

# Gigapixel Imaging

## An Overview of the Technology and Related Technical Challenges

Maria Shishkina

Cologne University of Applied Sciences

Cologne, Germany

mashishkina@web.de



**Figure 1.** Detail of the gigapixel image of Pieter Bruegel the Elder’s *The Fall of the Rebel Angels*, 1562, captured for Google Arts & Culture. The full-resolution version of the image is sampled at 72 dpi with 30.000x21.654 px total. [Public domain], via Wikimedia Commons. ([https://commons.wikimedia.org/wiki/File:Pieter\\_Bruegel\\_the\\_Elder\\_-\\_The\\_Fall\\_of\\_the\\_Rebel\\_Angels\\_-\\_Google\\_Art\\_Project.jpg](https://commons.wikimedia.org/wiki/File:Pieter_Bruegel_the_Elder_-_The_Fall_of_the_Rebel_Angels_-_Google_Art_Project.jpg))

### Abstract

Recent trends in the semiconductor technologies outline the possibility to acquire images with the resolution surpassing the physical limits of human perception. Resulting images lie in the gigapixel range; compared with the modern consumer electronics, the resolution of such sensors is higher by three orders of magnitude.

Here, we explore a recent gigapixel imaging solution, the specialized sensor model with a spectral sensitivity response and the spatial resolution nearing that of a photographic film. The pixels incorporated on an array converge to the theoretical minimum size limit – to the point where their respective full well capacitance is limited to a photon each – thus providing a non-linear binary response with the intensity distribution that adheres to the fundamental physical principles. To accommodate for the diffraction limits of optics, the sensor values are subjected to oversampling; temporal oversampling is possible as well to achieve better estimation of factual intensity values.

We present the image formation model of the quanta imaging sensor which should help simulate image capture in a sub-diffraction resolving light field. Later, we report on the

most commonly advised method used for the image reconstruction and the parameters needed for accurate estimation with present noise sources. Due to the compact design of the photoelements, quanta imaging sensor provides high readout speeds with high resolution and inherits most of the advantages of the current state-of-the-art CMOS.

**CCS Concepts:** • **Hardware** → *Single electron devices; Emerging optical and photonic technologies; Emerging architectures; Noise reduction; Digital signal processing; Sensor applications and deployments; Displays and imagers;* • **Mathematics of computing** → *Stochastic processes.*

**Keywords:** gigapixel imaging, photon-counting, jolt devices, binary image sensor model, SPAD, quanta imaging sensor

## 1 Introduction

### 1.1 Motivation

With every consecutive year, the information processing technologies advance in functionality and performance, leading to improvements in information capture, transmission, storage, and subsequent manipulations.

In modern imaging devices, one of the most significant performance metrics is the spatial resolution of the system,

expressed as the amount of resolvable information per unit of space. High resolution is crucial for distinguishing structures located within a small proximity to each other. Resolution improves with the shrinkage of the imaging elements, e.g. pixels in a sensor, and allows for more precise representations while potentially retaining the same system size in general.

Such developments are generally desired: while not necessarily as significant for the most conventional consumer imaging devices, improved accuracy and reconstruction of the data is pivotal for a number of domains. These include remote sensing for astronomical imaging, microscopy research in biological sciences, prognostics, and medical examinations, and many more. Super-resolution is equally important in industrial applications, where the amount of data available for interpretation can be used in computer vision tasks such as object recognition or traffic surveillance. [Anwar et al. 2020; Yue et al. 2016]

## 1.2 State of the Art

The current implementations tend to make use of the established digital sensor technology standard, the CMOS active pixel imaging sensor with intra-pixel charge transfer. Often described as a "camera-on-a-chip" technology, the conventional CMOS image sensor (CIS for short) has found use in both consumer and industrial applications due to its low-power consumption, high level of integration and small manufacturing sizes [Fossum 2018].

The highest resolution pixel-wise in current imaging devices is nearing 150 MP [Lawton and Artaius 2021], where it's partially associated with larger sensor formats, sacrificing the compact size in the process for insignificantly larger amount of picture elements. High prices render such cameras unattainable for an average consumer. Besides the actual acquisition process, existing images have to be reconstructed in higher resolutions as well.

**1.2.1 Software-Based Implementations.** Many existing image processing approaches can be employed to enhance spatial resolution in existing *low resolution* (LR) images. Based entirely in the software domain, such solutions require nothing but computational power and are utilized the most, to the point of being synonymous with the term "gigapixel imaging" itself.

The image recreation process is data-driven: the input can be a single image, or consist of a dataset of similar images with insubstantial differences from one another. Similarly, the methods employed can be based on the computer vision algorithms in the traditional sense or involve deep learning network architectures, which require an additional source of data in order to be outperforming the prior [Anwar et al. 2020].

**Image Stitching.** Image stitching techniques explicitly require abundant, spatially and temporally coherent image

datasets with overlapping regions. The most common application of such techniques consists in the production of high-resolution panoramic images, which can then be utilized for entertainment or preservation purposes [Jusof and Rahim 2014].

By now, the production of the panoramic gigapixel images is steadily made available to the consumers. A fine example is the GigaPan technology [Equipment and Company 2016], which stems from the collaboration between Carnegie Mellon University, NASA Robotics Group, and Google; since then, it has been spun off into a commercial service offering both hardware accommodations and the software, thus defining a clear, streamlined workflow. The solutions encompass a designated robotic mount facilitating continuous image capture through panning along its axis, a corresponding image stacking software, and even a web-based interface for sharing the composite image afterwards.

The camera mount is meant for use with any amateur DSLR camera, meaning that the exact resolution of the resulting panorama is dependent on the dimensions of the optical components and the sensor model. The demonstration of the process can be found online [Cremer 2012].

**Image Synthesis via Deep Learning.** Deep learning-based methods are best employed for a reconstruction of a single low-resolution image.

The super-resolution version of a single image is by definition an under-determined problem with a multitude of possible solutions. Classical computer vision algorithms don't fare well in such tasks, as the recovery of such images may lead to the propagation of wrong information. For an accurate reproduction of the scene, one may need reliable prior information which relates to the human perception in a meaningful way.

The objects depicted in the image can be upsampled based on their predicted taxonomy and feature sets, which is one of the main applications of the convolutional neural networks. Depending on the data involved in the training, the interpolated data could be refined in many ways; for instance, distinguished objects could have their corresponding surfaces refined based on their class, filling in the structures with already available high-resolution textures of the corresponding materials.

A number of network designs, varying widely in their complexity, depth, training data, and features, has been considered for the gigapixel imaging. An extensive overview on the available architectures can be found in [Anwar et al. 2020] and [Shackleton and Altahhan 2019].

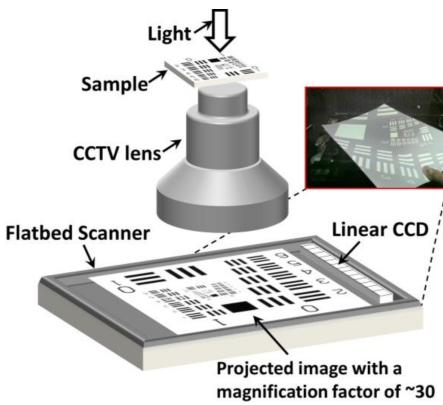
**Applicability.** One should note that neither image stitching nor image synthesis techniques are exclusive to the gigapixel imaging in their methodology; working with a sequence of images in particular is a challenge which quickly turns into an image registration problem, requiring integrated solutions regarding tonemapping, uneven illumination, projective distortions of the depicted objects etc. For

this reason alone, the strictly software-based solutions pose no special challenge to the gigapixel imaging and are therefore declared out of scope of this paper.

**1.2.2 Hardware-Based Implementations.** The computational complexity of the problem can be reduced by addressing it on a lower level of abstraction. Here, the focus lies on acquiring the super-resolution directly by means of the appropriate hardware architecture design. With this, the need for elaborate LR datasets lapses as well.

Several solutions focusing on the predominantly hardware-based domain exist.

**Scanning.** A popular imaging solution involves a flatbed scanner capable of relatively high scanning resolution (commonly provided in dots per inch, or DPI): with most consumer scanning devices supporting 300 dpi and above, this off-chip solution is both low-cost and widely available.



**Figure 2.** The proof-of-concept gigapixel microscopy setup, as demonstrated in [Zheng et al. 2014]. The scanner at hand can resolve 2400dpi, which corresponds to the pixel size of ca.  $10\mu\text{m}$ .

In gigapixel solutions tailored specifically to microscopy research, the scanners register magnified projections of the input samples [Zheng et al. 2012]; other papers attempt to enhance the scanning process by modifying the scanner device directly [Shin et al. 2020], or using alternative scanning methods (e.g. tilt scanning in [Ben-Ezra 2011]).

**Composite of Sensor Arrays.** An attempt to attain high resolution in the digital domain using existing technologies is a popular method in large scale applications.

Widely used in airborne and astronomical systems are precision-mosaicked sensor arrays. In this case, conventional MP-sized imaging sensors are brought together into an array and treated as a whole. Since the optical components of the system are mostly unchanged, this is a fixed-focus system.

Such cascading architectures are rather unwieldy, expensive, and is therefore mostly found in research facilities, where the rigidity of the architecture is unimportant.

This stacked sensor technology can be seen in airborne and astronomical systems such as ARGUS, LSST and Pan-STARR [Brady et al. 2012].

Alternatively, the system can employ multicameras, allowing for a multi-scaled approach. One of such gigapixel camera architectures is AWARE, developed by The Duke Information Spaces Project. AWARE employs arrays of microcameras, using a monocentric lens objective as the primary optical component. The first generation, AWARE-2, consisted of 98 microcameras on a monochromatic CMOS sensor, producing a 1.0 gigapixel image spanning a  $120\text{deg} \times 40\text{deg}$  field of view [Gong et al. 2017]. Since then, a variety of models have been developed, enabling color acquisition and ranging in optical resolution, size, and amount of microcameras. There are operational advantages to AWARE as well, since each camera can be independently controlled in acquisition settings [Brady et al. 2012]. Summarized, such multilens cameras are more compact and serve well in terrestrial applications. [Project [n.d.]]



