PIXIM
DIGITAL PIXEL SYSTEM

1395 Charleston Road | Mountain View, CA 94043 | www.pixim.com
| tel (650) 934 0550
| fax (650) 934 0560

White Paper

Digital Pixel System Technology

Platform Overview

Rev. 2.0

December 2005

Contacts

USA Home Office
Pixim Inc.
1395 Charleston Road
Mountain View, CA 94043
1-650-934-0550

Pixim Europe Office
Hochfeldweg 9A
82256 Fürstenfeldbruck
Germany
+49 (8141) 348 055

Pixim Japan Office
29-9 Nakacho
2-Chome, Musashino-shi Tokyo 180-0006 Japan
International: +81-422-53-6042
Local: 0422-53-6042

Introduction

Digital Pixel System® (DPS) technology from Pixim, Inc. is a breakthrough imaging system that greatly enhances video and still image capture in a wide range of applications. DPS technology's extended dynamic range, fast readout speed, system-on-chip integration capabilities and low power operation represent a substantial advance over existing image capture and processing technologies.

Solid-state image sensor technology dates back to the invention of the first charge-coupled device (CCD) in the late 1960s. The 1990s saw the introduction of the CMOS active pixel sensor (APS), followed by the development of the DPS platform. The invention of DPS technology grew out of more than eight years of research at Stanford University.

The Digital Pixel System couples the inherent advantages of Pixim's innovative pixel architecture with new advances in embedded processor design to provide manufacturers with advanced imaging systems that can be adapted easily into a wide variety of next-generation products.

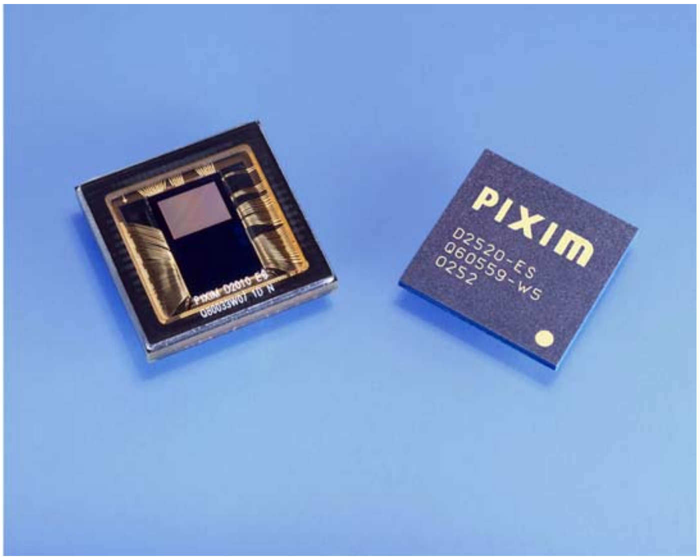

Figure 1. Pixim D2500 Video Imaging System

Image System Market Overview

Solid-state capture and display image systems are now an integral component in an enormous variety of electronic solutions. They can be found in high volume in markets as diverse as digital cameras, mobile phones, personal computers, factory automation, security, medicine and biotechnology. This reflects a convergence of technological advances in optics, semiconductors, electronics, and embedded processing.

Next-generation advanced imaging systems delivering the advantages of DPS imaging technology have entered the market. Markets that are currently benefiting from DPS technology include:

- Security and Surveillance — The wide dynamic range of DPS technology automatically adjusts to variable, real-world lighting conditions. The high-speed video capture rate of DPS enhances image detail. Integration of all sensor and processor functions in two chips allows compact cameras to be deployed in new security applications.
- Network / IP Cameras — In addition to security, IP-based cameras are being developed for broadband videoconferencing applications. Backlighting from a window can make it difficult to get high quality images with typical cameras. DPS technology provides an ideal solution.
- Machine Vision — DPS technology gives manufacturers enhanced observation and detection capabilities along automated assembly lines and in other production settings where reflections can cause problems. DPS technology holds great potential in manufacturing and industrial settings. The market for industrial imaging systems is pegged at more than $500 million annually.

DPS technology is also ideal for additional applications where operation under extreme lighting conditions is important, such as advanced automobile systems that help drivers avoid collisions or determine whether to deploy each of a vehicle's airbags, depending on the location of persons in the car.

Image System Types

Image sensors are typically categorized into three types: CCD, CMOS APS, and Pixim's Digital Pixel System technology. CCD sensors are an old technology, requiring complicated implementation systems and costly manufacturing processes that have limited their expansion in many markets. CMOS APS (Active Pixel Sensor) products were developed in the early 1990s as a result of emerging CMOS manufacturing technology and have been the principal alternative to CCD sensors at the low end of the market. DPS (Digital Pixel System) products are the result of a technology breakthrough pioneered in the mid-1990s at Stanford University.

DPS technology provides enhanced capabilities over existing technology including:

- Multi-sampling of an image within a single exposure for high frame rate and wide dynamic range
- Digital over-sampling for high resolution and image quality
- Integration of image processing functions to reduce power and area
- Temporal and spatial filtering for crisp, clear images
- Very low power consumption
- High signal-to-noise ratio

When compared to CCD and CMOS APS, the DPS platform provides significant improvements in image quality and increased flexibility for camera designers and manufacturers.

Overview of Digital Pixel System® Technology

The Digital Pixel System's unique architecture and tightly-coupled imaging software provide superior image quality even under widely varying light conditions and wide dynamic range scenes that have both dark and bright areas. DPS provides dramatically better wide dynamic range images than charge-coupled devices (CCDs) or CMOS active pixel sensors (APS) used in similar video camera applications.

Pixim's initial product line based on DPS technology consists of a highly integrated two-chip set consisting of a digital image sensor chip and a digital image processor chip. Layout of chip interconnections and user interface is straightforward. The user interface includes a customizable menu-driven on-screen display (OSD), switches and potentiometers, in any combination. Manufacturers define which of many settings are available for user control and which are fixed at the factory. Fixed settings can be locked so that they may not be identified or changed in the field, if desired.

Additionally, cameras based on DPS technology exhibit:

- Programmable exposure controls, configurable noise reduction, and greater dynamic range
- Digital video output and/or analog composite video outputs
- Reduced fixed pattern noise problems commonly associated with other sensors
- Digital pan / tilt / zoom, automatic white balance, and color correction, among other capabilities, using digital signal processing provided in the image processor chip

DPS Image Capture and Processing

DPS technology converts the quantity of light striking each picture element (pixel) to a digital value at the earliest possible point: at the pixel itself. An analog-to-digital converter (ADC) is designed into *each pixel*, and is operated simultaneously with all other ADCs in every pixel of the sensor. This pixel-level ADC architecture permits the use of many highly parallel low-speed circuits, operating close to where the photodiode signals are generated. This is key to optimizing the signal-to-noise ratio (SNR) for each pixel.

