

Digital Imaging Systems Evaluations: Matching the Analysis to the Imaging Requirements

M. A. Kriss, Consultant, Camas

Abstract

Digital imaging systems are now stable, if not mature. CMOS imaging sensors are closing the quality gap between themselves and CCD imaging sensors, and the buried multiple PIN-diode imaging sensors, Foveon X3, brings a potential of artifact free images. While the pixel count will always be the major image quality predictor, issues about speed, dynamic range, noise, sampling and compression artifacts, color fidelity, automatic white balance, and in-camera enhancement still remain and must be considered on a system-by-system basis. Uses can be classified into the following (but non-inclusive) categories: amateur; professional; graphic arts; catalogue; surveillance; satellite; astronomical; microscopy; military; real estate; insurance; etc. This paper will outline how to develop, by means of simple models, the most meaningful measures of image quality for different imaging uses by starting with the basic physics of the required imaging sensor, artifact suppression, speed and dynamic range, etc. The result presented will represent the concept, implementation and preliminary image quality results for a prosumer digital still camera.

Introduction

Over the last 12 years (between the introduction of the Apple Quick Take digital camera and the current six or more mega-pixel digital still cameras) the image evaluation tools developed primarily for conventional film systems have expanded to cover digital images as has the new image quality parameters associated with them. Conventional film systems relied on resolution (line-pairs per millimeter), the Modulation Transfer Function (MTF), RMS granularity and power spectrum, ISO speed, dynamic range, exposure latitude, toe contrast, and color reproduction accuracy and preference to define system quality. There were attempts to define system quality by use of SMT/CMT Acutance [1], System Quality Factor [2], Information Capacity [3,4], and experimental combinations of noise and sharpness [5]. The effect of non-linear nature of color films using Developer Inhibiting Releasing (DIR) Couplers on sharpness and color was explained as was how MTF measurements depended on both exposure and target modulation [6]. All these measurements and others provided the conventional film builder with a reliable set of measurements to evaluate a proposed film system. In some cases image simulations (by means of digital image processing) were used to predict the quality proposed systems[7,8].

The advent of digital imaging has introduced a series of new factors that influence image quality. All digital still cameras are sampled systems, which introduce aliasing. When the spatial frequency content of the image exceeds half of the sampling rate (Nyquist frequency) of the image sensor (the inverse of distance between pixel centers) the frequencies above the Nyquist frequency will be reproduced as low frequency artifacts and banding. The aliasing problems are compounded when a color filter array (CFA) is used to encode color on a single sensor chip. The aliasing artifacts and banding become colored and are pronounced. The use of a CFA also requires that some form of

interpolation algorithm be applied to the sampled image to form a viewable R-G-B image. The interpolation algorithms can introduce additional color artifacts into the image as well as a loss of sharpness [9]. Most digital cameras have some form of image enhancement and color corrections (white balance required by different taking illuminants, color corrections due to "poor" spectral sensitivities, etc.), which must be taken into consideration when evaluating image quality. An additional source of artifacts is the nature of the compression used to store the image. A six mega-pixel camera requires at least six to 18 mega bytes to store the image without compression. Most cameras use JPEG compression that utilizes 8-by-8 arrays (transformed from R-G-B to luminance and chrominance channels, YCC), Direct Cosine Transforms and quantizers that reduce the required storage to about one mega byte or lower. The greater the compression, the greater the tendencies for the basic 8-by-8 arrays to introduce blocking artifacts. Digital cameras will continue to add noise (due to point-to-point variations in the dark current) during prolonged exposures, thus the image quality (speed, dynamic range and exposure latitude) will vary with exposure time. All the above deals with the digital camera, but the nature of the printing device is equally important for overall image quality (sharpness and color). The impact of digital halftones on image quality of the final print is also very important and if a digital scanner scans the halftone print, moiré effects can be introduced.

Just as a series of image test targets were developed for conventional film imaging, a new set of targets has been introduced for digital photography including edge-trace MTF calculations, noise measurement patches and color reproduction charts. However, these targets cannot suffice for all imaging applications. A series of systems analysis have been developed for digital images using the Human Visual System and advanced theories of Color Constancy [10, 11, 12, 13]. While each of these quality evaluations addresses some of the issues of quality, none of them are universal. This requires one to consider a specific collection of quality measures (metrics) for any given imaging task. Taking this approach should lead to the development of a richer set of image quality metrics. To summarize, I will use a quote from Peter Engeldrum "The Universal Image Quality Model is not on the horizon [14]."