**Figure 3.** AWARE-2 camera design, via: [Cartwright 2018]

**Photon-Counting.** The latest advancements in the manufacturing of integrated circuits (ICs) have put us in the post-Moore's law era.

From 1970s on [Cerofolini and Mascolo 2007], the complexity of the circuitry would grow at a much higher rate than once theorized, while the physical dimensions of the semiconductors would scale down. By 2020, the possible pixel pitch in CMOS sensors can be as low as  $3\text{nm}$  [Hao et al. 2021], meaning that the sensor dimensions used in current photographic devices can now hold significantly larger pixel amounts, providing us with higher spatial resolution in an image. However, at the present time the pixel dimensions are approaching physical limits: by reducing the pixel area, the full-well [photoelectron charge] capacity of a singular pixel, too, is lowered. This has grave implications for the image quality, as the small difference between the signal and the noise decreases the dynamic range.

The solution to overcoming the trend of pixel shrinkage while maintaining the higher resolution has been proposed by Eric R. Fossum in 2005.

Sufficiently small pixels should only be able to store charge of one photoelectron alone. Each of those pixels can then assume values of either 0, meaning no photons have reached the light-sensitive sensor plane, or 1, signifying that at least one has struck the surface after all, resulting in registered exposure. The sensor response lies in a binary range. The actual evaluation of the sensor data, i.e. the image formation, happens post-capture via analysis of the Poisson distribution of the photons over multiple observations in time, and can be managed dynamically.

Photon-counting, or, quantum imaging sensors can ultimately represent a paradigm shift in photography. Since its introduction, the binary sensor has been applied and conceptualized in many ways.

## 2 Theoretical Background

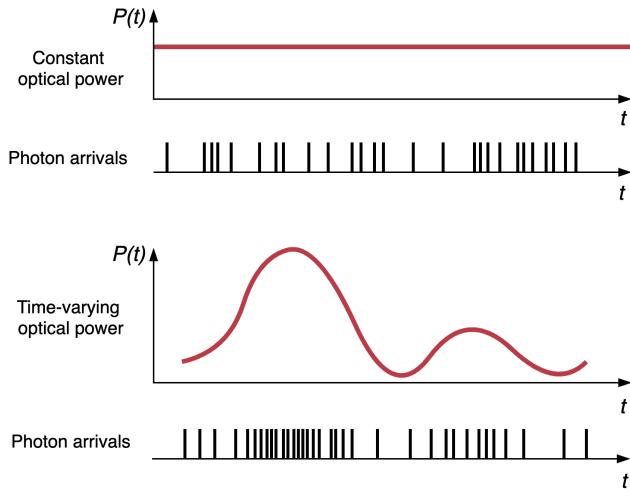
### 2.1 Light Propagation

The electromagnetic radiance is expressed through the flow of minute particles, or quanta, named *photons* in the context of light. The energy of a single photon,

$$W = h \cdot v \text{ [J]} \quad (1)$$

is a product of the frequency  $v$  [Hz], and the Planck constant  $h = 6.62607015 \cdot 10^{-34}$  J s, valid for every harmonically oscillating system.

Photons arriving at each pixel can be approximated by a Poisson distribution function and arrive at random times. The rate of the photon arrivals, i.e. the distribution of the photons, is however still dictated by the local light intensity, as seen in Figure 4.



**Figure 4.** Random photon arrival time, modeled with two optical power sources. Even for a constant source, the arrival is probabilistic. Adapted from [Saleh and Teich 1991].

### 2.2 Photoelectric Effect in Sensors

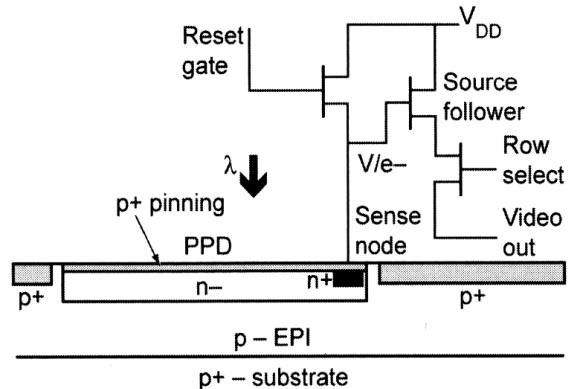
Imaging electronic devices consist of arrays of pixels, consisting of photodiodes accumulating charge through incoming light and integrated circuits which read out the charge and process it further.

For the accumulation of charge, the properties of the semiconductor materials are exploited. The unique electrical conductivity of semiconductors can be altered significantly through external sources of energy, e.g. thermally or by light illumination, and increases proportionally.

The incident photons impinge on the surface of the image sensor and are absorbed through the corresponding structures in the semiconductor material, creating electron-hole pairs as the electron is excited from the valence to the conduction-band. The photoexcited carriers remain within the sample; this is called *internal photoeffect*. [Saleh and Teich 1991]

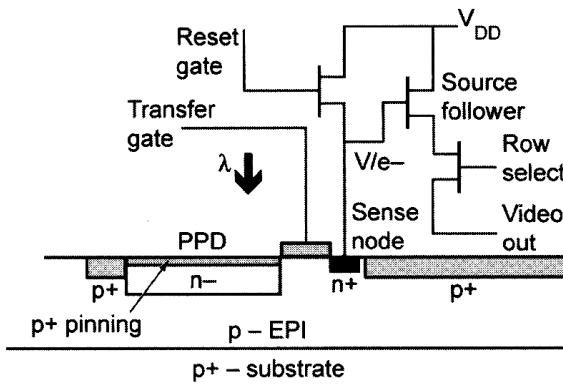
Electrical and optical properties of the semiconductors can be altered further through *doping*, an adding of local impurities for changed concentration of the carriers. In semiconductors, selective doping creates an electrostatic potential well. Subjected to the externally applied electric field, the freed photoelectrons drift through the material towards the potential well, where they are stored and integrated for the duration of exposure. This photoelectron charge is later transferred to the sense capacitance, thus producing measurable electric current through changes in the charge on the charge sense node (*floating diffusion FD*) and converting the signal to voltage, which is read out by the a source follower transistor. [Fossum 2020]

The outputs of the source-followers are amplified and fed to the signal processing circuits on a greater scale, allowing the separate pixels to be addressed column-wise. After the integration, the pixels are reset.



**Figure 5.** 3T PPD structure, via [Holst and Lomheim 2011]. Typically, the dark current is heightened through the direct n+ contact; here, this is prevented through p+-pinning structure on the pixel surface.

Many implementations for the photodiodes exist. In the simplest pixel circuit, three transistors (3T: source-follower, reset, and row select transistors) are needed for the integration process. Modern CIS technologies make use of 4T pinned photodiode structure, in which an additional transfer gate is positioned between photodiode and the sense node. This ensures reduction of dark current and complete charge transfer; the “pinning” of the photodiode allows for reduced sense node capacitance. Furthermore, with this design structure, the pixels can be reset simultaneously, allowing for same integration periods globally.



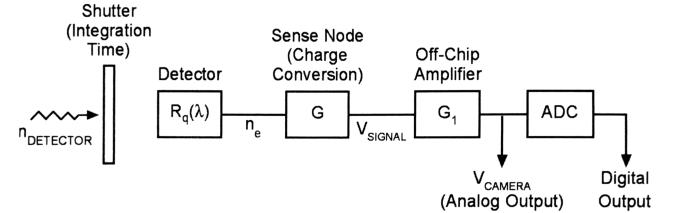
**Figure 6.** 4T PPD structure, via [Holst and Lomheim 2011]. The transfer gate provides isolation of the regions.

Only a fraction of the incident photon flux is absorbed and brought to a higher energy level. The ratio of the photoelectrons to the incident photons is given by *quantum efficiency* (QE), which often ranges from 50% to 80% [Fossum 2020]. The relationship between the changes in output voltage caused by the photoelectric effect to the number of photoelectrons is given by conversion gain (CG); in modern CIS devices, this typically amounts to  $100\mu\text{V}/e^-$  [Ma and Fossum 2015a].

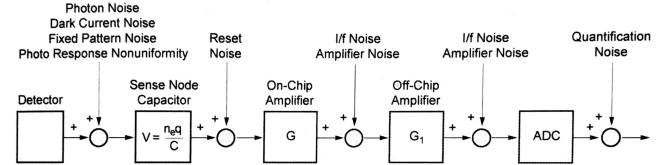
### 2.3 Noise Modeling in Digital Sensors

The noise, that is, unwanted deviations in the signal, can be comprised of the multiple sources which are subsequently added throughout the digital image acquisition chain. Deviations in the brightness are introduced right in the detection phase due to the discrete nature of electrons; such noise source is dubbed photon shot noise. Further fluctuations are caused during the readout of the electron charge and its subsequent amplification; for digital systems, the mapping of the voltage to discretized bit values results in the additional quantization noise. [Holst and Lomheim 2011]

Figures 7 and 8 show the signal transfer and the noise occurring during each step of the process, respectively. Each noise source has a different characteristics and can be attributed to either nonuniformities and physical flows of the sensor, or the natural nonstationary processes. The level



**Figure 7.** Signal transfer diagram in an imager, via [Holst and Lomheim 2011]. Both analog and digital outputs are considered.



**Figure 8.** Noise transfer in a digital sensor, via [Holst and Lomheim 2011]. Each subsystem introduces additional noise sources which affect the subsequent signal.

of detail in the noise transfer diagram is somewhat dependent on the application, and some of the sources caused by the circuitry can be minimized through the design of the architecture, as will be shown in the further chapters.

### 2.4 General Requirements in Gigapixel Applications

Several general requirements must be met in gigapixel imaging devices. Since the higher information efficiency is desired in both scientific and consumer-grade applications (see Section 1.1), a high-resolution camera must not differ drastically from current implementations in its imaging capability.

