High Dynamic Range Imaging at the Quantum Limit with SPAD-based Image Sensors

sensors

MDPI

Article

High Dynamic Range Imaging at the Quantum Limit with SPAD-based Image Sensors

Neale A.W. Dutton [1,*], Tarek Al Abbas [2], Istvan Gyongy [2], Francescopaolo Mattioli Della Rocca [1,2] and Robert K. Henderson [2]

[1] STMicroelectronics Imaging Division, Tanfield, Edinburgh, UK.
[2] School of Engineering, The University of Edinburgh, Edinburgh, UK.
* Correspondence: neale.dutton@st.com; Tel.: +44-131-558-4170

Academic Editor: name
Received: date; Accepted: date; Published: date

Abstract: This paper examines methods to best exploit the High Dynamic Range (HDR) of the SPAD in a high fill-factor HDR photon counting pixel that is scalable to megapixel arrays. The proposed method combines multi-exposure HDR with temporal oversampling in-pixel. We present a silicon demonstration IC with 96x40 array of 8.25µm pitch 66% fill-factor SPAD-based pixels achieving >100dB dynamic range with 3 back-to-back exposures (short, mid, long). Each pixel sums 15 bit-planes or binary field images internally to constitute one frame providing 3.75x data compression, hence the 1k frames per second (FPS) output off-chip represents 45,000 individual field images per second on chip. Two future projections of this work are described: scaling SPAD-based image sensors to HDR 1MPixel formats and shrinking the pixel pitch to 1-3µm.

Keywords: single photon avalanche diode; SPAD; high dynamic range; HDR; CMOS image sensor; CIS; single photon counting; SPC; HDR SPC; quanta image sensor; QIS; spatio-temporal oversampling

1. Introduction

Single photon sensitive image sensors offer the ultimate photo-sensitivity to a wide range of applications such as machine vision, scientific imaging, space, defense, and film cameras [1 – 5]. In all these use cases, high resolution imaging with both high sensitivity and wide dynamic range are required [4]. A range of pixel technologies serve these demands but have their individual limitations. Solid-state Electron Multiplied CCD (EMCCD) and non- solid state solutions such as photo-cathode based Intensified CCD (ICCD), and Intensified CMOS (ICMOS) either offer high sensitivity or Dynamic Range (DR) but not both, given the fundamental noise-floor and head-room limitations of the intensification process combined with the analogue signal chain as illustrated in Fig.1.(b) [6]. Furthermore, none of these technologies directly detects the incident single photons: instead indirectly sensing the electrons from the multiplication or intensification process (which itself brings a range of noise sources).

