Eastman Kodak Company

An Introduction to Electronic Cameras

By Kris S. Balch (1995)

An Introduction to Electronic Cameras

System Overview

This section will describe the fundamental digital technology used in electronic cameras. The camera technology will first be discussed with a look at the essential components needed for an image capturing system. We will close with a short overview into essential performance issues.

Imaging Chain

Let's begin by examining what building blocks are required for constructing an imaging system. For our purposes, we will describe this imaging system as an imaging chain. The analogy is that a chain has many links and each link must work together to bind as chain. If we remove a link or have a weak link, the chain will fail to hold together. This same analogy can be applied to the different links that make up the imaging chain. If their performance is not to the standards of other links, the overall system is degraded. Therefore, it is very important to understand each link's

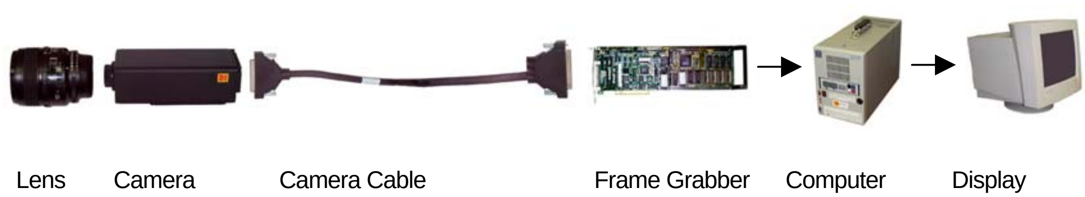


Figure 1 Typical Imaging Chain

contribution. In the most simplistic configuration, the links include: lens, camera, camera cable, frame grabber, image processor (computer), and display.

Lens

The lens is selected to appropriately frame the scene. The higher resolution cameras will typically use an F-mount lens although some cameras do use a C-mount. An F-mount lens has a larger percentage of area within the Field-of-View (FOV) that is maintained at a lower distortion tolerance than that of a C-mount lens. Some important issues concerning a lens selection are the f-stop and the focal length range. A fast lens (low f-stop number) will be very useful in low lighting conditions. However, we must be careful not to lose the depth-of-field (DOF) at a fast lens setting. The focal length is directly proportional to the FOV that we're trying to image. If we need to magnify the subject in the FOV, we select a lens that also has a macro capability.

Camera

The camera is essentially the eyes of your system. One wants the best vision possible within your system constraints. When selecting a camera, there are a number of important issues to consider. More will be discussed later under camera performance.

Camera Cable

The camera cable provides the connection between the camera and the frame grabber. Communicated through this cable are the timing control generated by the frame grabber, functional commands sent to the camera from the computer (via the frame grabber) that defines its operational mode, and image data coming from the camera. It is important that the cable be constructed so that sufficient shielding and proper grounding is provided. Otherwise, the cable may radiate and fail FCC requirements, the image data may be corrupted with cross talk between lines in the cable, and extraneous noise may be coupled into the image data from outside noise sources. One way to assure that a proper cable design has been executed is to check if the camera has passed FCC Class 15 regulation, CE and other regulatory requirements.

Frame Grabber

A simple frame grabber(1) composes an image file from the lines of data being sent from the camera and provides temporary memory in the imaging chain. There are many frame grabbers on the market with costs ranging from hundreds to thousands of dollars. The price range leads one to examine both obvious and subtle differences between frame grabbers. We will look at the display resolution, the accuracy of the pixel data, and the image frame rate speed.

The display capabilities of frame grabbers designed for high-resolution cameras should match the camera's resolution. Typically, these frame grabbers can interface with either non-standard or non-interlaced cameras. This allows for higher resolutions above standard video. Some low-end frame grabbers are only compatible with RS170, NTSC, and PAL video standards. The disadvantage in using these standard frame grabbers is the limited display resolution, speed (frames per second), and the fact that the video is interlaced (each frame split into odd and even fields).

The processing speed of real-time video imaging demands high throughput because images are large chunks of data. The PCI Bus is an excellent choice for a frame grabber because it supports at a high level of data throughput when coupled with a fast CPU. The PCI bus is essentially platform independent (PC, MAC, ect.) and the frame grabber can utilize system memory, lowering the cost of the frame grabber.

The accuracy of the image being displayed depends largely on the integrity of the camera data converted back to analog for display. The frame grabber's digital-to-analog (d/a) converter dictates this accuracy. Care should be taken in understanding the capabilities of the d/a converter. Converters that have missing codes or use a conversion scheme that produces less than 8 bits of information are not suitable for scientific and industrial imaging since the data can be corrupted. Therefore, select a frame grabber that demonstrates a true 8-bit conversion of data.

Other hardware features affecting image integrity are:

- timing circuits that introduce pixel jitter due to poor synchronization between the frame grabber and video timing
- a/d reference adjustments that allow tuning of the frame grabber input data range to match the video signal

Frame grabbers containing these features are going to cost more, but will produce images that are more accurate.

Computer

The computer is the intelligence in the imaging chain. It controls and powers the frame grabber. It provides most of the user interface to the camera through software. Many types of computers are supported by the various manufacturers of frame grabbers. These include: PCs, Macs, Sun Workstations, Dec, and others. Some performance parameters to consider are the hard drive storage capacity, the amount of DRAM memory, the CPU speed as well as the bus bandwidth. A short discussion on each of these parameters will be given in the context of a PC platform.

Hard Drive

Select a system with as large of a hard drive as you can afford because storing images will consume many megabytes. In addition, you want a drive that has a large sustained bandwidth for fast transfer of images (sustained rates of 1–4 MB/s). There are two popular types of hard drives, IDE and SCSI. Either one will be suitable for imaging applications. The storage capacity of IDE and some SCSI drives are now topping over 10 GB.

DRAM Memory

The computer should have at least 8 MB. However, 16 to 32 MB is more suitable for an imaging platform. Most memory today comes in 72-pin SIMM modules that easily plug into the sockets on the motherboard.

CPU and BUS Bandwidth

There are several popular bus architectures today. The VESA (Video Electronics Standards Association) local bus and the PCI (Peripheral Component Interconnect) have high bandwidths for moving imaging data. The PCI bus is arguably the best choice for frame grabbers. CPU speeds over 100 MHz is essential. As with hard drive capacity, obtain the fastest CPU that you can afford. Images consume bandwidth very fast in a computer.

Display

The last link in the imaging chain, the display, reverses the photon-to-electron process and produces an image from the electrical signal. Important parameters to consider concerning a display are the monitor's resolution (dot pitch), the refresh method and rate, and screen size. Other factors that may influence a decision are the way it looks as well as its price.

Dot pitch must be examined carefully. It is used to represent the distance between the centers of two adjacent dots within a color mask. Some manufacturers are reporting the dot pitch as an equivalent measurement. Shown in the figure is a graphic representation of how some dot pitches may be specified. Comparing monitor specs are more difficult with all the various methods of measurement. The most consistent method is from the center-to-center spacing from adjacent dots. A pitch below 0.31 will be good. (2)

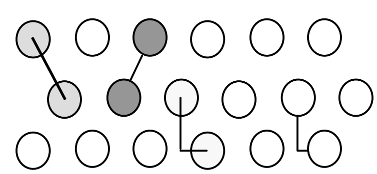


Figure 2 Dot pitch distance

There are two refresh methods, interlace and non-interlace. Non-interlace, also called progressive scan, is best for high-resolution imaging. Interlace monitors are lower performance displays due to their dependency of scanning every other line for each field. The non-interlace monitors have been designed to scan a frame in one sweep. The refresh rate is important because if it is too low, the screen will have a flickering effect. For a high-resolution monitor, the refresh rate should be above 60 Hz. The better monitors will be around 72 Hz or greater.

Screen size is measured on a diagonal from one corner to the opposite corner. The actual viewing area will be slightly less than the manufactures rated screen size. This is due to the measurement method. For high-resolution imaging, a screen size of 17 inch or greater will be very suitable for the application.

Camera Construction

Architecture

The basic construction of an electronic camera should be explained next. The fundamental building blocks for a camera include a mechanical housing, sensor, clock drivers, pre-amp, signal processing channels, analog-to-digital converter, timing controller, power supply (Bias Levels) and camera interface. There may be other building blocks depending on the amount of camera's functionality.

