly lower noise level to achieve an exposure range equivalent to that of a device with logarithmic response. Unlike motion picture camera negative film, which has a native response that is logarithmic, the CCD and CMOS sensors incorporated in current digital cinematography cameras all possess a native linear-type response to light.

Although these electronic cameras provide data output that is logarithmic, it is obtained by processing the linear output of the CCD or CMOS sensor with a logarithmic (or power “gamma”) transfer function at some later stage in the camera electronics or firmware. The primary benefit of this processing is minimization of contouring artifacts caused by insufficient quantization levels (due to low bit-depth), particularly at lower exposures levels. This processing does not duplicate the extended exposure range that would be obtained by capturing the image with a sensor having a native response that is logarithmic. Reference 12 describes a novel sensor design with non-linear response that provides extended dynamic range.

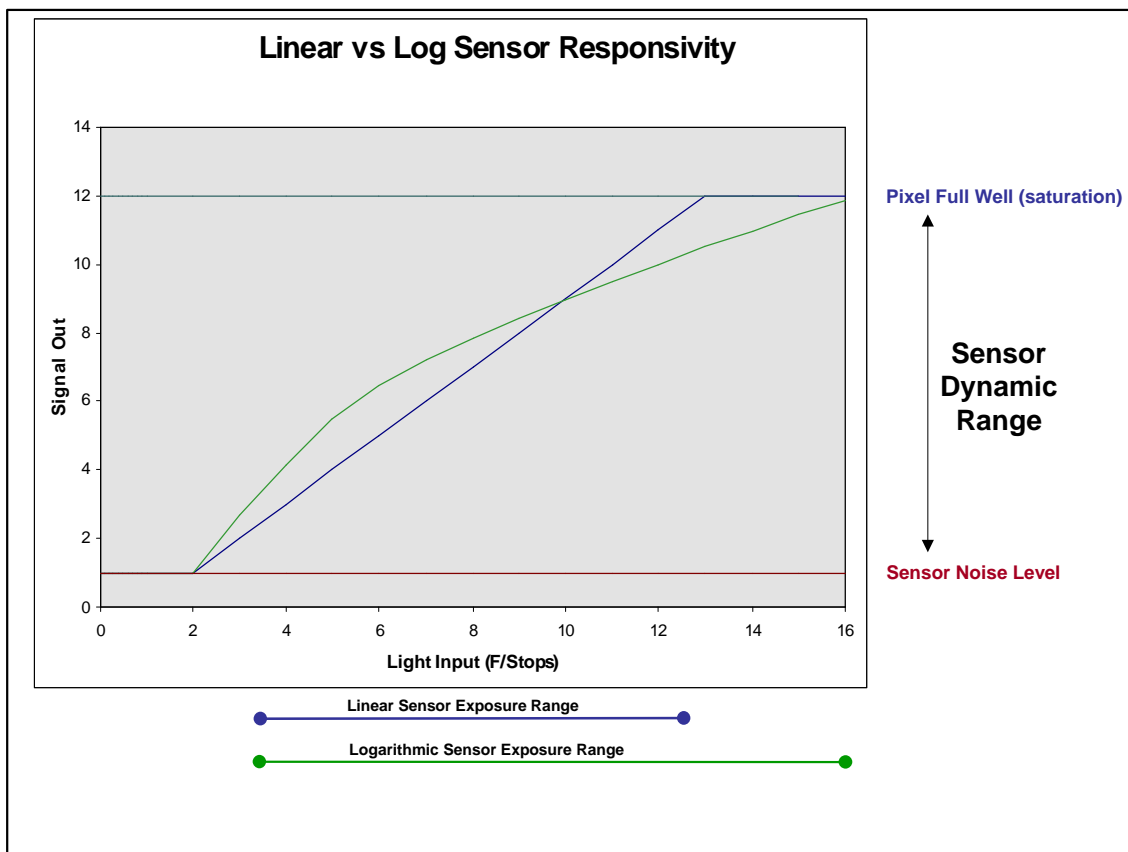
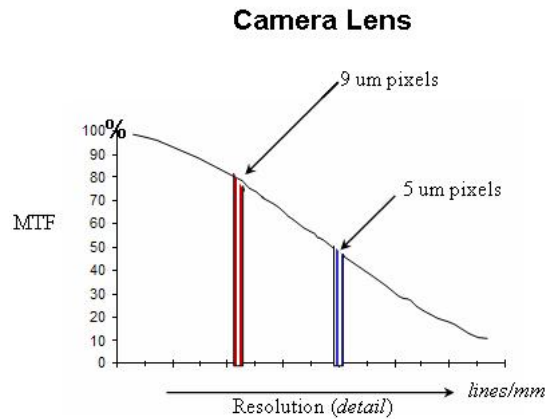


Figure 16: A sensor exhibiting a native logarithmic-type response produces a wider exposure range for a given pixel “well capacity” (or dynamic range) than a linear one, even when the former possesses a lower signal-to-noise level.

