

# Image Sensor Architectures for Digital Cinematography

*Regardless of the technology of image acquisition (CCD or CMOS), electronic image sensors must capture incoming light, convert it to electric signal, measure that signal, and output it to supporting electronics. Similarly, regardless of the technology of image acquisition, cinematographers can generally agree on a short list of capabilities that a capture medium needs in order to provide great images for big-screen feature films: capabilities such as Sensitivity, Exposure Latitude, Resolving Power, Color Fidelity, Frame Rate, and one we might call “Personality.” This paper will use such a list to evaluate image sensor technologies available for digital cinematography now and in the near future.*

## Image Quality: Many Paths to Enlightenment

The comparison of image sensor technologies for motion pictures is both difficult and complicated. The combination of an image sensor and its supporting electronics are analogous to a film stock; just as there is no single film stock that covers all situations or all cinematographers’ needs, there is no single sensor or camera that is perfect for every occasion. Every decision involves tradeoffs. The same sensor can even be more or less suitable for an application depending on the camera electronics that drive and support it. But no amount of processing can retrieve information that a sensor didn’t capture at the scene.

In designing the sensor and electronics for our Origin® digital cinematography camera, DALSA drew upon its 25 years of experience in CCD and CMOS imager design. Given the demands and limitations of the situation, we determined that the best image sensor design for our purposes was (and still is) a frame-transfer CCD with large photogate pixels and a mosaic color filter array. It is not the only design that could have succeeded, but it is the only design that *has* succeeded. No other design has demonstrated a similar level of imaging performance across the range of criteria we identified above. This is not to say that no other design will reach those performance levels; to bet against technology advancement would be short-sighted. On the other hand, the performance Origin can demonstrate today is several generations ahead of the best we’ve seen from other technologies and architectures, and Origin’s design team is forging ahead to improve it even more.

## Imaging Requirements: “what do cinematographers really want?”

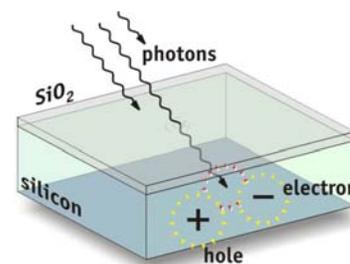
Individual tastes and rankings will vary, but most cinematographers would agree that any imaging medium can be judged by a short list of attributes including those described below.

### Sensitivity

Sensitivity refers to the ability to capture the desired detail at a given scene illumination. Also known as **film speed**. Matching imager sensitivity with scene lighting is one of the most basic aspects of any photography.

Silicon imagers capture image information by virtue of their ability to convert light into electrical energy through the photoelectric effect—incident photons boost energy levels in the

silicon lattice and “knock loose” electrons to create electric signal charge in the form of electron-hole pairs. Image sensor sensitivity depends on the size of the photosensitive area (the bigger the pixel, the more photons it can collect) and the efficiency of the photoelectric conversion (known as quantum efficiency or QE). QE is affected by the design of the pixel, but also by the wavelength of light. Optically insensitive structures on the pixel can absorb light (absorption loss); also, silicon naturally reflects certain wavelengths (reflection loss), while very long and very short wavelengths may pass completely through the pixel’s



photosensitive layer without generating an electron (transmission loss). (Janesick, 1)

Sensitivity requires more than merely generating charge from photogenerated electrons. In order to make use of that sensitivity, the imager must be able to manage and measure the generated signal without losing it or obscuring it with noise.

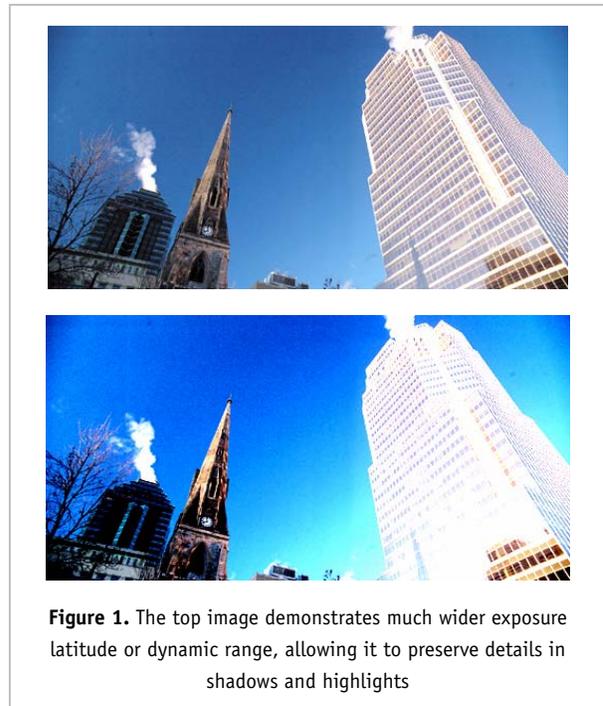
### Exposure latitude

Exposure latitude refers to the ability to preserve detail in both shadow and highlights simultaneously. Some of the most dramatic cinematic effects, as well as the most subtle, depend on wide exposure latitude. For film, latitude is described in terms of usable stops where each successive stop represents a halving (or doubling) of light transmitted to the focal plane. For example, at  $f2.0$  there is 50% less light transmitted than at  $f1.4$ ;  $f2.8$  transmits half as much as  $f2.0$ , and so on. Many film stocks deliver over 11 stops of useful latitude, while broadcast and early digital movie cameras have struggled to deliver more than eight.