Mechanical Housing

The mechanical housing needs to physically support the camera electronics, provide an accurate mounting structure for the sensor-lens assembly, dissipate heat from the camera electronics, and mount all cable/control interfaces.

Sensor

A sensor [a Charged Coupled Device (CCD)] is the heart of an electronic camera's design and defines some of the

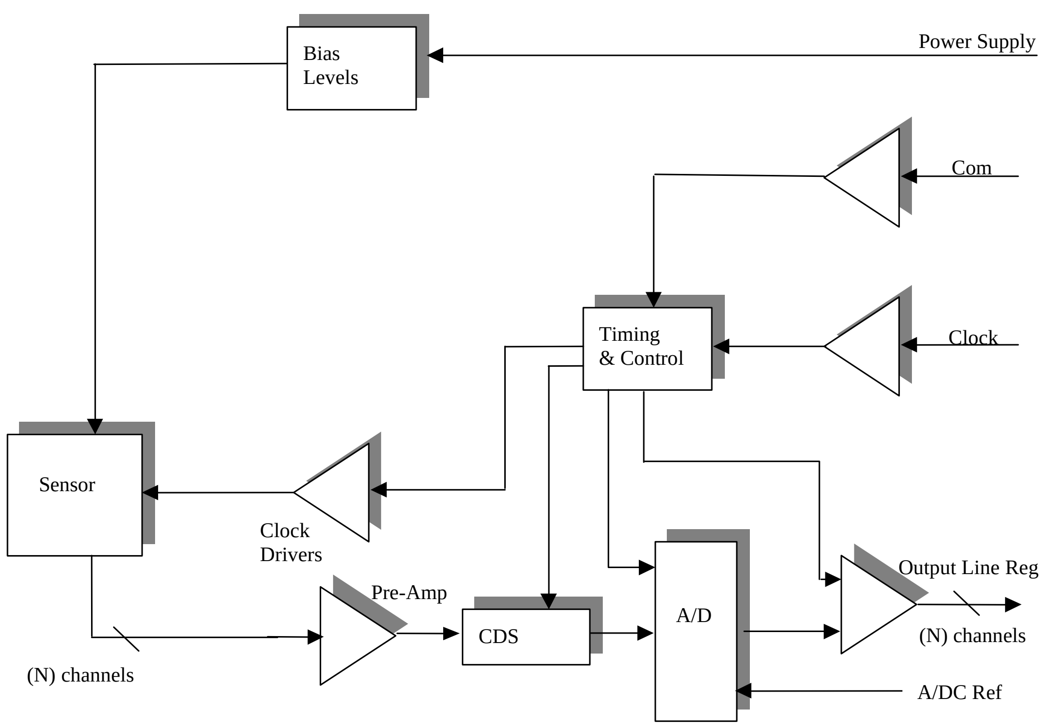


Figure 3 CCD Camera Building Blocks

most important performance attributes. The CCD performs the task of capturing photons of light energy striking its structure, which are then converted into electrical charges proportional to the amount of illumination seen. A sensor is divided structurally into a number of "capturing sites or wells," termed pixels, to collect the charges. Pixels may be either square or rectangular in dimensions. The more common designs will find the sensor's pixels arranged in neat rows and columns.

Using an analogy, each pixel site can be thought of as a water bucket. If each photon would transform into a raindrop, then each bucket collects an amount of water equal to the quantity of the incoming light at that point in the captured scene. Each pixel then contains one piece of information or detail about the scene. The more pixels in a sensor's structure therefore, the more details, or the more resolving power exists for the scene. This ultimately aids in the correct interpretation of image information. (3)

Sensor Driver and Pre-Amp

The clock drivers control the process of capturing photons on the sensor, converting them to an electron charge and moving the charge out of the sensor. The sensor is followed by a pre-amp. The pre-amp amplifies the signal off the sensor to a level sufficient for processing.

Signal Processors & Data Conversion

The signal processors are responsible for conditioning the signals from the pre-amp to the analog-to-digital conversion. The signal processing provides gain and offset that amplifies the signal and removes unwanted noise generated in the sensor. Typically, this block is a Sample & Hold that retains an averaged sample of the video. This video is then digitized by an A/D converter. After the A/D is a register, that buffers the digital signals for transmission out of the camera. This channel architecture is replicated if there are multiple outputs.

Output Line Register & Timing Controller

The Output Line Register buffers the image data for transmission over the camera cable. The timing controller is responsible for both high-level command interpretation and the generation of pixel-level timing signals needed by all other camera electronics.

Power Conversion

All special voltages necessary for the sensor clock drivers and A/D converters are derived from a single-voltage DC source. The voltages generated are used to translate the pixel-level timing generated by the timing controller through the clock drivers into precise timing and voltage drive signals.

Sensor

There are three sensor(3) architectures most commonly found in digital cameras: full frame, frame transfer, and interline. A short description for each structure will be given along with their performance characteristics.

Full Frame

In a full-frame sensor, a high percentage of each pixel site can be designed to capture and translate the incoming energy. Typically referred to as the "fill factor," this capturing and translating can approach 100%. However, a situation can occur, using our water bucket analogy, whereby the incoming light is so strong that the bucket fills and overflows (pixel blooming). Since the sensor is an array, adjacent buckets (pixels) are on the sides of the overflowing bucket (pixel). The only place for the excess overflow water to go is into the neighboring buckets. This overflow process, if excessive, can continue until all of the excess has been propagated to many adjacent pixels. This situation results from scenes that have bright, spectacular areas.

This is referred to as "blooming" of the sensor. The effect in an image is an area or vertical stripe, which saturates to full white level, indicating the pixels are saturated. An area of high illumination near an object of interest may be corrupted to the point that data cannot be accurately interpreted. Fortunately, if one is willing to give up some of the fill factor, some of the pixel site area can be configured to function as an overflow drain. This drain is able to bleed off the excess charge once the pixel's well (or bucket) is filled so as not to corrupt its neighbors. This overflow drain is an extremely important feature in preserving good data to interpret imaging information, or in an application where illumination cannot be totally controlled.

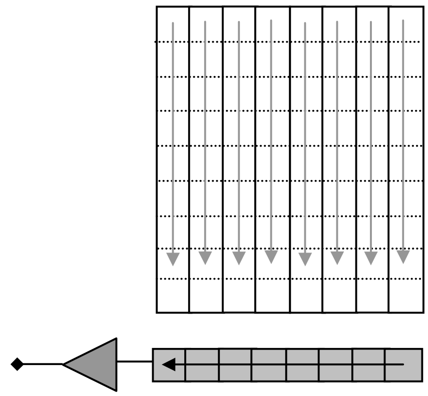


Figure 4 Full-Frame Sensor Architecture

The process of imaging with a full-frame sensor can be related to that of film in a camera. Film has no ability to capture an image at one moment in time and stop the incoming light until the film is advanced. If no means is employed to stop the light before the film is advanced it will either be blurred or fully exposed. The same situation exists for a full-frame sensor. In this case, the sensor's charge must be read off all the pixel sites to be ready for the next image capture. The answer is similar to that of the film camera; that is, some form of shuttering system must be incorporated. The shutter is opened for a specified amount of time to let the image's light fall upon the sensor for recording and then closed while the readout takes place.

Reading or collecting the sensor's data is handled through one or more readout registers located at the edge of the sensor. While different conventions exist, one of the more common is to have the first edge row of pixels transfer their values into an associated position in the adjoining register. Once there the charges are cascaded off the register in a serial fashion from one end. Once the last row of pixels has had its charges moved into the register, a transfer is set up to move all of the next to last row's pixel charges into the associated last row positions. This process prepares for the next cycle of moving charges into the readout register. In a similar fashion this cascading row movement continues until each remaining row has been shifted one row closer to the readout register.

Frame Transfer

A frame transfer CCD device can be similar to that of the full frame design just described. The major difference is that the sensor ends up as two adjoining sensor areas. The second area has a mask on the pixel's surface that blocks any light from adding charge to the storage site. In the sensor's operation, the light builds up charge in the unmasked area for a given amount of exposure time. Then, very rapidly, all the pixels have their collected charges transported to the corresponding pixel sites under the masked area. From that area, the charges can be read off the sensor to be translated into an image for viewing.

Before the next capture can begin, any accumulated charges in the unmasked pixel area must be drained off. The sensor is then ready to start the capture process again. Depending upon the specific architectural design of the sensor, smearing of the image can occur in this design because of the need to transfer the collected image rapidly to the masked portion. This can detract from the clarity of the final image and influence the ability to correctly interpret image data.

