Michael Vrhel, Eli Saber, and H. Joel Trussell

Color Image Generation and Display Technologies

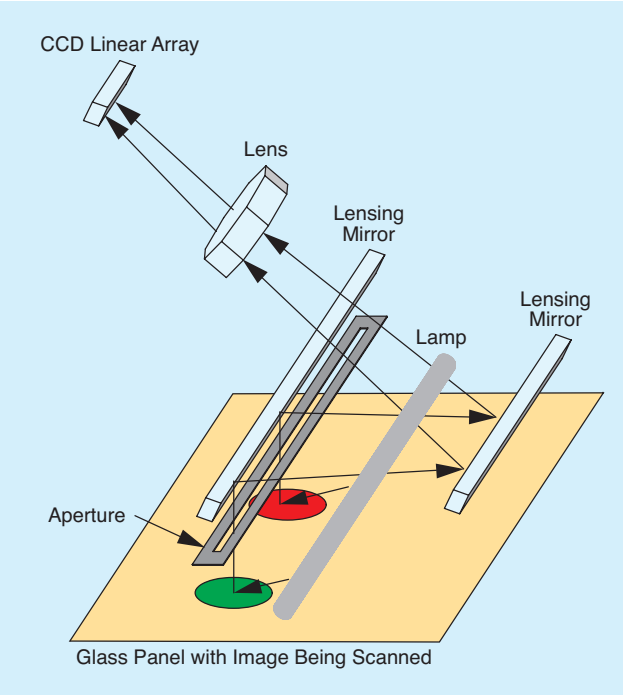
An overview of methods, devices, and research

Color image generation and processing has been growing at an unprecedented rate over the last two decades, and its effects are being felt throughout our everyday lives. Today, scanners, digital cameras, displays, and printers are available at relatively inexpensive prices for commercial and consumer applications.

With the overwhelming number of image input and output devices, it is critical for the imaging engineer to understand the underlying transformations, from physical properties to digital representation, that take place in the recording hardware. It is equally important to comprehend the transformations that will occur when a digital image is sent to the output hardware for display or printing. Generally speaking, recording and display devices are limited in terms of noise, dynamic range, metamerism, spatial resolution, gamut, and spectral response. To fully comprehend the meaning of the recorded digital data and the displayed image, it is necessary to know the limits of the recording system and display hardware in these terms. This knowledge enables the optimization of processing methods from a system viewpoint [32] yielding better results with reduced resources.

The goal of this article is to provide an overview of the transformations and limitations that occur in color imaging input and output devices. It is beyond the scope of the article to do an exhaustive survey of all devices. Instead, we will concentrate on two common recording devices and three common output devices. First we provide an overview of digital scanners and cameras, and then we discuss inkjet and laser printers. Finally, liquid crystal display (LCD) devices are presented.

1053-5888/05/$20.00©2005IEEE



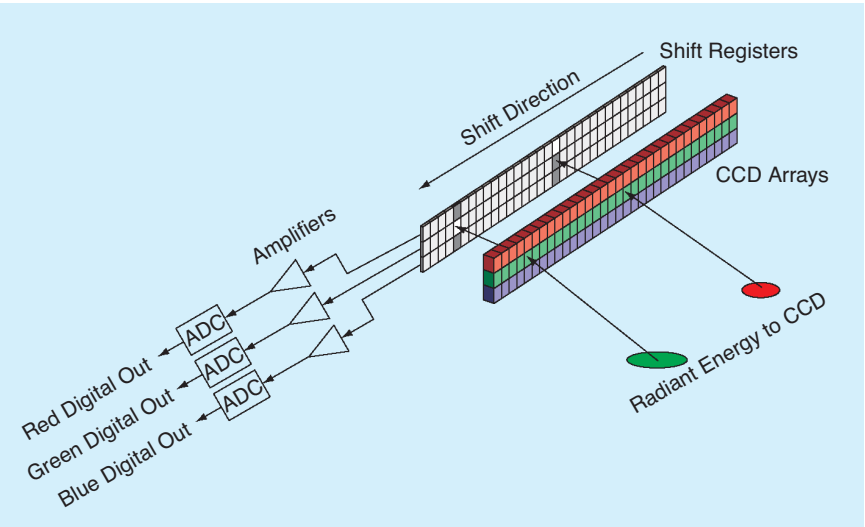
[FIG1] **CCD optical path for flat-bed reflectance scanner.**

SCANNERS

To effectively process images from scanners and digital cameras, it is necessary to understand the transformations that affect these images during recording. The most common desktop scanners record the image data using a row of sensors. Since images are two dimensional (2-D) in nature, it is necessary for either the paper or sensor (or both) to move in the orthogonal direction to the sensor to capture the entire image. (There are scanners that use 2-D sensors requiring no movement, as well as scanners using point sensors requiring 2-D movement.) The "paper moving" designs consist of both sheet-fed and flatbed scanners. The "sensor moving" designs include hand-held scanners, which require the user to move the instrument across the paper, as well as flatbed scanners. The primary sensor types are charge coupled device (CCD) arrays and contact image sensor (CIS) arrays. Currently, CCD arrays provide higher signal-to-noise ratio (SNR) levels, while CIS arrays result in scanners of very low cost due to simplistic optical designs. We will discuss both types in the following.

Figures 1 and 2 display the optical path and the initial processing of a CCD desktop scanner, respectively. In the CCD scanner shown Figure 1, a lamp is used to illuminate the image being scanned. The lamp consists of a small fluorescent tube that is mounted spatially close to the image and designed to emit a diffuse light. The reflectance from the image is passed through an aperture and focused onto the CCD array through a series of mirrors and lenses. Typically, a full page width (21.5 cm) is focused on a CCD array that may be only 28 mm wide. The mirrors, lens, and CCD elements are generally housed in a single enclosure, designed to move with a stepper motor. Most color scanners contain a single light source and a CCD array consisting of three rows that are covered with red, green, or blue filters. However, there is a variant that consists of three colored lamps [typically red, green, and blue (RGB)] and a single row CCD array.

The CIS design is displayed in Figure 3. In a typical CIS design, RGB light-emitting diodes (LEDs) are time multiplexed to illuminate an image row at an angle of 45°. The radiant power of the LEDs is directed to the image through the use of a plastic light pipe. The diffuse reflectance from the image is passed through a collection of plastic light pipes, which are spaced across the row. These light pipes (usually one piece) limit the viewing angle seen by each phototransistor. The spatial resolution is controlled by the light pipes and the number of phototransistors. Since there are no mirrors or lenses involved, this type of sensor is very inexpensive compared to its CCD counterpart, which requires expensive optical elements to reduce the scan line to the CCD array size.



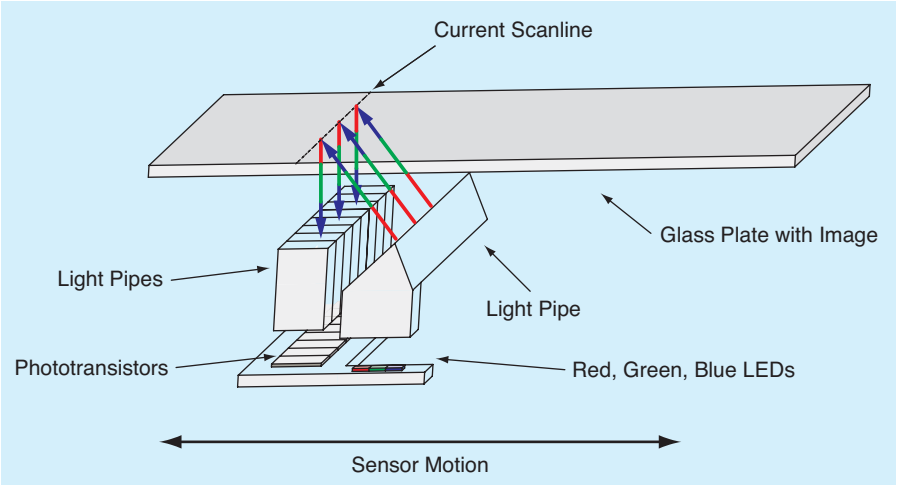
[FIG2] **CCD sensor details.**

In the CCD design, the sensor values are read into a shift register (Figure 2), amplified and then converted to digital values by an analog-to-digital convertor (ADC). The sensor performance is generally specified in terms of its dark signal level, saturation level, noise level, and photo response nonuniformity (PRNU). The ADC is primarily specified in terms of the number of bits used for quantization, the differential nonlinearity (DNL), integral nonlinearity (INL), and sampling rate [27], [28], all of which have a significant effect on the quality of the scanned image.