Combining Simulation and Quality Analysis

It is both costly and time consuming to develop digital imaging prototypes. Desktop computing power makes it feasible to simulate in an afternoon [15] what it took weeks to do in 1971 [7] and 1983 [8]. With suitable software (home grown packages or products like MatLab and Mathematica) it is possible generate "final images" from a proposed imaging system along with analytical assessments of the image quality (quality metrics). This paper will focus on the image quality metrics based on the imaging system's characteristics rather than any given image. However, a complete study of any proposed imaging system requires both.

Pro-consumer Digital Still Camera Study

From a marketing point of view consumer digital still cameras has become a mega-pixel war with the imaging pixels getting smaller and smaller. Smaller pixels may have better MTFs based on pixel geometry, but they tend to have reduced “photographic” speed, dynamic range and exposure latitude. Also, smaller pixels place a greater strain on the lenses where the lens point spread function must be smaller than (or equal to) the dimension of the pixel. In short, there are many system parameters that must be optimized to meet some quality criteria for the imaging system.

Final Print Quality Criteria

In this analysis the concept of CMT Acutance [1, 16] will be used to establish the sharpness criteria for the final print (4 inch by 6 inch viewed at approximately 16 inches or four viewing heights). Calibrated studies indicate that a CMT of 92 or above will produce a high quality (photographic print) when free of artifacts and visible noise. It is assumed the same will be true for a digital image printed onto photographic paper. One must now consider the various system parameters that will ensure a system CMT of 92 under the given printing and viewing conditions.

Designing the Consumer Camera

I will start by building a black and white Frame Transfer CCD camera that gives the required CMT Acutance of 92, and then move to an Interline CCD and then to a full color Interline CCD, adjusting the characteristics of the camera to retain the required quality and adjust the other parameters such as the optical blur filter and fill factor value to lower the impact of aliasing.

A full frame black and white camera that has 1300 lines and 1950 pixels per line (a 1.5 aspect ratio for a 2.535 mega-pixel camera) using 5 micron pixels and an $f/5.6$ diffraction limited lens will have a CMT value of 92.3 when printed on photographic paper by a laser printing device. This camera will show some signs of aliasing in that about 10% of the signal (high frequencies) is folded back into the region below the Nyquist frequency (5.12 c/mm on the paper). This means that images that contain a lot of periodic structure in them will introduce neutral banding or artifacts. To reduce the aliasing, a bi-refringent optical blur filter is used, which displaces the now split image a precise amount, thus reducing the aliasing. If one shifts the image one-half pixel (2.5 microns) the aliasing drops to about 3% of the signal, but the CMT Acutance drops below 92. If one uses a better lens, say an $f/3.5$ diffraction limited lens, one regains the sharpness, but the aliasing approaches the 10% level due to the added high frequency information. The best way to get back the lost sharpness is to increase the optimized sharpness enhancement built into the camera (a form of un-sharp masking). Now consider an Inter Line Transfer version of the camera because it is cheaper. Assume that the fill factor is 0.5. This results in a higher CMT value due to the higher MTF of the smaller pixel, but the aliasing increases as well. The optical blur filter is already optimized; hence one can lower the lens quality to that of an $f/8$ diffraction limited lens. This gives a CMT in excess of 92 and lowers the aliasing to 2.6 %.

The next step is to introduce a CFA (Bayer CFA) to make a color camera. In doing so, one just chooses an interpolation algorithm to