For components which are primarily influenced by optoelectronics, [Cossairt et al. 2011a] name following requirements:

- a *field of view* (FOV) of at least ca.  $40^\circ$ , which is equivalent to the angle of view for humans at infinity focus ( $\text{FOV} = 2 \times \arctan\left(\frac{\text{sensor size}}{2f'}\right)$ ,  $f' = 50\text{mm}$ ),
- an f-stop of ca. f/4-f/11 for operation under most common lighting conditions,
- and a lens design with low complexity to facilitate both portability and low manufacturing costs.

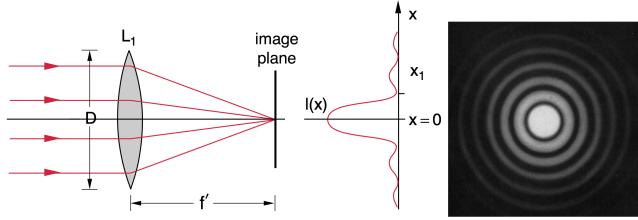
### 2.5 Diffraction-Limited Systems

In a typical imaging system, before arriving at the sensor, the incident light is collected and focused through an optical system first. Lenses, most common optical components, have a resolving power of their own, which has to match the resolution of the sensor and is fundamentally limited by the nature of light.

Upon encountering an opening (*aperture*), the incident light diffracts into a series of concentric waves propagated concentrically around the corners. Assuming a perfectly circular aperture, wavefronts belonging to one point in object space will produce a diffraction pattern with a bright spot in the middle, surrounded by concentrical rings of reduced intensities. The diameter of such pattern, *Airy disk*, is given through:

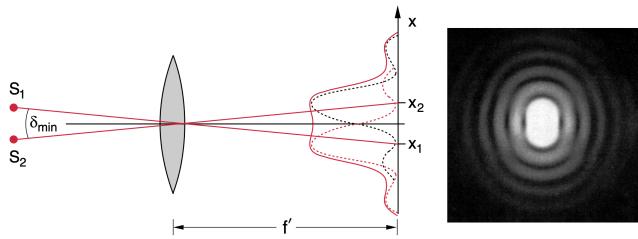
$$D_{\text{Airy}} \approx 2.44 \frac{\lambda f'}{D} = 2.44\lambda k$$

with  $\lambda$  being the wavelength of the incoming light,  $f'$  the focal length, and  $D$  the diameter of the entrance pupil, respectively; in lenses, which can be treated as a circular aperture of respective diameter, the *focal ratio* serves as the *nominal aperture*  $k$ , or *f-stop*.



**Figure 9.** Airy disk pattern caused through lens diffraction of a point source, via [Cagnet 1962], and the corresponding distribution of intensities, [Demtröder 2018].

For two distant neighboring points, the half-angle  $\delta'_{\min}$  subtended by the numerical aperture (represented by the lens) is too small. Thus, the two objects will not be perceived as separate entities; as their intensity functions coincide at their respective minimums, the disks will start to overlap.



**Figure 10.** The imaging subsystem is not capable of resolving two adjacent object points, leading to the interferences in the signal. Via [Cagnet 1962] and [Demtröder 2018].

The angular resolution of an optical system in an object plane is concomitant with a spatial resolution in an image representation, due to the small-angle approximations [Pedrotti et al. 2007]. Defined via *Rayleigh criterion*, the least resolvable distance between two points should correlate with the distance between the two peaks of the intensity distribution function (as seen in Figure 10):

$$\delta'_{\min} = 1.22\lambda \frac{f'}{D} = 1.22\lambda k \approx \rho'_{\infty}$$

The latter is defined as the resolution limit  $\rho'_{\infty}$  and can be approximated for the near field application, where the distance  $a'$  between the aperture and the sensor plane is larger than the focal length  $f'$ :

$$\rho'_{\text{near}} = 1.22\lambda \frac{f'}{D} \frac{a'}{f'} = 1.22\lambda k(1 - \beta')$$

The resolution limit presented earlier sets a limit for the peak lens performance, which varies with the fundamental physical quantities, like wavelength range and focal ratio. Further limitations are imposed through the geometric aberrations, caused by the physical imperfections in the optics itself.

The number of distinct resolvable points for a lens is defined by the *space-bandwidth product* (SBP) [Cossairt et al. 2011b] and ties the lens performance to the sensor size.

[Cossairt et al. 2011a] state two common approaches for enhancing the SBP. Those consist of either scaling the system up, or introducing further optical surfaces to the imaging system for better lens performance. The scaling can occur in either one of the subsystems, meaning an integration of a sensor with larger dimensions or scaling the lens. The second measure should lessen influence of geometric aberrations, leading to a diffraction-limited performance; however, the increased complexity of the optical system implies higher manufacturing costs and errors due to misalignment. Detailed description of scaling laws for the lenses is provided in [Cossairt et al. 2011b], who account for both diffraction-induced limits and geometric aberrations.

As the CIS pixel sizes are routinely surpassing the  $1.5\mu\text{m}$  extent, the optical system in a gigapixel application is doomed to have a lower resolving power than the sensor itself even if the lens performance is solely diffraction-limited; hence, the resolution in image space will be limited through optics. Resulting image representations contain spatial redundancies due to the light field being oversampled.

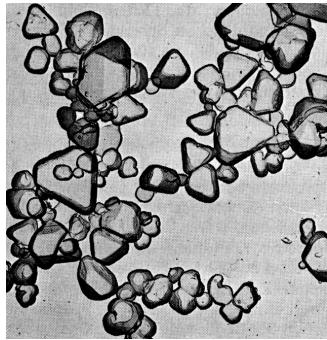
While many proof-of-concept solutions exist that aim to enhance the resolution on a solely optical level (i.e., multi-scale design and monocentric lenses), computational post-processing in gigapixel devices might turn out to be compulsory.

### 3 Photon-Counting Detectors

#### 3.1 Photon-Counting in Film Photography

The principal notions of the single-bit imaging sensors can be traced back to analog photography and demonstrate behaviors reminiscent to ones of the silver halide emulsion, a common component in the photographic film.

Akin to its digital counterpart, the photographic film is subjected to the photoelectric effect; in film, this is represented by the photographic emulsion coated on the transparent sheet of plastic. Such emulsion consists of microscopic crystals, or grains, of silver halides embedded in protective colloid like gelatin [Carroll et al. 1980], with the commonly considered compounds being silver bromide  $\text{AgBr}$ , silver chloride  $\text{AgCl}$ , silver iodide  $\text{AgI}$ , or some mixture thereof [Schröder 1981].



**Figure 11.** Electron microscopic structure of undeveloped silver bromide crystals, via [Jessien 1974].

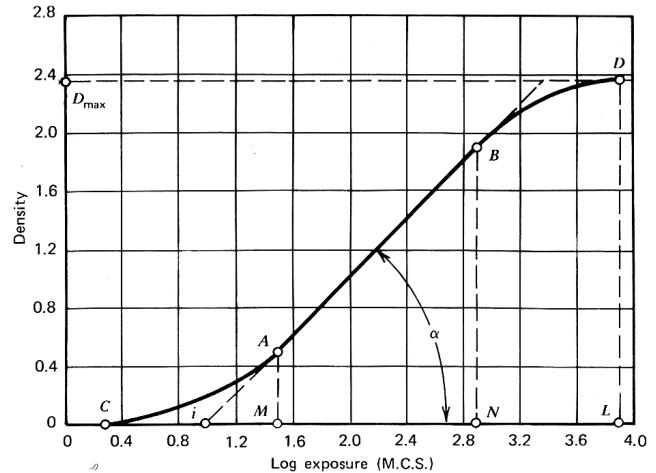
The process starts upon exposure of the film to the light. Film grains absorb the incoming radiation in form of the incident photons and dissociate into a group of free silver ions.

Once exposed to development agent, the silver ions in the film are reduced to metallic silver, rendering the exposed regions of the film sheet opaque black. The unaltered halides are later dissolved out in a chemical bath, so that the corresponding spots are transparent. Each crystal can either be struck by a photon or not, and so the film exhibits a binary response.

The exposed grains are spread randomly, fluctuate in size and differ in absorption rate – there is no uniformity regarding the required absorbed quanta as the particles arrive randomly. However, with the irradiance of the image areas varying, so does the stochastic photon distribution and therefore the concentration of silver particles per film area.

Fine granular structures are not perceivable by the human eye and are registered as continuous tone of gray [Yang et al. 2012] instead, matching the overall area density to the original light intensity information; the mean number of absorbed quanta per grain is given by the Poisson equation.

The photographic properties of the film emulsion are typically described through the ratio of the luminance transmitted in the medium to the luminance of the observed scene. The optical density is modelled as a function of logarithmic exposure and was first derived in [Hurter and Driffield 1890]. As seen in Figure 12, this relationship has a non-linear behavior.



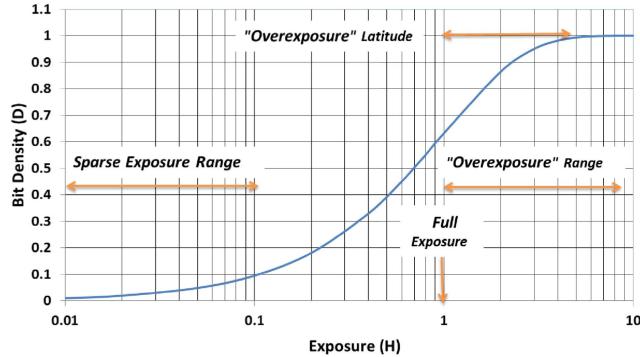
**Figure 12.** Exemplary characteristic curve of the film emulsion, via [Saleh and Teich 1991].

The  $D - \log H$  curve (also referred to as *characteristic curve*) can be split into three intervals based on the curvature of the response. For sufficiently small or large values of  $\log H$ , the function is asymptotic and results in under- or overexposure (the toe and shoulder portions of the curve, respectively [Hunt 2005]); in the middle interval, the slope is constant, and the relationship between the luminance and the photographic density is linear. In general, the S-shape of the curve is determined by the underlying randomness of photon arrivals. (This is different in conventional CIS sensors.)