In the electronic domain, exposure latitude is expressed as dynamic range, usually described in terms that involve the ratio of the device's output at saturation to its noise floor. This can be expressed as a ratio (4096:1), in decibels (72dB), or bits (12 bits).

It should be noted that not all of a device's dynamic range is linear. Above and below certain levels, device response is not predictable and its output may not be useful. When comparing device dynamic ranges specifications, note whether the value is given as linear—the linear segment is by far the most useful part of the dynamic range. Low noise and a large charge capacity, often contradictory goals, are crucial to delivering great dynamic range.

While extensive research goes into designing pixels to be as sensitive and as quiet as possible in low light, performance in bright light is also very important. Film stocks have been refined to respond to varied lighting with non-linear “toe” and “shoulder” regions for shadows and highlights; this is one of film's defining characteristics. Very few electronic imagers can offer similar performance. In contrast, we have all seen digital images in which extremely bright areas “bloom” or “blow out” the highlight details. The larger a pixel's charge capacity, the wider the range of illumination intensities it can manage. But to contain the brightest highlights without losing detail or blowing out the rest of the image, sensors need “antiblooming” structures to drain away excess charge beyond saturation. By their nature, CMOS pixels offer a high degree antiblooming; in CMOS designs there is almost always a drain nearby to absorb charge overflow. Some (but not all) CCDs also offer antiblooming, although antiblooming almost always involves a tradeoff with full-well capacity. For pixels that are already limited in charge capacity by small active area, good antiblooming performance can reduce exposure latitude significantly. The smaller the pixel, the greater the impact.



**Figure 1.** The top image demonstrates much wider exposure latitude or dynamic range, allowing it to preserve details in shadows and highlights

### Resolving power

Technically, the ability to image fine spatial frequencies through an optical system should be defined as “resolution” (Cowan, 1) but in the electronic domain “resolution” is too often used to mean mere pixel count. For clarity we will use the phrase “resolving power” here. Resolving power is measured in units such as line pairs per degree of arc (from the point of view of a human observer), line pairs per millimeter (on the imaging surface itself), or line pairs per image height (in terms of a display device, with viewing distances given).

Clearly, resolving power is quite different from pixel count. The performance of the pixels (and the lens focusing light onto them) has a huge impact on how much resolving power an imaging system has. Two related terms are **sharpness** and **detail**, both used to describe the amount and type of fine information available in the image, and both heavily influenced by the amount of contrast available at various frequencies in an image (Cowan, 1). Discussion of resolving power, contrast, and frequencies begs the inclusion of the technical term **Modulation Transfer Function** (MTF), which describes the geometrical imaging performance of a system, usually illustrated as a graph plotting modulation (contrast ratio) against spatial frequency (line pairs per unit). As MTF decreases, closely spaced light and dark lines will lose contrast until they are indistinguishably gray. Increasing the number of pixels in an imager will not improve its resolving power if the design choices made in adding pixels reduce MTF. This can happen if the pixels become too small, especially if they become smaller than the resolving power of the lens.

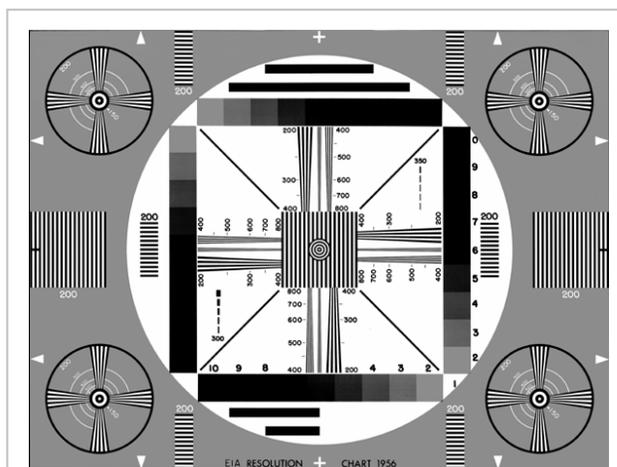
Some film negatives have been tested to exceed 4000 lines of horizontal resolving power. However, prints, even taken directly from the negative, inherit only a fraction of the negative's MTF (see ITU Document 6/149-E, published 2001). The image degrades during each generational transfer from negative to interpositives, internegatives, answer prints, and release prints. Clearly, electronic sensors for digital cinematography will need to be thousands of pixels wide, but exactly how many thousands is less clear. Whatever the display resolution, most cinematographers would prefer to capture as much detail as possible at the beginning of the scene-to-screen chain to have maximum flexibility in postproduction and archiving. The feature film industry has no consensus on sufficient resolution, but clearly "HD" (1920x1080) doesn't capture as much information as a 35mm film negative.