fill in the holes in the red, green and blue images. For the sake of this discussion a simple bi-linear interpolation is used (which will introduce some edge artifacts). Using the CFA reduces CMT Acutance to about 87.5 (a good but not excellent print), but the aliasing has grown to over 68%; that is, the aliased signal is 68% of the non-aliased signal. This results in serious color artifacts and banding. By increasing the blur filter shift to one full pixel, 5 microns, the aliasing is reduced to 40%, but still too high and the CMT Acutance value drops to about 85 (the edge of being a good print). To solve this problem one needs to increase the number of pixels. This can be done in two ways. One can lower the pixel size (keeping the sensor size fixed) or making the sensor larger with the same 5 microns pixels. If one reduces the pixel size to 3.25 microns, increase the sensor to 2000-by-3000 pixels (6 mega pixels), and slightly raise the enhancement level in the camera one gets a CMT Acutance of 92.5 and about 12.5 % aliasing. On the other hand, if one keeps the pixel size at 5 microns and increases the sensor size to a 2000-by-3000 device, the CMT Acutance is 94 and the aliasing is 28% (at the lower enhancement level). However, if the camera lens is degraded to an $f/11$ diffraction limited lens, the CMT Acutance is 92 and the aliasing drops to 15.3%. If one enlarges the pixels to 10 microns and decrease the lens quality to an $f/22$ diffraction limited lens one retains the CMT Acutance of 92 and 15% aliasing. The role of the camera lens is critical. For larger pixels, the effective MTF of the lens (in the print plane) is much greater than for small pixels (where the boundaries of the point spread function start to fall well outside the pixel). This means that it effectively “sees” more high frequency content and thus will show more potential for aliasing. Hence as the pixels get larger, a “poorer” lens is required to prevent aliasing, which is a very counter intuitive concept.

Based on the above one starts to see some critical trade-offs in design. It seems that making the pixels smaller gives slightly better results, but one has not yet considered the impact on speed, dynamic range or exposure latitude of the smaller pixels.

So far the focus has been on image sharpness quality, but one needs to consider speed, dynamic range and exposure latitude. Just as one needed some criteria for sharpness quality, one needs some criteria for speed, dynamic range and exposure latitude. For a color camera the following requirements will be set: an ISO speed for 100; exposure latitude of 1000:1; a dynamic range of at least eight bits or 256 clearly measurable levels (48 db).

The most important factors in the speed of an imaging sensor (CCD capacitor or photodiode) are the quantum efficiency and noise characteristics of the sensor. The major factors in the exposure latitude are the size of the sensor (active area), and the full well capacity, which is defined by the number of electrons that can be stored per unit area per volt. The dynamic range is defined by the full well capacity and the noise characteristics. CCDs can be illuminated from the front or rear, while Interline and CMOS sensors, which use photodiodes to image and store electrons, are normally front-illuminated.

It will be assumed that only front illumination takes place in a pro-consumer camera. Consider the fate of a photon. As it impinges on the sensor some of the light is reflected based on the index of refraction change at the surface of the sensor; multilayer coatings and reduce this reflection, but for the sake of discussion assume that 10% of the light is reflected, thus the maximum quantum

efficiency might be 90% (this number is of course wavelength dependent). In the case of a CCD there is a SiO₂ insulation layer that tends to absorb blue light more than green or red light. If a ploy-silicon is used for electrical contacts, it will also preferentially absorb blue light. Once in the silicon substrate, the blue light is absorbed first, creating electron-hole pairs. The electrons are captured quickly in the depletion layer near the surface of the silicon. The green and red light pass deeper into the silicon substrate before being absorbed and creating electron-hole pairs. The electrons generated deep within the silicon will diffuse up to the depletion layer where they are captured. However, as they diffuse they might recombine with a hole and be lost. Also, the diffusion back to the depletion layer smears the green and red images. Hence, when all is said and done, the far red and blue quantum efficiencies are less than the green and the green and red "physical" MTFs are lower than the blue MTF (these MTF losses were not included in the above analysis, but would lower CMT Acutance values cited). A CCD capacitor might have a maximum Quantum Efficiency of about 80% (but normally lower) and a CMOS or Interline device might have a photodiode with a maximum Quantum Efficiency of 90% (but normally lower).

The noise in a sensor is complex. The most fundamental noise is the Shot noise, which is equal to \sqrt{N} where N is the number of electrons stored in the potential well of the active pixel. The other noise sources are associated with the spatial variations in dark current, which is always being generated within the device, variations in transistor gains (pattern noise associated primarily with CMOS devices) and various readout noise such as kTC noise due to stray charge on the output capacitors. The kTC noise is greatly reduced by double-correlated sampling. The speed of a sensor is defined (by one measure) as the signal-to-noise ratio in the toe for the characteristic response curve [17]. Shot noise is not important here, but any residual kTC noise, pattern noise and dark current is critical. The signal is a function of the Quantum Efficiency, the light capture area (using micro-lenses in most cases) and the exposure time. The dark current is a function of the exposure time and the area of the sensor (larger areas will inject more dark current noise electrons). All of these issues have been modeled in the literature [18, 19]. One needs to specify the exposure time (due to dark current generation) when calculating the speed of a given sensor; for this paper an exposure time of 0.01 seconds will be used. The speed point will be defined as the exposure at which the signal-to-noise level is 30. The ISO speed will be defined as $Speed=0.8/E$, where E is the exposure (in Lux-seconds) for a signal-to-noise ratio of 30. This is equivalent to the early photographic speed measures and thus helps one relate to earlier film practices.