The DPS system uses the individual ADCs in each pixel to perform non-destructive correlated double sampling (CDS) at each pixel. DPS uses this capability to sample the growing light intensity at each pixel many times during each image capture period. This allows exposure level of each pixel to be determined by the *rate of change of charge collected* rather than only its absolute magnitude. Each pixel is also provided with an adjustable offset cancellation gain amplifier to assure uniform response throughout the sensor array. These innovations greatly reduce noticeable fixed pattern noise problems commonly associated with the column-level ADC used on APS sensors.

Because DPS sensors are digital, pixel readout is much faster and more accurate. Each sample of the digital image is captured in on-chip RAM. The high bandwidth provided by tightly coupled local memory is used to achieve its superior high dynamic range. This approach is not practical for CCD or APS sensors because of their reliance on analog readout circuitry. This is not a problem with DPS, which greatly benefits from the digital sampling performed on each pixel.

Digital Image Capture

Wide dynamic range is essential for capturing image detail at all light levels. The wide dynamic range achieved by DPS is realized by a patented non-destructive multi-sampling image capture capability, and advanced image processing algorithms.

Dynamic range is the ratio of the brightest image that can be captured by the imaging system to the darkest image that can be captured. Light intensity greater than the brightest possible image will cause the sensor to saturate, while light intensity less than the darkest possible image will not register on the sensor. Both of these conditions distort the image, hiding potentially vital information that lies outside the dynamic range of the sensor.

When an exposure begins, each pixel is charged at a rate that is proportional to the intensity of the light that strikes it. A stronger light source will charge a pixel more quickly than a weaker light source. Existing technology typically uses a single exposure time for all pixels. At the end of the exposure, the camera will sense the total charge accumulated in each pixel. But that means some pixels (the brighter ones) may be overexposed while others (the darker ones) may be underexposed. DPS overcomes this limitation as follows: with DPS, the light striking each pixel is sampled multiple times during the exposure period. DPS analyzes *how quickly* each pixel is being charged by the light striking it. This way, DPS measures light intensity by a combination of the *rate* at which the charge grows as well as the total charge accumulated during an entire exposure.

Specifically, the DPS system records the length of time required to *nearly saturate* each pixel. Pixels exposed to bright illumination will tend to saturate more quickly than other pixels. DPS determines for each pixel whether it will saturate before the next sample. If a pixel would saturate, then its elapsed exposure time is stored in memory, together with its current intensity of charge.

The advantage of this approach can be appreciated when one realizes that the entire range of each individual pixel, as well as the rate of change of the pixel charge, is used to form the resulting image, significantly increasing the dynamic range that is captured. Other technologies only measure the pixel value, not its rate of change.

DPS also provides improved color performance not available with other sensor technologies: the data recorded by each pixel is of very high quality, both in terms of accuracy and precision. High data quality allows the DPS image processing algorithms to render excellent fidelity for all colors and intensities.

DPS provides a fast global electronic shutter to capture bright lights and produce images that do not exhibit rolling shutter artifacts common in APS sensors.

Since multi-sampling is fundamental to DPS and is included in the basic firmware, no programming is required by developers to achieve this level of quality.

Optimal Signal to Noise Ratio

The *signal-to-noise ratio* (SNR) is a measure of a captured image's immunity to noise interference. Peak SNR is the ratio of the strongest recordable intensity without saturation to the background noise. A camera with higher SNR typically produces better (less noisy) video in darker scenes.

Photodiodes generally have better SNR when they are charged to more than half of their total capacity. But they cannot be charged beyond saturation. In other sensor technologies, high contrast portions of an image cause saturation of the photodiode and "blooming" in adjacent pixels.

The SNR in a DPS sensor is greater because each pixel is measured at its maximum value just prior to saturation. For example, for the photograph shown in Figure 2, the best exposure times are: T2 for the apple, T4 for the region below the cup, and T6 for the shadowed portion of the car because it is at these times that the photocurrent produced by these pixels is highest prior to saturation. Therefore, these are the times when the image is the most accurately represented by these pixels.

DPS reduces noise in the sensor in a number of ways. First, a negative feedback unity gain amplifier in each pixel eliminates any offset voltage, resulting in much greater uniformity throughout the sensor array. To