Another factor affecting resolving power is pixel size. At a given pixel count, bigger pixels mean fewer devices per silicon wafer (and therefore higher cost), so we are accustomed to designers making things ever smaller. Consumer digital camera sensors continue to make their pixels smaller to pack more pixels into the same optical format. There are good reasons for not following that route in digital cinematography imagers.

While they occupy more silicon, bigger pixels can provide a performance advantage, such as higher charge capacity (more signal). Fabricated with slightly larger lithography processes, they can handle larger operating voltages for better charge transfer efficiency and lower image lag. These signal integrity benefits must be traded off against power dissipation (battery life and heat), but properly designed, bigger pixels can deliver very low noise and immense dynamic range.

With larger pixels, a high pixel count creates a device considerably larger than the standard 2/3" format common in 3-chip HD cameras. But for the purposes of digital cinematography, this is actually a positive—an imager sized like a 35mm film negative allows the use of high-quality 35mm lenses, which help deliver good MTF. The 2/3" format is an artificial limiter (inherited from 1950s television standards) and should be just one consideration in the overall design of a camera system. In the still camera world, most professionals quietly agree that 5- and 6-megapixel sensors that have the same dimensions as their 3-megapixel predecessors (i.e. smaller pixels) exhibit higher noise. Pixel quality and lens quality have a greater effect on overall image quality than pixel count, above some minimum value.

Resolving power is further complicated by the challenges of capturing color.



**Figure 2.** An imaging system's resolving power can be tested with standard resolution charts such as this "EIA 1956" chart.

## Color fidelity

**Color fidelity** refers to the ability to faithfully reproduce the colors of the imaged scene. For cinematography, it is also vital to maintain the flexibility to allow color to be graded to the desired look in postproduction without adversely affecting the other aspects of image quality. The importance of predictable, stable color performance cannot be understated. Color digital imaging is complicated by the fact that electronic imagers are monochromatic. Silicon cannot distinguish between a red photon and a blue one without color filters—the electrons generated are the same for all wavelengths of light. To capture color, electronic imagers must employ strategies such as recording three different still images in succession (impractical for cinematography), using a color filter array on a single sensor, or splitting the incident light with a prism to multiple sensors. These approaches all have unique impacts on sensitivity, resolving power, and the design of the overall system. Since all electronic imagers share the same color imaging challenges, we will return to them after first touching on sensor architecture.

## Frame rate

Frame rate measures the number of frames acquired per second. The flexibility to allow variable frame rates for various effects is very useful. Television cameras are locked to a fixed frame rate, but like film cameras, digital cinematography cameras should be able to deliver variable frame rates. As usual, there is a tradeoff. Varying frame rates will have an impact on complexity, compatibility, and image quality. It will also have a considerable effect on the bandwidth required to process the sensor signals and record the camera's output.

## “Look” or “Texture” or “Personality”

Many people have their own way to describe the combination of grain structure, noise, color and sharpness attributes that give film in general (or even a particular film stock) its characteristic look. This “look” can be difficult to quantify or measure objectively (although it is definitely influenced by the other items on this list), but if it is missing, the range of tools available to convey artistic intent is narrowed. Electronic cameras also have default signature “looks,” but they can, in some cases, be adjusted to achieve a desired look. However from a system perspective, the downstream treatment of the image, either in camera electronics or in post, cannot compensate for information that was not captured on the focal plane in the first instance. Originating the image with the widest palette of image information practical is clearly the superior approach.

With these criteria in mind, we shall address the available electronic imaging technologies.

## Solid-State Imager Basics

All CCD and CMOS image sensors operate by exploiting the photoelectric effect to convert light into electricity, and all CCDs and CMOS imagers must perform the same basic functions:

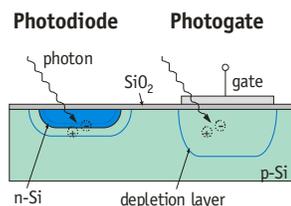
- generate and collect charge
- measure it and turn into voltage or current
- output the signal

The difference is in the strategies and mechanisms developed to carry out those functions.

### Generating and collecting signal charge

While there are important differences between CCD and CMOS, and many differences between designs within those broad categories, CCD and CMOS imagers do share basic elements.

Generating and collecting signal charge are the first tasks of a silicon pixel. The major categories of design for pixels are photogates and photodiodes. Either can be constructed for CCDs or CMOS imagers. Photodiodes have ions implanted in the silicon to create (p-n) metallurgical junctions that can store photogenerated electron-hole pairs in depletion regions around the junction. Photogates use MOS capacitors to create voltage-induced potential wells to store the photogenerated electrons. Each approach has its particular strengths and weaknesses.



Photogates’ major strength is their large fill factor—in a photogate CCD, up to 100% of the pixel can be photosensitive. High fill factor is important because it allows a pixel to make use of more of the incident photons and hold more photogenerated signal (higher full well capacity). The tradeoff for photogates is reduced sensitivity due to the polysilicon gate over the pixel, particularly in the blue end of the visible spectrum.