In addition to the reflectance scanners discussed previously, there are transmittance scanners used for digitiz-

ing slides, negatives, and transparencies. The major difference in these devices is that the illumination and sensor are mounted on opposite sides allowing for the light to travel through the media. The remaining components are the same as the ones discussed previously for reflectance scanners.

Once the digital data is obtained, a number of processing steps are performed to finalize the scan. From Figure 2, it is clear that the RGB linear arrays are spatially displaced with respect to one another. Hence, the red pixel at location $[x, y]$ will be read at the same time as the green pixel at location $[x, y + y_o]$ and the blue pixel at location $[x, y + 2 * y_o]$, where the magnitude and sign of y_o depend on the size of the displacement and the scan direction, respectively. It is necessary for the scanner firmware or driver software to reorder the recorded data to compensate or interpolate for this displacement. Other processing steps include optical compensation, color correction, and noise. These will be discussed in further detail in the following.



[FIG3] CIS optical path. The RGB LEDs are time multiplexed to illuminate a single scan line.

OPTICAL ISSUES

In the design of a CCD scanner, there is a tradeoff between the complexity and cost of the scanner optics and the irradiance falloff at the edges of the detector. Assuming an ideal lens and a Lambertian source, which implies that the radiance of the source (the image) is independent of the angle (or that the source is completely diffuse), then the irradiance at a point on the detector can be expressed as a function of the on-axis irradiance I_o and the off-axis angle θ (the axis in this case is the optical axis). The value of the irradiance at the off-axis angle θ is given by

$$I_\theta = I_o \cos^4(\theta). \tag{1}$$

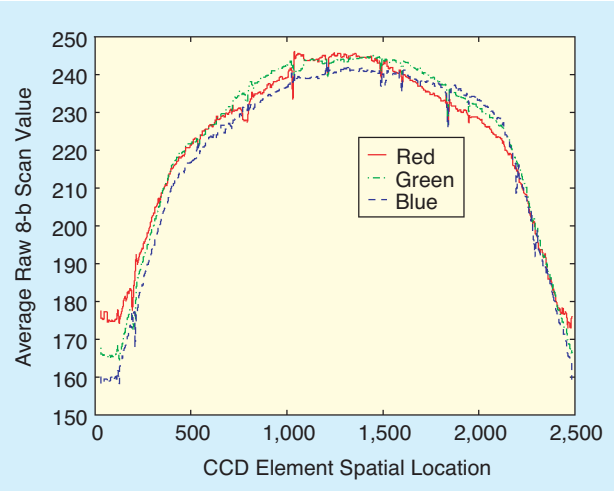
To reduce the falloff, it is necessary to make θ as small as possible. This is achieved by increasing the optical path, thereby increasing the size and/or optical complexity of the scanner.

To demonstrate the level of a typical falloff, we took a raw uncorrected image file of a constant white region from a CCD scanner and computed the average of the pixel rows. The average is shown in Figure 4 where the falloff, at its worst, depicted a 35% loss from the on-optical axis region. To correct this falloff, most scanners start by measuring a white sample and performing a dark measurement (light off) prior to scanning the image to define the dynamic range. The bias and scale of the recorded sensor values are then digitally adjusted to compensate for the optical falloff as well as any changes in the bias and gain of the system.

SPECTRAL RESPONSE AND COLOR CHARACTERIZATION

Due to the effect of metamerism, it is quite possible that spectra that appear the same to the standard human observer may look quite different to the digital scanner. This is due to the fact that the spectral response of the color scanner is not related by a linear transformation to the CIE color matching functions. This effect has led to a significant amount of research over the past decade on the design of color filters [3]–[6]. At best, a three-band scanner could provide true standard color values for only one type of illuminant. However, recent design methods have shown that it is possible to provide reasonable color accuracy for a number of illuminations by adding additional filter bands (i.e., more than three) [5]. In the limit of adding filters, the device will become a spectral scanner. Currently, digital still camera versions of such devices are used for research and art archival purposes [7].

The goal of color characterization was discussed in this issue's overview [29]. The characterization of a desktop scanner involves the measurement of a color target by both the scanner and a colorimeter. From these measurements, the scanner output values are



[FIG4] CCD falloff in a real CCD scan engine.

mapped to device independent color values by utilizing model-based techniques or multidimensional look-up tables (LUTs) [23]–[25].

NOISE

The noise sources found in a modern scanner can be generally classified into temporal or spatial types. Examples of temporal ones include dark noise, shot noise, noise from mechanical vibrations, and noise from illumination fluctuations. Examples of spatial noise sources include variations in CCD element sensitivity and dark current nonuniformity.

DUE TO THE EFFECT OF METAMERISM, IT IS QUITE POSSIBLE THAT SPECTRA THAT APPEAR THE SAME TO THE STANDARD HUMAN OBSERVER MAY LOOK QUITE DIFFERENT TO THE DIGITAL SCANNER.

Temporal noise sources can be reduced by averaging. On the other hand, spatial noise sources can be minimized by frame subtraction or gain/offset correction methods. However, these methods typically result in increased temporal noise. For example, consider the simple process of dark current removal. Let the dark measurement obtained with no illumination be denoted by

$$m_{\text{dark}}(j) = d(j) + n_1(j), \tag{2}$$

where j is the spatial location on the sensor element along a row, $n_1(j)$ is temporal noise, and $d(j)$ is the noise-free dark measurement. The variation of $d(j)$ as a function of j represents a spatial noise source. Furthermore, let the measured image row at the element j be denoted by

$$m_{\text{row}}(j) = r(j) + d(j) + n_2(j), \tag{3}$$

where $r(j)$ is the noise-free measurement, and $n_2(j)$ is the corresponding temporal noise. Ideally, the spatial noise could be removed by subtracting $d(j)$ from the equation. However, in practice, since $d(j)$ is not known, $m_{\text{dark}}(j)$ is subtracted instead, yielding

$$m_{\text{corrected}}(j) = r(j) + n_2(j) - n_1(j). \tag{4}$$

Since $n_1(j)$ and $n_2(j)$ are uncorrelated, the power of $n_2(j) - n_1(j)$ will be greater than the power of $n_2(j)$, implying that $m_{\text{corrected}}(j)$ has more temporal noise compared to $m_{\text{row}}(j)$. Ideally, the reduction in spatial noise is greater than the increase in temporal noise. In practice, the power of $n_1(j)$ is reduced through averaging, and the processing in the scanner driver or firmware reduces the spatial type noise sources yielding an output image that contains primarily temporal noise.