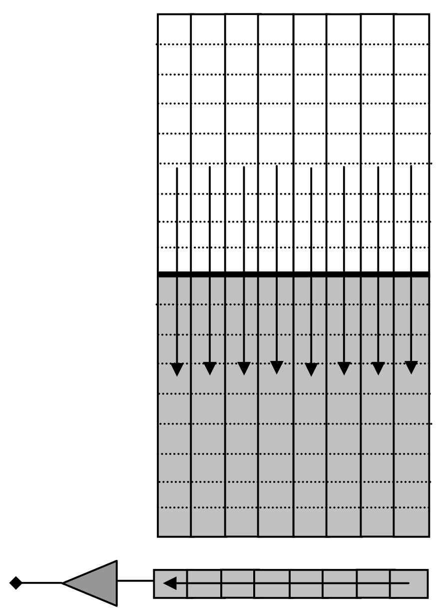


Figure 5 Frame Transfer Sensor Architecture

Interline Transfer

An interline transfer CCD, while sharing some common design features with the previous architectures, incorporates some significant differences. At the pixel site, part of the area is set aside with a mask to act as a local storage site at the end of an exposure. One of the performance trade-offs for this architecture is giving up some sensitivity because of smaller pixel area for the ability to perform electronic shuttering. Since the storage area is at the pixel site adjacent to the active area, the transfer of charge can be accomplished very fast, thereby stopping the exposure cycle. Because the transfer happens at the pixel site with a short transfer distance, little chance for smearing exists. With the charge stored safely at the pixel, a cascading movement of charge from one masked site to another is initiated to move these charges off the sensor, similar to the approach described for the full-frame architecture.

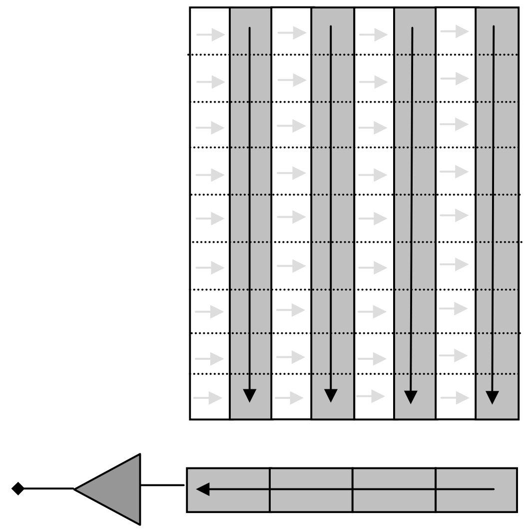


Figure 6 Interline Sensor Architecture

Additional area at the pixel site can be set aside to perform the blooming suppression, or overflow task, noted earlier. In an effort to make up for the loss of collection area, which impacts the sensitivity of the sensor, sensors from some foundries (i.e., Kodak) have processed a microlens on each pixel. A microlens can double or triple the amount of light directed onto the remaining active area. This compensates for much of the loss of pixel area. In terms of photographic measurement this can gain back almost one to one-and-a-half f-stops.

Charge Coupling with Pixels

A classical three-clock phase CCD gate structure will serve as the model for describing the charge transfer between pixels. A CCD sensor captures the optical image by tessellating the incident light into a two-dimensional array, X by Y picture elements (pixels). The light energy (photons) is converted into electrons. The electrons are collected as charge in a photosensing element. This element can be a photocapacitor, photodiode or photo-gate that is called a pixel. The control of the charge coupling within the sensor is illustrated in following figures.

A pixel consists of a photo-sense element formed under each gate, Figure 7(a). In the example, the gate with positive voltage forms a photosite for collecting charge. The gates with 0 voltage act as barriers. During charge accumulation, G1 is held at a positive voltage, which creates a potential well beneath the gate. Photons pass through the gate that is formed out of polysilicon material to create hole-electron pairs. The electrons are swept out of the region. he holes are captured in a depletion zone near the surface of the gate. The charge continues to accumulate (called integration) as photons create hole-electron pairs, until the cell's charge is read out.

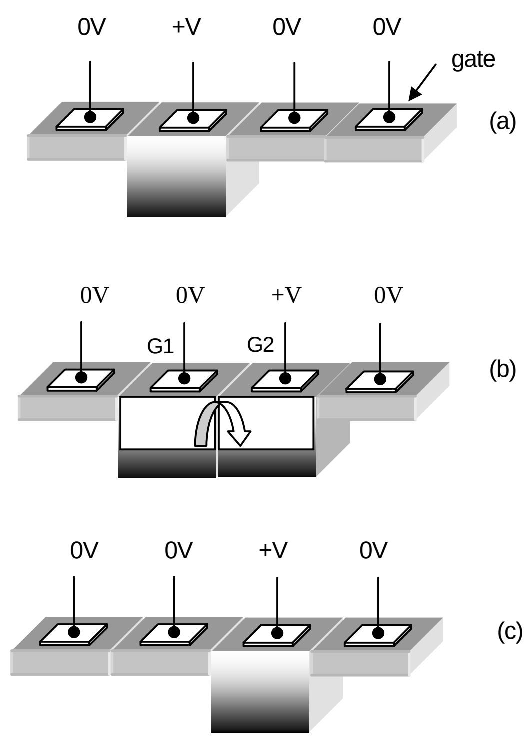


Figure 7 Charge Coupling with Pixels

The charge is read out by moving the charge from pixel to pixel, Figure 7(b). This is called charge coupling, hence the name charge coupled device (CCD). Charge is moved by clocking the adjacent gate (G2) to the same positive voltage as its neighbor. The charge flows into the newly formed well. To empty all the charge out of the previous well (G1), its gate voltage needs to be brought to 0 volts, Figure 7(c). This forces all the charge into the new well.

There is a last gate in this type of structure. It provides a node to dump the charge at the end of a column. It is biased to provide a DC potential barrier to isolate the photo-sense element from the output node. From this output node, the charge is moved to an amplifier that converts the charge into a voltage level. This is the clocking process in a 3-phase CCD structure.

For our next discussion about pixels and their mechanism for moving charge, we will address a true two-clock phase CCD example. The *KODAK DIGITAL SCIENCE* KAF-1600 Image Sensor (full-frame CCD) is manufactured by Eastman Kodak Company, Microelectronics Technology Division.[4] The device is built with an advanced true two-phase, two-polysilicon, NMOS CCD technology. This type of technology aids in the reliable fabrication of small pixel sizes. It also contributes to a higher short-free yield and aids in lowering the dark current without compromising charge capacity. The extremely low dark current makes the device ideal for low light imaging applications. The on-chip output amplifiers have been specially designed to perform at a high speed operation (45 MHz BW) at low noise levels (15 e-rms) to increase frame rate.

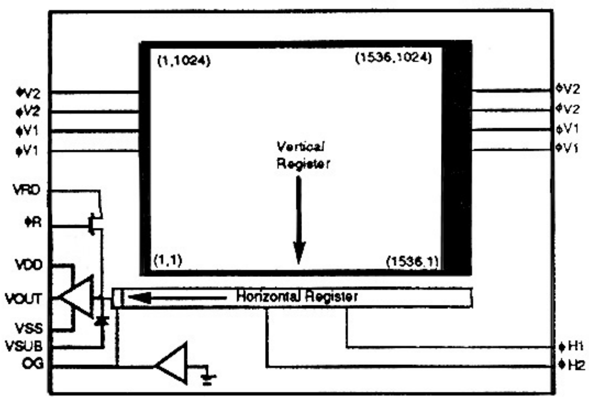


Figure 8 KAF-1600 Image Sensor Block Diagram

Referring to the full-frame block diagram the sensor consists of one vertical (parallel) register and one horizontal (serial) CCD shift registers and an output amplifier. All registers incorporate two-level polysilicon and true two-phase buried channel technology. The vertical register consists of 9.0 μm x 9.0 μm photocapacitor sensing elements (pixels), which also serves as the transport mechanism. T he pixels are arranged in a 1536(H) x 1024(V) array in which additional 16 columns and 8 rows of light shielded pixels are added as dark reference.

