Still Photography Throwdown: Silver Halide vs. Silicon

Mark D. Fairchild, Rochester Institute of Technology, Munsell Color Science Laboratory, Rochester, NY/USA

Abstract
Still photography has a rich history of technological innovations to record the light in a scene dating back to the early 19th century (at least). The recent decades have seen these technological innovations create a revolutionary shift in the materials, processes, and uses of still images. Most of the still photography world has completed a move from silver halide (AgX) technology that dominated the field for over a century to digital still cameras based on silicon (Si) sensors that arrived on the scene about two decades ago and have essentially supplanted AgX capture technologies in most applications. This research ponders the question of whether the image quality obtained with Si has also surpassed that of AgX in the context of typical consumer and professional photographic prints and soft displays. Four camera systems, two digital and two film based, were evaluated using both objective image quality metrics and psychophysical evaluation of prints and displayed images. The results show that a high-end digital SLR does indeed produce better images than an equivalent 35mm film system, but that a typical digital point-and-shoot camera has substandard quality that can be somewhat attributed to "too many megapixels" and "too much post processing" for the lens capabilities and sensor size. The conclusion is that indeed film is done, but there remain significant areas for improvement in digital systems. In particular improvements in printing techniques, lens-sensor matching, and noise reduction are called for.

Introduction

> **throw•down** |ˈθrōˌdoun|
> noun informal
> a performance by or competition between rappers, breakdancers, etc. *: a funky hip-hop throwdown.*[1]

Photographers have long been an interesting and innovative combination of artists, technicians, engineers, and scientists in order to both create compelling images and manage the systems required to do so. This is just as much the case today as it was in the mid-19th century when photographers like Timothy O'Sullivan, William Henry Jackson, John K. Hillers, and their contemporaries used large-format (up to 20x24 inches) cameras, glass plates, and the wet collodion process that required the plates to be coated with a mixture of collodion and potassium iodide, immersed in a sensitizing solution of silver nitrate, then exposed, processed and fixed before the plates dried; all in the field.[2] Today great photographers are also masters of color management, file systems, device characterization, image processing, and publishing. And photographic assistants have job titles such as *Digital Technician*.

Imaging scientists and photographers alike have debated the advantages and disadvantages of AgX and Si systems for as long as digital systems have been commercialized (and probably will for decades more) with the fundamental question being when will digital replace film. Of course, the answer has always depended on the application, but it appears that the day has come when digital systems have effectively replaced film systems in virtually all still photography applications. That presents the next question of whether or not digital systems actually produce comparable image quality, or has image quality taken a back seat to economics and convenience? This paper aims to provide a little more fuel for the fire by performing a comparison of film and digital systems. The application can be considered high-end consumer snapshots or low-end professional images. This market segment is often referred to as the *prosumer* market and is the place where art and technology often meet with the most interesting and revolutionary results.

Systems Evaluated
Four still-photographic systems were selected for this comparison. Two were film systems based on professional 35mm and medium format cameras with professional negative film and processing. The other two were digital systems. One a fairly recent, but not current, professional-level digital SLR and the other a current low-end consumer point-and-shoot camera. These systems are described below.

Hasselblad 120 (AgX)
The first film system was a Hasselblad 501c medium format (6x6cm negatives) camera with a Carl Zeiss Planar f/2.8 80mm T* lens (considered a *normal* lens for medium format). The film used was Eastman Kodak Co. *Portra 160nc* in the 120 format. This is considered a professional film and is commonly used fro portraiture. The *160* designation is the ISO speed rating and the *nc* designation indicates the natural color version of the film (as opposed to a different version with over saturated colors that are often preferred). Portra 160nc is also considered an extremely fine grain film, typical of its low ISO speed rating. By most professional photographic expectations, this was the "highest quality" system used in this research. The original price of this system around 1990 was about $2000.

The exposed film was sent to a professional photofinishing lab (A&I Photographic and Digital Services in Hollywood, CA) for development and optical printing to 5x5 inch proof prints. Prints were on FujiColor Crystal Archive paper. The same lab also scanned the negatives to produce digital images with a resolution 5078x5074 pixels across the frame and processed to positive images in the sRGB color space. The quality of the digital scanning was such that film grain was visible in the digital images. It is of interest to note that the lab selected was one of the few in the country that could perform both the required development and printing and digital scanning. The preferred local laboratory for professional films has recently gone out of business.