Photodiodes are slightly more complex structures that trade fill factor for better sensitivity to blue wavelengths. Photodiodes’ sensitivity is not reduced by poly gates, but this advantage is somewhat offset by having less photosensitive area per pixel. The additional non-photosensitive regions in each pixel also reduce photodiodes’ full well capacities.

CMOS pixels, whether photogate or photodiode, require a number of opaque transistors (typically 3, 4, or 5) over each pixel, further reducing fill factor. Each design has ways to mitigate its weaknesses: photogates can use very thin transparent membrane poly gates to help sensitivity (as Origin’s latest CCD does), while photodiodes (both CCD and CMOS) can use microlenses to boost effective fill factor. As we shall discuss later in this paper, these mitigators can bring additional tradeoffs.

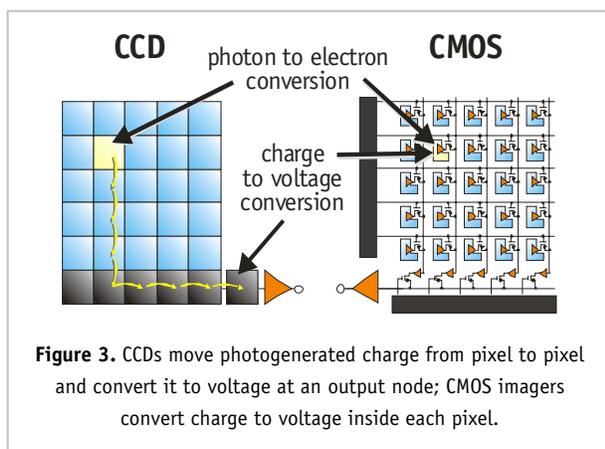
### In Retrospect

CCDs (charge-coupled devices) have been the dominant solid-state imagers since their introduction in the early 1970s. Originally conceived by Bell Labs scientists Willard Boyle and George Smith as a form of memory, CCDs proved to be much more useful as image sensors. Interestingly, researchers (such as DALSA CEO Dr. Savvas Chamberlain) investigated CMOS imagers around the same period of time, but with the semiconductor lithography processes available then, CMOS imager performance was very poor. CCDs on the other hand could be fabricated (then as now) with low noise, high uniformity, and excellent overall imaging performance—assuming the use of an optimized analog or mixed-signal semiconductor process. Ironically, as CMOS imagers have evolved, the quest for better performance has led CMOS designers away from the standard logic and memory fabrication processes where they began to optimized analog and mixed-signal processes very similar to those used for CCDs.

All foundry equipment and process developments are capital-intensive, and image sensors’ low volume (relative to mainstream logic and memory circuits) mean they are relatively high-cost devices, especially where high performance is concerned. CCD and CMOS imagers have comparable cost in comparable volumes. In performance-driven applications, the key decision is not CCD vs. CMOS; instead, it is individual designs’ suitability to task.

## Measuring signal

To measure accumulated signal charge, imagers use a capacitor that converts the charge into a voltage. With CCDs, this happens at an output node (or a small number of output nodes), which also amplifies the voltage to send it off-chip. To get all of the signal charge packets to the output node, the CCD moves charge packets like buckets in a bucket brigade sequentially across the device. This is one of the biggest differences between CCDs and CMOS imagers—CCDs move signal from pixel to pixel to output node in the charge domain, while CMOS imagers convert signal from charge to voltage in each pixel and output voltage signals when selected by row and column busses.



Within each broad category there are more differences. Among CCDs, interline transfer (ILT) sensors have light-shielded vertical channels connected to each pixel for charge transfer, like cubicles with corridors (see Figure 5). Full-frame CCDs don't need separate corridors—to move the charge they just collapse and restore the electrical walls between the pixel cubicles. Since CCDs use a limited number of output amplifiers, their output uniformity is very high. The tradeoff for this uniformity is the need for a high-bandwidth amplifier, since a cinematography imager will output many millions of pixels per second. Amplifier noise often becomes a limiter at high pixel rates. Optimizing amplifiers to meet these demands is a critical aspect of imager design.

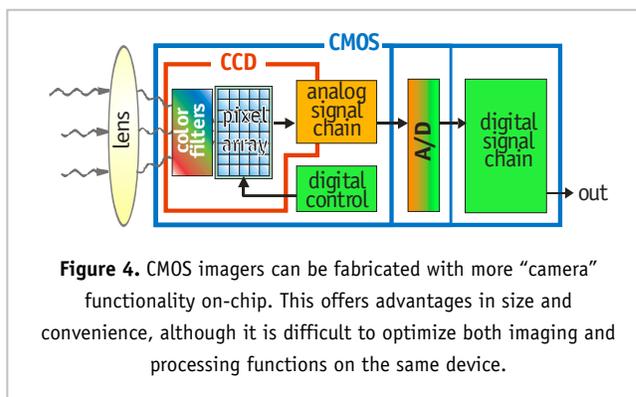
Each CMOS pixel converts its collected signal charge into voltage by itself, but beyond this fact there are differences in designs. From one amplifier per sensor to one amplifier per column, designs have evolved to place an amplifier in each pixel to boost signal (at the expense of fill factor). The more amplifiers, the less bandwidth and power required by each, but millions of pixels mean millions of amplifiers. Since amplifiers are ultimately analog structures, uniformity is a challenge for CMOS imagers and they tend to exhibit higher fixed-pattern noise.