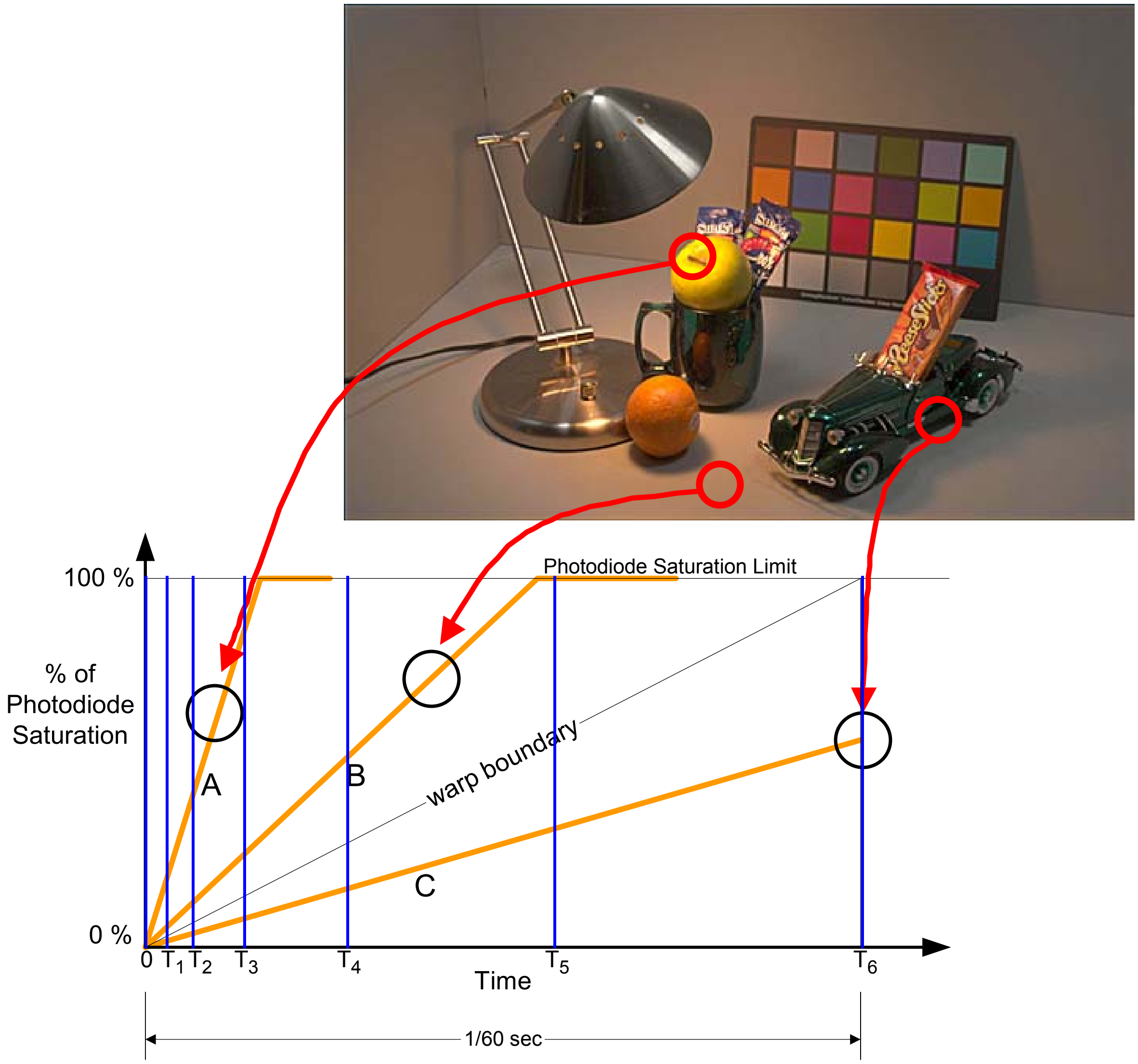

Figure 2. Optimal SNR Exposure
Different Light Levels are Captured at Different Times

minimize reset noise, each pixel value is read non-destructively at the beginning of the exposure, and this value is later subtracted from the final measured value for the pixel. This non-destructive method of correlated double sampling is unique — most other sensors must read the CDS value, reset the sensor, then capture a new value. But the random background noise change from the initial captured CDS value introduces distortion that DPS avoids. This reduces noise that would be detectable to the human eye.

Finally, the high quality of Pixim's pixels provides high signal levels and low noise levels, resulting in excellent inherent SNR characteristics. This can provide sensitivity on par with CCDs, and an order of magnitude sensitivity gain over APS sensors, while providing extended dynamic range and excellent image quality.

Instant Electronic Shuttering

Extremely fast electronic shuttering is possible from the DPS sensor array because readout is highly parallel. Image data is read from every pixel at the same time, so there are no rolling shutter effects and minimal motion artifacts.

Existing Image Capture Technology

CCD

CCD sensor technology has been dominant in image applications since the technology debuted in the late 1960's. However, its analog design is based on a "bucket brigade architecture" and requires a non-standard fabrication process that is both complicated and costly to optimize. (See Figure 3.)

It is difficult for commercially viable sensor arrays to integrate analog-to-digital conversion (ADC) for digital processing onto CCDs, or to customize them by adding logic blocks to the chip for specific applications. Their high voltage (+12 to +20 Volts) and power requirements shorten battery life and limit their applicability in small form factor, portable devices. Moreover, the need to add external supporting components (including analog front end, timing circuit, reference voltage, correlated double sampling, mechanical shutter and RAM) results in large printed circuit board real-estate requirements, contrary to the trend of miniaturization in weight and size.

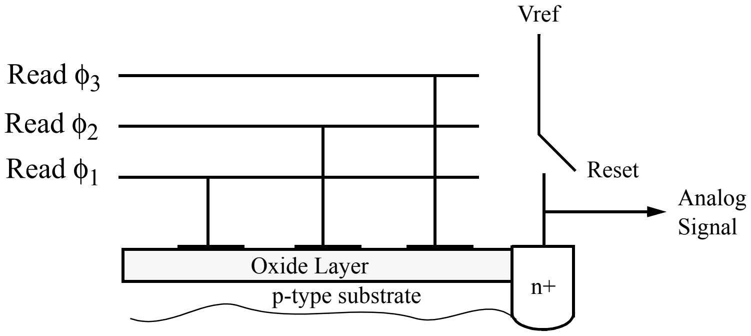

Figure 3. CCD Sensor

CMOS APS

CMOS sensors were developed in the early 1980's. Passive pixel sensor (PPS) image sensors were the first products in this family to come to market. The large feature sizes available in existing CMOS technology allowed only a single transistor and three interconnecting lines for each pixel. The speed and signal-to-noise-ratio of PPS was significantly lower than that of CCD sensors.

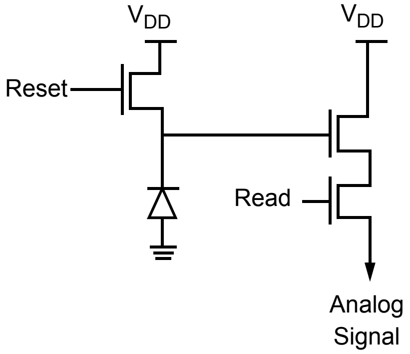

Figure 4. Active Pixel Sensor

In the 1990s, APS technology added an amplifier to each pixel. This increased sensor speed and improved the signal-to-noise-ratio, providing a big advantage over PPS sensors. (See Figure 4.)

When deep sub-micron CMOS technologies and micro-lenses appeared, APS became the alternative sensor technology. Its low power consumption and near-standard manufacturing process made it a competitor to CCD sensors for certain applications.

However, APS technology has inherent problems. Due to process variations that create nonuniformities in the column level ADCs and in-pixel amplifiers, large fixed pattern noise (FPN) at high resolutions typically yields limited sensitivity, less than is required for many applications, including security and the film industry. Human eyes are particularly sensitive to image edges, and the column-level ADCs amplify this noise.

Pixim DPS Image Sensing Technology

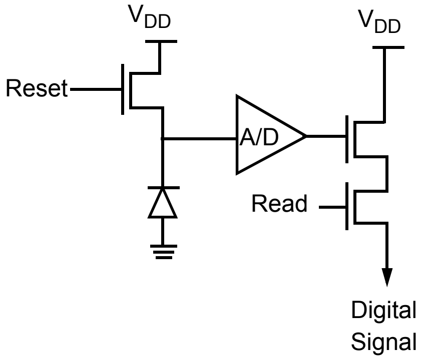

Figure 5. Pixim DPS Image Sensing

Because DPS technology integrates the ADC into the pixel, it can be manufactured using leading edge semiconductor processes. The pixel array has significantly higher noise immunity than APS sensors because DPS technology employs a digital readout from each pixel (see Figure 5). Additional image processing and camera functions implemented by Pixim provide a complete imaging solution in high-volume, commercially available chipsets.

DPS image systems integrate sensing, memory and processing functions into two chips. This is especially important for imaging systems that require significant processing, where quality of output is crucial, and where small size, low power and portability are important.

Until now, analog-to-digital conversion could only be integrated at the chip or column level (see Figure 6). Both approaches are common in APS sensor solutions. For the chip-level approach, a single conventional high-speed ADC is integrated with the sensor. For the column-level approach, one or more columns of the pixel array has a dedicated ADC. The ADCs are operated in parallel and, therefore, low-to-medium speed conversion techniques must be used (e.g., single-slope, algorithmic, successive approximation, or oversampling).