An image is acquired when incident light, in the form of photons, falls on the array of pixels in the vertical CCD register and creates electron-hole pairs (or simply electrons) within the silicon substrate. This charge is collected locally by the formation of potential wells created at each pixel site by induced voltages on the vertical register clock lines (¢V1, ¢V2). These same clock lines are used to implement charge coupling for readout. The amount of charge collected at each pixel is linearly dependent on light level and exposure time and nonlinearly dependent on wavelength until the potential well capacity is exceeded. At this point, charge will 'bloom' into vertically adjacent pixels unless drained.

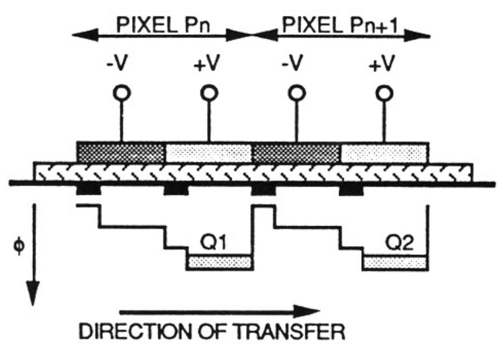


Figure 9 True 2-Phase CCD Cross Section

Integrated charge is transported to the output in a two-step process. Columns of charge are first shifted line-by-line into the horizontal CCD. Lines of charge are then shifted to the output pixel-by-pixel. Transfer to horizontal CCD begins when ¢V1 is brought high causing charge from the ¢V1 and ¢V2 gates to combine under the ¢V1 gate. ¢V1 and ¢V2 now reverse their polarity causing the charge packets to 'spill' forward under the ¢V2 gate of the next pixel.

The rising edge of ¢V2 also transfers the first line of charge into the horizontal CCD. A second-phase transition places the charge packets under the ¢V1 electrode of the next pixel. The sequence completes when ¢V1 is brought low while the horizontal CCD reads out the first line of charge using complementary clocking of ¢H1 and ¢H2 as shown. Vertical register clocking in this way is known as accumulation mode. The falling edge of tH2 forces a charge packet over the output gate (OG) onto the output node (floating diffusion) and sensed off-chip. The cycle repeats until all lines are read.

Charge packets received from the horizontal register are dumped onto the floating diffusion output node, whose potential varies linearly with the quantity of charge in each packet. A two-stage source-follower amplifier is used to buffer this voltage change to the outside world. The translation from electrons to voltages is called the output sensitivity or charge-to-voltage conversion. After the charge has been sensed off-chip, the reset clock (¢R) removes the charge from the floating diffusion via the reset drain (VRD). This, in turn, returns the floating diffusion potential to the reference level determined by the reset drain voltage.

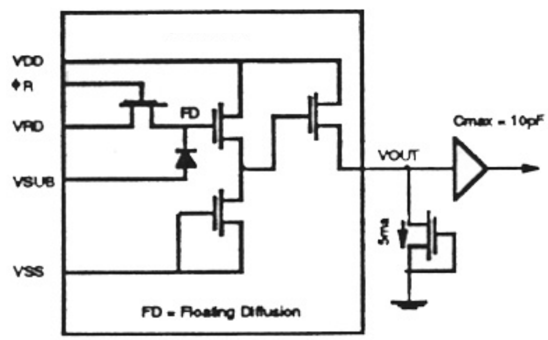


Figure 10 CCD Output Structure

User Interface

There are examples of well designed user interfaces all around us, light switch, radios, automobile accessory controls. Most interfaces are easily distinguishable between simple, and complicated. What factors separate the interfaces from those that are complicated from the simplistic are physical complexity, evolution of a design, and the analog/digital nature of the controls.

One could think of a VCR controller and imagine a device with 50 or more keys. This would be overwhelming to most users. Some elegant designs have only the most essential controls on the keypad, with the more detailed operations dedicated through on-screen interfaces. Many user interfaces for digital cameras have gone this direction.

The evolution of a design depends largely on the longevity of the technology. The telephone is a great example of an evolving technology that has become simpler, yet more sophisticated. While digital is new, analog electronic cameras have been around for years. It is safe to say that many user interfaces are still evolving for digital cameras. What is normally being expected as part of the interface is still to be established by customer requirements in imaging standards.

Digital has brought new capability into our world that previously was unavailable with analog. However, a human has a simple mental understanding of turning a knob vs. typing in a number. More user interfaces that are implemented in software for cameras are starting to use the representation of analog controls in the interface. This takes advantage of association of the analog operation, but in digital terms.

Now that some design philosophy has been given concerning a user interface, we should explore what control features are important to operating a digital camera.

The user interface can be defined in terms of camera controls. These controls have varying degrees of importance depending on the application. However, there are some fundamental controls that should be found with all digital cameras. These controls are:

- setting the exposure time to control the light
- changing the gain to improve sensitivity

- setting the black level to make an offset adjustment
- configure the mode of operation to define how an image is to be captured

Other controls that are useful but not essential are:

- enabling defect correction to remove bad pixels
- changing strobe output polarity
- detecting camera status
- defining external trigger polarities

These controls can be in the form of physical interfaces with knobs or switches. Most advanced cameras today are using software interfaces for control. This approach allows, in many instances, a path for updating the controls, more flexibility in the range of the controls, and a larger variety of commands. In addition, the Frame Grabber will often provide additional control of the camera through its interface including timing, pixel depth, and frame size.

In considering all aspects of the user interface, one should pay attention to the ease of use. Even if the application is one that does not require constant adjustment of the camera, an end user needs to feel comfortable in setting up the camera. Therefore, adherence to good ergonomic design with intuitive controls should be found in any camera's User Interface.

Factors to Consider

Image Quality

Image quality can be quantified in many measurable dimensions. However, the level of quality that is acceptable is often subjective to the individual and a particular application. Therefore, it is important to understand the fundamental terms used by an imaging specialist to describe image quality. From this information, you should be able to define, in your own terms, what image quality level is acceptable.

Resolution

There are many ways to express the resolution of a camera system. It is important to determine what factors must be taken into consideration to determine the resolution of a camera system. Shown below are the basic building blocks of a camera system that contributes to the system's overall resolution. The overall system and individual blocks can be expressed in terms of the capability to resolve a given object in the field-of-view as transfer function. A transfer function is simply the means of relating the output response to that of the input excitation.

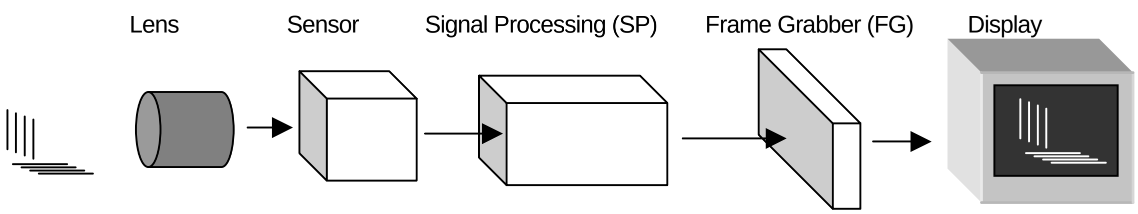


Figure 11 Elements of Resolution within the Imaging Chain

There are several common methods of expressing this transfer function. Two of the most widely used expressions are MTF (Modulation Transfer Function), and CTF (Contrast Transfer Function). MTF is used more often in describing a lens behavior since it is a continuous expression relating the modulation of a sinusoidal variation of light. CTF is used more extensively in describing digital cameras behavior since it is an expression of a transfer function response to a step in contrast difference. An explanation of the two needs to be given.