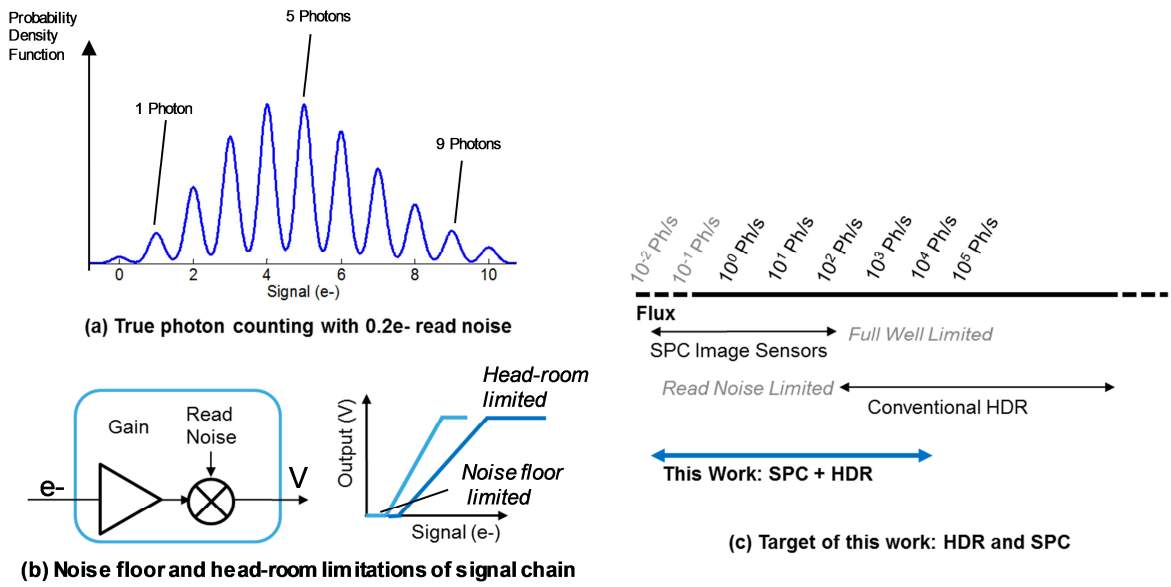

Figure 1. (a) Single photon counting is realised in DSERN pixels with <0.30e- RN. Example is 0.2e- RN with clearly visible peaks from counted photons during an integration time (b) Image sensor analogue signal chains impose both head-room and noise floor limitations on dynamic range (c) Target of this work is high dynamic range in a photon counting regime.

Truly counting single photons with >90% certainty is realised below 0.3e- read noise (RN), with an example Deep Sub Electron Read Noise (DSERN) pixel response illustrated in Fig.1.(a) [6]. There are two approaches offering promising solutions for high resolution photon counting image sensors. Firstly, CMOS image sensor (CIS) technology with active pixel sensor readout (APS) has been recently employed in DSERN CIS pixels offering <0.30e- RN with no electron multiplication process [7][8]. Secondly and the focus of this work, CMOS SPADs offer single photon counting either using analogue techniques or digital logic. However, both CIS and SPAD pixels encounter the same fundamental trade-off between pixel size, dynamic range and sensitivity. Recent research into spatio-temporal oversampling of photon-counting image sensors overcomes the tradeoff of pixel size to full-well or maximum photon count through a range of different techniques [4][6][9].

This paper explores this trade-off using CMOS SPADs and complementary pixel circuits, in an advanced deep sub-micron (DSM) imaging CMOS technology [10], with temporal oversampling to achieve both Single Photon Counting (SPC) and High Dynamic Range (HDR) as shown in Fig. 1.(c) with a scalable and compact pixel architecture capable of realising megapixel imaging arrays in the future. We expand on our original works [1][2] with greater detail and further measurements.