First lets assume that one has a 3.25 micron pixels used in the sharpness study above. Further lets assume that one has a PIN Photodiode with a peak Quantum Efficiency of 80% (at about 550 nanometers). It will be further assumed that there are a base noise of 15 RMS noise electrons and a dark current of 0.5 nano-amperes per cm². The speed calculation gives a value of about ISO 55, which is slow. If one uses the 5 micron pixels, the speed becomes ISO 136, but without a color filter array. With the filter array (green) the speed drops back to ISO 61. To achieve the ISO of 100 one needs to enlarge the pixel size to 6.5 microns. By lowering the base noise to 5 RMS electrons (this does not include dark current noise) one can use a 6 micron pixel to achieve a speed of ISO 100 with a color filter array. Note that if the exposure time increases to

0.1 seconds that the ISO speeds drops to 84 and at a 1 second exposure it fall to and ISO speed of 48. This drop in speed as a function of exposure is due to the increase in noise due to the dark current. Unlike conventional films, digital still cameras have an exposure dependent ISO speed. This is reflected in how the programmed exposure profiles are set up. For lower light levels, the exposure will be fixed at, say, 1/30 of a second and the aperture will opened to meet the get the required exposure.

Now consider the issue of exposure latitude and dynamic range. The model used to calculate the ISO speed uses a signal-to-noise ratio of 10 to determine the first usable image signal and this defines the lower boundary of the exposure latitude. The upper range of the exposure latitude is determined by the full well capacity of the active area of the pixel. From above one has a 6 micron pixel ($36 \times 10^{-8} \text{ cm}^2$), which with the aid of a micro-lens can gather in all the light impinging on the pixel. However, one has used a fill factor of 50%, which means the full well capacity will be based on an area of $18 \times 10^{-8} \text{ cm}^2$. The exact nature of the PIN photodiode fabrication and applied reverse bias voltage will determine the full well capacity, but a reasonable number is about 2×10^{11} electrons/cm². This will result in a full well capacity of about 35,000 electrons and an exposure range of 2.55 Log Exposure or 354:1. This is considerably less than the required 1000:1. Hence, one is again forced to increase the size of the pixel and/or increase electron density by means of changes in fabrication or higher gate voltage. Lets assume that one can get 3×10^{11} electrons/cm² and increase the overall pixel size to 8.57 microns. This will give the required 1000:1 exposure latitude and a full capacity of 108,000 electrons per pixel. One must now recalculate the impact of this larger pixel on the CMT Acutance and potential for aliasing. The adjustments to the camera require an f/16 diffraction limited lens, which retains a CMT value above 92 and introduces a potential for aliasing of about 21%.

One has now met all the requirements set forward with the exception of the dynamic range of the sensor. There are several ways to calculate the dynamic range of the sensor, but the commonly used one is to divide the full well potential, 108,000 electrons, by the lowest noise level (including dark current), 7 RMS electrons. This would give a dynamic range of 20 Log (108,000/7) db or about 84 db. This would be equivalent to 14 bits when an A/D converter is used to turn the analog signal into a digital one. A more exacting method might be to start with the full well capacity and subtract the square root of the number (Shot noise) to get to the next "measurable" level: $108,000 \rightarrow 107,671$ and continue this process until one reaches the lowest noise level of 7 RMS electrons. If one takes this approach one gets 651 levels or 565 db or between 9 and 10 bits. Even this more conservative measure meets the state requirements of 8 bits.