In a well-designed scanner, the dominant noise source is due to the actual fluctuations in light measured at the sensor. This fluctuation is a quantum phenomenon, whose collective energy is made up of individual photons of various energies. Hence, what is observed visually (or through a given sensor) is the average behavior of a stochastic process. Consequently, the intensity of the radiation is the average number of photons reaching the detector, and its variation is defined as photon noise or shot noise. The distribution of photons is modeled by a Poisson distribution due to the counting nature of the physical process and the assumed independence of the samples. Since the mean and variance of a Poisson distribution are equal, this noise will be signal dependent. Due to the high light levels, however, the number of photons is so large that the noise distribution can be well approximated by a Gaussian with a signal dependent mean.

Nonadditive noise can be present in scanned image data due to actual fluctuation in the light source. This is true for LED-based sources, as well as tungsten and inert gas lamps. These fluctuations are the reason high-quality color measuring instruments will typically use a second detector that monitors the light source. This extra detector makes it possible to compensate for the lamp fluctuations. However, inexpensive desktop scanners do not have such a sensor. Hence, the lamp variation is manifested in a smaller sensor voltage variation when a dark sample is scanned, as compared to a white sample. Consequently, the variations are represented as multiplicative noise.

Modern desktop scanners often claim 16-b resolution/channel (usually marketed as 48 b). While it is relatively common to collect data to that level of resolution with a pipelined ADC, limitations in the mechanical and optical design produce an image with far fewer useful bits. For demonstration purposes, the SNR of a popular 48-b desktop scanner was estimated by collecting 100 images of a mid-level gray. This scanner used a Toshiba 2903(7) CCD and an Exar 9816 ADC (a 16-b, three-channel ADC). The image was of size 267×239 at 300 dpi in 48-b mode. For the 100 images, the average SNR was found to be 37.9 dB. Using the standard formula for relating bits to SNR, this results in the same effective SNR as an ideal 6-b/channel system and is significantly less than the 16-b/channel advertised. It should be noted, however, that some accuracy was lost due to adjustments of the image data for effects such as spatial noise patterns, optical falloff, and colorimetric characterization.

RESEARCH DIRECTIONS AND CHALLENGES

CIS scanners with LED-based light sources are becoming very common due to their low cost and low power consumption. However, using only three LED colors to illuminate an image introduces possible metameric problems that depend on the media, colorants of the image, and reference light source. Optimization of the LED spectra over a number of reference illuminants and media sources with a constraint on realizability is an open extension of the existing work on color filter design. Moreover, the current filter design algorithms deal exclusively with additive noise models. As discussed, there are nonadditive and signal dependent noise

sources. It would be of great interest to quantify their effect on the optimization of the scanner spectral response.

In almost every consumer product, cost is the number one issue. Lower cost is often used as a tradeoff for quality as in the case of the CIS versus CCD sensors. The inherent challenge, which will continue as new sensors are developed, will be to extend the accuracy of these low-cost sensors to match that of the low-noise CCD arrays.

LACK OF ILLUMINATION CONTROL MAKES CHARACTERIZATION OF DIGITAL STILL CAMERAS CHALLENGING.

DIGITAL STILL CAMERAS

All practical digital still cameras (DSCs) use either CCD or complimentary metal-oxide-semiconductor (CMOS) 2-D sensor arrays. Compared to CCDs, the CMOS sensors have a lower SNR but have the advantage of using a fabrication process that allows the integration of other components like ADCs into the sensor. Typically, the sensor is covered by a color filter array (CFA), so that each pixel records the output of only one of the color filters. Hence, at each spatial location, the resulting color channels have a single recorded value (red, green, or blue for an RGB camera) and must be further processed using appropriate interpolation schemes to generate an image, where each pixel location is defined by three color values. This process is called demosaicking and is discussed in detail later in this issue [33].

There are digital color sensors that achieve color separation by using the fact that the photon penetration depth into silicon is wavelength dependent [8]. Silicon has the property that long wavelengths penetrate deeper than short wavelengths. These sensors do not require demosaicking.

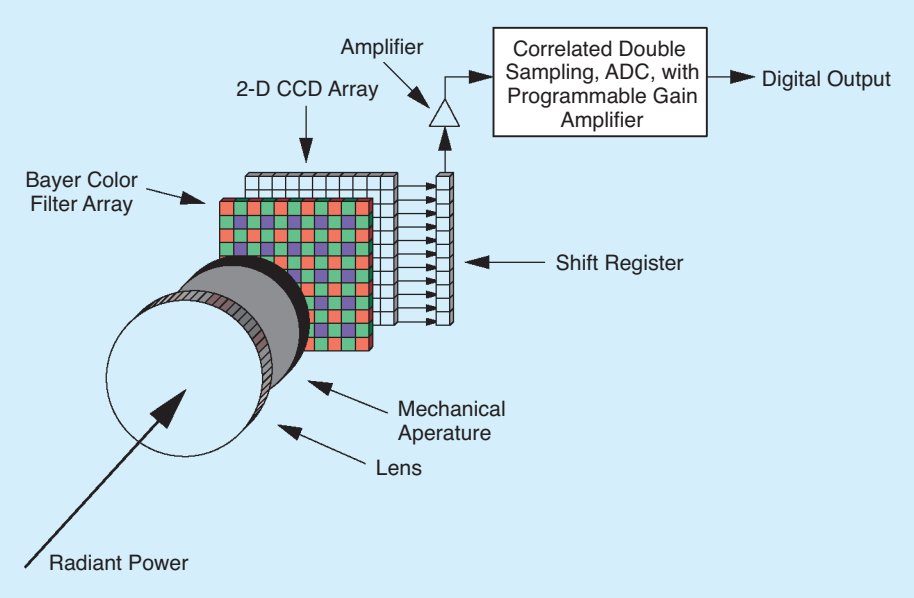
Compared to a scanner, the DSC electronics are smaller, use less power, and are required to work in more demanding environments with various illuminations. The image formation and digitization process is shown in Figure 5. The radiant energy passes through the lens and aperture, through the CFA, and is measured by the 2-D sensor arrays. In addition to imaging the scene onto the sensor, the lens functions as an antialiasing filter. In the CCD based camera, the charges are shifted across the array into a shift register and processed through the ADC to obtain a digital value for each element.

Colorimetrically, the primary difference from a scanner is that the DSC has no control over the illumination source. This makes generic characterization of the device a much more complex and challenging task [25]. Presently, most devices possess the ability to do some type of compensation for the illumination, as discussed in this issue [31].

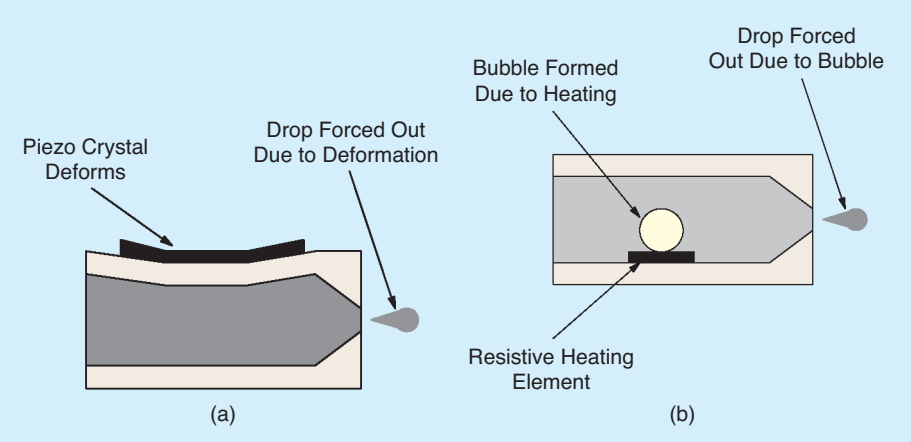
The imaging engineer who is processing images from a digital camera should realize that the noise sources, optical limitations, and quantization that were discussed in the scanning section will effect the DSC image data, as well as the interpolation of pixels for demosaicking, interpolation for image size, possible data compression, color balance adjustments that depend on camera settings, and the lighting environment. In addition, most cameras allow the adjustment of exposure time and system gain (similar to the ISO film speed). These settings will also effect the dynamic range and noise level of the recorded image. The details of the image formation for a DSC are discussed in this issue [31].