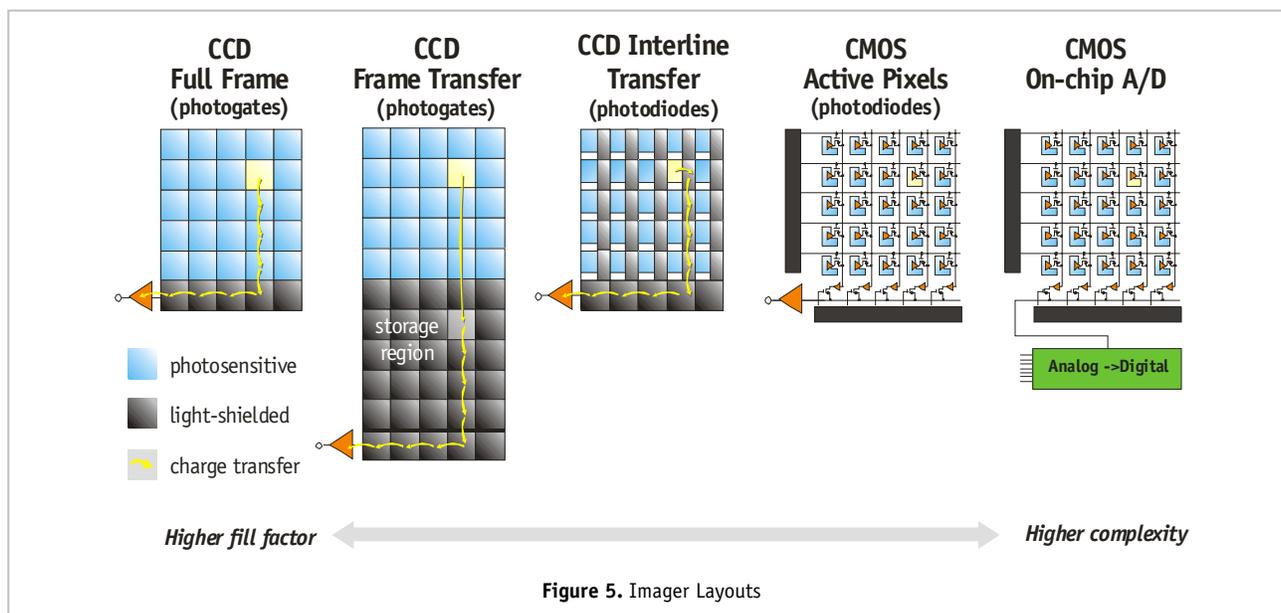
## Outputting signal

CCDs' bucket brigade operation outputs each pixel's signal sequentially, row by row and pixel by pixel. CMOS pixels are connected to row and column selection busses. These opaque metal lines impact fill factor, but allow random access to pixels as well as the ability to output sub-windows of the total imaging region at higher frame rates. This can be useful in industrial situations (motion tracking within a scene, for example), but has limited use in digital cinematography for the big screen.

Most imagers output analog signals to be processed and digitized by additional camera electronics, but it is also possible to place more processing and digitization functionality on-chip to create a "camera on a chip." This has been demonstrated with CMOS imagers and is in theory possible with CCDs as well, although it would be impractical. The analog process lines that have been honed and optimized for CCD imager performance are not well suited to additional electronics. Adding more functionality would require extensive process redevelopment and add a lot of silicon to each device, translating into considerable expense. It would also most likely reduce imaging performance and cause excessive power dissipation since CCDs tend to use higher voltages than CMOS imagers. CCD camera designers have tended to adopt a modular approach that separates imagers from image processing, finding it more flexible and far easier to optimize for performance.

In contrast, designers have taken advantage of the smaller geometries and lower voltages used in CMOS imager fabrication to implement more functionality on-chip. The convenience is clear from a system integration perspective: smaller overall device, usually a single input voltage, lower system power dissipation, digital output. But the convenience has tradeoffs. The chip becomes larger and much more complex, dissipating more power, generating more substrate noise and introducing more non-repairable points of failure to affect device yield. As always it is difficult to optimize both the imaging and processing functions at the same time, especially for the level of performance demanded in cinematography. The most commercially successful CMOS imagers to date have not integrated A/D and image processing on-chip; rather, they have optimized for imaging only and followed the modular camera electronics approach.





## Designs in More Detail

### Full Frame CCDs

CCD “full frame” sensors (not to be confused with the “full frame” of 35mm film) with photogate pixels are relatively simple architectures. They offer the highest fill factor, because each pixel can both capture charge and transfer it to the next pixel on the way to the output node (this is the “charge coupling” part from “charge coupled device”). High fill factor (up to 100%) tends to offset their slightly lower sensitivity to blue wavelengths and allows them to avoid the tradeoffs associated with microlenses.