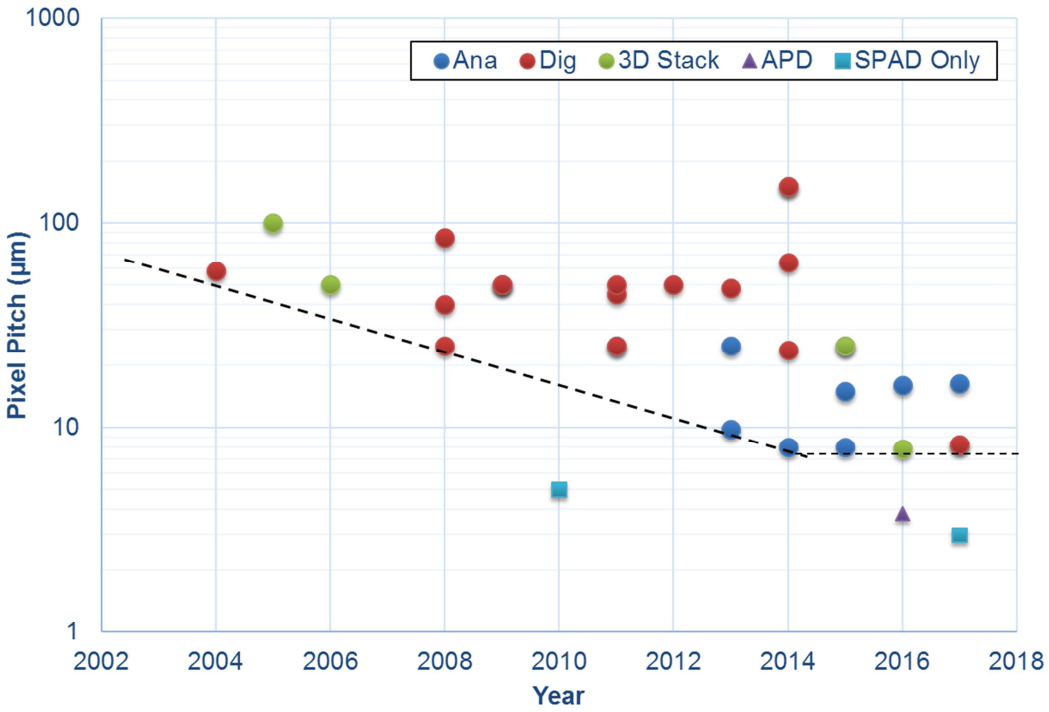

Figure 2. Pixel pitch compared by year of publication for a variety of CMOS SPAD image sensor pixel architectures (analogue or digital monolithic, and all–digital 3D stacked) with a dotted line indicating the trend on reducing pixel pitch now limited at 7.83 µm [11]. Three notable exceptions sit outside this trend, two SPAD diode test structures (SPAD only) at 5 µm [12] and 3 µm pitch [13] and 3.8 µm APD pixel [14].

2. Background

To achieve high resolution megapixel arrays using avalanche based pixels, the pixel pitch must be sufficiently small and competitive with the state of the art. For comparison, the leading example of pixel pitch for a CIS SPC pixel is 1.1 micron with 0.22e- DSERN with a full well in the region of 200e- in 3D stacked implementation [15]. In contrast avalanche-based pixels are larger for two primary reasons: the pixel circuit is more complex and the APD or SPAD structure itself does not scale readily with technology node. Addressing the latter, scaling down the diode structure is the first pixel design challenge as the device structure requires careful design of the planar high electric field region and guard ring regions providing a transition zone between high and low field regions. Recent examples can be used to illustrate the scaling of APDs and SPADs to achieve compact pixels. Figure 2 shows the recent chronological trend in decreasing pitch of avalanche-based pixels. The pixels are compared for monolithic designs where the majority of smaller pixels are based on analogue circuits due to fewer transistors. This work, [2] and [11] are the first to employ advanced 40nm CMOS to reduce the pixel pitch of digital photon counting pixels. The black dotted line indicates the trend of pitch reduction. Three data points sit outside of the trend: two lead the field for SPAD pitch (without image sensor pixel circuits) at 5µm [12] and 3µm [13], whilst [14] is the first and a remarkable example of a high resolution APD back-illuminated image sensor at 3.8µm pitch, although with a full-well of only 1 photon. The latter reports two modes of operation: 40dB dynamic range single photon mode and 60dB dynamic range CIS mode [14]. Yet neither mode has inherently high dynamic range for SPC.

The CMOS SPAD is an ideal detector for HDR SPC as it has an intrinsic DR greater than 100dB: capturing photon flux with count rates from ~1 count/second to >10M count/second for passive recharge (~140dB dynamic range) and >100M count/second for active recharge (~160dB dynamic range). Until recently no SPAD pixel designs, to the knowledge of the authors, had fully utilised the intrinsic HDR of the photo-detector; a first implementation of time-gated pixel with HDR photon counting is very recently published [16].