By having separate ADCs for each pixel operating in parallel, the ADCs can operate at very low speed, operating at a few thousand samples per second. This lessens noise and reduces power requirements.

The large bit stream used to read the ADCs is supported by on-chip RAM. These features enable much faster and more accurate readout characteristics.

The DPS sensor consists of an $m \times n$ digital pixel sensor array, ADCs and RAM. A separate chip incorporates digital signal processing (DSP), and I/O. The sensor core is powered by a single low-voltage power supply. Manufacturing is simpler than for CCDs and greater miniaturization is realized. These advantages will continue to increase in the future in accordance with Moore's Law.

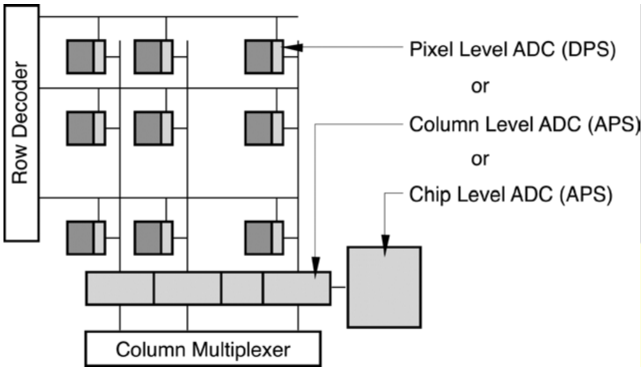

Figure 6. ADC Integration Options

Pixel-Level ADC

The DPS pixel-level ADC architecture permits the use of low-speed conversion; the ADCs operate close to where the photodiode signals are generated. This optimizes signal-to-noise ratio (SNR) and power consumption beyond any sensors currently available. The large number of independent, small ADCs significantly reduces noticeable fixed pattern noise problems commonly associated with the column-level ADC variation on APS sensors. Finally, because pixel signals are available to ADCs at all times, the number and timing of images—as well as the number of bits from each image—can be freely chosen. This offers important advantages, such as the ability to optimize the image capture and processing to the scene characteristics. All of these features have been implemented in Pixim DPS systems.

Pixel Readout

Each pixel's ADC feeds into a RAM-speed bit-serial readout. The multichannel bit-serial design reduces the ADC output data rate, allowing it to use standard Nyquist-rate conversion instead of oversampling. It facilitates DPS dynamic range enhancement via multiple sampling, and reduces nonuniformity by globally distributing the ADC control and clock signals, and by performing auto-zeroing.

The implementation consists of a two-dimensional array of pixel blocks, a row decoder and column sense amplifiers and latches. (See Figure 7.) Each pixel block comprises one or more photodetectors sharing an ADC channel. The captured analog pixel values are digitized in parallel, one bit at a time. Each latched set of bits, using the row decoder and column sense amplifiers, forms a bit-plane that is read out in a manner similar to a standard digital memory. As depicted in Figure 8, a set of bit planes constituting an m-bit frame of the data is collected during a single exposure window. In this example, the digitized value of the uppermost right-hand corner pixel is 1001…1. Note that this image output format is quite different from the raster-scan format commonly used in CCD and APS devices.

DPS Image Processing

In addition to superior image capture in the DPS sensor chip, Pixim DPS technology incorporates signal processing and camera support functions such as exposure metering. The DPS processor chip processes digital pixel information from the sensor. Software supplied by Pixim in the processor performs exposure

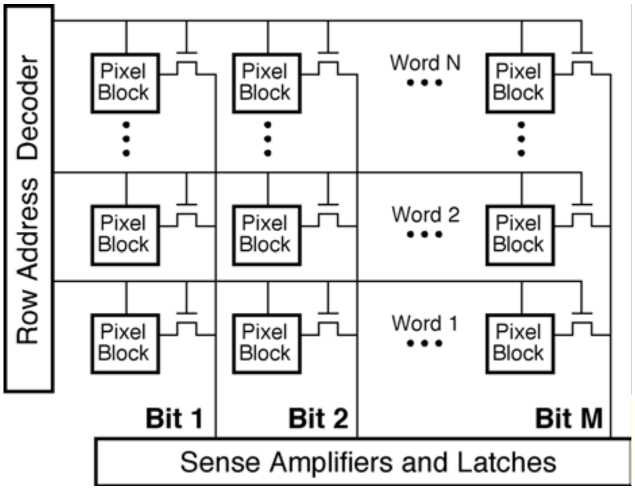

Figure 7. Block Diagram of a DPS Sensor with Pixel-Level ADC

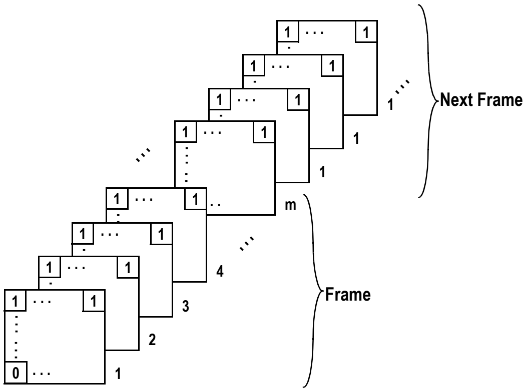

Figure 8. Output Image Format

control, color processing, and many other camera functions, including user interfaces. The result is a complete imaging system in two chips.

Exposure Control System

Exposure Control is divided into three blocks, as illustrated in Figure 9. The **Scene Analysis** block is similar to the photographer's light meter, returning information on the amount and quality of light in the scene. The Settings Table interprets the scene information to adjust it for optimal viewing. The Transition Control block keeps the video image smooth. These blocks can be tested with known scenes for accuracy.

Scene analysis results in three estimates:

- Light Level Estimate
- White Balance Estimate
- Scene Range Estimate

Calculation of these estimates is discussed in greater length below.

The **Settings Table** block is a collection of functions and values that calculate the best image system settings for each result from the scene analysis block. This is akin to the photographer reading the exposure time and *f*-number from the light meter or a table.

These settings can be verified independently by manually selecting the scene values and observing or measuring the quality of the video output. Because it is a video system, transitions between settings must be smooth and not oscillate. This is achieved by the **Transition Control** block, which acts as the steady hand of a camera operator, smoothly changing settings while providing a pleasurable picture. The overall operation of this system and specific detail on the scene analysis is reviewed in more detail in the following sections.