Full frame CCDs provide an efficient use of silicon, but like film, they require a mechanical shutter. This is a non-issue in digital cinematography if the camera is designed with the rotating mirror shutter required for an optical viewfinder. Without a shutter, however, images from a full frame CCD would be badly smeared while the sensor read out the image row by row.

With the highest full well capacity, photogate full frame architecture provides a head start on high dynamic range. CCD designs and fabrication processes have been optimized over the years to minimize noise (such as dark current noise and amplifier noise) in order to preserve dynamic range. Minimizing amplifier noise, especially at high bandwidth operation, is very important since all pixels pass sequentially through the same amplifier (or small number of amplifiers). This sequential output is a limiter to frame rate—the amplifier can run only so fast before image quality begins to suffer.

To some eyes, the antiblooming performance of full frame sensors (via vertical antiblooming structures that preserve fill factor) provides a softer, more film-like treatment of extremely bright highlights. This is an aspect of imager “personality” that is difficult to define or measure and is open to interpretation.

### Frame Transfer CCDs

A variation of the full frame CCD architecture is the frame transfer design, which adds a light-shielded storage region of the same size as the imaging region. This sensor architecture performs a high-speed transfer to move the image to the storage region and then reads out each pixel sequentially while it accumulates the next image’s charge. This design improves smear performance and allows the sensor to read out one image while it gathers the next; the tradeoff is the cost of twice as much silicon per device and more complex drive electronics which can increase power dissipation.

Frame transfer CCDs have many of the same strengths and limitations as full frame CCDs: high fill factor, and charge capacity, slightly lower blue sensitivity, high dynamic range, and highly uniform output enabled (and limited ) by a small number of high-bandwidth output amplifiers.

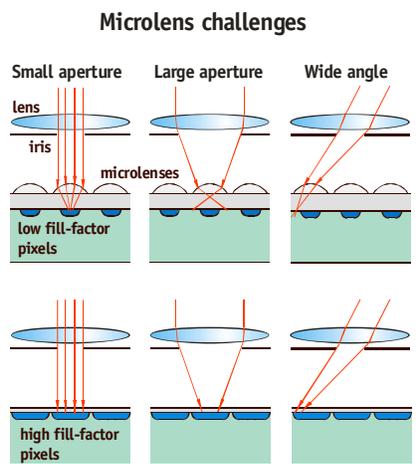
Origin uses a large frame-transfer CCD with large pixels. Combined with the high fill factor, the large pixel area and transparent thin poly gates allow the latest Origin sensor to offer ISO400 performance in the camera. The huge charge capacity and advanced, low-noise amplifiers also allow tremendous dynamic range—more than 12 linear stops plus nonlinear response above that (courtesy of vertical antiblooming and patent-pending processing). Origin’s sensor uses multiple taps to enable high

frame rates, and while these taps must be matched by image processing circuits in the camera, DALSA deemed this an acceptable tradeoff for being able to deliver 8.2 million pixels with very high dynamic range at elevated frame rates of up to 60fps.

### ILT CCDs

Interline transfer CCDs use photodiode pixels. Sensitivity is good, especially for blue wavelengths, but this is offset by low fill factor due to the light-shielded vertical transfer channels that takes the pixel's collected charge towards the output node. The advantage of the shielded vertical channels is a fast and effective electronic shutter to minimize smear, but this is not a critical feature for digital cinematography.

To compensate for lower fill factor (typically 30-50%), most ILT sensors use microlenses, individual lenses deposited on the surface of each pixel to focus light on the photosensitive area. Microlenses can boost effective fill factor to approximately 70%, improving sensitivity (but not charge capacity) considerably. The disadvantage of microlenses (besides some additional complexity and cost in fabrication) is that they make pixel response increasingly dependent on lens aperture and the angle of incident photons. At low f-numbers, microlensed pixels can suffer from vignetting, pixel crosstalk, light scattering, diffraction (Janesick, 2), and reduced



While microlenses help fill factor, they do not alter an ILT pixel's full-well capacity. Lower full-well capacity means that while their overall noise levels are comparable, ILT devices generally have lower dynamic range than full-frame CCDs.

Like other CCDs, ILTs have a limited number of output nodes, and so their output uniformity is high and their frame rates are limited accordingly.

### 3T CMOS

The first "passive" CMOS pixels (one transistor per pixel) had good fill factors but suffered from very poor signal to noise performance. Almost all CMOS designs today use "active pixels," which put an amplifier in each pixel, typically constructed with three transistors (this is known as a 3T pixel). More complex CMOS pixel designs include more transistors (4T and 5T) to add functionality such as noise reduction and/or shuttering. In some senses, the comparison between 3T and 4/5T CMOS imagers is similar to the comparison between full-frame and ILT CCDs. The simpler structures have better fill factor (although the full-frame CCD's fill factor remains much higher than the 3T CMOS pixel), while the more complex structures have more functionality (e.g. shuttering).