In film, the grains may fluctuate in size, ranging from 0.048 to  $1.71\text{ }\mu\text{m}$  depending on the application; on average, one grain of the developed emulsion corresponds to ca. 3-4 electrons [Carroll et al. 1980; Fossum 2013]. Feature sizes of QIS are similar in their capacitance; the jots in the single-carrier implementation of QIS are meant to store around 3 carriers as well. Thus, both digital sensors QIS and analog film deliver quantized, spatially discrete values obtained off similar area sizes. So, the behavior in binary imaging sensors is, too, dominated by the photon distribution. Resulting characteristic curve is shaped comparably, demonstrating slightly larger overexposure latitude [Fossum 2013].

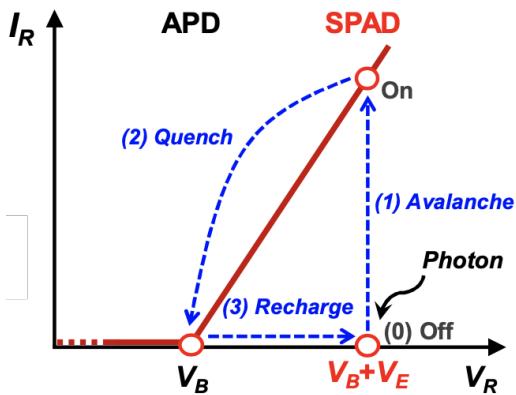
### 3.2 SPAD Matrices

The binary response of the sensor defines the need for a subsequent signal amplification process to generate a sufficiently large voltage signal from a single photon. In one of the currently available gigapixel imaging technologies, this is achieved via the integration of the *single-photon avalanche diode* (SPAD) structures, which implement gain mechanisms internally ([Saleh and Teich 1991]); this internal amplification process is referred to as *electron avalanche multiplication* [Ma et al. 2017]. Since each detected photon is converted



**Figure 13.** Characteristic curve of the quanta imaging sensor, via [Fossum et al. 2016].

into a cascade of moving carrier pairs, even weak light can produce a clearly detectable current.



**Figure 14.** SPAD operation process, demonstrating changes in voltage during the avalanche effect and subsequent quenching. Via [Charbon et al. 2018]

In avalanche photodiodes, the junction electric field is large enough such that the available carriers are accelerated, causing *impact ionization*, i.e., producing further photoelectrons through collision. Ergo, high electrical voltage over 20V is required [Gnanasambandam et al. 2019]; some prototypes, like the one described in [Amri et al. 2016], define even higher supply voltages of ca. 22–27V for biasing above breakdown. SPAD is operated in Geiger mode [Bérubé et al. 2015]. For SPAD arrays, higher breakdown voltage is desirable in general, as it allows for higher photon detection efficiency and improved timing resolution [Bérubé et al. 2015].

The first proposal for a fully integrated CMOS SPAD array has been reported in [Rochas et al. 2003]. [Dutton et al. 2016] allude to other SPAD detectors which have been developed since, noting the variety in both general form (single-point detectors, line sensors, large array photomultipliers, and image sensors) and pixel designs; subsequently, it is stated that

no single SPAD-based pixel architecture has been deemed dominant to date.

The fast detection times in SPADs are associated with the rapidity of the impact ionization process. However, each photon detection event must be followed by finite recovery time, in which the diode is restored to the operative level voltage-wise, freeing excess charge carriers. This process is called *quenching*. During this time, device does not respond to further incident photons. To enable quenching, additional in-pixel circuitry is needed. Due to the larger amount of transistors needed per single photodetector, the fill-factor of SPAD-based pixels is limited (<40%), and so is the quantum efficiency (<30%) [Ma et al. 2017].

Generally, the SPAD devices aren't entirely compatible with the CMOS technologies due to high operating voltage requirements; the necessity for a large electric field and additional readout circuitry results in the larger dimensions of the structure, lower spatial resolution and higher power dissipation [Ma et al. 2017]. Furthermore, the generation of electrons is based on thermal reactions, rendering the detectors significantly more susceptible to dark current.

However, significant progress has been achieved in recent years [Amri et al. 2016; Sarbazi et al. 2015]. Performance of SPAD-based architecture using the developments in 3D stacking technologies is described in detail in [Charbon et al. 2018].

### 3.3 The Pixels as Jots

The avalanche-mode operation in SPAD-based photodetectors limits the pixel pitch and the manufacturing yield and delivers highly undesirable amounts of dark current, which can be somewhat reduced through cooling. In fact, the abstinence from the avalanche photodiode technology is stated to be one of the design criteria in photon-counting applications. Quanta imaging sensor (or QIS for short), a different implementation of the binary image sensor, attempts to overcome such limitations [Ma and Fossum 2015b; Ma et al. 2017]. The principal advantages of CMOS arrays (low-voltage operation, low power consumption, ease of integration of on-chip operation control, the possibility of random access to the image data, etc., [Holst and Lomheim 2011]) must be retained.

In QIS, the photodetectors are intentionally made small as to be capable of storing charges of several electrons. Placed together in a structure, these become *jot devices*, with "jot" denoting the "smallest thing", a word of Greek origin [Fossum 2013].

In some of the available research ([Chan et al. 2016]), jots are referred to as subpixels, or, partitions of conventional CMOS memory bit-cells; other sources ([Fossum et al. 2016]) may use the term "*SDL pixels*", referring to the pitch values within sub-diffraction limit (200–500nm).

The jot photoelements have a *small full-well capacity* (FWC, or storage capacity density, often provided in photoelectrons per cm<sup>2</sup>), which ranges from one to around 100 or

200 carriers – compared with the state-of-the-art conventional imaging sensors where the full well is first reached at ca.  $4000e^-$  [Gnanasambandam and Chan 2020a], the values are low. Subsequent circuitry discriminates the output to one of the binary states [Fossum 2013], based on the pre-defined, semi-arbitrarily set threshold  $q$ .

Instead of converting the number of photons impinging on the sensor surface to voltage, either 1 or 0 are transferred and quantized. Precision of quantization is not of essence. The thresholds are set lower due to the relatively small FWC, and the binary response is triggered quickly. Thus, unlike conventional CIS pixels (described as "buckets" accumulating and integrating [larger] charges of multiple photoelectrons), jots require lesser integration times and can be readout at rates as high as 1000 frames, or fields, per second.

[Ma et al. 2014] envision QIS to have structures consisting of 0.1...10.0 Gjots; due to the fast readout speeds, dimensions of this order must lead to data rates of 0.1...10 Tb/s.

Research regarding the quanta imaging sensor technology is currently in progress. A prototype of camera featuring QIS has already been produced, albeit with a lesser spatial resolution and pixel pitch than sought-after. Representing a crucial paradigm shift, this imaging technology has full potential to stand in as the third generation of digital sensors, replacing CMOS. The benefits, advantages, and the detailed methodology behind QIS will be discussed in the remainder of this paper.

## 4 Quanta Imaging Sensor

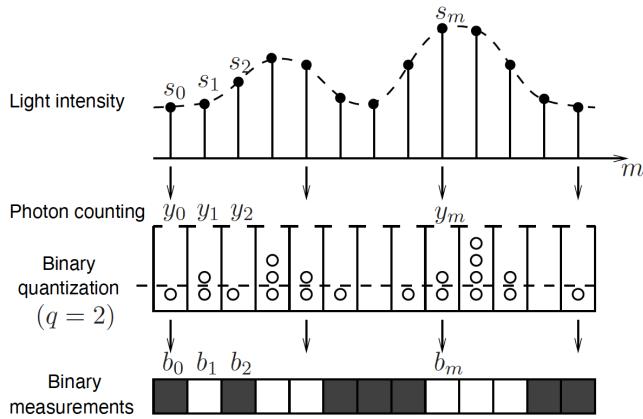


Figure 15. Binary sensor model, via [Yang et al. 2012].

**4.0.1 Jot Data Structure.** The output of the simple QIS design is binary in nature. Combined output of the QIS jots is a series of bit planes, observed over time. Since both spatial and temporal sampling take place, the output can be conceptualized as a cube, as shown in Figure 16.

Unlike CIS arrays which produce images of fixed sizes, quanta imaging sensor allows the output pixel image size to be programmable and dynamically adjustable [Fossum 2013];

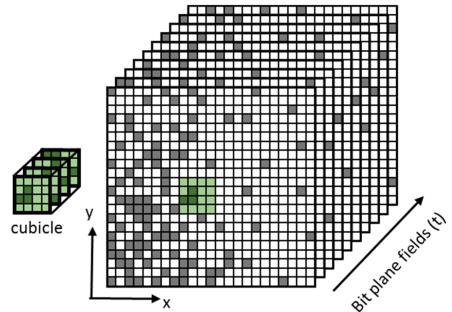


Figure 16. Jot data cube, via [Fossum et al. 2016]

the image reconstruction mostly occurs post-acquisition, and the bit plane can be divided into smaller jot cubicles of arbitrary sizes which should then form a single pixel in the conventional sense from temporal and spatial data.

It is understood that the forming of the image pixels does not require uniformly sized jot cubicles; for separate pixels, these can have different dimensions in  $x$ ,  $y$ , and  $t$ . It follows that selected cubicles of the adjacent pixels may overlap, aiding formation of multiple CIS pixels at once [Fossum et al. 2016].

### 4.1 Distinctions

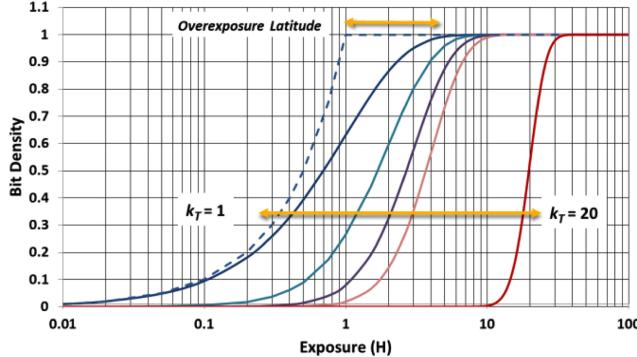
The similarity of the QIS response to the sensitivity of the photographic film emulsion has been established in the previous sections; in both cases, the shape of the characteristic curve is determined by the photon arrival statistics.