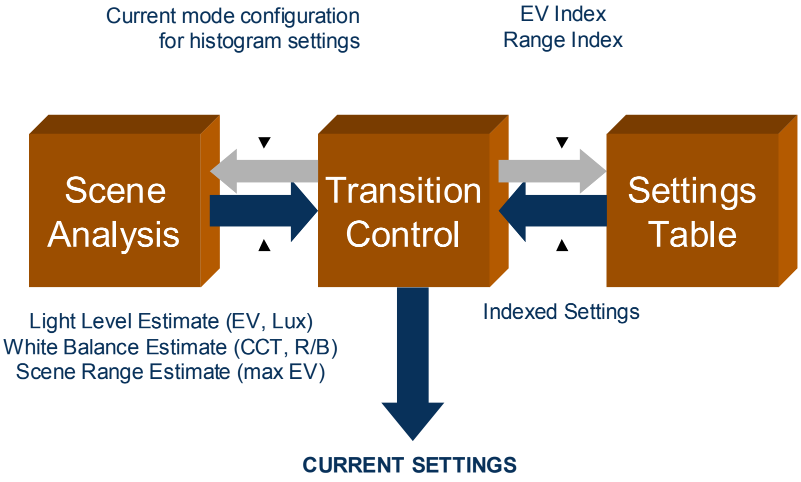

Figure 9. Exposure Control

Exposure Control Objectives

The primary objective of CCTV surveillance cameras is to achieve a form of telepresence. The imaging system, from the camera to the display, should allow the user to recognize, interpret and judge remote events and objects as if he or she were physically there. A successful imaging system accomplishes this by reducing the wide range of possible inputs to a representation that is nearly constant with respect to the reflectance of objects in the scene, and independent of the light sources illuminating the objects.

The reproduction of images by preservation of relative reflectances requires that the intensity of the scene illumination be estimated and factored out. The estimation can be made manually or within the camera by measuring the illumination of the scene. Once the light at the subject is estimated, the exposure system is adjusted by allowing more or less light, so as to obtain a record that is independent of the illumination.

Light Level Estimate

The accuracy of any system that estimates the level of illumination from the scene is heavily dependent on scene contents and the methods used to calculate the estimate. For instance, if the entire scene consists of a wall with even illumination, and no other objects are visible, there will be no way to differentiate a very dark wall with bright illumination from a very light wall with dark illumination. Both would be returning the same amount of light to the camera and would be reproduced the same way, equal in video levels and so forth. Although this reproduction will not be objectionable as long as this wall is the only object in the scene, the introduction of any new objects, particularly those which we know a priori (e.g. a person’s face), will immediately show one exposure setting to be too bright and the other too dark.

Although specific instances of this problem will always remain intractable, in the 1930’s Ralph Evans of Eastman Kodak introduced the concept of “gray world”, based on the observation that white objects are very unlikely in nature, since they get dirty, and truly dark objects are also unlikely as they also get dirty and have specular reflectances. In general, the observation is that any large collection of reflective objects will have the mean probability of reflectance at about 20% (or 18% more exactly). A distribution depicting this is shown in Figure 10.

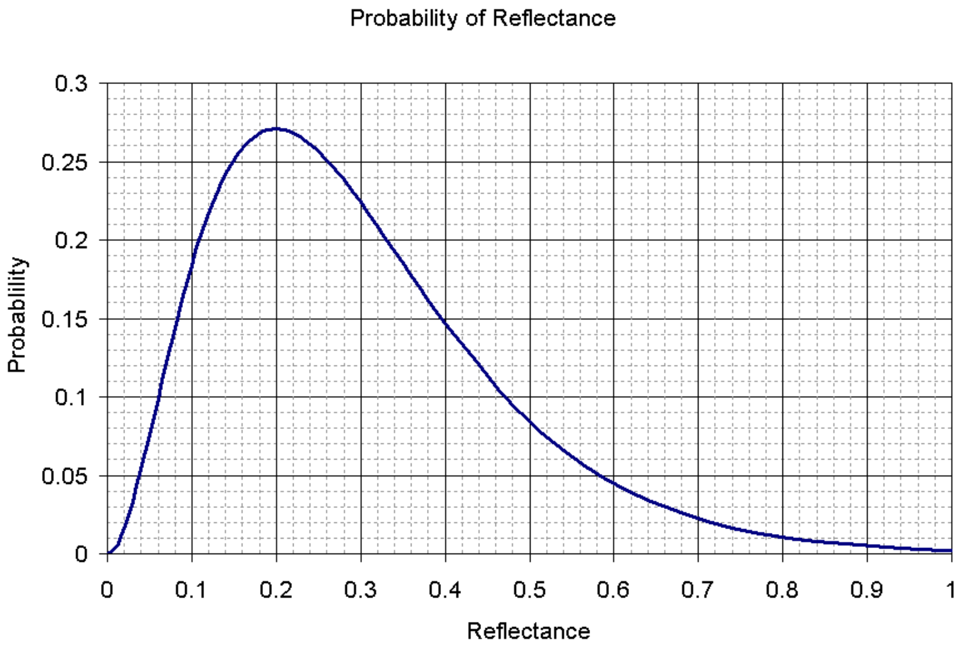

Figure 10. Reflectance Probability of Natural Objects

Scene Range Estimate

This "gray world" assumption, as it is commonly known, is the guiding principle behind most automatic exposure systems. A number of fortuitous coincidences makes the reproduction of this gray world relatively simple. First, it is well known from psychophysics that the visual perception of reflectance follows a power law (Weber's Law) that makes the perceived lightness of an object roughly the logarithm of the object reflectance. In practice, this means the 18% mean reflectance of our gray world is perceived as 50% object lightness, right in the middle of the perceptual scale. This has been argued to be the result of evolutionary adaptation.

The second fortuitous coincidence is that the amount of light emitted by cathode ray tubes (CRT) is related to the driving voltage (also a power law) roughly approximated by a 2.2 exponent (the gamma). This means that mid-level video voltage (50 IRE) is transformed into a low light emission (approximately 20% of the range), which in turn is perceived as 50% lightness, with the overall result being that *video levels are roughly proportional to perceived lightness*.

Because of these coincidences, the main task of the exposure control system is simplified to estimating the mean reflectance of a scene and adjusting the exposure (gain) so it corresponds to 20% of the signal range. The signal is then gamma-corrected (using 0.45, the inverse gamma of the CRT), encoded and transmitted. The net result of this sequence of transformations is that a video image produced by a CRT display preserves approximately the relative reflectances of objects in the scene and is independent of the level of illumination.

Video systems are expected to operate in a wide range of illumination levels, from < 10 Lux in a Photography dark room up to > 50,000 Lux in bright daylight. This wide range of illuminations, roughly 100,000:1, called the intra-scene range, requires that different mechanisms be employed to modulate the amount of light collected by the sensor. Most CCTV lenses have irises that can be adjusted from F1.2 to F22, resulting in a range of modulation of approximately 300:1. The remaining 300:1 range has to be accounted for by adjusting the exposure time, electronic gain, filters or other means.