RESEARCH DIRECTIONS AND CHALLENGES

Applications of digital cameras that require accurate color have received a fair amount of research attention over the last few years [9]–[11]. The majority of the published research, however, has dealt with the demosaicking problem [33]. It is likely that this trend will continue for some time, yielding more robust and practical schemes. There has been only limited work on colorimetrically characterizing DSCs [25] due to its dependency on the environment and the proprietary nature of the algorithms. Complex and computationally expensive methods for estimating and compensating for the illumination in an image have also been published in the literature [34], [35]. Most of these algorithms are described as color constancy algorithms, since the goal is to have the colors of the image remain constant, regardless of the illumination spectrum. Since the problem is, in general, ill-posed, illumination estimation and correction will likely remain an active area of research for many years to come.



[FIG5] **Image digitization process in CCD digital still camera.**



[FIG6] **Methods for dot-on-demand inkjet printing: (a) pressure scheme and (b) thermal scheme.**

In photography, accurate colorimetric reproduction is often not the primary goal. Instead, it is desired to create an image that is "visually pleasing." This goal is typically achieved through the use of a rendering function [36], [37], which creates more pleasing colors (e.g., greener grass and bluer skies). Such problems rely heavily on the psychological aspects of color.

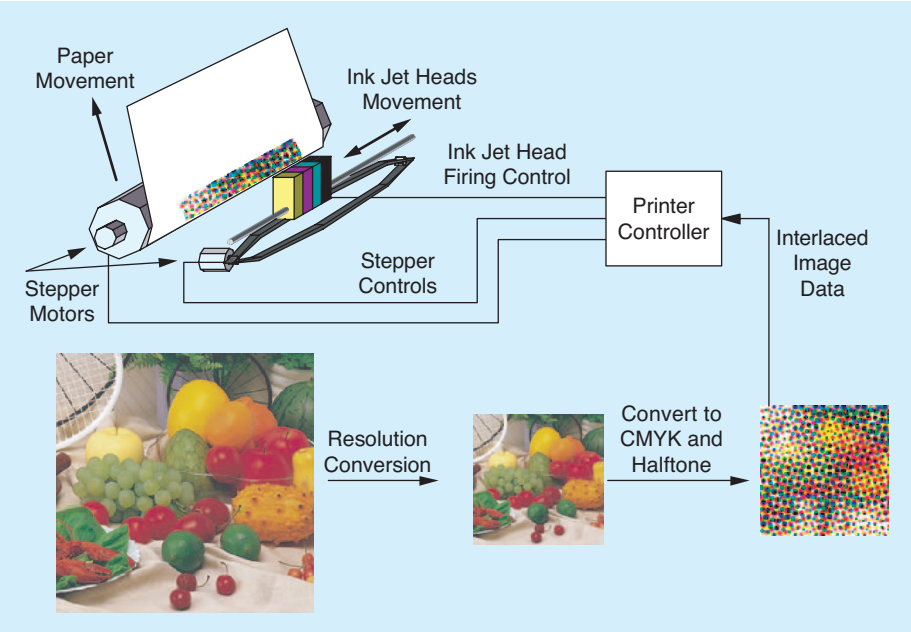
The trend for higher resolution CCD and CMOS devices will most certainly continue. Since the cost of the sensor is directly related to the area of the silicon die, the push is to increase resolution by decreasing the footprint of each imaging element. This typically results in lower SNR, which leads to the need to develop improved noise reduction methods.

INKJET PRINTING

Pressure and thermal schemes are currently the two primary approaches for drop-on-demand inkjet printing heads; we will not discuss continuous flow printers, since they are not used for printing images. Pressure and thermal scheme methods are illustrated in Figure 6. In the pressure scheme, a piezo crystal is electrically stimulated, causing a deformation of the microchamber and thereby forcing a drop of ink out onto the paper. In the thermal scheme, a heating element causes the formation of a bubble, which forces ink from the microchamber onto the paper.

The printers are designed with the inkjet heads mounted in line on a movable stage. Depending on the design and the settings, the heads may spray ink when going from left to right and/or right to left. The direction the head is moving at the time the dot is sprayed can affect the dot shape and placement.

Each head contains a number of micronozzles, where each micronozzle corresponds to one of the units in Figure 6. The line spacing is fixed mechanically, and the timing of the spray of each dot as the heads move along the line is such that each color is printed on the same rectangular grid. In practice, the inkjet head passes each pixel location more than once. This occurs so that only a portion of the required ink needs to be placed on the paper in each pass, thereby reducing ink coalescence problems. Ink coalescence occurs when the ink has not had sufficient time to be absorbed into the paper before additional ink is placed on the same or nearby location. Hence, the ink viscosity, ink surface tension, and paper type greatly affect the rate at which dots can effectively be placed on the paper.



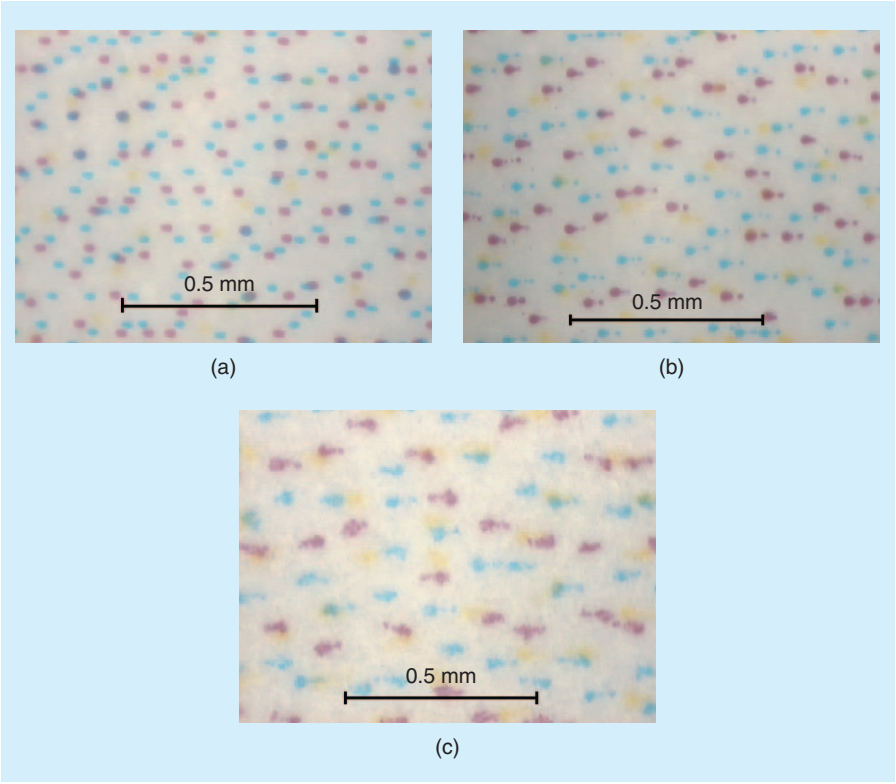
[FIG7] **Ink-jet printing process.**

Inkjet engines are bilevel, i.e., they can either place a dot on the paper or not. To create the appearance of a continuous tone image, it is necessary to perform halftoning, which is covered in this issue [30]. There are also multilevel devices, which will basically place a second dot on top of a dot with the same color. This has the effect of making a darker and/or larger dot. In this case, multilevel halftoning methods are necessary. In practice, inkjet engines typically use error diffusion as a halftoning technique, due to the well defined and repeatable dot shape [1].