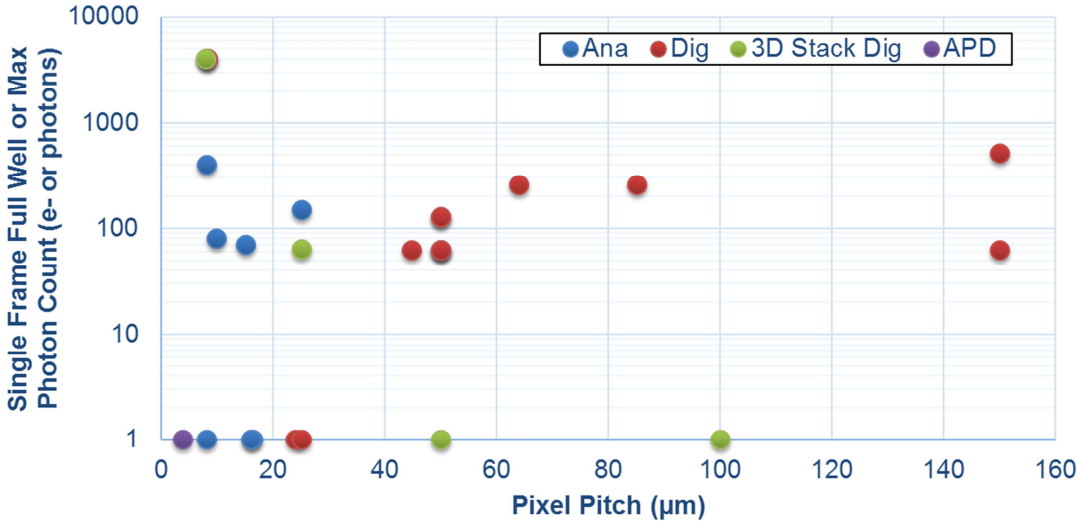

Figure 3. Maximum photon count of SPAD pixels grouped by architecture in comparison to the pixel area.

The pulsing output of the SPAD (or current avalanche of an APD) poses the second pixel design challenge: how to count these pulses in a compact pixel pitch whilst attaining this maximum dynamic range. The three main SPAD image sensor pixel architectures may be considered in relation to the problem: digital counter, analogue counter and 1-bit memory. The simplest architecture is the all-digital ripple counter which is well explored in the literature [17][18]. The counter bit depth is proportional to pixel area and scales readily with CMOS technology node. Alternatively analogue counters (either based on switched current sources [19][20] or charge transfer amplifiers[21][22]) can realise approximately 100 counts in a compact form with reasonable PRNU. The main limitation to increasing the maximum count is the addition of the noise of the analogue signal chain [6]. Furthermore, like all precision analogue circuits in deep sub-micron (DSM) CMOS it does not scale easily with technology node. The 1-bit memory, based on dynamic [23] or static [24] memory structures are the same size as the analogue counter, but record only a single SPAD count so has the lowest maximum count of the three architectures. Yet, it scales much more readily with DSM CMOS technology scaling. Fig. 3 highlights the maximum count in comparison to pixel pitch of the three architectures. The most promising in the context of HDR SPC is the digital ripple counter for scaling and functionality.

To overcome the dynamic range limitations of the maximum count of the pixel, two techniques can be combined: HDR imaging and oversampling. Dynamic range enhancement for CIS is well known for over 20 years [25][26]. The most applicable HDR technique is multiple in-pixel memories (or storage nodes) with independent global shuttering providing the benefit of capturing HDR images simultaneously (without multiple sequential exposures) with reduced motion artefacts, for example, for suppression of LED or indoor lighting flicker [27]. Fossum describes HDR oversampling of photon counting image sensors in his theoretical paper on the Quanta Image Sensor (QIS) [4]. Individual binary images (referred to as 'bit planes' or 'field images') are captured for multiple exposure times then oversampled temporally or spatially to form a HDR frame image. Employing these two concepts together is the basis of our silicon test chip trialing a pixel design capable of HDR. While the sensor used in this work has a limited resolution (96×40), its 3D-stacked counterpart [11] is scalable to megapixel arrays and allows pitch reduction with the progress in decreasing technology nodes.