Nikon D2x (Si)
The expected next-best system was the Nikon D2x digtial SLR. This is a high-level professional camera representing the top of Nikon's previous generation. This camera was introduced in 2006. The current equivalent is the Nikon D3x, introduced in 2009. The D2x uses a single DX-format 12.4 Mpixel CMOS sensor. The DX format is smaller than a 35mm frame at approximately 16x24mm and therefore results in traditional 35mm lenses functioning with an effectively-higher focal length. The lens used was a Nikon 17-55mm f/2.8 ED-IF AF-S DX Zoom-Nikkor lens. This is a very high-quality lens with minimal flare and an equivalent focal-length range of 25.5mm-85.5mm for 35mm format. This camera system has been well characterized colorimetrically and used in previous research on high-dynamic-range imaging.[3] While now 4 years old, it still represents the high end of digital photography capabilities. Images were recorded in full-resolution RAW format and post processed using the Apple Aperture software and RAW decoding to produce AdobeRGB images. Prints were made using Apple Inc.'s Aperture print service to represent a typical professional workflow. Prints were on FujiColor Crystal Archive paper. The D2x was set to an ISO speed rating of 160 to match the film cameras. The original retail price of this camera was about $5000 and the lens about $1200.

Nikon F3 135 (AgX)
The second film system evaluated was a professional 35mm SLR, the Nikon F3. This camera represented the state-of-the-art for professional photographers in the late 1980s and the one used in this study was purchased in 1987 and kept in pristine condition. It was paired with the same Nikon 17-55mm f/2.8 ED-IF AF-S DX Zoom-Nikkor lens used on the D2x digital camera. The lens was manually-focussed when on the F3 camera. It was confirmed that this particular DX-format lens did not produce any significant vignetting for the 35mm format. The film used was Eastman Kodak Co. Portra 160nc in the 135 format (the same film used in the medium-format camera). Since the same lens was used, it would be reasonable to expect image quality comparable to the D2x, or perhaps slightly lower since the effective resolution of good 35mm film is on the order of 6 Mpixels, or about half that of the D2x. The original retail price of this camera in the 1980s was about $1200.

The exposed film was sent the same professional photofinishing lab for development and optical printing to 4x6 inch proof prints. The same lab also scanned the negatives to produce digital images with a resolution 5035x3339 pixels across the frame and processed to positive images in the sRGB color space. The quality of the digital scanning was such that film grain was clearly visible in the digital images (moreso than in the medium format scans as would be expected). Prints were made using unmarked paper with properties similar to FujiColor Crystal Archive.

Nikon COOLPIX S220 (Si)
The final camera in the study, and the one with the lowest initial expectations, was a current consumer-level point-and-shoot digital still camera, the Nikon COOLPIX S220 (price about $125). The S220 is notable more for its interesting image processing (*e.g.*, smile and open-eye detection) than its image quality which was generally stated as fairly low due to sensor noise in many reviews. This camera includes a 10 MPixel CCD sensor with a size of 1/2.33 inch. The lens is a 3x Optical Zoom-NIKKOR Glass Lens, f/3.1 - f/5.9. All images from this camera were recorded in sRGB using the highest resolution and best quality JPEG format available. Prints were made using Apple Inc.'s iPhoto print service to represent a typical consumer workflow. Prints were on FujiColor Crystal Archive paper. The S220 has an automatic ISO setting.