In film, around 3-4 carriers were needed to elicit the chemical reaction. Since the binary value is obtained through comparisons against a certain threshold, the choice of threshold can influence the final response. Figure 17 shows the response curves plotted for the number of carriers  $k_T$  needed for the single-bit pixel output of "1" ( $k_T = 1, 2, 3, 4, 20$ ). It is evident that the linearity of the QIS response curve changes with the larger thresholds; similarly, the overexposure latitude is shifted towards larger exposures. Hence, the single-bit jot devices can be applicable under multiple varying lighting conditions, enabling low-light and high dynamic range imaging applications, where the acquisition using multiple threshold values would provide different values for linearity.

For this reason, it is convenient to distinguish between the single-bit jot devices with full-well capacitances in the classical sense, of only one electron (*single-carrier*), and multiple electrons (*multi-carrier*) which can be dynamically adjusted in their responses based on the threshold.

### 4.2 Hardware Design

To facilitate the transition to the QIS as the future dominant sensor technology in the future, strong compatibility with a CIS fabrication line must persist.



**Figure 17.** Single-bit QIS response curves in dependence of the amount of carriers, via [Fossum 2013]

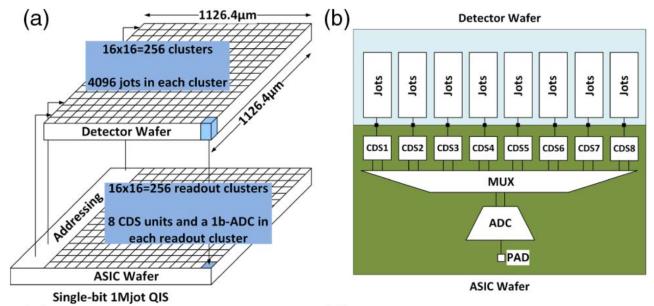
CMOS sensor serves as a suitable basis for quanta imaging due to its low voltage specifications, high quantum efficiency, high spatial resolution, and more. [Ma et al. 2017]. However, due to the source-follower output voltage, the read noise is too high to enable photon detection. Current state-of-the-art CIS demonstrate read noise in the ranges of  $1\ldots 2e^-$  r.m.s. [Ma et al. 2015]. Due to the binary nature of the signal, such amount of noise has grave consequences for the single-bit QIS, resulting in high possibilities of false misquantization. At low light levels, desired read noise in single-bit QIS should amount to  $0.3e^-$  r.m.s. at best, delivering error-free estimation with  $0.15e^-$  r.m.s. under especially sparse exposures ( $0.01 < H < 1$ ). For such low bit error rates, conversion gain of at least ca.  $1mV/e^-$  is needed, which is five times higher than the values demonstrated in the state-of-the-art CIS devices [Ma et al. 2014]. For this reason, current CIS technologies must be modified in order to enable photon-counting.

During the charge-to-voltage conversion, the signal transfer is corrupted by noise due to the voltage fluctuations and the parasitic capacitances in the circuitry, which reduce the transfer efficiency. Higher conversion gain ensures that the estimations are accurate, and the transferred electrons are capable of producing signal above the noise floor. This could be achieved by increasing the voltage, however, higher operating voltage would render the sensor incompatible with most baseline CMOS processes. [Ma et al. 2017] Otherwise, one can attempt to minimize the read noise, which is caused by the capacitance of the floating diffusion. In available works, FD is minimized via specialized jot structures, where the doping profiles of the pixels are modified via implantations. Jot candidates based on their doping properties are covered in detail [Ma et al. 2017; Masoodian et al. 2016]. On average, the doping profiles of the jots have lead to the reduction of the noise to ca.  $0.21e^-$  r.m.s. with the conversion gain nearing  $350\mu V/e^-$  with 15% variation [Ma et al. 2017].

Most of the publications referenced in this work explore prototypes for jot devices which were fabricated in conformance with the 45/65nm 3D stacked CMOS process with

backside-illumination (BSI), which guarantees high fill factor and quantum efficiency [Ma and Fossum 2015b].

Figure 18 demonstrates the architecture of existing 1Mjot clusters with vertical integration. The jot device is vertically separated from the readout circuitry, which reduces the total area and allows faster readout. Presented array can be implemented with analog and digital readout circuitry; Such clusters can be arranged in parallel, thus forming a chip of significantly larger dimensions without any significant delays in speed. In this particular jot device, analog clusters were integrated together with the digital clusters.



**Figure 18.** Simplified architecture of the 1 Mjot stacking QIS chip (left) and one digital cluster (right), via [Ma et al. 2017]



**Figure 19.** QIS test chip featuring 20 1Mjot arrays (left), via [Deng and Fossum 2019].

Within a single-bit QIS, ADC acts as a simple comparator with an exposure-appropriate threshold  $q$ . Since the response is binary, the bit depth of only 1 is required for quantization.

#### 4.3 Multi-Bit QIS

The field readout rate of the sensor can be traded off for the increased bit depth of the ADC, thus allowing to count more photoelectrons per each readout cycle. Such implementation of the sensor, the multi-bit QIS, can be seen as the middle ground between the conventional CIS and the single-bit photon-counting sensor [Yin et al. 2021]. The carrier numbers, which were previously reduced to single-bit values, can now be mapped to their respective quantities in

a certain range. Upon saturation, the maximum number of the converter is reported instead [Fossum 2013].

Through lowered framerates and increased electron resolution, the output data transfer rate can be lowered. The increase in complexity of the ADC is negligible in comparison with the lowered complexity of the calculations; since the represented number of carriers is defined to be in the range of  $1\dots 200e^-$ , the resolution of the ADC does not need to be overly high, presenting significant improvements with the bit depth of  $n = 6$ . Reduction of the framerate for a multi-bit QIS occurs by the factor  $2^n - 1$ ; assuming the framerate of 1000 fps for single-bit QIS, the framerate of a multi-bit QIS will amount to  $\frac{1000\text{fps}}{2^{n-1}}$ , ensuring reduction by the factor of 63 in case with the 6b ADC.

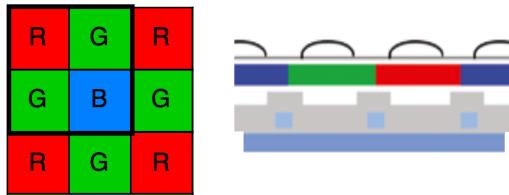
Another advantage provided by the bit depth selection is the ability to adjust linearity and compression between the input photon flux and the output signal.

The prototype in Figure 19 features 20 1Mjot arrays, allowing easy and flexible scaling in overall resolution due to the cluster-parallel architecture mentioned earlier. However, with each integrated arrays, the power dissipation is multiplied accordingly. The choice of ADC architecture for multi-bit QIS in terms of power consumption has been evaluated in [Yin et al. 2021], where the successive approximation register ADCs were deemed too be the least power-consuming option in relation to the occupied area.

**4.3.1 Color Imaging.** Spectral information in digital sensors is typically obtained using *color filter arrays* (CFAs), semi-transparent masks integrated on top of the sensor. These form a 2D periodical pattern with the three primary colors. Thus, each pixel is limited to irradiance related to the wavelength range of the respective color.

Later, the intensity values in each 2D kernel are reconstructed and interpolated, or *demosaicked*, to form a single pixel with three colour channels.

The most widely used CFA configuration is Bayer RGB pattern, though many other arrangements exist [Krig 2014]. Figure 20 shows the Bayer pattern and its exemplary position on a sensor plane.



**Figure 20.** Bayer RGB pattern and the common integrated CFA arrangement (right, via [Krig 2014]).

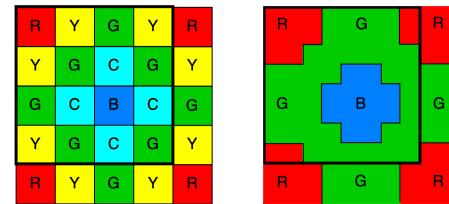
Following problems pertain through inclusion of color information via CFA:

- **Aliasing.** Through demosaicking, combined spatial resolution is lowered, and under sampling rates below the Nyquist frequency, *Moiré artifacts* appear.
- **Reduced sensor sensitivity.** Part of light is blocked out and never reaches the sensor plane, and the net signal is more susceptible to noise.
- **Crosstalk.** The optical or electrical charge in the pixels may leak towards the adjacent pixels. In CFAs, where charges depend on wavelengths and change rapidly, this results in a desaturation of the color measurements.

While these problems appear in monochromatic sensors as well, here, the errors persisting translate to false color values, which is more disturbing for the human perception [Hunt 2005].

Not all effects are equally pronounced, however, attempts to alleviate one factor may result in higher impact of the others. For one, QIS sensors are unaffected by aliasing due to their small feature sizes, while conventional CMOS sensors may require designated algorithmic suppression. At the same time, smaller pixel pitches lead to severe crosstalk; even more so when the adjacent colors are different [Elgendi and Chan 2019], which is the case with the Bayer pattern. In binary sensors, the leakage means false positives, resulting in a substantially lower SNR.

[Anzagira and Fossum 2015] attempt to mitigate the crosstalk through the extensions of Bayer pattern.



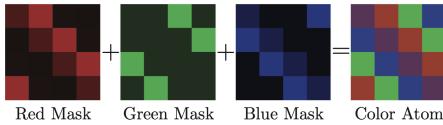
**Figure 21.** CFA Proposals made in [Anzagira and Fossum 2015].

Figure 21 depicts two of the proposals. On the left, the initial mosaic is expanded through secondary colors cyan, magenta, and yellow. These are positioned between the respective primary colours so that the spectral overlap is smaller. For increased light sensitivity, some of the green pixels may also be replaced with white. The second proposal expands the active color regions of the filter, leveraging the high spatial resolution of the QIS. The overlaps now occur in the regions where the pixels are halved, which can be accounted for during computations.

[Elgendi and Chan 2019] focus on more than merely crosstalk mitigation, proposing design criteria for sub-diffraction limit CFAs. Paired with a demosaicing algorithm, this should

maximize spatial resolution between the channels while improving the sensitivity in general. Similarly, the demosaicking process should minimize the noise power after the light passes the filter.