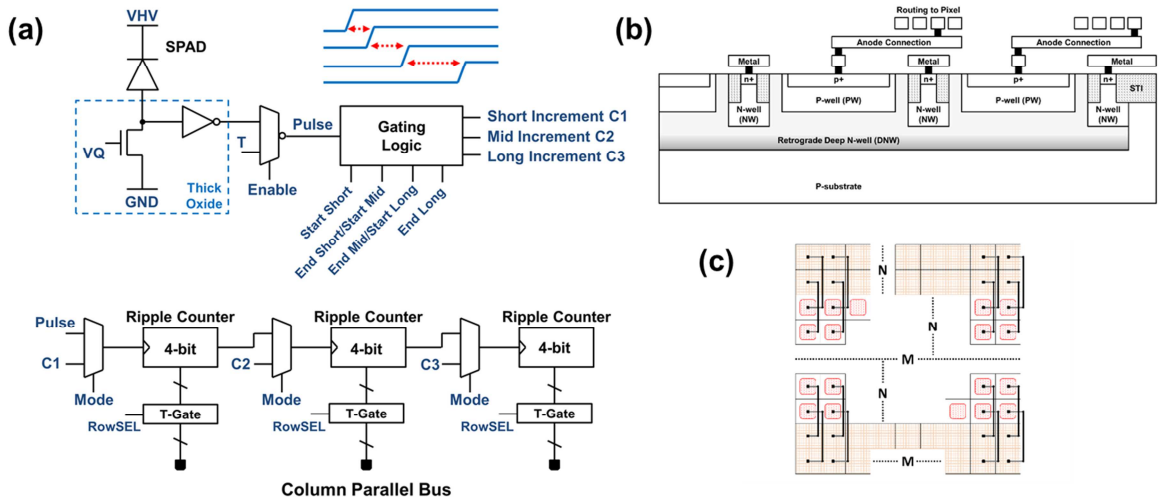

Figure 4. (a) Pixel Schematic with dual modes of linear 12b counter or three 4b counters with individual exposure controls (default mode in this work). (b) Cross-section of SPADs shared in a global well (c) Top view of SPADs in global well with pixel circuits placed at the edge.

3. Silicon Design

This section describes the design and architecture of our demonstrator IC for HDR SPC in FSI technology. The pixel schematic is shown in Fig. 4.(a). It consists of a SPAD with a single passive quench passive recharge (PQPR) NMOS transistor, 4 time gating front-end D-type flip flops (in the gating logic block) and 12 toggle flip-flops configurable either as 12b ripple counter for linear counting mode or three individual 4b ripple counters for SPC HDR mode (the default configuration in this work). In linear counting mode, only the first time-gate flip-flop is employed. For HDR mode, the four time-gating D-type flip flops are positive edge triggered where the time-gate integration window is between the rising edges of two gating signals. This provides three contiguous exposure windows 'short', 'mid' and 'long' each with independent in-pixel capture and storage. The time-gating technique is described in [2] and [11] and provides zero loss of sensitivity around the transition of the three time-gate windows. To provide good matching between the time-gating signals, each is routed through a clock tree at the edge of the array with a driver and line per column.

As described in detail in [2] and shown in Fig.4.(b) in this monolithic implementation the 8.25 µm x 8.25 µm SPAD structure has a p-well (PW) to deep n-well (DNW) junction with retrograde guard ring. The cathode is a shared global DNW permitting 66% fill factor of the anode (considering only the imaging array). The anode is routed to the matching pixel circuit at the edge of the array. The pixel circuits are placed outside of the imaging array and are pitch matched at 8.25 µm x 8.25 µm. In this manner, the pixel circuit is ready for a future 3D-stacked implementation [11].

