What To Do With Sub-Diffraction-Limit (SDL) Pixels?

-- A Proposal for a Gigapixel Digital Film Sensor (DFS)*

Eric R. Fossum†

Abstract

This paper proposes the use of sub-diffraction-limit (SDL) pixels in a new solid-state imaging paradigm – the emulation of the silver halide emulsion film process. The SDL pixels can be used in a binary mode to create a gigapixel digital film sensor (DFS). It is hoped that this paper may inspire revealing research into the potential of this paradigm shift.

The relentless drive to reduce feature size in microelectronics has continued now for several decades and feature sizes have shrunk in accordance to the prediction of Gordon Moore in 1975. While the "repeal" of Moore's Law has also been anticipated for some time, we can still confidently predict that feature sizes will continue to shrink in the near future.

Using the shrinking-feature-size argument, it became clear in the early 1990's that the opportunity to put more than 3 or 4 CCD electrodes in a single pixel would soon be upon us and from that the CMOS "active pixel" sensor concept was born. Personally, I was expecting the emergence of the "hyperactive pixel" with in-pixel signal conditioning and ADC. In fact, the effective transistor count in most CMOS image sensors has hovered in the 3-4 transistor range, and if anything, is being reduced as pixel size is shrunk by shared readout techniques. The underlying reason for pixel shrinkage is to keep sensor and optics costs as low as possible as pixel resolution grows. Camera size has recently become highly important in the rapidly expanding camera-phone market place. Concomitant with the miniaturization of camera components such as sensors and optics will be the miniaturization of optical system components such as actuators for auto-focus and zoom in megapixel camera phones.

We are now in an interesting phase in the development of image sensors – for both CCDs and CMOS active pixel sensors. The physical dimension of the pixel is becoming comparable to the diffraction limit of light at the wavelengths of interest. A perfect lens can only focus a point of light to a diffraction-limited spot known as an Airy disk and the Airy disk is surrounded by higher order diffraction rings. The Airy disk diameter D_A is given by the equation:

$$D_A = 2.44\ \lambda\ F\#$$

Where λ is the wavelength and F# is the F-number of the optical system. For example, at 550 nm, and F-number of 2.8, the Airy disk diameter is 3.7 um. Yet, pixel sizes in megapixel image sensors are at this size or smaller today. We refer to pixel sizes smaller than the 550 nm Airy disk diameter as **sub-diffraction-limit** (**SDL**) pixels.

Today it is possible to build a 6-T SRAM cell in less than 0.7 um^2 using 65nm CMOS technology. It is not difficult to imagine the feasibility of making pixels in the 0.25 um^2 to 1.0 um^2 size range in the near future. However, significant issues exist for a 0.25 um^2 pixel even though it might be tempting to make a 2Mpixel sensor with 1mm diagonal using such a small pixel. For example, full-well capacity in a low-voltage deep-SDL pixel would probably drop to a few hundred electrons leaving a maximum SNR of 10-20:1 and very poor dynamic range. The number of photons collected by the pixel would also be small leading to totally unacceptable noise levels.

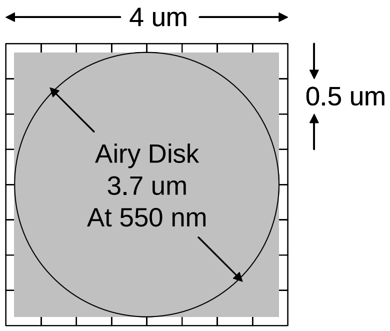
---

* Portions of this work will appear in the book Image Sensors and Digital Signal Processing for Digital Still Cameras, J. Nakamura, ed., Taylor and Francis Books (2005) ISBN:0849335450.

† This work was self-sponsored and performed in 2004. The author is currently affiliated with Siimpel Corporation, 400 E. Live Oak Avenue, Arcadia, CA 91006 www.siimpel.com Fossum@siimpel.com The author is also affiliated with the University of Southern California, Department of Electrical Engineering-Electrophysics, Los Angeles, CA.

Furthermore, the resolution of the sensor would be well beyond the diffraction limit. In fact, over 40 pixels would fit inside the Airy disk and well over a billion pixels (gigapixels) could fit on a single chip.