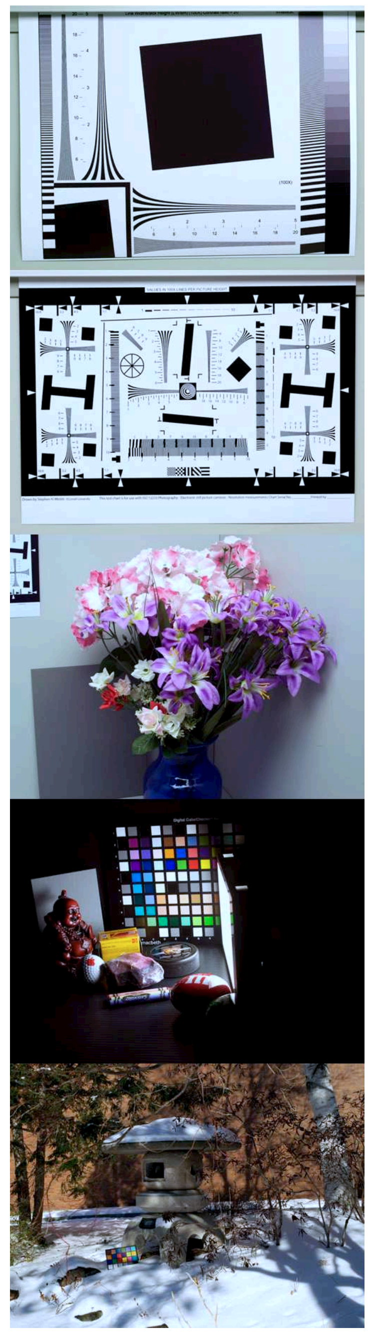

Figure 1. The five scenes used in all experiments. Names from top to bottom are Slant Edge, Resolution Chart, Flowers, Shadow Cow, and Tojo Garden.

Experimental Procedures

Images were captured with each of the four camera systems for five different scenes. The scene content is illustrated in Fig. 1. The first was a printed low contrast (contrast ratio = 20) resolution target with a large slant-edge target from Imatest. This scene is called *Slant Edge* and was used for an objective spatial frequency response (SFR) function computation and also included in the psychophysical observations. The second image was also a resolution test chart, a printed ISO 12233 chart created by Stephen Westland at Cornell University. It was named *Resolution Chart* and only used in both objective and psychophysical evaluations.

Three pictorial scenes were also captured for psychophysical evaluation (but also used in computations of a no-reference quality metric). These images were called *Flowers*, *Shadow Cow*, and *Tojo Garden*. *Flowers* is a still-life of artificial flowers in a daylight viewing booth. *Shadow Cow* is an image of a small high-dynamic-range scene with variety of items designed to compare the HDR capabilities (or lack thereof) of the the systems. There was a cow lurking in the shadows. It also included a Munsell Digital ColorChecker SG in the background. *Tojo Garden* is an image of an outdoor winter scene that also included a traditional Munsell ColorChecker chart. *Tojo Garden* presented a difficult challenge in retaining highlight detail.

All exposures were made with an f/8 aperture to enhance depth of focus and minimize and differences due to small focus errors. A tripod was used with all cameras and exposures. Scenes were metered with an external spot meter to determine optimal exposure and then bracketed with 5 exposures representing nominal exposure and -2, -1, +1, and +2 stops. The Hasselblad was only bracketed plus-and-minus one stop due to limits in its exposure range and the amount of film available. For the *Slant Edge*, *Resolution Chart*, *Flowers,* and *Shadow Cow* scenes, the optimal exposure was 1/8 sec. The *Tojo Garden* scene had an optimal exposure of 1/500 sec. For this scene the Hasselblad was set to f/22 to allow bracketing since its maximum shutter speed was 1/500 sec.

Objective image quality metrics were obtained by computing SFR functions for each camera using the slanted edge procedure[4] in the public-domain *mitre_sfr* software. The images were first converted to grayscale by averaging the linearized RGB channels and then cropped to the appropriate slanted edge area. Size calibrations of the images were completed to compute the results in terms of object-plane cycles per mm and allow conversion to effective cycles per degree for the imaging systems. Objective no-reference image quality metrics were also computed for grayscale versions each image using the public domain *iqm_macIntelCmd* software. This software uses a power-spectrum approach weighted by a human contrast sensitivity function to produce the IQM metric. The logarithm of the IQM metric has been shown to correlate well with visual quality assessments.[5] Both programs are available from the MITRE Corporation.[6]

In addition to the objective image quality computations outlined above, three objective psychophysical experiments were also completed to assess the quality of the captured images.

The first experiment was performed in a laboratory setting using projected images. Nineteen observers, experienced in psychophysical experiments on color and image quality performed the experiment. Images from the four capture systems were arranged on a single 1024x768 image display with each image, similarly cropped, filling 25% of the display area. The four images were labelled A, B, C, and D and presented simultaneously on a well-characterized Panasonic D6000 projector in sRGB mode. All of the images were first cropped, then resized to 500x372 pixels (a significant downsampling for all cases), processed through the *Auto Color* and *Auto Contrast* enhancements of *Adobe Photoshop CS4* to somewhat equate general white balance and exposure differences, then converted to sRGB (if necessary) and saved as lossless JPEG images. All 5 scenes were evaluated. Observers were asked to complete 3 rank order tasks in succession. The first task was to assess *Overall Image Quality* using whatever criteria they chose. The second was to assess *Color Quality* and the third was to assess *Sharpness*. In each case images were ranked from best (1) to worst (4). Observers recorded their own results on a score sheet. Results were simply analyzed by computing the rank histograms for each image capture system since the statistical significance of the scale differences was obvious.