Presented in Figure 22, the filter has a diagonal structure. The auxiliary demosaicking algorithm works on frequency selection and was found to be universally applicable, performing well in both CIS and QIS imagers.



**Figure 22.** CFA pattern proposed in [Elgendi and Chan 2019]. The patterns are positioned diagonally, with gradations, as to reduce the cross-talk between the pixels.

In both cases, the filter kernel is larger, and the spatial resolution of the sensor is traded off for better performance in presence of crosstalk.



**Figure 23.** Quanta imaging sensor, as seen in a Gigajot camera. Via: [Gnanasambandam et al. 2019]

The strictly algorithmic processing of the color image in QIS has been described and applied in [Gnanasambandam et al. 2019], who present the QIS Pathfinder camera module developed at Gigajot Technology with a Bayer pattern color filter array implemented on-chip.

#### 4.4 Compared Performance

Table 1 shows some of the most significant operating parameters for the two photon-counting sensor implementations. We have collected these specifications based on the state-of-the-art implementations of the devices reported in the references section, taking the values from the most recently published sources in case of ambiguities.

It is evident that while SPAD-based arrays have higher dark current through thermally-induced avalanche gain, their read noise is still rather low, and the framerate is higher than the one in QIS by several orders of magnitude. The reasons for this behavior, i.e. the amount of circuitry required to

**Table 1.** Comparison of QIS and SPADs

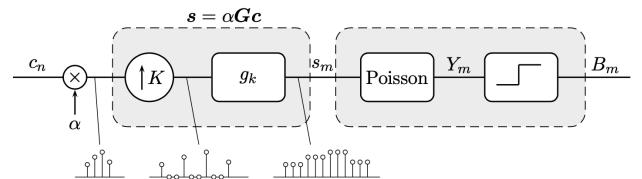
Parameters	QIS	SPAD
Pixel pitch	$1.1\mu\text{m}$	$5\text{-}10\mu\text{m}$
Read noise [RT]	$0.21e^-$ r.m.s.	$<0.15e^-$ r.m.s.
Dark current [RT]	$0.16e^-$ r.m.s.	$>10e^-$ r.m.s.
Framerate	1040 fps	97 000 fps
Quantum efficiency	70-80%	30-50%
Fill factor	>90%	<70%
Full well capacity	$1\text{-}200e^-$	N/A
Operating voltage	2.5/3.3V	15-25V

implement quenching, were mentioned in detail in previous sections. The read noise in QIS is nearing the ideal, relatively error-free value of  $0.15e^-$  r.m.s., while the dark current is completely negligible.

The contrast in approaches is noticeable in almost every other specification. Compared to QIS, SPADs need high electrical voltage to accelerate the photoelectrons, while providing lesser quantum efficiency. Such differences provide a clear distinction between the applications of both photon-counting sensors: SPAD-based imaging is significantly more functional in applications that value acquisition rate above all and need better temporal resolution, e.g. Time-of-Flight imaging, while QIS is better suitable for applications that prioritize accuracy and low-light sensitivity. Low-light imaging applications of QIS may include astronomy, surveillance, medical examinations, and many others. With sufficiently large oversampling, the single-bit QIS can have substantially higher dynamic ranges, being able to acquire scenes containing both bright and dark regions [Yang et al. 2012] and benefitting from the non-linear binary curve. According to [Fossum 2020], about 73 CMOS QIS pixels can fit into the area of one SPAD pixel, rendering the technology especially desirable in microbiology purposes.

## 5 Image Formation in a QIS

The image formation process in a QIS imaging sensor consists of two fundamental processing blocks and can be seen in Figure 24. For reasons of simplicity, we present the model based on the single-bit sensor.



**Figure 24.** QIS image formation model, via [Chan et al. 2016].

The first block comprises operations necessary to perform spatial oversampling of the diffraction-limited light field. In the second block, the upsampled light undergoes the binary sensing, where the exposure values are filled in using the one-bit quantized Poisson statistics.

It should be noted that the temporal oversampling is omitted in this particular figure, though it is foreseen by definition.

### 5.1 Spatial Oversampling of the Light Field

Assuming  $N$  discrete spatial coordinates, i.e. the amount of observable pixels per unit space, the light intensities are a vector of coefficients in discrete space  $\mathbf{c} = [c_0, \dots, c_{N-1}]^T$ . For convenience, we consider the variables to be acquired in a one-dimensional space, i.e., a vector  $\mathbf{c}_{N \times 1}$ .

The signal is assumed to be normalized to prevent scaling ambiguity such that  $c_n \in [0, 1]$ . In order to achieve the proper range scaling for the observations, the inputs are upscaled via multiplication with the factor  $\alpha > 0$ , representing the *sensor gain*. Afterwards, the existing data is ready to be spatially upsampled with the specified degrees of freedom needed for the modelling.

Depending on the total amount of jots  $M$ , every observation  $c_n$  is multiplied by the *spatial oversampling factor*  $K \triangleq \frac{M}{N}$ ; every diffraction-limited pixel is subsequently spread outwards. It is per definition that the  $K \in \mathbb{Z}^+$ .

The missing intensity values are filled in via lowpass filtering with a discrete, non-negative kernel  $\mathbf{g}$ ; a common choice for an interpolation kernel is the box-car function, which is defined such that its coefficients are

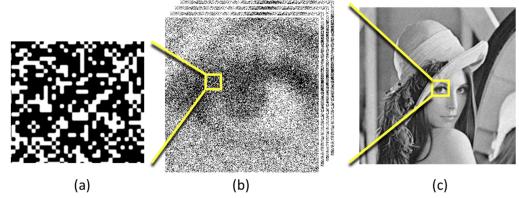
$$g_k = \frac{1}{k} \text{ for } k = 0, 1, \dots, K - 1. \quad (2)$$

Although other lowpass filtering operators are available ([Yang et al. 2012] name cardinal B-splines and squared sinc function), boxcar function is assumed to be the most accurate emulation of the oversampled light field: due to the microlenticular lenses placed in front of the jots in most functioning prototypes, it can be safely assumed that the light is focused on each jot.[Elgendi and Chan 2018]

Assuming the sampling is performed using the boxcar filtering function, one can describe the entirety of the oversampling process as a sequence of linear operations; the aforementioned operations can therefore be concatenated and rewritten in the form of  $\mathbf{s} = \alpha \mathbf{G}_c$ , corresponding to a matrix-vector multiplication with the filter response function  $\mathbf{G}_c$  of the interpolation filter in the Fourier domain.

### 5.2 Binary Sensing

The oversampled light field, already provided in discrete form, has to be approximated further by simulating the light fluctuation throughout the jot device and the subsequent thresholding.



**Figure 25.** Oversampled light field simulation. Via: [Fossum et al. 2016]

The fluctuations, i.e. the photon shot noise, adhere to the Poisson distribution. In measurements over an arbitrary period of time  $t$ , the probability of obtaining  $y_m$  photoelectrons is based on exposure  $s_m$  like

$$\mathbb{P}\{Y_m = y_m; s_m\} = \frac{s_m^{y_m} e^{-s_m}}{y_m!} \quad (3)$$

The quantization process is a mapping of detected photoelectrons into the binary range, based on the non-negative threshold  $q$ :  $B : \mathbb{Z}^+ \cup \{0\} \mapsto \{0, 1\}$ .

$$B(y) \triangleq \begin{cases} 1 & \text{for } y \geq q \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

For a threshold of  $q = 1$ , the probabilities of no photon arrivals are therefore estimated as

$$\mathbb{P}\{B_m = 0; s_m\} = e^{-s_m} \quad (5)$$

leaving us with  $\mathbb{P}\{B_m = 1; s_m\} = 1 - e^{-s_m}$  otherwise.

### 5.3 Error Rates

In the presented model, only natural photon shot noise is considered [Chan et al. 2016; Yang et al. 2012]. Generally, the signal transfer chain is much closer to the equation

$$\mathbf{x}_{\text{QIS}} = \text{ADC}\{\text{Poisson}(\alpha \cdot \mathbf{x}_{\text{true}} + \boldsymbol{\eta}_{\text{dark}})) + \boldsymbol{\eta}_{\text{read}}\} \quad (6)$$

and the photon noise is less dominant under sparse exposures.

Generally, there is also read noise, introduced by the readout process through the readout circuit transistors (resulting in flicker noise,  $1/f$ ) and reset on capacitors. These are critical in low-light imaging conditions, where the number of photoelectrons is sparse. The readout noise is often referred to the equivalent number of input photoelectrons.[Fossum 2016].

In a normalized uncorrupted signal

$$U_{\text{sig}} \triangleq \frac{V_{\text{sig}}}{CG} \quad (7)$$

the value is only dependant on the number of photoelectrons readout, i.e. the *electron number*, and is Poisson distributed. There is equivalence between the interpretation of the electron number and the ratio of the voltage to conversion gain  $CG$ .

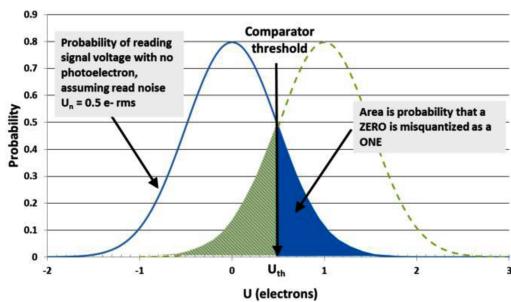
However, assuming additional read noise, normalized as

$$u_n \triangleq \frac{v_n}{CG} \quad (8)$$

the probability distribution function depends not just on obtaining a certain amount of photoelectrons, but the noise  $u_n$  as well. Since the read noise is best approximated through Gaussian, the total probability is a convolution of the Poisson distribution for quanta exposure  $H$  (equivalent to  $s_m$  in previous chapters) and a normal distribution with read noise  $u_n$ .

$$\mathbb{P}[U] = \sum_{k=0}^{\infty} \frac{\mathbb{P}[k]}{\sqrt{2\pi u_n^2}} \exp \left[ -\frac{(U - k)^2}{2u_n^2} \right] \quad (9)$$

When  $U_{sig}$  is quantized with a 1b analog-to-digital-converter, it is subjected to a comparison with a threshold  $U_{th}$ , commonly 0.5 [Fossum 2013]. With fluctuations caused by read noise, the binary response is prone to errors. We are presented with two options: either there are no electrons, but the 0 is misquantized as 1 due to the added noise, or there is a positive response of 1 misquantized as 0. The sum of these false positives and false negatives in ratio to total number of jots is the *Bit Error Rate* (BER).