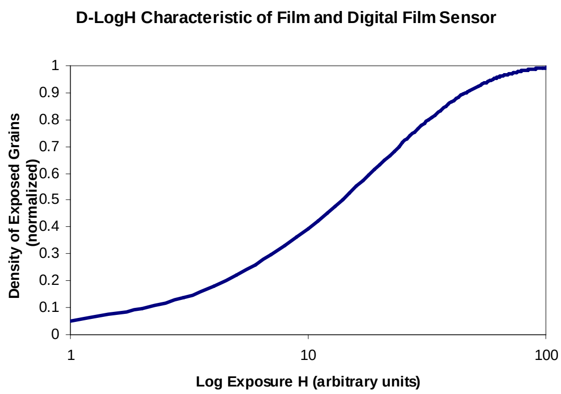
Despite the seemingly impractical nature of SDL and deep-SDL pixels one might instead postulate that since it will be possible to build such tiny pixels, they will be built. The question then is, what will we do with them?



One strategy is to employ oversampling concepts. In essence, we highly oversample the optical resolution of image. For SDL pixels, this can help with color aliasing problems from color filter arrays, and the diffraction effect can probably be used to eliminate the need for anti-aliasing optical filters. For deep-SDL pixels, improved resolution of the optical image can probably be achieved using digital signal processing. However, unless something can be done about the SNR achievable in SDL and deep-SDL pixels, the effective use of digital signal processing will be limited.

In this paper we wish to examine the emulation of film. In film, silver halide (AgX) crystals form grains in the sub-micron to the several micron size range. A single photon striking the grain can result in the liberation of a single silver atom. This grain is effectively tagged as "exposed" and constitutes the latent image. In the subsequent wet chemical development process, the one silver atom results in a "runaway" feedback process that chemically liberates all the silver atoms in the grain. This leaves an opaque spot in the film since the grain has been converted to silver metal. Unexposed grains are washed away. The image intensity is thus converted to a local density of silver grains.

The probability that any particular grain is exposed under illumination grows linearly at first but only eventually approaches unity. The quantitative calculation is beyond the scope of this paper, but this process gives rise to film's special D – log H contrast curve, where D is density and H is light exposure. The smaller the grain size, the lower the probability that the grain will be struck by a photon in a given exposure, and the "slower" the film speed will be since more light is required to ensure a high probability that all grains are struck by photons. However, the spatial resolution of the image is determined by grain size, with smaller grain sizes and slower film having higher image resolution.



In a developed film image, the grains are binary-like since they are either exposed or not exposed. The local image intensity is determined by the density of exposed grains, or in digital parlance, by the local spatial density of "1"'s.

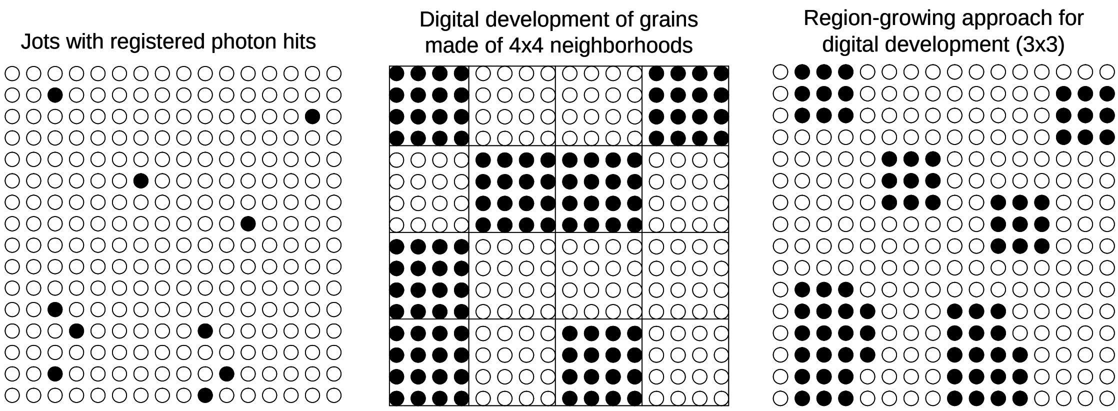
It is easy to transform this concept into a digital-film sensor (**DFS**) by emulating this process. Say we have an array of deep-SDL pixels where each pixel is just a fraction of a micron in size. The conversion gain needs to be high and the readout noise low so that the presence of a single photoelectron can be determined. (Actually, several photoelectrons could contribute to pushing the output signal above some threshold, but ultimately single photoelectron sensitivity would be desired). From the discussion above, it is evident that a pixel that only needs to detect a single photoelectron has much lower performance requirements for full-well capacity and dynamic range than an analog pixel in a conventional image sensor. We call these specialized pixels "**jots**"‡.

Implementation of a jot might be done is several ways. A brute force approach would be to make a conventional active pixel with very high conversion gain (low capacitance). Other approaches include using avalanche or impact ionization effects to achieve in-pixel gain, and the possible application of quantum dots and other nanoelectronics devices. Stacked structures are also possible, especially since performance requirements are reduced. Dark current is still critical.