White Balance Estimate

The light sources will also vary in color temperature. The CIE chromaticity coordinates of most light sources in common use fall close to the Plankian locus despite large differences in spectral distribution. This follows from the human eye's evolutionary adaptation to the different phases of daylight and the primordial use of incandescent materials for artificial lighting. Consequently, lights far from the Plankian locus do not look white and are not used for general illumination. The notable exceptions are outdoor night lights such as Sodium and Mercury vapor, which are economical and in wide use despite their poor color. Within this range of useful light sources, which is roughly from 2,000K to 10,000K CCT (Correlated Color Temperature), the maximum ratio of visible radiation at the short and long ends of the spectrum does not vary more than 5:1.

Another important variation in the incoming light field is the physical distribution of the illumination. This distribution is dependent on the spatial arrangement of light sources, the proximity of objects to the sources and the angle of illumination.

Dynamic Range Considerations

The issues described above are well understood and most video and imaging systems routinely adjust the exposure successfully for the wide range of intra-scene illuminations required. Yet, a new complication is introduced when multiple illuminations are used in the same scene with widely different light levels. This type of scene would be typical of interiors (200 lux) mixed with exteriors (50k lux and above) viewed together when looking at a window or a door of a building. Since the probability of reflectance of the objects is not different indoor and outdoors (the same gray world statistics apply) the range of light levels that will be represented in the scene will increase by the ratios of the levels of illuminations. In our 200-lux to 50,000-lux example, this ratio of illumination is 250:1. While the eye has the remarkable ability to instantaneously adjust to such disparate light levels, previous-technology electronic image sensors cannot capture such range in a single image.

To estimate the required range we must first look at the requirements of the interior scene. In an office, the reflectance of most objects will fall within a narrow 64:1 reflectance ratio. Although illumination within offices and other living spaces is mostly diffuse and even, objects might be closer or further from the light sources and at different angles of illumination. This will increase the range of reflected lights (luminances) likely in a scene, but in general a total ratio of 256:1 is all that is required to capture such scenes. This range, which can be covered by an 8-bit ADC, is typical of most CCD-based CCTV cameras. Since the same objects are equally likely to occur inside or outside, the required range of our high dynamic range scene will be the product of the ratio of illuminations and the ratio of luminances with a given light level, resulting in a total range of about 65,000:1, or about 16 bits of data range for the image capture.

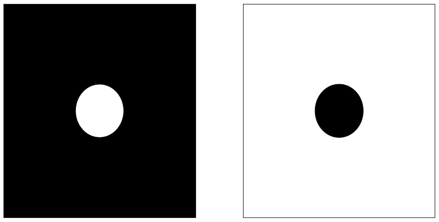

Figure 11. The Simple Flare Experiment

One factor limiting the ability to reproduce such scenes is that as light passes through optical systems, a small amount of it is dispersed and reflected internally, creating a flare, or veiling glare (or “light leakage”) that will decrease the ratio of illuminations impinging on the image sensor. The amount of this flare is dependent on the quality of the optics, the opto-mechanical design and the amount of optical energy in the scene. One simple experiment to demonstrate this limitation is illustrated bellow in Figure 11.

In this experiment, the luminance of the white areas in both squares is the same, and likewise the luminance of the dark areas is the same, the only significant difference being the relative size of the bright and dark areas. Using the same lenses and imaging each object individually, it will be observed that the ratios of luminance will not be preserved. The dark circle will become lighter than the corresponding dark background since the white background would have “leaked” into it due to the lens flare. This represents a practical limit to what can be achieved with a high-dynamic range imaging system. Internal reflections and the flare typical of CCTV lenses (around 5%) will reduce the 65,000:1 scene in our example to a mere 2,500:1 (12 bits) if most of the scene was the exterior.

Exposure of High Dynamic Range (HDR) Scenes

The practical limits described above make it possible to set principles and guidelines for the exposure adjustment of HDR scenes:

1. The main portion of the scene is that which occupies the majority of the image area (> 50%).
2. The exposure is set for the main portion of the scene.
3. The dynamic range of the scene is the ratio of exposures between the main portion and the secondary portion.

The implementation of this scheme is fairly simple and robust. The following two metrics are tracked in the DPS signal processing system:

- The estimated light level for the median (50% of pixels)
- The estimated light level for 98% of the pixels

The dynamic range is the ratio between these two estimates, and it suffices to determine the correct exposure, the range required for the capture, and the tone correction needed to compress the wide range into the rendering range. A typical reproduction curve is shown in Figure 12.

- Scenario A — The camera and scene are both indoors or both outdoors. In this case, the range will be small, and the main exposure will be correct for all parts of the scene. With reference to Figure 12, there will be a small roll-off portion, and most of the image will be in the linear range.

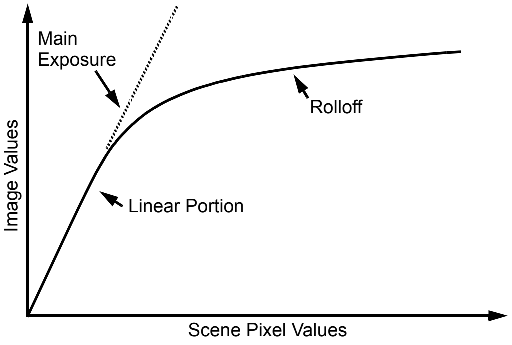

Figure 12. HDR Capture Scheme

- Scenario B — The camera is indoors, and less than 50% of the pixels receive light from a bright exterior scene. The main exposure will be set for the indoor scene, and the range and tone correction are adjusted to accommodate the wide dynamic range of the scene. In this scenario, there will be a small linear portion, proportional to the size of the range, and a large roll-off section to accommodate the extended range.
- Scenario C — The camera is indoors, and most of the pixels are illuminated by bright outdoor light. The exposure will be the same as in Scenario A above, and the dynamic range will be limited. This is justified since the outside flare will dominate the darker part of the scene.
- Scenario D — The camera is outdoors, looking indoors, and more than 50% of pixels capture the indoor scene. This is same as scenario B.
- Scenario E — The camera is outdoors, looking indoors with less than 50% of pixels capturing the indoor scene. This is the same as scenario C.

Pixim provides exposure and range biases for the user to adjust for special cases of scenarios C and E.

Gain Control

Because DPS can read pixels non-destructively multiple times during one field integration, the dynamic range is augmented by capturing different light levels at different times. In Figure 13, the three lines marked A, B, and C show how the photo current of three different photodiodes (pixels) grows with time. Brighter pixels have the highest rate of increase. (This is shown graphically in Figure 2.)