The second experiment was identical to the first, but completed with the printed images. The same tasks were completed by the same 19 observers. The experimenter recorded observer responses that were stated orally by the observers. The observations were completed in an ISO-standard graphic arts viewing room with D65 (rather than D50) illumination as illustrated in Fig. 2. Each print was kept in its natural size, thus the cropping of the four images was not identical.

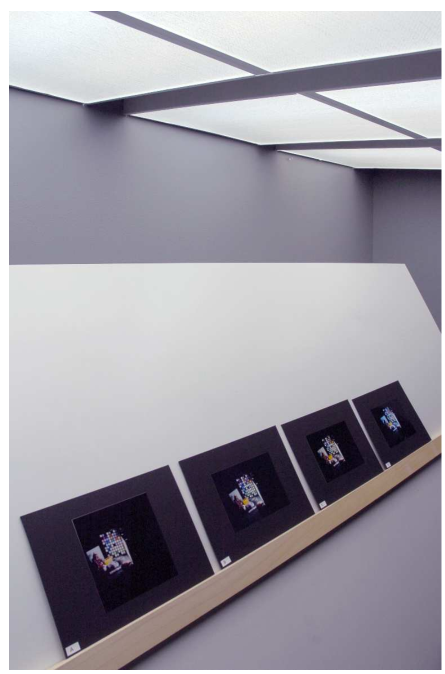

Figure 2. Experimental setup for printed image evaluation. The dedicated print viewing room includes ISO standard daylight illumination and gloss-free viewing arrangement.

The third experiment was completed online. The same digital images as used in the first experiment were used. These were tagged with the sRGB ICC profile such that any browsers including color management could accurately render the images. The images were incorporated into an interface using Amazon Mechanical Turk.[7] This interface allows the collection of human response data from large populations of observers for a very low cost and with a very fast turnaround time. In this case observers were asked to rank only overall image quality with three questions. (1) Which image is best, (2) Which image is second best, and (3) Which image is worst. Observers were paid $0.01 for each response. A total of 837 rank-order responses across the five scenes were obtained in one week (on average 167 per scene). The validity of results was checked to assure there are no meaningless automated responses and only observers with previous accuracy ratings greater than 97% were allowed to complete the tasks.

Results and Discussion

As is well established in image quality psychophysics, there are many attributes that observers attend to when judging overall quality.[8] This research did not aim to be exhaustive in evaluating image quality, but rather to focus on a few important attributes and metrics to provide a sense of the relative performance of the imaging systems. The author (and photographer) thought that the results were completely obvious once the images were obtained and was surprised by the degree of variability in the responses. Perhaps that is the most important result of these experiments; an image that is scene as absolutely horrible by one observer can be ranked as the best of the set by another experienced observer. This reminds us all that there is not a single aesthetic of *best image quality* to strive for in designing imaging systems.

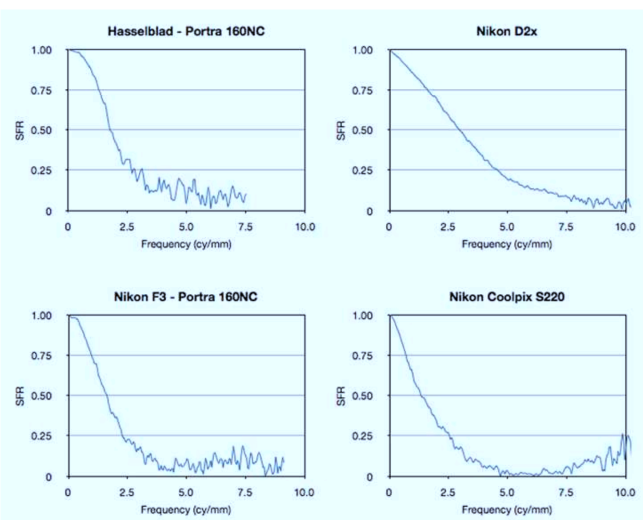

Figure 3. Spatial frequency response functions derived via slant-edge analysis for each of the four systems. Frequency in object-plane cycles per mm can be converted to cycles per degree by multiplying by 230.