Color inkjet engines commonly use three or four colors: cyan, magenta, yellow, and black (CMYK). Six-color engines are becoming more prevalent, where a light

magenta and light cyan are introduced. These light colors reduce the visibility of halftone artifacts in high reflectance regions, where the lower reflectance magenta and cyan dots would be highly visible. The transition between light and dark colorants can be handled by the design of one-dimensional (1-D) LUTs.

Figure 7 displays a conceptual diagram of the transformation from the digital source image to an inkjet print. In this process, the image is first adjusted to the proper resolution for the printer, converted to CMYK values, and halftoned. The data is interlaced, since the printer will be using all colors at the same time. Hence, the print driver must indicate to the controller the image dimensions as well as factors that will control the number of passes the heads should make over the same image area to avoid coalescence problems. With this information, and the binary halftone data, the controller can determine if it should place a dot on the paper as it moves the inkjet heads and advances the paper.



[FIG8] Dot patterns created by ink-jet engines. (a) Glossy paper. Engine creates single dots. (b) Glossy paper. Engine creates satellite dots. (c) Same engine as (b), but ink-jet coated paper type was used.

In any real system, there will be various sources of noise that affect the quality of the printing process. Some factors affecting the variability include the halftoning method, variation in dot size, shape, and location, as well as chemical and physical interaction of the inks and media. Obviously, the ink must have a consistency that permits it to be sprayed through a small opening. This type of ink has had problems with fading. However, newer, more expensive, archival inks are becoming available.

Figure 8(a) and (b) displays dot distributions from two different inkjet devices. Figure 8(b) and (c) displays the effect of changing paper type from (b) glossy to (c) coated inkjet. Note that the dots shown here are asymmetric. They are longer in the direction that the inkjet cartridge moves. The dot shape of a particular printer can be used in the model of error diffusion or other halftone methods to obtain improved performance.

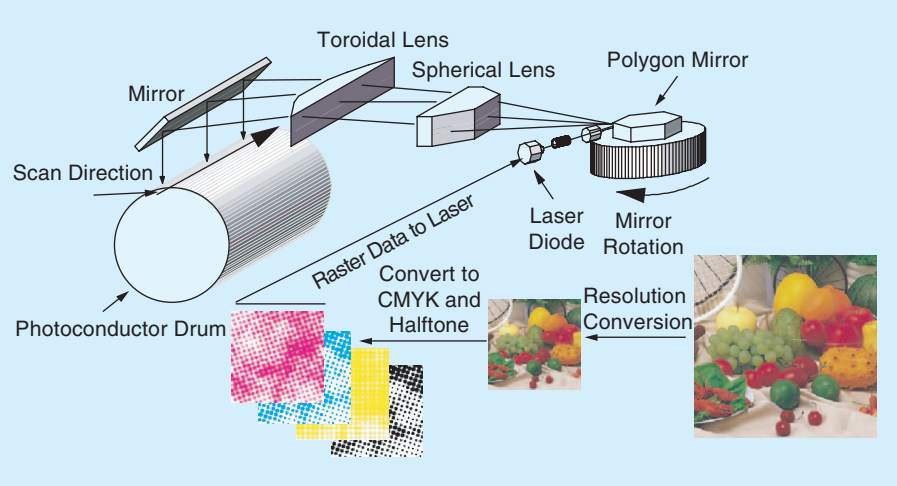
RESEARCH DIRECTIONS AND CHALLENGES

Halftoning is, of course, a critical component of inkjet printing. As discussed in this issue [30], color halftoning remains an active area of research. A primary area of interest will be the development of methods that compensate for dot shape, size, and placement. Thus far, there has been very little published research on multilevel color halftoning or halftoning with more than four inks, operations that inkjet printers are currently capable of performing. The color characterization of devices that use more than four inks has also received only limited attention in the literature and will likely be a source for future research.

Finally, there has been a fair amount of work on embedding color measuring instrumentation within high-end inkjet printers to determine the adjustments necessary for media/ink changes. Since the number of measurements in such a system needs to be relatively small, the success of these methods requires additional research on mathematical characterization models.

LASER PRINTING

Chester F. Carlson created the first electrophotographic image in 1938. Today, the electrophotographic process is used in laser printers and copiers. (This is also known as xerography, which comes from the greek and means "dry writing." The Haloid Company purshased the rights to Carlson's invention and eventually changed their name to The Xerox Corporation.) As in the inkjet printing process, the digital image must be halftoned. Traditionally, digital screen halftoning methods have been used for laser printing. The transformation from digital data to print is conceptually shown in Figures 9 and 10. In Figure 9, the RGB image data is spatially scaled to the output resolution of the printer, converted to CMYK color space, and halftoned. The halftoned data is used to modulate the laser. The polygon mirror and the lens assembly cause the laser to scan across the photoconductor drum as shown. Unlike the inkjet engine, which prints all four colors at the same time, the laser engine processes each color plane in a sequential manner. Figure 10 displays each of the primary steps:



[FIG9] From digital RGB data to exposing the photoconductor drum.

- initial charging of the photoconductive material (there may also be a cleaning step which is not shown)
- exposing the photoconductive material with a laser, spatially related to the negative of the image to be printed; exposure to light removes the positive charge
- presenting negatively charged toner particles to the photoconductive material; those areas that were not illuminated will now have toner
- electrostatic transfer of the toner from the photoconductive material to paper
- heat fusion of the toner to the paper.

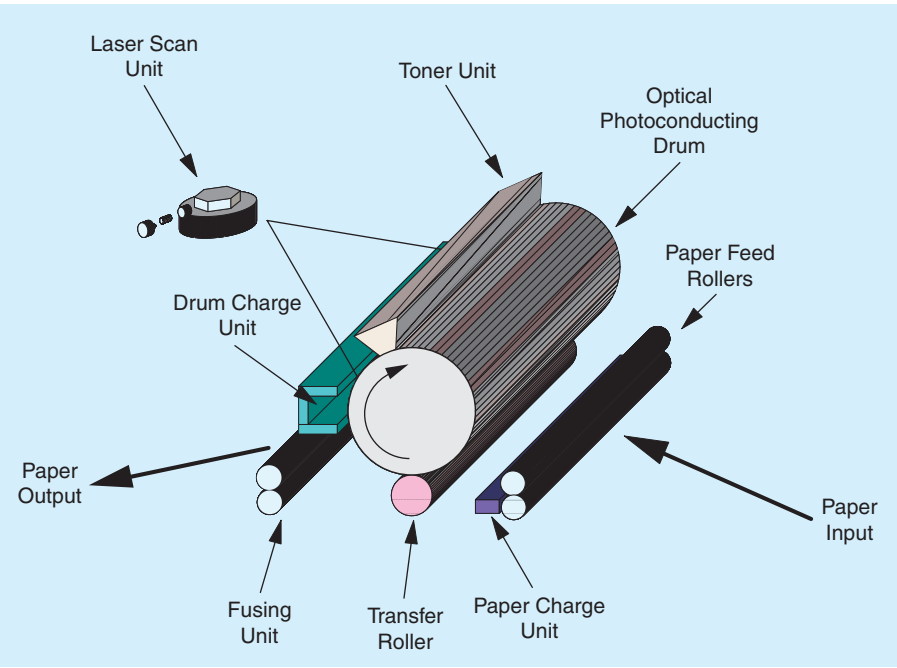
The first four steps listed here are performed for each toner color (e.g., CMYK) prior to fusing the toners to the paper. This repeated application of each of these steps can introduce distortions in the printed image.