Using the above results (8.57 micron pixel) with a 2000-by-3000 sensor with a Bayer CFA, one can realize a CMT Acutance of 92 by using a f/16 diffraction limited lens and little more sharpness enhancement, but one is left with a potential of 21% aliasing. The question now arises, what does 21% potential aliasing mean. In the analysis of aliasing it is assumed that the "image" is a flat spectrum where all spatial frequencies are equally represented. Hence, the 21% represents a worst case scenario or the potential for aliasing. The actual aliasing is scene dependent ranging from an image of a uniform wall with no aliasing to an image of a black and white test chart containing vertical and horizontal bars of

different frequencies, which will show low frequency, color banding. Figure 1 shows the cascaded MTFs of the camera lens, the optical pre-filter and the geometric sensor. For this camera system the Nyquist frequency (in the plane of the paper) is approximately 5 cycles per millimeter. The entire signal beyond the Nyquist frequency will be “folded” back into the region below the Nyquist frequency causing aliasing and color banding (due to the CFA). The actual visual impact will be modified by the subsequent interpolation MTF, the enhancement MTF, the printing MTFs and the visual MTF. It is clear from Figure 1 from where the 21% aliased signal comes. However, “real scenes” do not have flat power spectra and the question to be resolved is “for a normal scene, what is the power spectrum and how much information is beyond the Nyquist frequency?” Figure 1 also shows a “typical” scene spectrum [20] ($1/(1+5 f)$, where f is the spatial frequency). The amount of signal available for aliasing is now greatly reduced, however this can vary greatly when man-made structures with strong vertical or horizontal structure content are imaged with good lenses. Hence, normal photography will not introduce strong aliasing with the given camera design, but when the camera is used with a lens with a MTF greater than the $f/16$ diffraction limited response and images with strong periodic structure are captured, one should expect color banding and artifacts.

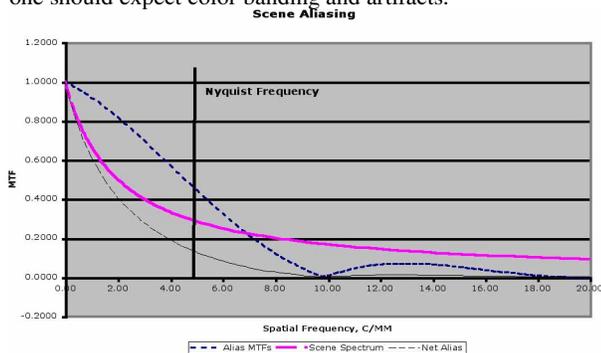


Figure 1. A “typical” scene spectrum cascaded with the MTFs of the camera lens, optical pre-filter and sensor.

As a “reality check” images were taken using a 6 mega-pixel Nikon D70 that has the Bayer CFA and about 7.5 micron pixels. Nikon claims an ISO speed of 200, but the metering is based on an ISO speed of 100, so they may be using electronic gain to get ISO values of 200 to 1600. The images show aliasing at high frequencies when using test charts and can be seen in the fine details of buildings, which is in agreement with the above. As another check, Kodak makes a full frame CCD camera back for a Leaf camera. There are 6726-by-5040 7.2 micron pixels and a Bayer CFA is used. Using the same “fabrication technology” as above, such a camera (for a 5-by 6.666 inch print viewed from 20 inches) gives a CMT Acutance of 98 with a potential for aliasing of about 5%. If the lens quality is increased to $f/5.6$ diffraction limited lens (a very good lens), the CMT Acutance increases to 99 and the potential for aliasing increases to 15%. Hence, one can expect some aliasing issues when studio scenes have fine detail (like the fine detail in a fabric). Taking into consideration that a Full Frame CCD will have a lower QE (as outlined above) and that narrower CFA filters might used for better color reproduction (at the cost of speed), the ISO speed can drop to about 61 which is close to the base speed of 50 given in the literature. The literature quotes exposure latitude of 12 stops or roughly 4000:1, while the calculations give about 11 stops or about 2000:1. The full well capacity should be near 150,000 electrons for a dynamic range of

about 768 levels or between 9 and 10 bits using the conservative method outlined above. Assuming about 10 RMS electrons base noise, the output would be set between 14 and 15 bits using the more standard measure.

Conclusions

As stated by Engeldrum [14], there may not be a universal quality metric. Furthermore, it may not be a practical or even theoretical goal when trying to evaluate digital imaging systems. It is better, in the author’s opinion, to clearly state the required specifications for any given imaging system (based on usage) and then to use simulations, analytical models and system analysis to hone in on the required physical attributes. This short paper has taken this path for a pro-consumer digital still camera with the focus on sharpness, aliasing, ISO speed, exposure latitude and dynamic range. The results are reasonably consistent with known camera systems, and using more advance physical models for CCDs and CMOS sensors one can refine the modeling. The issues surrounding choice of CFAs, interpolation algorithms and artifact classification (for visibility) and color were not attempted due to space limitations.