In-pixel amplifiers boost the pixel's signal so that it is not obscured by the noise on the column bus, but the transistors that comprise amplifiers are optically insensitive metal structures that form an optical tunnel above the pixel, reducing fill factor. As a result, most CMOS sensors use microlenses to boost effective fill factor. The tradeoffs involved with microlenses are more pronounced with CMOS imagers since the microlenses are farther from the photosensitive surface of the pixel due to the "optical stack" of transistors. As with ILT CCDs, this can affect resolving power and color fidelity.

Fill factors can also be increased by using finer lithography in the wafer fabrication process (0.25 $\mu$ m, 0.18 $\mu$ m...), but this comes with its own set of tradeoffs. While a reduction in geometry reduces trace widths, it also makes shallower junctions and reduces voltage swing, making it more difficult to gather photogenerated charge and measure it—voltage swing is a major limiter to dynamic range because the noise floor stays fairly constant. Smaller geometries also make devices more susceptible to other noise sources. Narrowing traces does not reduce the height of the optical stack either, so all the aperture-dependent microlens effects still apply to finer lithography. And once again, standard logic and memory semiconductor processes do not yield high-performance imagers. Imagers require customized, optimized analog and mixed-signal semiconductor processes; ever-smaller imager-adapted processes are very costly to develop. The tradeoffs involved in using smaller geometries will not be worthwhile for all applications.

Where frame rates are concerned, CMOS can demonstrate good potential. Higher frame rates are possible because pixel information is transmitted to outside world largely in parallel as opposed to sequentially as in CCDs. With more output amplifiers, bandwidth per amplifier can be very low, meaning lower noise at higher speeds and higher total throughput. On the other hand, the outputs have lower uniformity and so require additional image processing. Imaging processing is often a bandwidth limiter for imaging systems attempting to perform high precision calculations in real time for high frame rates.

In-pixel amplifiers let 3T CMOS pixels generate useful amounts of signal, but their noise performance still lags behind CCDs, thus limiting dynamic range.

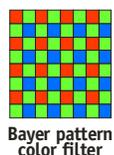
## 4T/5T CMOS

To improve upon 3T performance, designers have tweaked fabrication processes and/or added more transistors. Pinned photodiodes, a concept originally developed for CCDs, use additional wafer implantation steps and an additional transistor to improve noise performance (particularly reset noise), increase blue sensitivity, and reduce image lag (incomplete transfer of collected signal). The tradeoffs are reduced fill factor and full-well capacity, but with their much better noise performance, 4/5T CMOS pinned photodiodes can deliver better dynamic range than 3T designs.

Other designs add a transistor that can allow global shuttering or correlated double sampling (but not at the same time). Global shuttering avoids image smear or distortion of fast-moving objects during readout, while CDS reduces noise by sampling each pixel twice, once in dark and again after exposure. The dark signal is subtracted from the exposure signal, eliminating some noise sources. CDS is used widely in electronic imaging, but a 5T CMOS imager can perform it in-pixel instead of using camera electronics.

## The Complications of Color

One of the factors complicating electronic image capture is the fact that electronic imagers are monochromatic. Silicon cannot distinguish between a red photon and a blue one without color filters—the electrons generated are the same for all wavelengths of light. To capture color, silicon imagers must employ strategies such as recording three different images in succession (impractical for any subject involving motion), using a color filter array on a single sensor, or splitting the incident light with a prism to multiple sensors.

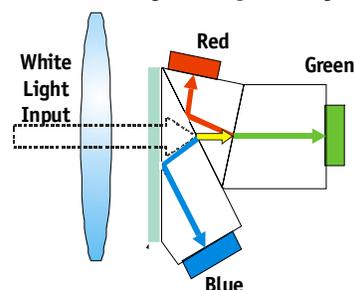


Bayer pattern color filter

A color filter array (CFA) mosaic such as a Bayer pattern allows the use of a single sensor. Each pixel is covered with an individual filter, either through a cover glass on the chip package (hybrid filter) or directly on the silicon (monolithic filter). Each pixel captures only one color (usually red, green, or blue), and full color values for each pixel must be interpolated by reference to surrounding pixels. Compared to a monochrome sensor with the same pixel count and dimensions, the mosaic filter approach lowers the spatial resolution available by roughly 30%, and it requires interpolation calculations to reconstruct the color values for each pixel. However, a mosaic filter's great strength is its optical simplicity: with no relay optics it provides the single focal plane necessary for the use of standard film lenses. The best mosaic filters provide excellent bandpass transmission,

separating the colors with a high degree of precision and providing very stable color performance over time with minimal crosstalk. Of course it goes without saying that inferior filters, inferior sensors, or inferior processing algorithms will give inferior images. But modern demosaic algorithms work extremely well, and all of the best professional digital SLR and studio cameras use mosaic filters. Since lenses govern what an imager “sees,” the importance of the single focal plane and standard lensing should not be underestimated.