Spatial Frequency Response

The slanted edge SFR functions are plotted in Fig. 3. For each imaging system they are plotted on equivalent axes regardless of the system's pixel count or magnification. Frequency is presented in cycles per mm at the object plane and can be converted to cycles per degree for visual reference by multiplying by 230. These images were captured from a distance of about 1.3 meters, thus an object-plane frequency of about 0.25 cycles per millimeter exceeds the visual resolution limit. Clearly each system resolves detail that cannot be perceived by the naked eye. The 50% SFR level for each system is 1.76, 3.00, 1.66, and 1.36 cycles/mm respectively for the Hasselblad 120, Nikon D2x, Nikon F3 135, and Nikon CP S220 systems. The best SFR function is clearly obtained from the Nikon D2x system. The two film systems produced similar SFR functions. The S220 digital camera produced an SFR similar to the film systems, but showed significant aliasing that was not present in the scans of film. Visual inspection of the film scans confirmed that the limiting factor was the imaging system and not the digital scan. The similarity in performance of the two film systems was probably due to the Hasselblad lens being a little "softer" than the Nikon lens.

IQM Analysis

The MITRE IQM metric was computed for the full frames of each of the five scenes and each of the four systems. The results are plotted in Fig. 4 using an offset logarithmic scale such that better quality has higher scores. The logarithmic scale roughly correlates with visual assessments. It can be seen that the IQM metric is scene dependent, but on average the Nikon D2x had the highest quality, followed by the F3-35mm system and the S220, and then the Hasselblad. Interestingly, the S220 scored well because it's high pixel count produced a lot of "information", unfortunately much of that information was aliasing and noise that was not discriminated by the IQM metric. It is thought that the Hasselblad performed less well due to the soft-focus nature of the lens. On a correlated 10 point visual scale, the average results for the D2x would score a 5, the F3 and S220 a 4.5, and the Hasselblad a 3.5. The D2x would receive a 9 out of 10 for the *Shadow Cow* image.

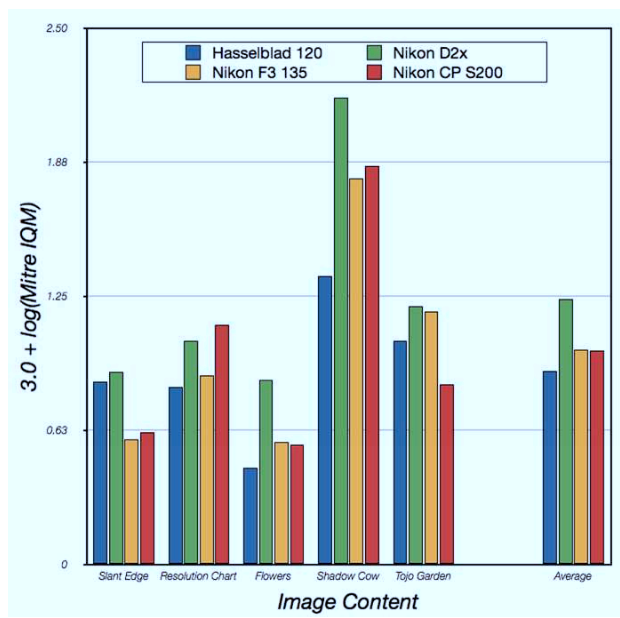

Figure 4. MITRE IQM no-reference image quality metric for each of the four systems and each of the five image contents and average results. Note logarithmic units to correlate generally with perceived quality results.

Projected-Image Psychophysics

Figure 5 summarizes the psychophysical rank order results for the three tasks using projected digital images. Each plot essentially shows the probability distribution of responses in each of the four rank positions. The overall image quality results show the D2x to clearly rank first with the S220 last and the two film systems effectively tied in the middle. The same is true for color quality while the sharpness results show the S220 performing worse (likely due to aliasing) with the other three systems essentially tied (likely due to the significant downsampling).