The integration is global shutter and the rolling all-digital readout is through conventional row-wise timing. Top and bottom readout is employed and the data for each row is sequentially serialised, and each array-half is transmitted off-chip by a single I/O pad at 21.97Mb/s rate at 1kFPS. Here we define a frame as the full 12b data per pixel whether it represents a single 12b linear mode exposure or an in-pixel summation of 15 binary fields at 3 different exposures (4b per exposure) in HDR mode. The data and frame rates are kept intentionally moderate, to understand how this architecture scales as a building block to very high resolution arrays where data-rate and power will be primary limiting factors.

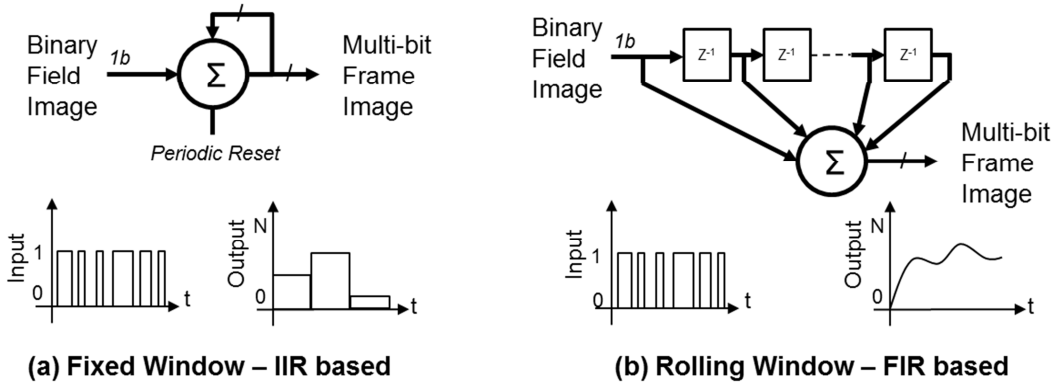

Figure 5. The two primary methods of temporally summing single photon bit planes: (a) Fixed averaging based on an IIR filter with periodic reset (i.e. a simple counter). (b) Rolling averaging based on FIR filter.

In our previous research [6][9][23][28] into oversampled photon counting, a single bit represents the detection of a photon (the image from the sensor is referred to as a field image or a bit-plane); however, this unary encoding of photon counting is not a power efficient method of data transmission. To address this limitation, some degree of in-pixel summing provides data compression and a power saving. Fig. 5 describes the two primary methods of temporally summing bit-planes that can be employed: fixed time window integration (achieved by infinite impulse response (IIR) type filter but periodically reset) [23] and rolling window (finite impulse response (FIR) filter based) averaging [9]. The downside to fixed window summing is the loss of temporal resolution and output frame rate whereas the FIR rolling average provides temporal resolution at the input bit-plane frame rate but comes with higher power, data rate and area costs. Fixed summing in-pixel is easily implemented, and provides data compression. It is clear that a trade-off is made of data compression versus temporal resolution and frame rate. Furthermore, this problem is intensified when implementing HDR with multiple exposures. In this work, a compromise is chosen to sum up to 15 bit planes in pixel for each of the HDR exposures using each 4b counter. In effect this is a 3.75 times data compression and power saving (15 unary bits to 4 binary bits), at the cost of a 15 times reduction in temporal resolution by fixed IIR filter summing.