Multiple-chip prism systems produce images in separate color channels directly. The imagers are uncomplicated—each sensor is devoted to a single color, preserving all its spatial resolution. The



prism, on the other hand, is not simple. Aligning and registering the sensors to the prism requires high precision. Misaligned or imprecise prisms can cause color fringing and chromatic aberration. In theory, for pixels of the same size,

prism systems should allow higher sensitivity in low light conditions, since they should lose less light in the filters. In practice, this advantage is not always available. Beamsplitting prisms often include absorption filters as well, because simple refraction may not provide sufficiently precise color separation.

The prism approach complicates the optical system and limits lens selection significantly. The additional optical path of the prism increases both lateral and longitudinal aberration for each color's image. The longitudinal aberration causes different focal lengths for each color; the CCDs could be moved independently to each color's focal point, but then the lateral aberration would produce different magnification for each color. These aberrations can be overcome with a lens specifically designed for use with the prism, but such camera-specific lenses would be rare, inflexible, and expensive.

Most 3-chip systems have used small imagers, but experimental systems have been built by NHK (Mitani, 5) and Lockheed Martin that use large format, high resolution sensors in a 3-chip prism architecture. Both require huge “tree trunk” custom lenses whose bulk and cost make them impractical for most applications. Three-chip prism systems also require three times the bandwidth and data storage capacity, creating challenges for implementing a practical recording system.

Yet another approach for deriving spectral information seeks to use the silicon itself as filter. Since longer (red) wavelengths of light penetrate silicon to a greater depth than shorter (blue) wavelengths, it should be possible to stack photosites on top of each other to use the silicon of the sensor as a filter. This is the

architectural approach of the Foveon “X3” sensors. The idea is not new—Kodak applied for patents on this approach in the 1970s, but never brought it to market. In practice, silicon alone is a relatively poor filter. Prisms and focal plane filters have far more precise transmission characteristics. Another challenge of this approach is that the height of each pixel’s “optical stack” not only reduces fill factor, it tends to exaggerate undesirable effects such as vignetting, pixel crosstalk, light scattering, and diffraction. For example, red and blue photons may enter at an angle near the surface of one pixel, but the red may not be absorbed until it enters a different pixel. Again, these effects are most prominent with small pixels and wide apertures and are exaggerated by microlenses. Additionally, the extensive circuitry required for stacked photosites introduces more noise sources to the imager. Any solutions to these challenges will add complexity to the system design, particularly for higher performance applications. As a point of perspective, the stacked photosite approach has not gained traction in the professional digital photography market.

## Summary

An image sensor is just one component in a system. A camera cannot improve the output of a poor sensor, but it can degrade the output of a good one. A good sensor cannot save a bad camera, although a good camera must start with a good sensor. Camera system design, like sensor design, involves tradeoffs, and there is no “right” design, only one that meets the needs of an application and its audiences.

Regardless of the technology of capture (CCD or CMOS), electronic image sensors for digital cinematography must deliver high performance in sensitivity, exposure latitude, resolving power, color fidelity, and frame rate with an agreeable “Personality.” They must be designed with their situations and systems of use in mind—lenses are good examples of non-sensor, non-electronic system elements that affect sensor performance (and design) considerably.

DALSA has designed leading-edge CCD and CMOS imagers for 25 years. Given the demands and limitations of the situation, we determined that the best imager design for our purposes was (and still is) a frame-transfer CCD with large photogate pixels and a color filter array. It is not the only design that could have succeeded, but it is the only design that *has* succeeded. No other design has demonstrated a similar level of imaging performance across the range of criteria we identified above. This is not to say that no other design will reach those performance levels; to bet against technology advancement would be short-sighted. On the other hand, the performance Origin can demonstrate today is several generations ahead of the best we’ve seen from other technologies and architectures, and Origin’s design team is

forging ahead to improve it even more. When we look to the future of digital cinematography, we see a clear, bright, colorful vision—one with high sensitivity, variable frame-rates and tremendous exposure latitude, of course.

## References and More Information

*35mm Cinema Film Resolution Test Report*, Document 6/149-E, International Telecommunications Union, September 2001.

Cowan, Matt. *Digital Cinema Resolution—Current Situation and Future Requirements*, Entertainment Technology Consultants, 2002,

Hornsey, Richard. *Design and Fabrication of Integrated Image Sensors*, course notes, University of Waterloo.

Janesick, James. *Dueling Detectors*, OE Magazine, February 2002.

Litwiller, Dave. *CCD vs. CMOS: Facts and Fiction*, Photonics Spectra, January 2001.

Mitani, Kohji, et al. *Experimental Ultrahigh-Definition Color Camera System with three 8M-pixel CCDs*, SMPTE 143rd technical Conference and Exhibition, New York City, November 2001.

Theuwissen, Albert. *Solid-State Imaging with Charge-Coupled Devices*, Kluwer Academic Publishers, Dordrecht, 1996.

Theuwissen, Albert, and Edwin Roks. *Building a Better Mousetrap*, OE Magazine, January 2001.

### DALSA Corp.

605 McMurray Rd.  
Waterloo Ontario, Canada  
N2V 2E9  
dc@dalsa.com  
[www.dalsa.com/dc](http://www.dalsa.com/dc)