References

1. M.A. Kriss, Theory of the Photographic Process, 4th Edition (T. H. James, Editor), Macmillan Publishing co. Inc. New York, 1977, pp. 592-635.
2. E. M. Granger and K. N. Cupery, *Photogr. Sci. Eng.* **7**, p.173 (1963)
3. R. C. Jones, *J. Opt. Soc. Am.* **48**, p. 425 (1959).
4. M. A. Kriss, J. O’Toole and J. Kinard, *Image Analysis and Evaluation*, (R. Shaw, Editor), SPSE Conference Proceedings, p. 122 (July 19-23, 1976, Toronto, Canada).
5. Bartelson, C.J. *J. Photogr. Sci.* **30**, p. 30 (1982)
6. M. A. Kriss, C. N. Nelson and F. C. Eisen, *Photogr. Soc. Eng.*, **18**, p.131 (1974).
7. M. A. Kriss, N. Nail and R. Mickelson, *J. Opt. Soc. Am.*, **61**, p. 644 (Abstract, 1971).
8. M. A. Kriss and J. Liang, *SMPTE J.*, **92**, p. 804 (1983).
9. R. Ramanath, W.E. Snyder, G.L. Bilbro, and W.A. Sander III, *J. Electron. Imaging*, **11**: 3, p306 (2002).
10. S. Daly, *Digital images and human vision*, MIT Press, Cambridge, Mass., pp 179-206 (1993).
11. M. Fairchild, *Color Appearance Models, Second Edition*, Wiley-IS&T Series in Imaging Science and Technology, West Sussex, England (2005)
12. M.D. Fairchild and G.M. Johnson, *Journal of Electronic Imaging*, **13**, p. 126 (2004)
13. X. Zhang and B.A. Wandell, *SID Digest of Technical Papers*, **27**, p. 397 (1996).
14. P. Engeldrum, *PICS 1999: Image Processing, Image Quality, Image Capture, Systems Conference*, Savannah, Georgia, April 1999, Volume 2.
15. J. E. Farrell, F. Xiao, P. Catrysse, B. Wandell, *In Proceedings of the SPIE Electronic Imaging 2004 Conference*, Vol. 5294, Santa Clara, CA, January 2004
16. M. A. Kriss, *Handbook of Photographic Science and Engineering*, Second Edition (C. N. Proudford, Editor), IS&T, Springfield, Virginia p. 645, (1997).
17. M. K. Kriss, *Proceedings of the International Congress on Imaging Science*, Sept. 7-11, 1998, Antwerp, Belgium, p. 341.
18. A. J. P. Theuwissen, *Solid State Imaging with Charge-Coupled Devices*, Kluwer Academic Publishers, Dordrecht, The Netherlands (1995)
19. J. R. Janesick, *Scientific Charge-Coupled Devices*, SPIE Press, Bellingham, Washington (2000).
20. G. J. Burton and I R Moorhead, *Applied Optics*, 26(1),p.157 (1987).

Biography

Dr. Kriss received his B.A. M.S. and Ph.D. in Physics from the University of California at Los Angeles. He joined the Eastman Kodak Research Laboratories (KRL) in 1969 where he focused on the understanding and simulation of the image structure of color films. Starting in 1979 he turned his attention to image processing of film and digital images. He headed the Image Processing Laboratory in the Physics Division of KRL. From 1985-1988 Dr. Kriss helped establish a Kodak research presence in Japan. Upon returning in 1988, Dr. Kriss headed up external research for the Physics Division and took over the Image Algorithm Laboratory in 1990. Dr. Kriss retired from Kodak in late 1992 and assumed a consulting job with the Center for Electronic Imaging System (CEIS) at the University of Rochester. Dr. Kriss was later to assume the role of Executive Director of the CEIS and as an Adjunct Professor taught courses on digital imaging in the Department of Electronic and Computer Engineering. Dr. Kriss joined Sharp Laboratories of America in late 1999 as the Manager of the Color and Imaging Group. He retired in 2004. Dr. Kriss is a Fellow of the IS&T and received the Davies Medal from the Royal Photographic Society in 1999. Dr. Kriss continues to be active in IS & T and SPIE conferences and gives short courses on digital imaging. He is the Editor of the Wiley-IS&T Series on Imaging Science and Technology.