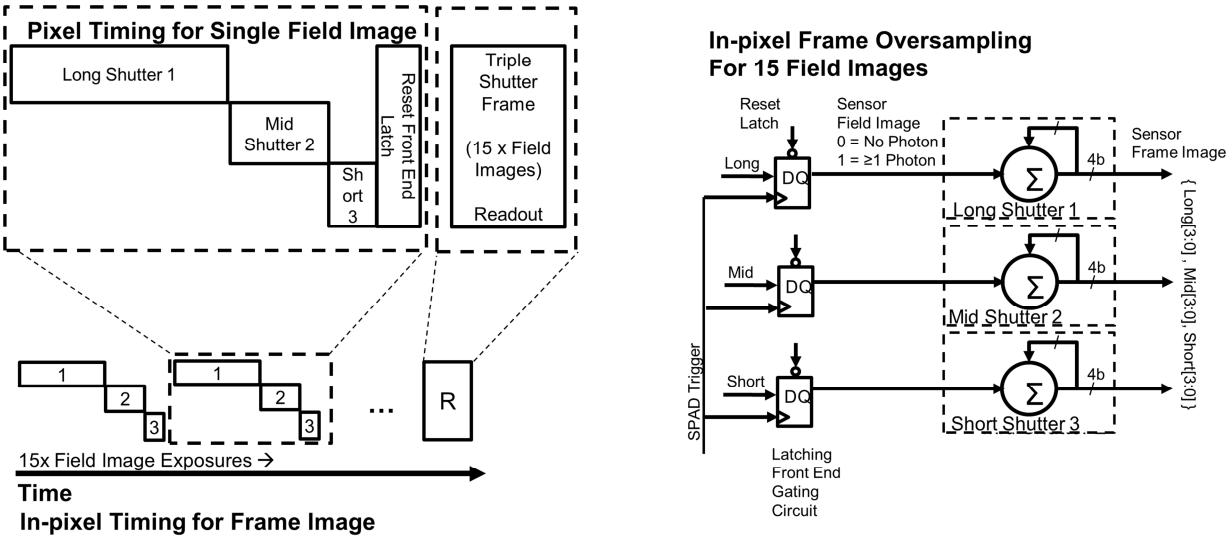

Figure 6. Left: HDR SPC pixel timing diagram with three field images with different shutter times (long, mid, short) each capturing a single bit (representing ≥1 photon), 15 field images are summed in pixel before column parallel readout. Right: pixel functionality block diagram showing the latching front end and the 3 ripple counters temporally summing 15 field images per readout frame.

The pixel timing is illustrated on the left of Fig. 6. To create the HDR image, three exposures (short, mid and long) are captured. Ideally for conventional HDR timing [27], the exposures are interleaved to minimise motion blur, but due to the front end circuit these are captured back to back. However, this effect is considered to be minimal in our implementation of SPC HDR as the three exposures are captured within micro-seconds of each other. QISs capture 1b per field image (representing ≥ 1 photon) and, here in the HDR QIS, 1b is captured for each exposure. After the three exposures are completed, the front end latching circuit is reset for the next field image and the in-pixel ripple counters are incremented as shown on the right of Fig. 6. Once 15 field images have been summed in-pixel (compressed to 4b) to constitute one frame, the 12b data represents 45 field (or bit plane) exposures. The data is readout via all-digital column parallel readout at 1000 frames per second (FPS) i.e. the sensor operates at 45000 fields per second (FiPS).