Originally, in image copiers, the device used an optical formation of the image to expose the photoconductor. Today, it is common for a digital scanner to be used to record the data and a laser print engine to be used to create the copy. This newer design makes it possible to digitally manipulate the image data for operations such as color characterization, resolution conversion, digital halftoning, and segmentation.

There are a number of variants to the system shown in Figure 10. These include devices that image all four color planes on the photoconductive drum prior to the paper transfer and devices that transfer all four toners to an intermediate belt (transfer belt) prior to doing a one-time transfer onto paper. Both of these methods will have better registration between the color planes compared to the system that performs four individual toner transfers to paper. There are also variants that have the laser add charge to the photoconductor at the locations that require toner.

Figure 11 displays magnified dot patterns for two different tonal levels from gray level patches printed by a color laser printer. Compared to the inkjet samples discussed previously, the dot shape from the laser device does not have a well-defined edge. This poor dot shape and its unstable nature at small resolutions is the reason error diffusion halftone algorithms are rarely used for laser engine applications.

Note in the figures that each dot consists of very small toner particles. It is not possible to control the placement of individual toner particles, but it is possible to control the approximate (within some statistical range of course) number of particles that make up a single dot by modulating the pulse width and/or amplitude of the exposing laser. In effect, that is the difference between (a) and (b) in Figure 11, where the number of toner particles that make up the dots in (b) are significantly greater. From this figure, it is clear that the parameters of interest in modeling an electrophotographic system include the dot location (position), shape, splatter, and their statistical variation.

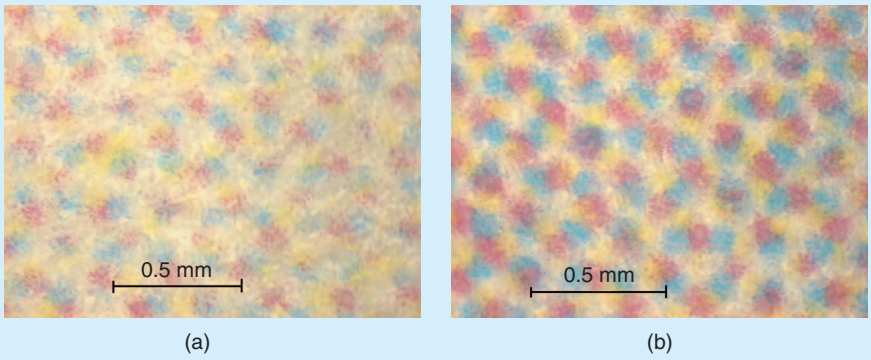


[FIG10] Details of using charge on photoconductor drum to pick up toner, transfer to paper, and fuse to paper.

RESEARCH DIRECTIONS AND CHALLENGES

There are a number of variations of electrophotographic systems. Modeling these devices to achieve improved halftoning is a challenge and has received minor research attention in the published literature [2]. Most of the research and development work is proprietary.

Similar to the inkjet systems, there has been work on embedding color measuring instrumentation within these devices to determine the adjustments necessary to compensate for media/toner changes, as well as changes due to device drift. Such a system may require colorimetric characterization (or at least some adjustment of the characterization mapping) of the device with a small number of measurements.



[FIG11] **Magnification of tonal levels created by electrophotographic printer. (a) Light gray. (b) Medium gray.**

LIQUID CRYSTAL DISPLAYS

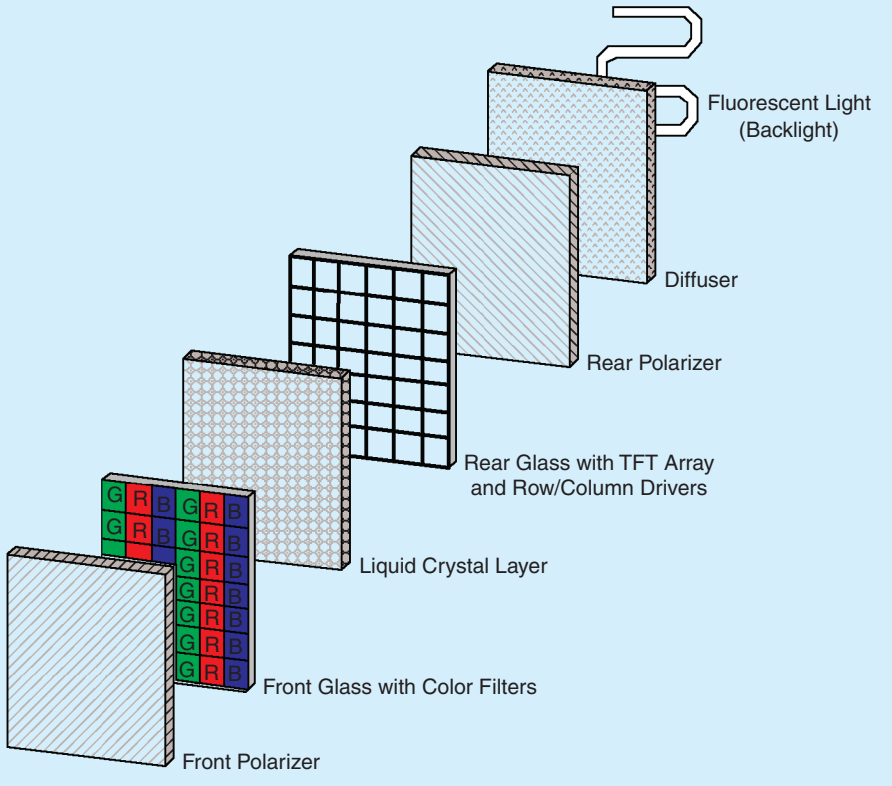
Since it is relatively new, flat panel technology is rapidly changing [14], [18], [12]. Early flat panel displays suffered from low luminance, poor color, and significant dependence on viewing angle. With the solution of these problems, it is likely that flat panel displays will bring the demise of CRT displays. Most flat panel displays have the advantage of being thinner, lighter, and less power hungry than a CRT. In addition, in typical viewing conditions, flat panel displays have a larger dynamic range compared to CRT displays.

A common flat panel technology in laptop PCs is the LCD. Liquid crystals were discovered by Friedrich Reinitzer in 1888. They were not considered to be of much use until 1963, when Robert Williams discovered their electro-optic features. Two years later, George Heilmeier invented an LCD [19].