Modulation Transfer Function

The Modulation Transfer Function (MTF) was derived from an information theory as a mathematical principle describing the ratio of signal-to-noise. MTF is usually expressed as a graphical curve (Figure 13) showing the ability of an imaging sensor to spatially reproduce the scene consisting of a black and white bar pattern incident on the surface of sensor. The black and white bar pattern is used to measure the MTF, since it has a sinusoidal contrast function varying at different spatial frequency as shown below.[5]

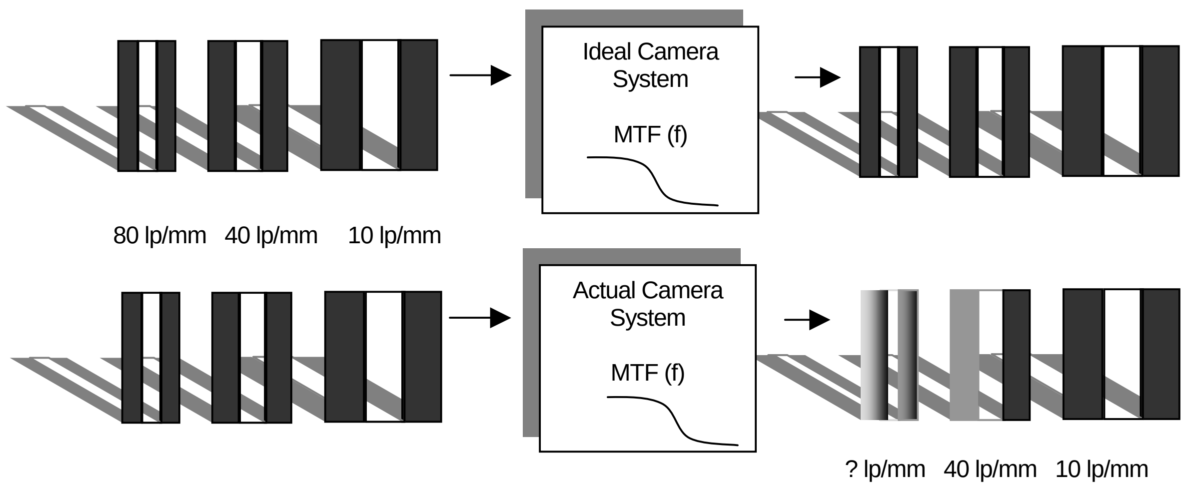


Figure 12 Modulation Transfer Function

The ideal camera system can resolve every frequency pattern of the black and white transitions. If the imaging system was sampling at a continuous rate, its transfer function would have a MTF (f) = 1. However, an actual

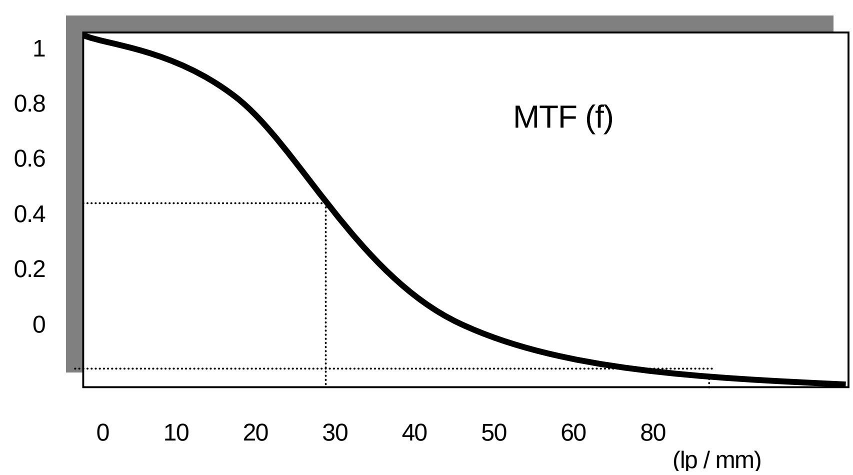


Figure 13 Normalized MTF Response

camera system is a sampling system and it may not be capable of resolving certain spatial frequencies as shown above. In this case, the edge of the black-to-white-to-black transition of the 80 lp/mm target can't be accurately determined. Therefore, the edge of the transition is indeterminate.

The MTF is expressed as a spatial frequency for a sinusoidal varying light stimulus as shown below;[6]

$$MTF(f) = \frac{S(f)_{max} - S(f)_{min}}{S(f)_{max} + S(f)_{min}} \times 100 \qquad \{1\}$$

$S(f)_{max}$ = maximum height of the sinusoidal signal

$S(f)_{min}$ = minimum height of the sinusoidal signal

Each element in the system (Figure 11) can be expressed in terms of MTF as shown below.

$$\text{Sys MTF}(f) = \text{Lens MTF}(f) \times \text{Sensor MTF}(f) \times \text{SP MTF}(f) \times \text{FG MTF}(f) \times \text{Display MTF}(f) \qquad \{2\}$$

Attempting to image a scene with spatial frequencies near the pixel's theoretical limit will cause aliasing, moiré patterns, or beating frequencies within the image. Aliasing occurs when sampled data at frequencies above one half the sampling frequency (Nyquist) can't be distinguished as lower frequency signals. In addition, the highest spatial frequency imaged is restricted by the diffraction limit of the optics.

Contrast Transfer Function

The Contrast Transfer Function (CTF) is similar to MTF except the scene imaged is not sinusoidal, but that of a square-wave function. CTF is more widely used with CCDs since it is easier to implement.

$$CTF(f) = \frac{B(f)_{max} - B(f)_{min}}{B(f)_{max} + B(f)_{min}} \times 100 \qquad \{3\}$$

$B(f)_{max}$ = maximum output an adjacent pixel

$B(f)_{min}$ = minimum output an adjacent pixel

Each element in the system (Figure 11) can be expressed in terms of CTF as shown below.

$$\text{Sys CTF}(f) = \text{Lens CTF}(f) \times \text{Sen CTF}(f) \times \text{SP CTF}(f) \times \text{FG CTF}(f) \times \text{Display CTF}(f) \qquad \{4\}$$

An approximate relationship between MTF(f) and CTF(f) with the following equation :[7]

$$MTF(f) = \frac{\Pi}{4}\left[CTF(f) + \frac{CTF(3f)}{3} - \frac{CTF(5f)}{5} + \frac{CTF(7f)}{7} \ldots\ldots\right] \times 100 \qquad \{5\}$$

Array Size

The resolution of a camera is largely influenced by the architecture of the sensor and the light-sensing element called a pixel. Pixels can vary in size from 5 microns to as large as 70 microns. However, most CCDs will have a pixel range from 6 micron to 24 micron. The construction of a CCD can be expressed as (X) pixels in the horizontal direction by (Y) pixels in the vertical direction. The number of pixels in the horizontal and vertical directions determines the resolving capability in the field-of-view. In addition, the size of an individual pixel determines the physical size of the sensor array. The larger the pixel, the larger the array. There are economical limitations in the fabrication of large sensor arrays. Therefore, most CCDs will tend towards a smaller pixel that allows for more devices per wafer. In addition, the smaller the pixels the smaller the process design rule, which means the device can move charge faster.

Pixel Theoretical Limiting Resolution

In principle, the pixel has a theoretical limiting resolution of 1/ [2 x pixel size] in units of line pairs per millimeter. As an example, a 9 microns square pixel will have a theoretical limiting resolution of 55 lp/mm. However, the use of the theoretical resolution would assume an ideal imaging system that has a unity transfer function. This assumption means literally, what you put in is exactly what you get out. In practice, an ideal imaging system is not possible. Therefore, to determine the actual resolution of a camera one must examine the elements that contribute to the reduction of the theoretical resolution.

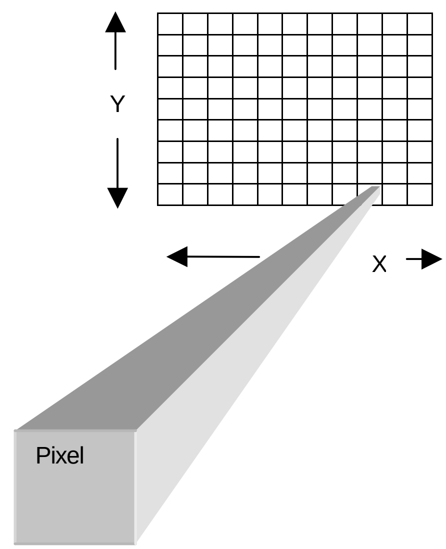


Figure 14 2-D Sensor Array

Sensitivity

Sensitivity[8] for a CCD camera may be defined as the minimum light level that an image sensor gives useful output. For analog video cameras, sensitivity is usually defined as the required illumination in lux incident to the camera that creates a mean video output level. It is commonly accepted that 30 video IRE units in amplitude, or 0.210 video volts for PAL, is the mean level with black being represented as 7 IRE units. An IRE unit is a scale for measuring full video output that is normalized to 100 units. Therefore, if 1 volt was your full scale video out, 1 IRE unit would equal 0.01 volt. Another way of expressing sensitivity for a digital camera is the normal response, often expressed as a digital number (dn) that corresponds to light illuminating the sensor (lux-sec). The sensitivity for a camera will be determined by the optical responsivity, the charge integration time and the signal-to-noise (SNR) ratio.