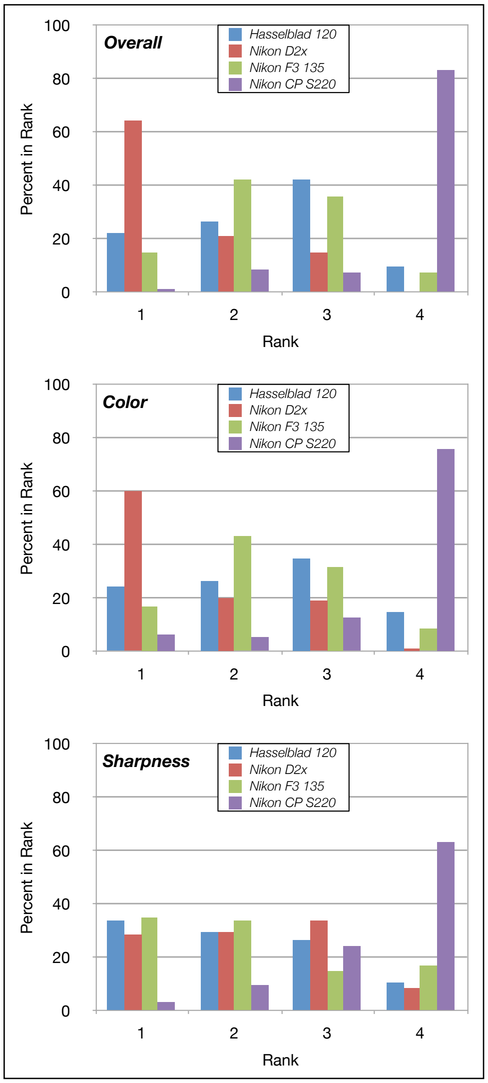

Figure 5. Psychophysical rank order results for projected images with 19 observers and averaged across all five images.

Printed-Image Psychophysics

Similar results are illustrated in Fig. 6 for the printed image psychophysics. However the rank order of the imaging systems is differs from the first experiment. In this case, the rank of the four systems is clearly delineated with the F3-35mm film system ranked best, followed by the Hasselblad-120 system, the D2x, and then the S220 with worst quality. It appears that the direct optical prints of the film systems were of superior quality than the digital prints which were on the same photographic paper, but subject to unknown image processing steps. This shows room for improvement in digital systems, since the closed-system capture-print film system still produced better prints by default. There is little doubt that better prints could have been made from the D2x digital captures and this is less likely for the digitized film captures. Similar results, with slightly less delineation were obtained for the color quality and sharpness assessments.

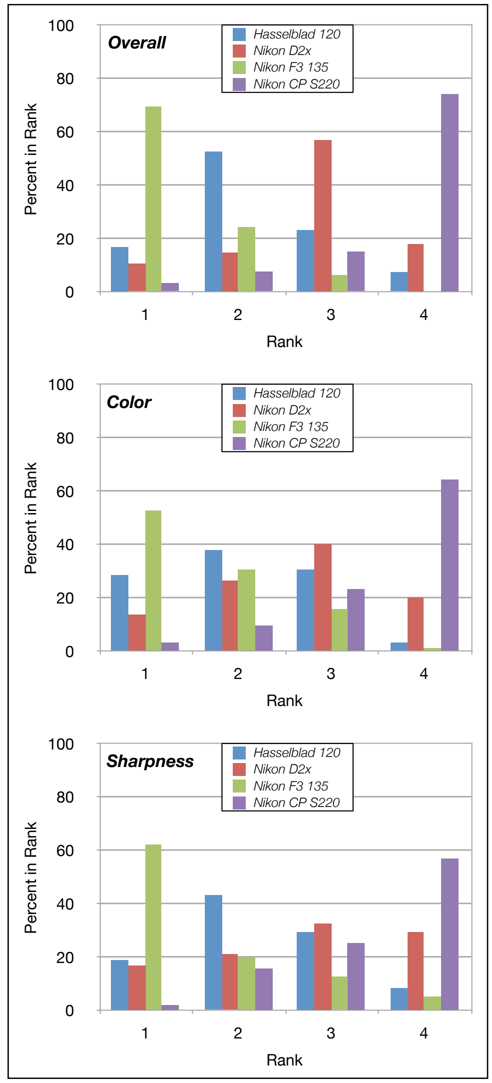

Figure 6. Psychophysical rank order results for printed images with 19 observers and averaged across all five images.