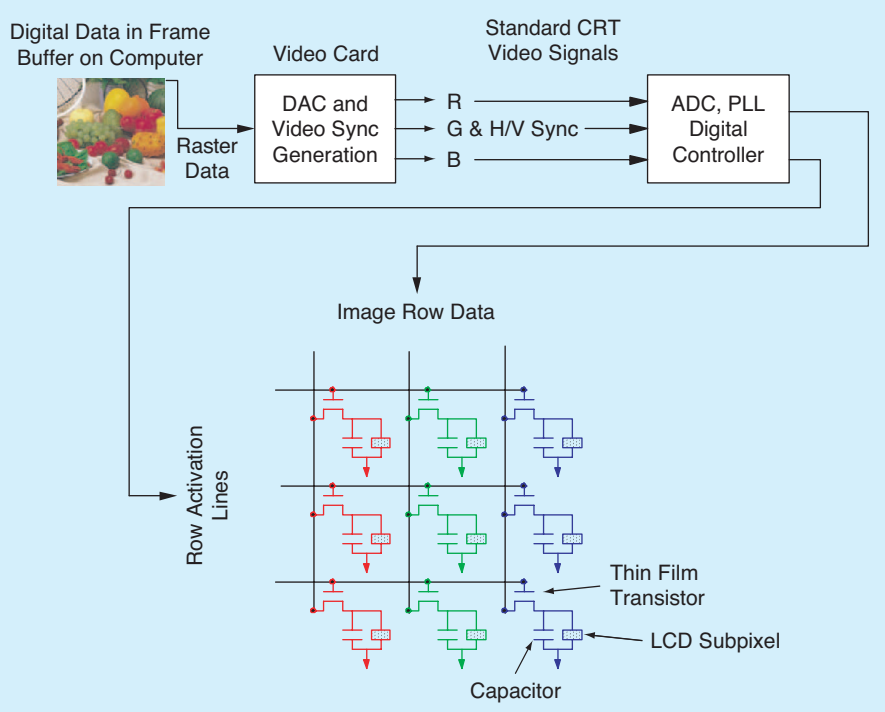
The active matrix version of the LCD is the most common flat panel device used to view color images. The components of the device and the image production process are shown in Figure 12. The system consists of

- a fluorescent light source (there are displays available that use LED light sources), which may be at the top, side, or directly behind the display screen
- a diffusing element, which scatters the light from the lamp
- a rear polarizing element, which transmits light of only one polarization
- a glass element with thin film transistors (TFTs)
- a liquid crystal layer, which provides control for the amount of light that will reach the viewer
- a front glass element with color filters
- a front polarizer element, which transmits only light that is perpendicular in polarization to that of the rear polarizer.

To reduce power consumption and improve efficiency, the fluorescent backlight typically has peaks in the RGB spectral



[FIG12] **Typical components for active matrix LCD display.**



[FIG13] **Typical interface for LCD.**

regions. Ideally, these are close to the filter peak transmittances. To obtain a reasonable image when driven by the same digital signals that produce an image on a CRT, the combination of backlight and color filters are designed to produce chromaticities close to those of the CRT phosphors [12].

The purpose of the liquid crystal element is to change the polarization of the light. In one state, the crystal changes the light's polarization (it actually twists the light) allowing it to be transmitted to the viewer. In the other state, it does not change the polarization causing it to be blocked by the front polarizer. There are several methods used to modulate the light polarization. Early devices used twisted nematic (TN) liquid crystals. This approach is quickly being replaced by multidomain vertical alignment (MVA) and in-plane switching (IPS) technologies, which offer improvements in terms of brightness, contrast, response time, and viewing angle.

The amount of light that reaches the output is a nonlinear function of voltage. The nonlinearity, which defines the crystal transmittance as a function of voltage, depends on the type of crystals that are used in the display. Gray (or color) gradations are achieved by partial voltage application.

In the active matrix LCD, each color element of a pixel is controlled by a single thin-film transistor (TFT). The use of a transistor with a capacitor for charge holding enables the creation of larger displays (since the refresh cycle can be longer) and faster switching liquid crystal material enabling the viewing of video on LCD displays. Due to the legacy of CRT displays, the typical interface of a LCD display with a desktop computer is shown in Figure 13, where the image is converted from digital data to an analog video signal and then back to a digital signal for the LCD controller. The sampling rate of the ADC is dependent on the refresh rate of the incoming video signal and the number of pixels in the LCD. The conversion from digital to analog and back to digital will create distortions. A digital standard for direct digital interface to the display digital controller is being established [21]. During the refresh cycle, the digital controller will activate an image row and apply the image row data as shown in Figure 13.

Other flat panel technologies include field-emission displays [17], organic LED displays [13], and plasma displays [16]. Each of these have their own advantages, which may be in manufacturability, flexibility, efficiency, CRT-like images, and resolution. There are flat panel systems that are not emissive devices like the aforementioned systems but are reflective. These systems have the advantage of working well in bright environments (e.g., outdoors). Many of these display systems are being developed as alternatives to hardcopy [15].

RESEARCH DIRECTIONS AND CHALLENGES

The colorimetric characterization of CRT devices has traditionally been performed through the use of a mathematical model and a small number of measurements. As new displays such as field-emission, plasma, organic LEDs, and traditional LCDs with LED light sources are introduced, it will be of interest to develop models that make the characterization of these devices possible with very few measurements. There will be continued development on hardware methods to create displays with larger gamuts, dynamic ranges, and resolution. There is also significant development in creating more sophisticated portable displays [20]. Since these devices use relatively expensive wireless bandwidth, there will be a need for compression schemes and processing methods [22], which take into account the properties of these display devices. System methods such as those described in this issue [32] would likely be useful.

AUTHORS

Michael Vrhel is a distinguished engineer at Conexant Systems, Redmond, Washington. He graduated summa cum laude from Michigan Technological University with a B.S. in 1987 and the M.S. degree in 1989 and a Ph.D. in 1993, from North Carolina State University, all in electrical engineering. From 1993 to 1996, he was a National Research Council, research associate at the National Institutes of Health (NIH) Bethesda Maryland, where he researched biomedical image and signal processing problems. In 1996, he was a Senior Staff Fellow with the

Biomedical Engineering and Instrumentation Program at NIH. From 1997 to 2002, he was the senior scientist at Color Savvy Systems Limited, Springboro, Ohio. From 2002 to 2004, he was the senior scientist at ViewAhead Technology, in Redmond, Washington. He has two patents and several pending. He has published more than 40 refereed journal and conference papers. He is a Senior Member of the IEEE and a member of the SPIE. He is a guest editor for *IEEE Signal Processing Magazine*.

Eli Saber is an associate professor in the Electrical Engineering Department at the Rochester Institute of Technology. From 1988 to 2004, he was with Xerox. He received the B.S. degree from the University of Buffalo in 1988 and the M.S. and Ph.D. from the University of Rochester in 1992 and 1996, respectively, all in electrical engineering. From 1997 until 2004, he was an adjunct faculty member at the Electrical Engineering Department of the Rochester Institute of Technology and at the Electrical and Computer Engineering Department of the University of Rochester. He is a Senior Member of the IEEE and a member of the Electrical Engineering Honor Society, Eta Kappa Nu, and the Imaging Science and Technology Society. He is an associate editor for *IEEE Transactions on Image Processing*, *Journal of Electronic Imaging*, *IEEE Signal Processing Magazine*, and guest editor for the special section on color image processing for *IEEE Signal Processing Magazine*. He is chair of the Technical Committee on Industry DSP Technology. He was finance chair for the 2002 International Conference on Image Processing and general chair for the 1998 Western New York Imaging Workshop.

H. Joel Trussell received the B.S. degree from Georgia Institute of Technology in 1967, the M.S. degree from Florida State in 1968, and the Ph.D. degree from the University of New Mexico in 1976. He joined the Los Alamos Scientific Laboratory in 1969. During 1978-1979, he was a visiting professor at Heriot-Watt University, Edinburgh, Scotland. In 1980, he joined the Electrical and Computer Engineering Department at North Carolina State University, Raleigh, where he is professor and director of graduate programs. His research has been in estimation theory, signal and image restoration, and image reproduction. He was as associate editor for *IEEE Transactions on Acoustics, Speech, and Signal Processing* and *IEEE Signal Processing Letters*. He was a member and past chair of the Image and Multidimensional Digital Signal Processing Committee of the IEEE Signal Processing Society and was on the Board of Governors. He is a Fellow of the IEEE and the corecipient of the IEEE-ASSP Society Senior Paper Award and the IEEE-SP Society Paper.