Large Frame CCD Sensor (e.g. 35mm size)



Small Frame CCD Sensor (e.g. 2/3-inch HD video)



NOTE: Image Detail for the same number of pixels is produced *sharper* with the larger pixel size after going through the lens.

Figure 17: MTF image sharpness advantage of a 35 mm size sensor as compared to a 2/3 -inch video sensor for a typical lens.

Summary

This paper quantifies the image quality trade-offs associated with digital cinema cameras employing tri-sensor and single-sensor designs. When evaluating different systems, we assume that the color, tone, and noise characteristics are optimal and equal, and we employ the same high quality lenses, digital sampling, and digital projection components in each case. As a result, the quality differences reported herein arise solely from variations in sensor design, spatial interpolation method, and the effect of anti-aliasing filters.

We have shown that images captured with single-sensor designs that use a Bayer CFA with adaptive interpolation can approach the sharpness of images captured with a tri-sensor camera. However, images captured with the single-sensor cameras are far more likely to contain undesirable aliasing, reconstruction, and CFA artifacts, leading to lower overall quality than tri-sensor cameras. Increasing the capture resolution from 2K to 4K provides only a modest benefit with single-sensor cameras, because CFA artifacts, rather than resolution, are driving the image quality.

This result may seem surprising, particularly given the wide spread acceptance of Bayer CFA sensor solutions for consumer still photography, but one must only consider the far more demanding reproduction magnifications and viewing angles encountered in typical motion picture viewing environments. For example, the classic consumer print of 4 x 6 inches, viewed at typical distance of 16 inches, corresponds to a motion picture viewing distance of four picture heights. This is at least twice the viewing distance where critical motion picture quality assessments are generally made.

Finally, it is tempting to look at the results produced by the 4K tri-sensor camera with 4K projection and conclude that this is the ultimate system. When considering this, it is important to remember that we assumed ideal color, tone, and noise characteristics in our analysis, but this is not yet the case in practice. Moreover, all of our analysis assumed a two picture-height viewing distance, and additional capture fidelity (more than 4K resolution) may be beneficial at higher projection magnifications and/or closer viewing distances.

References

1. R. B. Wheeler, "Use of System Image Quality Models to Improve Product Design," IS&T PICS Conference, 2000.
2. B.W. Keelan, "Characterization and Prediction of Image Quality," IS&T PICS Conference, 2000.
3. B. Keelan (ed.), Handbook of Image Quality: Characterization and Prediction, Marcel Decker, Inc., New York, 2002.
4. ISO/DIS 20462-1, ISO TC 42/SC /WG 18, International Standard – "Photography - Psychophysical experimental methods to estimate image quality – Part 1: Overview of psychophysical elements," November 15, 2005.
5. ISO/DIS 20462-3, ISO TC 42/SC /WG 18, International Standard – "Photography - Psychophysical experimental methods to estimate image quality – Part 3: Quality ruler method," November 15, 2005.
6. R. Thiel, P. Clark, R.B. Wheeler, P.W. Jones, M. Riveccie, J-F Dupont, "Assessment of Image Quality in Digital Cinema Using the Motion Image Quality Ruler Method," SMPTE Technical Conference, October, 2006.
7. G. Fielding and P.W. Jones, "Temporal and Chromatic Sampling Artifacts in Digital Intermediates," SMPTE Technical Conference, October, 2006.
8. L.L. Thurstone, "A Law of Comparative Judgment," *Psych.Rev.* 34, 1927, pp. 273-286.
9. B.E. Bayer, "Color Imaging Array," U.S. Patent No. 3971065, 1975.
10. J.E. Adams Jr. and J.F. Hamilton, "Adaptive Color Plane Interpolation in Single Sensor Color Electronic Camera," U.S. Patent No. 5652621, 1997.
11. D.R. Cok, "Signal Processing Method and Apparatus for Producing Interpolated Chrominance Values in a Sampled Color Image Signal," U.S. Patent No. 4642678, 1987.
12. C. Parks, "Extended Dynamic Range Image Sensor with Fixed Pattern Noise Reduction," U.S. Patent Appl. No. 10/809073, 2004.