At the start of the exposure period, the jot is reset to a logical '0'. If it is then hit by a photon during exposure the jot is set to a logical '1' immediately or upon readout. (This is reminiscent of using memory chips as image sensors except we seek much higher sensitivity). Due to the single-bit nature of the "analog-to-digital" conversion resolution, high row-readout rates can be achieved, allowing scanning of a gigapixel sensor with perhaps 50,000 rows in milliseconds and enabling multiple readouts per exposure.

The read out binary image is **digitally developed** to a conventional image of arbitrary pixel resolution, using a two-step process, trading image intensity resolution for spatial resolution. We define a "grain" in the DFS as being composed of a neighborhood of jots, say 4x4 jots (or even a single jot). In the first step of development, if any jot in this grain has been hit by a photon and is a logical '1', the grain is considered exposed and all jots in the grain are set to '1'. The digital development process allows the flexibility of setting a grain size during readout to adjust the effective ISO of the DFS. Or, the grain size can be selected to optimize image quality after the exposure. Alternatively this first step of digital development can be performed as a region-growing image processing function. In any case, the development is a sort of jot area-amplification process. This first step of digital development might be used in very high jot-count image sensors under low light conditions and corresponds to large-grain film emulsions for very high film speed although the necessity of the grain construct is not clear to me at this time.

Unlike film where the grain boundaries are fixed during an exposure, it is possible to conceive of an imaging process where the jots are read out several times during a single exposure. The exposures can be added (logically 'OR'd) together so that the grain construct is both spatial and temporal. Alternatively, the neighborhood mapping function can be different for each readout which would be like dithering the grain position in a film emulsion during exposure, and perhaps even varying the grain size during the exposure.



In the second step of digital development, the grains which form a binary image need to be converted to a conventional digital image that contains pixels with intensity values between, say, 0 and 255. In this case, a local density of exposed grains is mapped into a pixel image. The more exposed grains in a neighborhood, the higher the pixel value. Neighborhoods can overlap or be

‡ After the Greek and Latin words iota or jota, meaning the smallest thing.

distinct. If they overlap, this second step is a sort of blurring convolution process followed by subsampling. In a sense this step is not unlike digitizing a conventional image on film. At high magnification, the image would appear to be binary due to the presence or absence of silver grains, but at the lower magnifications used for digitizing film, the image appears as a continuous gray tone that can be digitized into an array of pixels.

Color can be handled in a manner analogous to today's color image sensors. Jots could be covered with color filters. Red (R), green (G), and blue (B) jots could be treated separately and later the digitally-developed images combined to form a conventional RGB image. R, G, and B jots need not appear at the same spatial frequency, and since the deep-SDL nature of the jot pitch results in blurring from diffraction effects, color aliasing is not expected to be an issue.

Like film, we expect this jot-based DFS to exhibit D-log H exposure characteristics since the physics and mathematics is nominally identical to film. That is, the dynamic range could be large and the exposure characteristics more appealing for photographic purposes.

The DFS imaging process proposed above is not clearly superior to today's imaging techniques, although it might be. Rather, it is a proposal on how to use deep-SDL pixels and introduce a paradigm shift in solid-state image sensors. In a sense, it is an extreme extrapolation of CMOS active pixel sensor technology where pixel sizes are measured best in nanometers, conversion gain becomes extremely large, charge-handling capacity minute, and pixel resolution increased by orders of magnitude.

The DFS imaging process has many open questions that need to be addressed. For example, how is shot noise manifested in this type of sensor? How does shot noise translate into SNR as one varies grain size? What is the optimum trade off between grain size, sensitivity and SNR? What is the conversion characteristics of digital film – is it exactly D-log H? Does it become easier to suppress dark current in a DFS compared to a conventional sensor, or does it work out to be exactly the same? How does one optimally trade off the conversion between the binary image and the digitized pixel image? Is there a need for multi-jot grains or can the same effect be accomplished without the grain construct? How many jots does it take to become equivalent to, say, a 3 Mpixel CMOS APS in image quality? Is there a way to use this DFS approach today to reduce optics cost or size? How does one implement a jot? How does one read out jots at high speed and low power? To what extent can neural processing techniques be used to convert the jot-based image to conventional image representations? Does the DFS approach make sense for certain applications compared to conventional CMOS or CCD image sensor approaches?

The author hopes that some of these questions will be intriguing enough to stimulate research into the DFS concept at the university or industrial level.