Online Psychophysics

The online results for a significantly larger observer pool (of unknown experience) and just for overall image quality are shown in Fig. 7. Despite having approximately an order-of-magnitude more observers, the results are very similar to those obtained with the same digital images and projection system in the controlled laboratory experiment. The Nikon D2x digital SLR system was clearly ranked first, the two film systems were essentially tied for second and third, and the point-and-shoot Nikon CP S220 system was clearly ranked last in overall image quality. It is reassuring to know that a high-quality digital SLR and lens do perform significantly better than a point-and-shoot digital still camera with similar numbers of Mpixels costing nearly two orders-of-magnitude less.

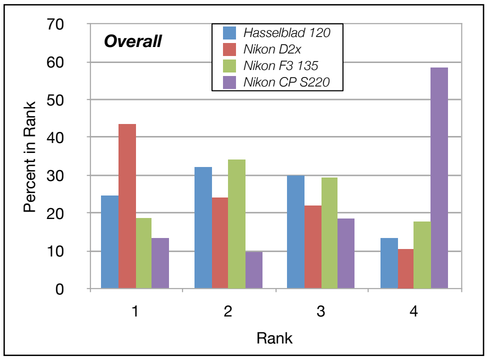

Figure 7. Psychophysical rank order results for online images for 837 observations and averaged across all five images. Only overall image quality was ranked in the online experiment.

Summary Results

The results of this *throwdown* are fairly clearcut. With five different assessments made, the D2x digital SLR was the clear winner on four and the F3 35mm film system using the same lens was the winner of one. Thus, it can be concluded, at least in this limited context, that digital systems have indeed matched and likely exceeded the quality of film for still photography. The one loss for the digital system was for the printed images, which points to a place for significant improvement, the automated printing of digital still photographs. The point-and-shoot Nikon CP S220 camera was competitive with the film systems in the objective metrics, but the visual assessments revealed the truth that objective metrics don't tell the whole story and illustrated the importance of the lens and noise minimization. It is clear that the quality of this type of camera could be significantly improved with fewer Mpixels, on the order of 6.[9] However, that would require image quality to become a better marketing tool than sometimes meaningless specifications.

Conclusions

This paper describes a set of experiments aimed at providing a small bit of data toward settling the long-lasting digital vs. film arguments with respect to low-to-mid-level still photograph capture and printing. The results clearly confirm that digital systems have evolved to the point of exceeding equivalent film cameras in almost every dimension. This, along with the added flexibility of digital systems, can lead to the final tolling of the death knell for film in this application (if it hasn't already). The results also illustrate the importance of the lens in image capture and perhaps will provide some guidance to systems designers to realize that small cameras and cell-phone cameras cannot possible have the lens quality or sensor size required to support high pixel counts. A return to overall system design for the lens and sensor combination could result in significantly improved image quality (and reduced data loads) for all.

Acknowledgements

This research was supported by the RIT Munsell Color Science Laboratory and inspired by my daughter, Acadia Reniff Fairchild, who realized I have a room full of "museum pieces" and asked the seemingly simple question of whether we could still get some film and try to use some of them.

References

1. Oxford University Press, *New Oxford American Dictionary, 2nd Ed.*, Oxford (2005).
2. B.Newhall, *The History of Photography*, Museum of Modern Art, New York (1982).
3. M.D. Fairchild, The HDR photographic survey, *IS&T/SID 15th Color Imaging Conference*, Albuquerque, 233-238 (2007).
4. P.D. Burns, Slanted-edge MTF for digital camera and scanner analysis, *IS&T PICS Proceedings*, Portland, 135-138 (2000).
5. N.B. Nill and B.H. Bouzas, Objective image quality measure derived from digital image power spectra, *Optical Engineering* **31**, 813-825 (1992).
6. www.mitre.org/tech/mtf
7. mturk.amazon.com
8. M. Pedersen, N. Bonnier, J.Y. Hardeberg, and F. albregsten, Attributes of image quality for color prints, *J. Electronic Imaging* **19**, 011016-1-13 (2010).
9. 6mpixel.org

Author Biography

Mark D.Fairchild is Associate Dean for Research and Graduate Education of the College of Science at Rochester Institute of Technology. He is also Professor of Color and Imaging Science and Coordinator RIT's M.S. and Ph.D. programs in Color Science. He is a Fellow of IS&T and was recently presented with the society's Raymond C. Bowman award for mentoring future researchers in the field. He started his path toward imaging science with aspirations to be a photojournalist and has recently rediscovered the power of making images in addition to making images better.