A digital camera's sensitivity can be compared to that of an analog camera by converting the dn/lux-sec to equivalent video IRE units/lux-sec. A method of comparing the two can be computed as equivalent ISO speed or comparing the ratio of equivalent faceplate illumination, for a given output. However, both should be compared under the same exposure time (i.e., 1/30 sec or 1/60 sec) and the lens needs to be operated at infinity. Close-up lens conjugates will not work for the following equation.

The following equation for converting a scene illuminance level to faceplate luminous exposure for a typical photographic scene (lens focused at infinity) is given as:

$$E_{g\ \text{for 23 IRE units}} = \frac{8(CxT)}{(K \times \text{f-stop}^2)} \qquad \textbf{\{6\}}$$

where:

Eg	= minimum luminous exposure in lux-sec for 23 IRE units
C	= scene illuminance in lux
T	= exposure time
K	= specific lens constant (i.e., 3.3 to 4.0), use 3.4 nominal
f-stop	= lens setting @ for Eg
H_{sat}	= Eg x 4 (quarter well illumination at 23 IRE units)
Eg/T	= faceplate illuminance for 23 IRE units
Sx	= equivalent ISO speed = $78/H_{sat}$

A dn-to-IRE unit conversion factor is needed for digital cameras as follows:

8-bit cameras = 256 counts, 1 IRE = 2.56 counts or 1 count = 0.391 IRE units
10-bit cameras = 1024 counts, 1 IRE = 10.24 counts or 1 count = 0.0976 IRE units

Therefore, let's compare a given analog camera sensitivity to that of a digital camera. The following information is taken off two manufactures actual data sheets:

Camera Type	**Sensitivity**	**Settings**
Analog (NTSC)	20.0 Lux @ 23 IRE units	AGC = on
		Shutter = none
		Gamma = 0.45
		Gain = normal
		F# = 1.4
		Exposure = 1/60
		Color Temp= 2856 Kelvin
Digital, 8-bit	720 dn/ Lux –second at faceplate	AGC = none
		Shutter = 1/60
		Gamma = 1.0
		Gain = 0 dB
		F# = Not Applicable,
		Exposure = Not Applicable
		Color Temp= 6400 Kelvin

$$E_{g\ (\text{analog camera})} = \frac{8(20 \times 1/60)}{(3.4 \times 1.4 \times 1.4)} = 0.390156 \text{ lux-seconds}$$

and the faceplate illuminance level = $E_{g\ (\text{analog camera})}$ /T = 0.390156/1.0166 = 23.5 ≈ 24

We must convert the 8 bit digital camera's counts/lux-sec faceplate sensitivity to equivalent video IRE units/Lux-seconds given at 23 IRE units

0.391 x 720 = 282 IRE units/lux-second

Eg (digital camera) = 23/282 = 0.0816 lux-seconds

Next, we must compute the digital camera's faceplate illumination level for 0.0816 luminous exposure, for the same exposure time (i.e., 1/60 second).

1/60 = 0.0166, therefore 0.0816/0.0166 = 4.897 lux faceplate illumination for equivalent 23 IRE units

Comparing equivalent sensitivity levels:

$$\frac{\text{Analog Camera}}{\text{Digital Camera}} = \frac{24}{4.897} = 4.9$$

The conclusion is that the digital camera is 4.9 times more sensitive than the analog camera for equivalent faceplate illumination levels. The sensitivity is a very important factor for obtaining clear images. An inexperience user may confuse motion blur with poor depth-of-field. If the sensitivity of the camera is not high enough for imaging an object for a given scene, the lens aperture must be opened up. This reduces the depth-of-field for the object to remain in focus. As the object moves, it could take a path outside the area that is in focus. This would then give the appearance of an object with motion blur. In reality, the object is out of focus.

Exposure

Many factors influence the exposure[9] or the amount of light required to produce the best image possible. Without sufficient light, the image may be:

- underexposed, detail is lost in dark regions
- unbalanced, poor color reproduction
- blurred, due to the lack of depth-of-field

The time that light is exposed to the imaging sensor depends on several factors. These factors include, lens f-stop, frame rate, shutter time, light levels, reflectance of surrounding material, imaging sensor's well capacity, and the sensor's signal-to-noise (SNR) ratio. All these factors can significantly impact the image quality. An often overlooked factor is the exposure time.

Exposure Time

Exposure time, shutter rate, and shutter times are interchangeable terms. The exposure time for a mechanical shutter is set in terms of number of degrees that it is open or in units of time. The exposure time for electronic sensors is either the inverse of the frame rate (if no electronic shutter exists), or the time that an electronic shuttered sensor is exposed in microseconds. Below are the relationships for defining the exposure time.

- mechanical shutter = (shutter's revolutions per second x shutter angle)/ 360
- no shutter = 1/frame rate
- electronic shutter = period of time that the sensor is exposed

The exposure time determines how sharp or blur-free an image is regardless of the frame rate. The exposure time needed to avoid blur depends on the subject's velocity and direction, the amount of lens magnification, the shutter speed or frame rate (which ever is faster), and the resolution of the imaging system.

A subject will be blurred in an image if the velocity is too high during the integration of light on the sensor. If a sharp edge of an object is imaged, and the object moves within one frame more than 2 pixels or a line pair, the object may be considered blurred. This is because multiple pixels are imaging an averaged value of the edge. This creates a smear or blur effect on the edge. To get good picture quality, the shutter rate should be 10x that of the subject's velocity.

The lens magnification can influence the relative velocity of the subject being imaged. The velocity of an object moving across a magnified field-of-view (FOV) increases linearly according to the magnification level. Likewise, if an object is viewed far away, the relative velocity in the FOV is less than that viewed next to the object.

Electronic cameras use electronic, or mechanical shutters that operate as fast as 10 microseconds (1/100,000 of a second). This shutter speed is fast enough to provide blur-free images of many high-speed events. The shutter controls the amount of light that is exposed to the sensor by the cycle rate of the shutter and the time that the shutter is open. The cycle time is set by the frame rate. The shutter then determines the exposure time. If no shutter capability exists for the imaging sensor, then the frame rate will be the effective exposure time.

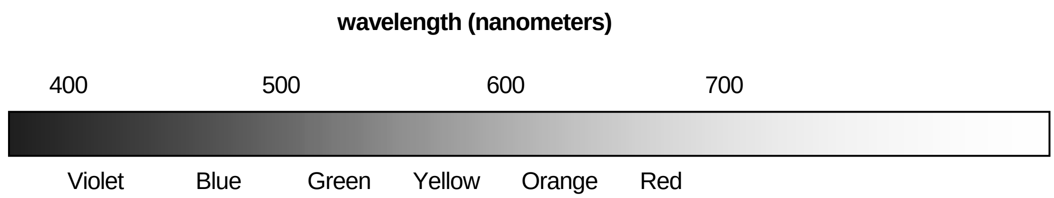
Color

Color[9] images are dependent upon the full complement of visible light frequencies in the 400-750 nanometer range being present. A color image is split into its three primary components of blue, green, and red for representation and storage. Depending upon the sensor architecture and color representation scheme chosen, some digital cameras may require as many as three sensors, one for each color, to capture a scene in color. This has a direct impact on the cost of such cameras.

Secondly, once an image is represented in its three primary color components, the image's file size likewise is increased by a factor of three. This requires additional capacity for archiving images. The use of compression algorithms and other color capturing schemes will become important in the practical and economical use of color in future systems.

Understanding color is difficult, but necessary, even for monochrome imaging. The color of light is determined by its wavelength. The longer wavelengths are hotter in color (red). The shorter wavelengths are cooler (blue). Shown below is the visible spectrum of light.

Color perception is a function of the human eye. The surface of an object either reflects or absorbs different light wavelengths. The light that the human eye perceives is unique in that it produces a physiological effect in our brain. What is red to one person may be perceived slightly differently by another person. Terms that further describe the color of an object are hue, saturation, and brightness. Hue is the base color such as red, blue violet, yellow, and others. Saturation is the shading that varies from a basic color to that of a different shade. An example of a hue would be green, and a saturated color would be lime (light green). Brightness also known as luminance is the intensity of the light. The subject of color would take an entire book to fully explain the science. However, studying a color chart can give the user some insight into the composition a color scene.