**Figure 26.** Bit Error Rate in single-bit sensors, via [Fossum 2013].

In Figure 26, the misquantizations are shown for the single-bit case with a threshold exactly in the middle. There are only two possible states, and for each state, there's a possibility of an error. The false estimations are formed in the overlapping section.

The false negatives and false positives are obtained through the integration of the probability distribution function, which leads us to expressing these as an error function:

$$P_{fn} = \frac{1}{2} \operatorname{erfc} \left[ \frac{1 - U_{th}}{\sqrt{2}U_n} \right] \quad (10)$$

and

$$P_{fp} = \frac{1}{2} \operatorname{erfc} \left[ \frac{1}{\sqrt{8}U_n} \right] \quad (11)$$

When the threshold is put exactly in the middle, the false positives and false negatives are equally probable. The total bit error rate with the threshold of  $U_{th} = 0.5$  amounts to

$$BER = \frac{1}{2} \operatorname{erfc} \left[ \frac{1}{\sqrt{8}U_n} \right] \quad (12)$$

and is only acceptably small with read noise of  $0.30e^-$  r.m.s. and below.

## 6 Conclusion

Commercial realization still needs to overcome several technical challenges to fulfill demands listed in previous chapters. The pixel pitch is far from the ideally theorised pitch of 200...500 nm range, although it should be said that for existing dimensions, some of the requirements (e.g. low read noise) have already been fulfilled. Similarly, existing chips contain multiple 1Mjot chips, which, while combined together, will increase the resolution, but also result in higher power dissipation.

The artistic use of the cameras may suffer some limitations regarding integration of further subsystems: for instance, the use of the synchronized flash devices would mean that nearly each pixel would simultaneously detect a photon, rendering the estimation impossible. Similar effect could theoretically be observed in auxilliary optical elements like polarizing filters, which would per definition cut the amount of incident light in half. Additionally, the gigajot device may not be ripe for the mainstream market just yet due to the low support of conventional color imaging. However, the problem of color imaging capacities in QIS has been adressed in detail in the recent years, concentrating on denoising the data in low-light, or photon-starved, regime [Elgendi and Chan 2019; Gnanasambandam and Chan 2020b] – with further progress made, one may be able to accomodate both sparse and overly present illumination, especially via multi-bit QIS.

Founded under the aegis of the IEEE community, the International Roadmap for Devices and Systems™ (<https://irds.ieee.org>) publishes annual reports on recent advancements and trends in the manufacturing of semiconductor devices, computing systems, and the electronic industry in general.

While the photon-counting devices have not been covered in 2020, the ‘‘Beyond CMOS’’ roadmap describes many emerging FET technologies that could minimize their power consumption, and states the need for the circuits that follow the probabilistic/stochastic computational paradigm. Likewise, AI implementations in the integrated circuits are an active area of research [Hao et al. 2021] and might alleviate the computational expenses of quanta image post-processing altogether. [Elgendi and Chan 2018] denote the need for a threshold adjustment module and refer to per-pixel FPGAs, a project currently pursued by MIT Lincoln Lab.

In terms of hardware, main challenges of the [photonic] quanta imaging remain to be design of internal high-speed and low-power addressing and readout circuitry [Masoodian

et al. 2016]; for monitoring the current progress, we urge the readers to familiarize themselves with the roadmaps.

Many proposals for the jet device implementation and the conformance with the baseline CIS process should signify high probability of establishment of gigavision acquisition process in the future – at least in the scientific applications. As the feasibility of the QIS has been demonstrated in working prototypes, it can indeed become a third generation digital sensor with time, as previously envisioned.

## References

- Emna Amri, Yacine Felk, Damien Stucki, Jiaju Ma, and Eric R. Fossum. 2016. Quantum Random Number Generation Using a Quanta Image Sensor. *Sensors* 16, 7 (2016). <https://doi.org/10.3390/s16071002>
- Saeed Anwar, Salman Khan, and Nick Barnes. 2020. A Deep Journey into Super-resolution: A survey. *ACM Computing Surveys (CSUR)* 53 (2020), 1 – 34. arXiv:[1904.07523](https://arxiv.org/abs/1904.07523) [cs.CV]
- L. Anzagira and E. Fossum. 2015. Color filter array patterns for small-pixel image sensors with substantial cross talk. *Journal of the Optical Society of America. A, Optics, image science, and vision* 32 1 (2015), 28–34.
- Moshe Ben-Ezra. 2011. A Digital Gigapixel Large-Format Tile-Scan Camera. *IEEE Computer Graphics and Applications* 31, 01 (jan 2011), 49–61. <https://doi.org/10.1109/MCG.2011.1>
- D.J. Brady, M.E. Gehm, R.A. Stack, D.L. Marks, David Kittle, Dathon Golish, Esteban Vera, and Steven Feller. 2012. Multiscale gigapixel photography. *Nature* 486 (06 2012), 386–9. <https://doi.org/10.1038/nature11150>
- Benoit-Louis Bérubé, Vincent-Philippe Rhéaume, Samuel Parent, Luc Maurais, Audrey Corbeil Therrien, Paul G. Charette, Serge A. Charlebois, Réjean Fontaine, and Jean-François Pratte. 2015. Implementation Study of Single Photon Avalanche Diodes (SPAD) in 0.8  $\mu\text{m}$  HV CMOS Technology. *IEEE Transactions on Nuclear Science* 62, 3 (2015), 710–718. <https://doi.org/10.1109/TNS.2015.2424852>
- Michel Cagnet. 1962. *Atlas optischer Erscheinungen - Atlas of optical phenomena*. Springer, Berlin, Heidelberg.
- Burt Haring Carroll, George Clinton Higgins, and Thomas Howard James. 1980. *Introduction to Photographic Theory - The Silver Halide Process*. J. Wiley, New York.
- Jon Cartwright. 2018. Gigapixel camera pushes resolution limit. <https://physicsworld.com/a/gigapixel-camera-pushes-resolution-limit>
- G. F. Cerofolini and D. Mascolo. 2007. *A Hybrid Route from CMOS to Nano and Molecular Electronics*. Springer US, Boston, MA, 1–65. [https://doi.org/10.1007/978-0-387-49965-9\\_1](https://doi.org/10.1007/978-0-387-49965-9_1)
- Stanley H. Chan, Omar A. Elgendi, and Xiran Wang. 2016. Images from Bits: Non-Iterative Image Reconstruction for Quanta Image Sensors. *Sensors* 16, 11 (2016). <https://doi.org/10.3390/s16111961>
- Edoardo Charbon, Claudio Bruschini, and Myung-Jae Lee. 2018. 3D-Stacked CMOS SPAD Image Sensors: Technology and Applications. In *2018 25th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*. 1–4. <https://doi.org/10.1109/ICECS.2018.8617983>
- Oliver S. Cossairt, Daniel Miau, and Shree K. Nayar. 2011a. Gigapixel Computational Imaging. In *2011 IEEE International Conference on Computational Photography (ICCP)*. 1–8. <https://doi.org/10.1109/ICCPHOT.2011.5753115>
- Oliver S. Cossairt, Daniel Miau, and Shree K. Nayar. 2011b. Scaling law for computational imaging using spherical optics. *J. Opt. Soc. Am. A* 28, 12 (Dec 2011), 2540–2553. <https://doi.org/10.1364/JOSAA.28.002540>
- Jeff Cremer. 2012. Machu Picchu 16 Gigapixels. Retrieved July 1, 2021 from <https://vimeo.com/53364475>
- Wolfgang Demtröder. 2018. *Experimentalphysik 2 - Elektrizität und Optik*. Springer-Verlag, Berlin Heidelberg New York.
- Wei Deng and Eric R. Fossum. 2019. 1/f Noise Modelling and Characterization for CMOS Quanta Image Sensors. *Sensors* 19, 24 (2019). <https://doi.org/10.3390/s19245459>
- Neale A. W. Dutton, Istvan Gyongy, Luca Parmesan, Salvatore Gnechi, Neil Calder, Bruce R. Rae, Sara Pellegrini, Lindsay A. Grant, and Robert K. Henderson. 2016. A SPAD-Based QVGA Image Sensor for Single-Photon Counting and Quanta Imaging. *IEEE Transactions on Electron Devices* 63, 1 (2016), 189–196. <https://doi.org/10.1109/TED.2015.2464682>
- Omar A. Elgendi and Stanley H. Chan. 2018. Optimal Threshold Design for Quanta Image Sensor. *IEEE Transactions on Computational Imaging* 4, 1 (2018), 99–111. <https://doi.org/10.1109/TCI.2017.2781185>
- Omar A. Elgendi and Stanley H. Chan. 2019. Color Filter Arrays for Quanta Image Sensors. arXiv:[1903.09823](https://arxiv.org/abs/1903.09823) [cs.CV]
- GigaPan Systems | Panoramic Photography Equipment and Software Company. 2016. *What is a Gigapan?* Retrieved July 1, 2021 from <http://gigapan.com/cms/about-us>
- Eric R. Fossum. 2013. Modeling the Performance of Single-Bit and Multi-Bit Quanta Image Sensors. *IEEE Journal of the Electron Devices Society* 1, 9 (2013), 166–174. <https://doi.org/10.1109/JEDS.2013.2284054>
- Eric R. Fossum. 2016. Photon Counting Error Rates in Single-Bit and Multi-Bit Quanta Image Sensors. *IEEE Journal of the Electron Devices Society* 4, 3 (2016), 136–143. <https://doi.org/10.1109/JEDS.2016.2536722>
- Eric R. Fossum. 2018. CMOS Image Sensors and the Quanta Image Sensor. In *2018 International Conference on Optical MEMS and Nanophotonics (OMN)*. 1–5. <https://doi.org/10.1109/OMN.2018.8454568>
- Eric R. Fossum. 2020. The Invention of CMOS Image Sensors: A Camera in Every Pocket. In *2020 Pan Pacific Microelectronics Symposium (Pan Pacific)*. 1–6. <https://doi.org/10.23919/PanPacific48324.2020.9059308>
- Eric R Fossum, Jiaju Ma, Saleh Masoodian, Leo Anzagira, and Rachel Zizza. 2016. The quanta image sensor: Every photon counts. *Sensors* 16, 8 (2016), 1260.
- Abhiram Gnanasambandam and Stanley H. Chan. 2020a. HDR Imaging With Quanta Image Sensors: Theoretical Limits and Optimal Reconstruction. *IEEE Transactions on Computational Imaging* 6 (2020), 1571–1585. <https://doi.org/10.1109/tci.2020.3041093>
- Abhiram Gnanasambandam and Stanley H. Chan. 2020b. Image Classification in the Dark using Quanta Image Sensors. arXiv:[2006.02026](https://arxiv.org/abs/2006.02026) [eess.IV]
- Abhiram Gnanasambandam, Omar Elgendi, Jiaju Ma, and Stanley H. Chan. 2019. Megapixel photon-counting color imaging using quanta image sensor. *Optics Express* 27, 12 (Jun 2019), 17298. <https://doi.org/10.1364/oe.27.017298>
- Qian Gong, Esteban Vera, Dathon R. Golish, Steven D. Feller, David J. Brady, and Michael E. Gehm. 2017. Model-Based Multiscale Gigapixel Image Formation Pipeline on GPU. *IEEE Transactions on Computational Imaging* 3, 3 (2017), 493–502. <https://doi.org/10.1109/TCI.2016.2612942>
- Yue Hao, Shuiying Xiang, Genquan Han, Jincheng Zhang, Xiaohua Ma, Zhangming Zhu, Xingxing Guo, Yahui Zhang, Yanan Han, Ziwei Song, et al. 2021. Recent progress of integrated circuits and optoelectronic chips. *Science China Information Sciences* 64, 10 (2021), 1–33.
- Gerald C. Holst and Terrence S. Lomheim. 2011. *CMOS/CCD Sensors and Camera Systems - JCD Publishing*, Winter Park, FL.
- R. W. G. Hunt. 2005. *The Reproduction of Colour*. John Wiley Sons, New York.
- Ferdinand Hurter and Vero C. Driffield. 1890. *Photo-chemical investigations and a new method of determination of the sensitiveness of photographic plates*. Society of Chemical Industry, London.
- Eyke Jessien. 1974. *ABC der Fotochemie: Fachausdrücke in Wort und Bild* (2 ed.). Agfa-Gevaert AG, Leverkusen-Bayerwerk.
- Muhammad Jafni Yusof and Hj. Ruslan Abd. Rahim. 2014. Revealing visual details via high dynamic range gigapixels spherical panorama photography: The Tempurung Cave natural heritage site. In *2014 International Conference on Virtual Systems Multimedia (VSMM)*. 193–200. <https://doi.org/10.1109/VSMM.2014.7136690>
- Scott Krig. 2014. *Image Capture and Representation*. Apress, Berkeley, CA, 1–37. [https://doi.org/10.1007/978-1-4302-5930-5\\_1](https://doi.org/10.1007/978-1-4302-5930-5_1)