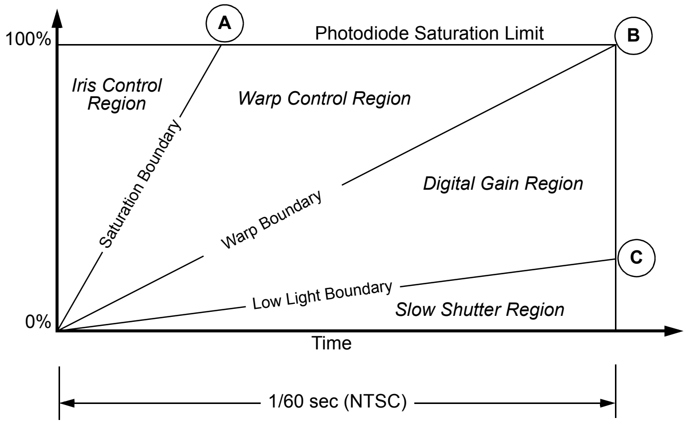

Figure 13. Gain Control Regions

As we see in this example, analog to digital conversion of the brightest pixels (line A) must start the earliest to avoid saturating the photodiode. Analog to digital conversion is performed from the brightest pixels to the least bright pixels with no sample and hold, so there is a race in the brightest parts of the scene to capture all the information before saturation. This sets an upper boundary to the brightest pixel that can be captured. Lower light levels—as represented by line C, that do not saturate the pixel during the entire exposure period—are converted at the end of the exposure period.

Gain Control Regions

Current DPS products have three distinct exposure capture regions in as illustrated in Figure 13. The region marked Digital Gain Control in Figure 13 is so dark that no pixels achieve the photodiode saturation limit during the exposure interval, and the image striking the sensor is underexposed. To compensate, an analog camera would use an analog gain stage to increase the image brightness. However, DPS signal processing implements digital gain control by multiplying the image within the image processor.

Gain levels in the Digital Gain Control region are always positive, because the intensity of the underexposed image must be boosted. Up to 60 dB of digital gain is supported.

Warp Control

As the scene becomes brighter, the photodiodes in the most intense parts of the image will come closer to saturation. The upper boundary of the Digital Gain Control region is called the warp boundary (see Figure 13). At this boundary, the most charged pixels equal but do not exceed the maximum ADC value at the end of the exposure period, so the entire image fits within the measurable range of the ADC. The precision of the ADC is 10 bits, so the brightest nonsaturating pixels produce an ADC value of 1023. Pixels that reach a value of 1023 at the end of the exposure period are said to be charging at warp one. This is also the gain setting corresponding to 0 dB, or about 133 lux of scene illumination with an F1.2 lens in NTSC mode.

One might expect that images more intense than warp one would begin to show image saturation, and in an ordinary camera, this would be true. To prevent the brightest pixels from saturating in an ordinary camera, the exposure period would have to be shortened, or the amount of light entering the camera would have to be lessened by stopping down the aperture. However, either approach will lessen the amount of light striking the darkest pixels. If the darkest regions of the image are less than the noise floor, or less than the least significant bit of the ADC, information in the darkest parts of the scene will be lost.

Pixim DPS technology overcomes this limitation of ordinary cameras. Pixels charging faster than warp one are sampled earlier in the exposure period, just prior to their saturation point, while pixels charging slower than warp one are sampled at the end of the exposure period. (See Figure 2). DPS records the photocurrent just prior to saturation and also records the elapsed time when the measurement is taken. The intensity recorded for each pixel is the product of its warp number and the value of the ADC.

For example, suppose the exposure period is 1/60th of a second (16.67 ms), and the elapsed time when the brightest pixel reaches saturation is 1/120th of a second (8.3 ms). The DPS algorithm records an ADC value (1023) for this pixel and also records the elapsed time (1/120th second). If somehow we could allow this pixel to continue charging past its saturation point, and our ADC had enough precision, this pixel would reach a value of 2046 at the end of the exposure period. We can compensate for the limitations of our equipment by extrapolating the ADC value, as follows.

The warp value is the ratio of the shortest exposure time to the nominal exposure time. For this example:

$$\frac{1}{60} \div \frac{1}{120} = \frac{0.01667}{0.00833} = 2\,.$$

Multiplying the warp value (2) times the ADC value (1023) gives us the estimated intensity value for this pixel at the end of the exposure period (2046). The validity of this estimate depends upon the intensity of light striking this pixel not changing throughout the exposure period, because from the moment the pixel saturates, we are no longer actually measuring it. This is not a concern for ordinary scenes and lighting, though camera or scene motion can cause artifacts, as they would with any camera. Any liability here is compensated by the rapidity of the capture process, which for Pixim DPS is quite high indeed.

Darker pixels (that is, pixels that do not reach saturation even at the end of the exposure period), are sampled at the end of the exposure period, thereby giving them the longest possible time period to capture intensity information. Thus, warp allows DPS to extend the dynamic range beyond what would otherwise be available, to capture useful intensity information from both very bright pixels and very dark pixels.

Alternately, we can think of warp as scaling the exposure time. For instance, a warp of 64 means that for a 1/60 sec exposure, the first capture was made at $1/(64 \cdot 60) \approx 1/4000$ sec. Because the smallest representable non-zero value is the LSB of the 10-bit ADC for the full exposure time (1/60 sec), the total dynamic range available is therefore equivalent to 1024 * 64, or 16 bits.

Iris Control

Light levels above the highest warp go beyond the ability of DPS to represent. The upper boundary of the Warp Control region is called the saturation boundary (see Figure 13). Images with greater intensity than this require the control of the light level with a manual or automatic iris. The transition between the warp control and the auto-iris control is determined by a user-adjustable property of the DPS signal processing system, which typically is set to –20 dB by default.

The DPS System

Beyond the fundamental DPS technology, Pixim has integrated functionality into the silicon and software to produce a comprehensive systems solution. Developers can build high-performance cameras with straightforward hardware design and simple configuration using the Pixim Configuration Language (PCL) to control the hardware interfaces, develop menus, and control the imaging behavior of the camera. No signal processing software development is necessary. Pixim also provides complete reference designs and a variety of camera configurations for very quick time to market.

DPS Technology System Design Application — Security Cameras

Pixim's initial product based on DPS technology was the D2000 video imaging system, a two-chip set targeting the high-resolution color camera security market. The product consists of the D2010 Digital Image Sensor, D2020 Digital Image Processor, and the real-time software and image science needed to control the camera and get high-quality images.

The D2500, Pixim’s second generation chipset shown in Figure 1 has a similar architecture and enables security camera manufacturers to develop high quality, high performance cameras for the video surveillance market. Besides the unmatched image quality and dynamic range made possible by DPS technology, capabilities were designed into the chipset that allow manufacturers to develop full-featured cameras with few additional components.

Benefits for security cameras include:

- High quality images - excellent resolution and colors allows better object and person identification
- Wide dynamic range - can see the entire scene in mixed lighting situations
- Low light performance - see more clearly in darker scenes
- Full feature set - lower power and cost for the camera
- High Integration - smaller camera, lower power, quick time to market

The block diagram of a D2500 design for security cameras is shown in Figure 14. A complete full-featured surveillance camera can be developed with the two Pixim chips, an EEPROM to store the control program, a crystal, power supply, and lens. A surveillance camera using this chipset will exhibit the benefits of DPS technology outlined earlier in this white paper. Camera manufacturers can customize the interface and develop a variety of cameras from a single hardware and software design.