Color	Wavelength (nm)
Ultraviolet	390 - below
Violet	390 - 430
Blue	460 - 480
Green	490 - 530
Yellow	550 - 580
Orange	590 - 640
Red	650 - 850
Infrared	650 - above

Source	Color Temperature
candle	2000 K
sunlight, daybreak	+ 2000 K
tungsten	2900 K
sunlight, noon	4400 K
sunlight, afternoon	5400 K
HMI	5800 K
sunlight, evening	6500 K
sunlight, cloudy	6800 K
sunlight, cloudless	+10,000 K

Figure 15 Visible Light Spectrum

Color temperature is a common way of describing a light source. Color temperature originally derived its meaning from the heating of a theoretical black body to a temperature that caused the body to give off varying colors that ranged from red hot to white hot. This term was developed by Lord Kelvin and his name is associated with the unit measure. Listed above is a table of values for common lighting sources.

Color versus Monochrome

To understand the strengths and weaknesses of both color and monochrome[9] for video applications, some background must be discussed. There are various methods of producing color in electronic imaging. Shown below are three of the most common color methods in electronic imaging. The color wheel is used in still imaging where

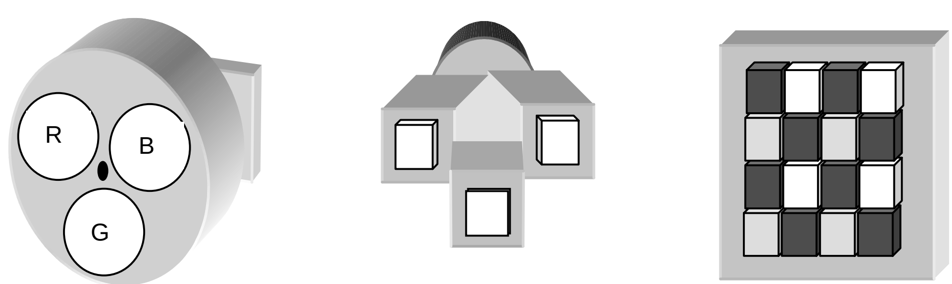


Figure 17 Methods for Generating Color with CCDs

the subject does not move during the imaging with the three color filters. This technique is suitable for copy stand photography or subjects that are not moving. However, it is not suitable for any subject that is moving. This unsuitability is due to the motion differences between each successive image. Using three imaging sensors with different color filters and a beam splitter, true color reproduction is possible. True color means that all primary colors and saturation levels are possible. This technique is costly since all the electronics are tripled to support the three imaging sensors. The alignment of the three sensors must be very precise. Otherwise, misregistration will occur on the colors. The last technique is a cost saving compromise. Color filter arrays (CFA) provide a more cost-effective means for producing colors (only one imaging device). There are individual color filters deposited on the surface of each pixel. There is some combination of red, blue, and green or complementary color scheme. Each pixel is isolated to a certain color spectrum. Although the pixels are filtered, the raw data must be interpolated for solving the missing pixels in each color plane.

Now that the main methods for producing color have been discussed, we need to review why color is better than monochrome. Generally, monochrome images are better in image quality. Monochrome cameras are more sensitive due to the lack of color filtering. The resolving capability is better than with CFAs. The reason is no interpolation is involved. The disadvantage of a monochrome image is the loss of color differentiation. For the human eye, small change in gray levels is more difficult to observe than a change in hue or saturation. Color is valuable for differentiating shades in certain imaging applications.

Quantum Efficiency

Quantum efficiency (QE) is defined as the ratio of the number of electronic charge collected divided by the number of incident photons.[10] The intrinsic properties of silicon and related materials used in fabrication of sensors have an affect on the absorption properties of light at certain frequencies. Therefore, a QE curve is extremely because it shows the efficiency at a given wavelength (Figure 18). Photons are absorbed by the polysilicon at the short wavelengths. The photons at the long wavelengths pass through the polysilicon and into the substrate unimpeded. The photons in the visible wavelengths are collected by the pixels and converted to electrons, with approximately 30% to 40% efficiency.

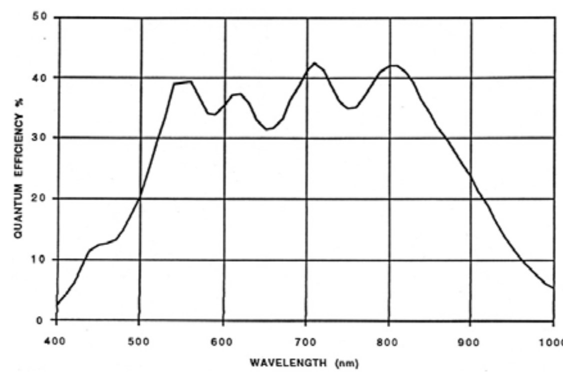


Figure 18 Typical CCD QE Curve

Spectral Response

Spectral response is defined as the ratio of the output current of the sensor to the incoming light power.[10] The visible region in which the eye perceives illumination ranges from 400 nanometers to approximately 750 nanometers wavelength. The interline architecture exhibits good responsivity within this region, but starts to decline in the red wavelengths. The full-frame and frame transfer architectures, on the other hand, can be a little weak in the blue response but retain good responsiveness up until around 850-900 nanometers. Some sensors overcome this low spectral response in the blue region by coating each pixel with a light emitting phosphor. This extends the blue response. The frame-transfers exhibit virtually the same response as the CCDs. Therefore, the differences in sensor architecture and material composition result in different response curves to spectral frequencies of light.

Output Format, Analog vs Digital

The factor that determines whether a camera is analog or digital is the nature of its output video signal.[3] Analog has traditionally been the basis of output since the first electronic cameras were introduced. By continuously varying the voltage levels at very high frequency, all the information, needed to describe an image, can be transmitted from the camera. If direct display is desired, this analog signal can be inputted into an analog monitor that directly converts the voltage into an image.

Most 30-fps analog cameras have interline sensor architectures. These low-resolution cameras with their inherent electronic shutters are packaged in compact housings. To analyze one of these images requires capturing the image with a frame grabber. There the signal must be digitized. Absolute pixel boundaries are less precise than a digital camera image. This is because of the signals timing. The storage options of analog images may take one of two alternative storage techniques. One is an analog recording medium such as a VCR tape. The other storage media is electronic memory. Analog images are converted to digital images by a frame grabber. These images then are stored on disk or DRAM memory.

Video Format

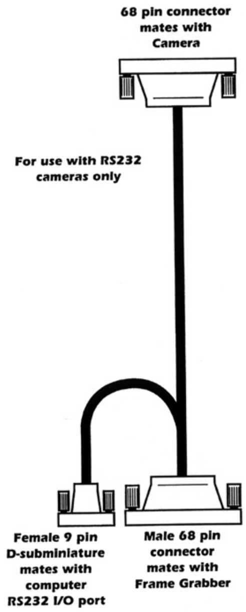
To complete the discussion of output format, a short review of interlaced vs progressive scan readout is needed. To support direct viewing of the captured image on an analog monitor, the video signal is read in an interlaced fashion. Under this convention, the sensor is exposed within 1/60 of a second intervals. On the first interval, all of the odd-numbered rows of pixel data are sent from the camera and refreshed upon a monitor. The following interval will select all the even-numbered rows. Each interval or partial image is called a field. Therefore, in 1/30 of a second the entire image has been refreshed upon a monitor with a slight difference in time between the segments as described. The eye responds in a way that integrates these rapid changes and smoothes the transitions of each update.

In a progressive scan convention, the entire sensor will be exposed within an exposure interval. All rows of pixel data are then read off the sensor sequentially to be refreshed upon a monitor. If done at a frequency of around thirty times per second or greater, the resulting images are presented in a smooth flow of changes upon the screen, without the eye detecting the updating. In this case, all of the image's data being displayed was taken at the same moment in time, and this lends itself better for image processing techniques.

Digital cameras in contrast take the analog signal from the sensor and convert it into a digital representation. The signal will still be a function of voltage level, but in this case little, or no voltage, may be used to indicate a "0" and a high voltage a "1." Sequential groupings of 0's and 1's are used to convey the image information. This technique is much less susceptible to extraneous electrical interference, or noise, which can lead to poor image quality when sending the signal over longer distances. Many storage/handling options are available such as computer hard drive, floppy disk, removable hard disk, CD-ROM, etc.