- Rod Lawton and James Artaius. 2021. The 12 highest resolution cameras you can buy today: ultimate pro cameras. <https://www.digitalcameraworld.com/buying-guides/the-10-highest-resolution-cameras-you-can-buy-today>
- Jiaju Ma and Eric R. Fossum. 2015a. A Pump-Gate Jot Device With High Conversion Gain for a Quanta Image Sensor. *IEEE Journal of the Electron Devices Society* 3, 2 (2015), 73–77. <https://doi.org/10.1109/JEDS.2015.2390491>
- Jiaju Ma and Eric R. Fossum. 2015b. Quanta Image Sensor Jot With Sub 0.3e-r.m.s. Read Noise and Photon Counting Capability. *IEEE Electron Device Letters* 36, 9 (2015), 926–928. <https://doi.org/10.1109/LED.2015.2456067>
- Jiaju Ma, Donald Hondonwga, and Eric R. Fossum. 2014. Jot devices and the Quanta Image Sensor. In *2014 IEEE International Electron Devices Meeting*, 10.1.1–10.1.4. <https://doi.org/10.1109/IEDM.2014.7047021>
- Jiaju Ma, Saleh Masoodian, Dakota A. Starkey, and Eric R. Fossum. 2017. Photon-number-resolving megapixel image sensor at room temperature without avalanche gain. *Optica* 4, 12 (Dec 2017), 1474–1481. <https://doi.org/10.1364/OPTICA.4.001474>
- Jiaju Ma, Dakota Staskey, Arun Rao, Kofi Odame, and Eric R. Fossum. 2015. Characterization of Quanta Image Sensor Pump-Gate Jots With Deep Sub-Electron Read Noise. *IEEE Journal of the Electron Devices Society* 3, 6 (2015), 472–480. <https://doi.org/10.1109/JEDS.2015.2480767>
- Saleh Masoodian, Arun Rao, Jiaju Ma, Kofi Odame, and Eric R. Fossum. 2016. A 2.5 pJ/b Binary Image Sensor as a Pathfinder for Quanta Image Sensors. *IEEE Transactions on Electron Devices* 63, 1 (2016), 100–105. <https://doi.org/10.1109/TED.2015.2457418>
- F. Pedrotti, L. Pedrotti, W. Bausch, and Hartmut Schmidt. 2007. *Optik für Ingenieure - Grundlagen*. Springer Berlin Heidelberg, Wiesbaden.
- Duke Information Spaces Project. [n.d.]. AWARE2 Multiscale Gigapixel Camera. <https://disp.duke.edu/research/aware2-multiscale-gigapixel-camera>
- A. Rochas, M. Gosch, A. Serov, P.A. Besse, R.S. Popovic, T. Lasser, and R. Rigler. 2003. First fully integrated 2-D array of single-photon detectors in standard CMOS technology. *IEEE Photonics Technology Letters* 15, 7 (2003), 963–965. <https://doi.org/10.1109/LPT.2003.813387>
- Bahaa E. A. Saleh and Malvin Carl Teich. 1991. *Fundamentals of Photonics*. Wiley, New York.
- Elham Sarbazi, Majid Safari, and Harald Haas. 2015. Photon detection characteristics and error performance of SPAD array optical receivers. In *2015 4th International Workshop on Optical Wireless Communications (IWOW)*, 132–136. <https://doi.org/10.1109/IWOW.2015.7342281>
- Gottfried Schröder. 1981. *Technische Fotografie: Grundlagen und Anwendungen in Technik und Wissenschaft* (1 ed.). Vogel, Würzburg.
- Andrew M. Shackleton and Abdulrahman M. Altahhan. 2019. A Comparison Study of Deep Learning Techniques to Increase the Spatial Resolution of Photo-Realistic Images. In *Neural Information Processing*, Tom Gedeon, Kok Wai Wong, and Minho Lee (Eds.). Springer International Publishing, Cham, 341–348.
- Ryung Shin, Woojae Choi, Taekyung Kim, Donghyun Kim, Runjia Han, Kyungjin Lee, Nahyun Won, and Shinill Kang. 2020. Gigapixel confocal imaging using a massively parallel optical probe array with single directional infinite scanning. *Scientific Reports* 10, 1 (1 Dec. 2020). <https://doi.org/10.1038/s41598-020-64602-3> Funding Information: This work was supported by National Research Foundation of Korea (NRF) Grant funded by the Korean Government (MSIT) (No. 2015R1A5A1037668). Publisher Copyright: © 2020, The Author(s).
- Feng Yang, Y. M. Lu, L. Sbaiz, and M. Vetterli. 2012. Bits From Photons: Oversampled Image Acquisition Using Binary Poisson Statistics. *IEEE Transactions on Image Processing* 21, 4 (Apr 2012), 1421–1436. <https://doi.org/10.1109/tip.2011.2179306>
- Zhaoyang Yin, Yibing M. Wang, and Eric R. Fossum. 2021. Low Bit-Depth ADCs for Multi-bit Quanta Image Sensors. *IEEE Journal of Solid-State Circuits* 56, 3 (2021), 950–960. <https://doi.org/10.1109/JSSC.2020.3045430>
- Linwei Yue, Huanfeng Shen, Jie Li, Qiangqiang Yuan, Hongyan Zhang, and Liangpei Zhang. 2016. Image super-resolution: The techniques, applications, and future. *Signal Processing* 128 (05 2016). <https://doi.org/10.1016/j.sigpro.2016.05.002>
- Guoan Zheng, Xiaoze Ou, and Changhuei Yang. 2012. Towards Giga-pixel Microscopy, In Conference on Lasers and Electro-Optics 2012. *Conference on Lasers and Electro-Optics 2012*, CTu3J.1. [https://doi.org/10.1364/CLEO\\_SI.2012.CTu3J.1](https://doi.org/10.1364/CLEO_SI.2012.CTu3J.1)
- Guoan Zheng, Xiaoze Ou, and Changhuei Yang. 2014. 0.5 gigapixel microscopy using a flatbed scanner. *Biomedical optics express* 5, 1 (2014), 1–8.

## A Online Resources

This work is accompanied by a presentation which additionally demonstrates the image formation model of QIS and the following reconstruction. We generate synthetic data in accordance with the methods listed in Section 5 using grayscale images, provided by the USC-SIPI Image Database (<http://sipi.usc.edu/database/>).

Image reconstruction with variable threshold sizes has been covered by [Elgendi and Chan 2018], who provide executable MATLAB code on their page: [https://engineering.purdue.edu/ChanGroup/project\\_QIS.html](https://engineering.purdue.edu/ChanGroup/project_QIS.html). The code is copyrighted by 2017 Statistical Signal and Image Processing Lab, Purdue University, with granted permission to use, copy, modify, and distribute this software and its documentation for educational, research and non-profit purposes, without fee, and without a written agreement.