REFERENCES

[1] N. Damera-Venkata, B.L. Evans, and V. Monga, "Color error-diffusion halftoning," *IEEE Signal Processing Mag.*, vol. 20, no. 4, pp. 51–58, July 2003.

[2] T.N. Pappas, J.P. Allebach, and D.L. Neuhoff, "Model-based digital halftoning," *IEEE Signal Processing Mag.*, vol. 20, no. 4, pp. 14–27, July 2003.

[3] G. Sharma, H.J. Trussell, and M.J. Vrhel, "Optimal nonnegative color scanning filters," *IEEE Trans. Image Processing*, vol. 7, no. 1, pp. 129–133, Jan 1998.

[4] M. Wolski, J.P. Allebach, C.A. Bouman, and E. Walowit, "Optimization of sensor response functions for colorimetry of reflective and emissive objects," *IEEE Trans. Image Processing*, vol. 5, no. 3, pp. 507–517, Mar. 1996.

[5] M.J. Vrhel, H.J. Trussell, and J. Bosch, "Design and realization of optimal color filters for multi-illuminant color correction," *J. Electron. Imaging*, vol. 4, no. 1, pp. 6–14, Jan. 1995.

[6] P.L. Vora and H.J. Trussell, "Measure of goodness of a set of colour scanning filters," *J. Opt. Soc. Amer.*, vol. 10, no. 7, pp. 1499–1508, July 1993.

[7] P.L. Vora, J.E. Farrell, J.D. Tietz, and D.H. Brainard, "Image capture: Simulation of sensor responses from hyperspectral images," *IEEE Trans. Image Processing*, vol. 10, no. 2, pp. 307–316, Feb. 2001.

[8] R. Merrill, "Color separation in an active pixel cell imaging array using a triple-well structure," U.S. Patent 5,965,875, 1999.

[9] Y. Komiya, K. Ohsawa, K.Y. Ohya, T. Obi, M. Yamaguchi, and N. Ohyama, "Natural color reproduction system for telemedicine and its application to digital camera," in *Proc. ICIP 99,* 24–28 Oct. 1999, vol. 3, pp. 50–54.

[10] Y.V. Haeghen, J.M.A.D. Naeyaert, I. Lemahieu, and W. Philips, "An imaging system with calibrated color image acquisition for use in dermatology," *IEEE Trans. Med. Imag.*, vol. 19, no. 7, pp. 722–730, July 2000.

[11] M. Herbin, A. Venot, J.Y. Devaux, and C. Piette, "Color quantitation through image processing in dermatology," *IEEE Trans. Med. Imag.,* vol. 9, no. 3, pp. 262–269, Sept. 1990.

[12] G. Sharma, "LCD displays vs. CRTs: Color-calibration and gamut consideration," *Proc. IEEE*, vol. 90, no. 2, pp. 605–622, Feb. 2002.

[13] S. Forrest, P. Burrows, and M. Thompson, "The dawn of organic electronics," *IEEE Spectr.*, vol. 37, no. 8, pp. 29–34, Aug. 2000.

[14] K. Werner, "The flowering of displays," *IEEE Spectr.*, vol. 34, no. 5, pp. 40–49, May 1997.

[15] G.P. Crawford, "A bright new page in portable displays," *IEEE Spectr.*, vol. 37, no. 10, pp. 40–46, Oct. 2000.

[16] H. Uchiike and T. Hirakawa, "Color plasma displays," *Proc. IEEE*, vol. 90, no. 4, pp. 533–539, Apr. 2002.

[17] S. Itoh and M. Tanaka, "Current status of field-emission displays," *Proc. IEEE*, vol. 90, no. 4, pp. 514–520, Apr. 2002.

[18] D.E. Mentley, "State of flat-panel display technology and future trends," *Proc. IEEE*, vol. 90, no. 4, pp. 453–459, Apr. 2002.

[19] H. Kawamoto, "The history of liquid-crystal displays," *Proc. IEEE*, vol. 90, no. 4, pp. 460–500, Apr. 2002.

[20] J. Kimmel, J. Hautanen, and T. Levola, "Display technologies for portable communications devices," *Proc. IEEE*, vol. 90, no. 4, pp. 581–590, Apr. 2002.

[21] Digital Display Working Group, "Digital Visual Interface DVI," Rev. 1.0, Apr. 1999 [Online]. Available: http://www.ddwg.org/

[22] J. Luo, K.E. Spaulding, and Q. Yu, "A novel color palettization scheme for preserving important colors," in *SPIE*, 2003, vol. 5008, pp. 409–418.

[23] H.R. Kang, *Color Technology for Electronic Devices*. Bellingham, WA: SPIE Press, 1997.

[24] M.J. Vrhel and H.J. Trussell, "Color device calibration: A mathematical formulation," *IEEE Trans. Image Processing*, vol. 8, no. 12, pp. 1796–1806, Dec. 1999.

[25] R. Bala, "Device characterization," in *Digital Color Imaging Handbook*, Gaurav Sharma Ed., Boca Raton, FL: CRC, 2003.

[26] J. Hynecek, "Theoretical analysis and optimization of CDS signal processing method for CCD image sensors," *IEEE Trans. Electron. Devices*, vol. 39, no. 11, pp. 2497–2507, Nov. 1992.

[27] E. Barnes, "Integrated solutions for CCD signal processing," *Analog Dialogue,* vol. 32, no. 1, pp. 6–8, 1998 [Online]. Available: http://www.analog.com/library/analog Dialogue/archives/32-1/contents.html

[28] K. Buckley, "Selecting an analog front end for imaging applications," *Analog Dialogue*, vol. 34, no. 6, pp. 40–44, 2000 [Online]. Available: http://www.analog.com/library/analogDialogue/archives/34-06/imaging/index.html

[29] H.J. Trussell, E. Saber, M.J. Vrhel, "Color image processing," *IEEE Signal Processing Mag.*, vol. 22, no. 1, pp. 14–22, Jan. 2005.

[30] F.A. Baqai, J.H. Lee, A.U. Agar, and J.P. Allebach, "Digital color halftoning," *IEEE Signal Processing Mag.*, vol. 22, no. 1, pp. 87–96, Jan. 2005.

[31] R. Ramanath, W.E. Snyder, Y. Yoo, and M.S. Drew, "Color image processing pipeline," *IEEE Signal Processing Mag.*, vol. 22, no. 1, pp. 34–43, Jan. 2005.

[32] R. Bala and G. Sharma, "System optimization in digital color imaging," *IEEE Signal Processing Mag.*, vol. 22, no. 1, pp. 55–63, Jan. 2005.

[33] B.K. Guntruk, J. Glotzbach, Y. Altunbasak, R.W. Schafer, and R.M. Mesereau, "Demosaicking: Color filter array interpolation," *IEEE Signal Processing Mag.*, vol. 22, no. 1, pp. 44–54, Jan. 2005.

[34] S. Tominaga and B.A. Wandell, "Natural scene-illuminant estimation using the sensor correlation," *Proc. IEEE*, vol. 90, no. 1, pp. 42–56, Jan. 2002.

[35] M.S. Drew, J. Wei, and Z.N. Li, "Illumination-invariant image retrieval and video segmentation," *Pattern Recognit.*, vol. 32, no. 8, pp. 1369–1388, Aug. 1999.

[36] K.E. Spaulding, G.J. Woolfe, and E.J. Giorgianni, "Optimized extended gamut color encodings for scene-referred and output-referred image states," *J. Imaging Sci. Technol.*, vol. 45, no. 5, pp. 418–426, 2001.

[37] K. Parulski and K. Spaulding, "Color image pprocessing for digital cameras," in *Digital Color Imaging Handbook*, Gaurav Sharma Ed., Boca Raton, FL: CRC, 2003. SP