Interface Standards

Imaging societies have a guiding influence on imaging standards. The most influential organizations have committees composed of users, consultants, and manufacturers. These committees work to develop various standards for the entire imaging chain. Organizations that may be of interest for standards activity include the Automated Imaging Association [AIA]. Shown in Figure 19 are the 68 pin camera interfaces standard. (http://www.automated-imaging.org), the International Society for Optical Engineering [SPIE] (http://www.spie.org, or http://optics.org), the American National Standards Institute [ANSI] (http://www.ansi.org). The ANSI/AIA organizations have developed a standard for camera interfaces that are being used by over 24 frame grabber manufacturers (please reference ANSI/AIA A15.08/3-1992 & ANSI/AIA A15.08/2-1992). Conforming to these standards will assist the use of compatible interfaces in the imaging chain.



Interface cable pin out for cable that mates with 68 pin Frame Grabber connector

SIGNAL NAME	PIN	SIGNAL NAME	PIN
GROUND	1	GROUND	35
MSB (+)	2	MSB (-)	36
MSB-1 (+)	3	MSB-1 (-)	37
MSB-2 (+)	4	MSB-2 (-)	38
MSB-3 (+)	5	MSB-3 (-)	39
MSB-4 (+)	6	MSB-4 (-)	40
MSB-5 (+)	7	MSB-5 (-)	41
MSB-6 (+)	8	MSB-6 (-)	42
MSB-7 (+)	9	MSB-7 (-)	43
Not used	10	Not used	44
Not used	11	Not used	45
GROUND	12	GROUND	46
Not used	13	Not used	47
Not used	14	Not used	48
Not used	15	Not used	49
Not used	16	Not used	50
Not used	17	Not used	51
Not used	18	Not used	52
Not used	19	Not used	53
Not used	20	Not used	54
Not used	21	Not used	55
SER CNTRL OUT(+)	N/C	SER CNTRL OUT(-)	56
SER CNTRL IN(+)	N/C	SER CNTRL IN(-)	57
Not used	24	Not used	58
FRME ENA(+)	25	FRME ENA(-)	59
LINE ENA(+)	26	LINE ENA(-)	60
Not used	27	Not used	61
Not used	28	Not used	62
PIX DATA STRB(+)	29	PIX DATA STRB(-)	63
EXPOSE(+)	30	EXPOSE(-)	64
MC0	31	Not used	65
MC1	32	Not used	66
MC2	33	Not used	67
GROUND	34	GROUND	68

Figure 19 AIA Interface Standard

Applications

It is beyond the scope of our discussion to cover, in detail, any one application. However, there are some key points that need to be mention about a few major applications.

Machine Vision

Electronic shuttering:	eliminates the need for strobes, fill lighting is sufficient lowers the maintenance since there are no mechanical shutter blades to wear reduces motion blur for fast moving objects
Anti-blooming:	reduce hot spots in the image
High Frame Rates:	faster machine processes can be imaged with the higher rates
Resolution:	small pixels in high-resolution 2-D arrays provide more measurement detail greater FOV allows faster subsampling an image for decision making

Medical

High Bit Depth for Pixels:	required for x-ray imaging to discriminate subtle detail needed to replace traditional film applications

Aerial

Electronic shuttering:	reduces ground motion blur in airborne applications
Anti-blooming:	reduce hot spots in the image and uncontrolled light source intrusions
Resolution:	high resolution 2-D arrays provide more detail to capture ground detail large arrays reduce the number of image tiles required for digitizing huge photographic images

Scanning

High Frame Rates:	faster document capture is highly desirable
Aspect Ratio:	4:3 or 16:9 can provide a better match for scanning film
High Bit Depth for Pixels:	needed to replace traditional film applications

Example of an Imaging Application

An example of a new application for electronic imaging can be described in Automotive Traffic Control.[3] For years, cameras have been triggered to automatically photograph vehicles committing various traffic violations. The practice has its origins in Europe in the 1960s. It has been implemented in the Far East and most recently in the United States. The captured photo incorporates, or is merged, with other relevant data to document the incident. Lawbreakers are identified by reading the vehicle's license plate from the photo and then performing a match with that number in the registry files to find the vehicle's owner. Citations, which usually include a copy or representation of the captured photo, are then prepared and sent out by the appropriate authority for prosecution within the normal legal system.

While able to perform the intended task, this approach has its operational disadvantages. A service person must routinely visit each camera site and retrieve the film cartridge for processing. Sometimes the film cartridge had been fully exposed, with the likelihood that some violators were missed once this occurred. Otherwise, the cartridge had unused film that had to be wasted while developing the exposed portion.

In today's environmentally conscious atmosphere, there exists the ongoing concern over the use of chemicals to develop the film. In addition, the cameras require routine maintenance with age due to the film advancement and shuttering mechanisms. Once the film has been processed, the issues of data management arose including handling, reading, and storage. The physical element of this process can create a sizable task to support efficient access and retrievability of the photo evidence if required for prosecution. However, given all this, the biggest attribute of film remains its high resolving power for deriving information for prosecution of the violator.

In recent years, electronic camera technology has entered the marketplace. These cameras have begun to address many of the drawbacks associated with film-based cameras. As their usage broadens, they will contribute significantly to moving increasing numbers of vehicles safely and efficiently through the roadway system. Initially these cameras were installed for surveillance of traffic flows along the busiest expressways and highways. From a control center traffic controllers could monitor activity, looking for congestion or emergency situations.

More recently, such camera technology is beginning to replace loop detectors to control traffic signals as has been demonstrated in Oakland County, Michigan (USA). Captured images are analyzed for the presence of vehicles within predetermined image zones. When coupled with a control computer and traffic pattern history, signal light times can be optimized for vehicle throughput. Secondary benefits such as improving fuel economy and minimizing pollution can result.

The newest version of electronic camera, the digital camera, is now coming into the marketplace. In addition, the introduction of megaresolution digital cameras has begun. Initial enforcement applications include weight-in-motion, speed, and electronic toll collection.

For years, film-based cameras have photographed vehicles committing various violations. With the arrival of digital image capture technology, the movement to displace these cameras has begun bringing improved operating benefits. Digital traffic cameras are tailor-made for license plate recognition requirements, which is at the core of violation enforcement. Effective implementation will lead to moving an ever increasing volume of vehicles safely and efficiently through the ITS roadway infrastructure, while minimizing manpower enforcement resources. Practitioners-traffic management experts, governmental transportation authorities, police technical advisory staffs, and prosecuting legal representatives must be knowledgeable of the new technology's strong points and limitations in order that the full potential of this technology can be realized.

Conclusion

We have covered the fundamental digital technology used in electronic cameras. This was an introduction to the technology. To fully understand electronic imaging, more reading is recommended. There are several excellent sources for information through SPIE or various technical book publishers. We wish to thank Gary Erickson, Al Febraro, Mike Menadier, William R. Balch and Wendy Telford for their help on the paper.

(1) “Tutorial: Frame Grabbers and Software for Scientific Applications”, Feb. 6, 1997, I-Cube, Inc.
(2) “Don’t Fall for Monitor Sellers ‘Dot Pitch”, Cynthia Morgan, Windows Magazine, August 1997.
(3) “The Digital Camera Role in Future Violation Enforcement”, Gary Erickson, Eastman Kodak Company, MASD.
(4) KAF-1600 1536 (H) x 1024 (V) Pixel Full-Frame CCD Image Sensor Performance Specification, Rev 0, 1991. Eastman Kodak Company, Microelectronic Technology Division.
(5) SMPTE Pixels, Pictures, and Perception: The Differences and Similarities Between Computer Imagery, Film and Video, “Electronic Imaging Seminar – Film Recording Part I”, Thor Olson, April 8, 1995.
(6) “Development of the modulation transfer function and contrast transfer function for discrete systems, particularly charge-coupled devices”, John C. Feltz, SPIE Optical Engineering, Aug. 1990, Vol. 29 No. 8.
(7) “CCD Transfer Method – Standard for Absolute Performance of CCDs and Digital CCD Camera Systems”, Short Course 46, Electronic Design of High Performance Digital CCD Cameras, SPIE July 29, 1997James Janesick, Pixel Vision.
(8) Application Note: “Comparing Sensitivity to RS-170 Video Camera”, Bob Gazerro, Revision 1. June 1997, Eastman Kodak, MASD.
(9) High Speed Photography and Photonics, “High Speed Vidoegraphy”, Kris Balch, pg. 99, Focal Press, 1997.
(10) “Solid-State Imaging with Charge-Coupled Devices”, pg. 136, 139, Albert J. P. Theuwissen,.