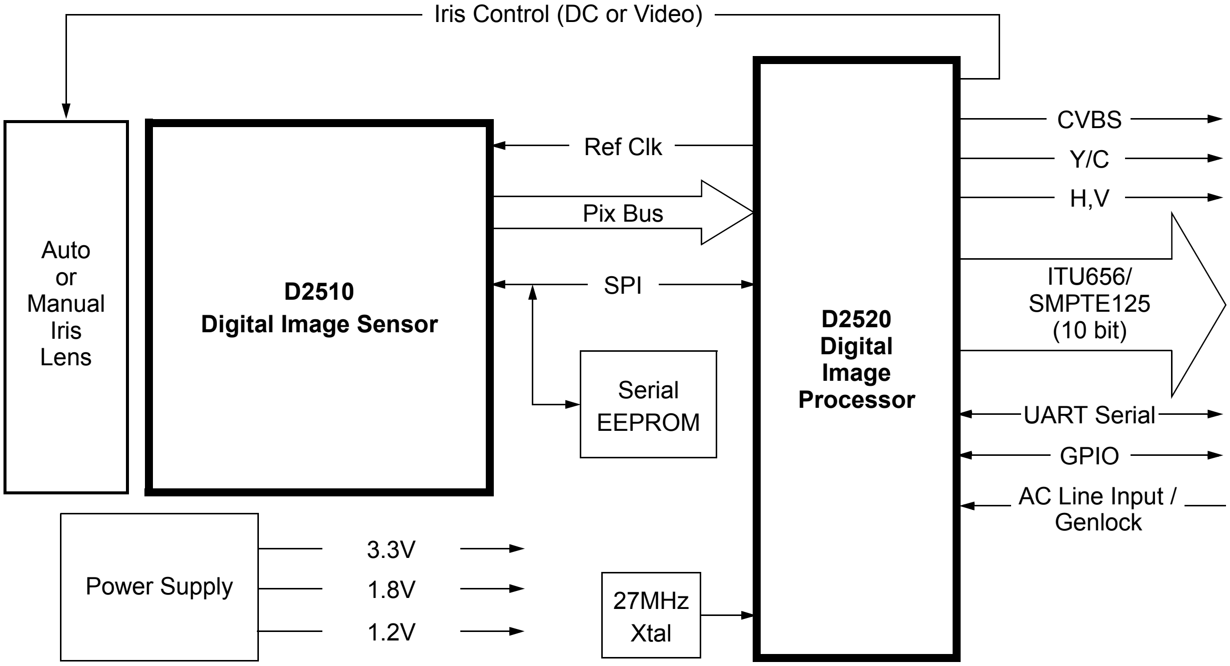

Figure 14. DPS Image Sensor and Processor Block Diagram

Conclusion

Pixim has developed an innovative technology, the Digital Pixel System®, that allows a new level of performance in cost-effective video image capture and processing. The high-speed capture enabled by Pixim's innovative design allows high frame rates, wide dynamic range, and accurate color reproduction at full video rates. The D2000 and D2500 product families are bringing tremendous benefits to the fast-growing security camera market. Other markets that value the image quality, dynamic range, and other benefits of DPS technology will benefit in the future as new products from Pixim are introduced.

Glossary

Image sensor capabilities are evaluated based primarily on the following characteristics:

Dynamic Range: the ratio of the brightest image that can be captured by the imaging system to the darkest image that can be captured. Technically, it is *the ratio of the maximum non-saturating signal to the standard deviation of the noise under dark conditions.* A sensor with a higher dynamic range can detect a wider range of scene illumination than one with a lower dynamic range, producing images with greater detail and contrast levels. Greater contrast levels, such as 16-bit, result in superior images in high contrast scenes that intrinsically require a higher dynamic range. Thus, a sensor's dynamic range determines the contrast levels of an image and is commonly expressed in bits or $\mathrm{dB} = 20\log(n)$, where n is the number of bits. For example, the dynamic range for traditional CRTs and printers is about 10 bits or 60 dB.

Resolution: *The total number of pixels in the sensor array*, usually expressed as x-dimension followed by y-dimension (e.g., a VGA sensor is 640×480 pixels, a total of 307,200).

Noise: Usually expressed as signal-to-noise ratio (SNR) and measured in dB. SNR is a parameter that measures the sensor's *immunity to electrical interference*. Peak SNR is the ratio of the strongest recordable intensity without saturation to the background noise. Some important noise sources are: shot noise, due to the discrete nature of electrons and therefore a function of both exposure and dark current; fixed pattern noise, resulting from asymmetries between pixel and analog-to-digital conversion (ADC) circuits due to process variations; thermal noise introduced when resetting the charge on the pixel; amplifier $1/f$ and white noise; quantization noise; switching noise, created by switching transients coupled through the substrate; and variations in pixel responsivity. The evaluation of the various noise sources depends on the application.

Quantum Efficiency (QE): *The average number of photoelectrons produced per photon at a specific frequency*. For an ideal device, one photon always produces one photoelectron; $QE = 100\%$. Commercially available sensors generally have a range of $20\% < QE < 80\%$.

Sensitivity: A measurement of the *ratio of photon energy incident on the sensor to the sensor's electrical energy output*. For commercial cameras it is sometimes specified by the minimum illumination (lux) at which a camera can be operated, or by ISO or ASA film speed. This characteristic is the complex result of quantum efficiency, noise and assumptions about minimum acceptable picture quality. Semiconductor process characteristics, use of micro-lens structures, CMOS layer deposition, and photodiode fill factor are just a few of the factors that affect sensitivity.

Pixel Size: *The physical dimensions of a pixel*, including its photodetector and supporting circuitry. It is typically measured in microns (µm). Generally, square pixels are more flexible and conducive to quality image processing in imaging applications than rectangular or hexagonal pixels.

Fill Factor: A measure of the *percentage of the total pixel area that collects incident light with a photodiode to the total size of the pixel*. The use of photodiode micro-lenses can substantially improve the effective pixel fill factor.

Readout Speed: *The rate at which the sensor supplies data for further processing*. It is commonly expressed in Mbytes/second or frames/second. CMOS APS sensors use analog transistors to readout data, whereas DPS sensors feature digital readout, resulting in much faster and more accurate readout characteristics. The implementation of pixel-level digital readout has been patented by Stanford University from research conducted since the beginning of 1993.

PIXIM

DIGITAL PIXEL SYSTEM

1395 Charleston Road | Mountain View, CA 94043 | www.pixim.com
| tel (650) 934 0550
| fax (650) 934 0560