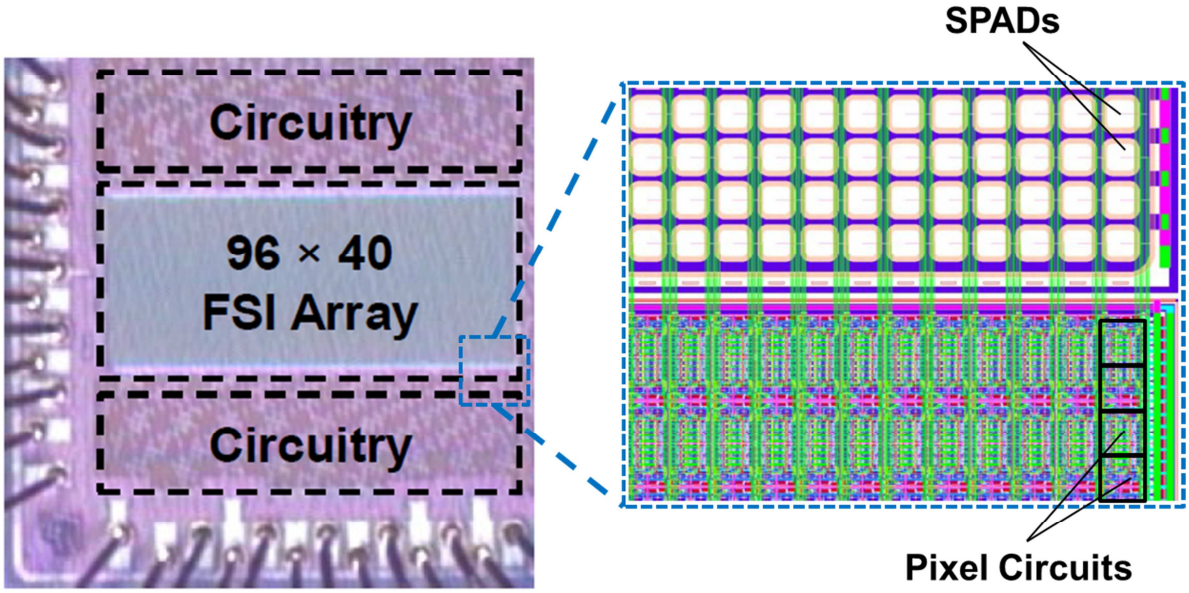

Figure 7. Left: Photomicrograph of the fabricated IC. Right: Top layout view showing the SPADs shared in a global well and pitch-matched pixel circuits in a bank below.

4. Measurements Results

The 96x40 imager was fabricated in STMicroelectronics 40nm FSI imaging technology. A photomicrograph and layout view is shown in Fig. 7. The test chip measures 1.0mm × 1.0mm. The SPADs are in a single global shared well and the pixel circuits, at the same pitch, are at the periphery. This test array allows the oversampled HDR capability to be evaluated.

Fig. 8 illustrates the photon transfer curve (PTC) of a single pixel in linear counting mode to confirm that the photon counting mechanism of the SPADs and the image sensor is entirely shot-noise limited. The red-line is a model of shot-noise limited SPC and there is minimal deviation of measured results from the ideal model.

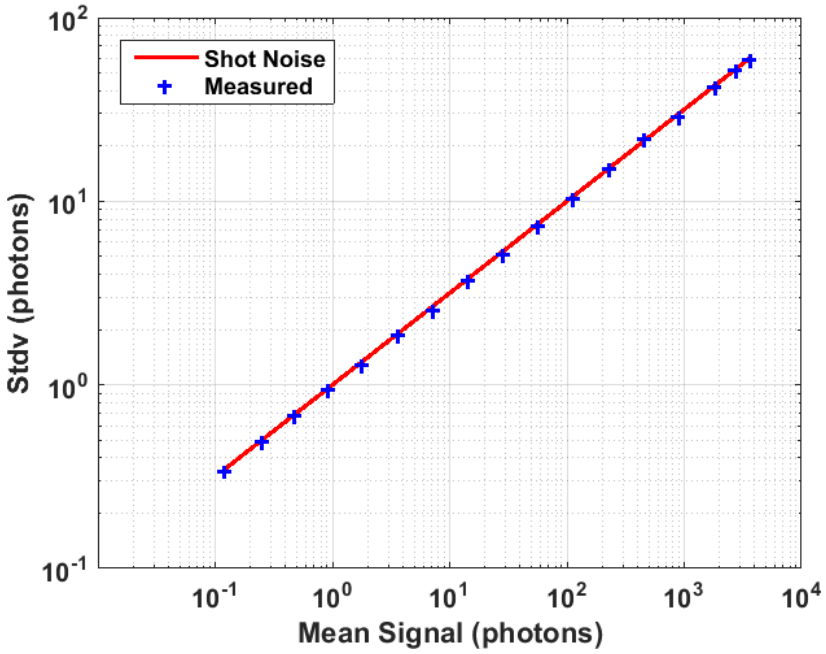

Figure 8. Photon transfer curve for single pixel in linear counting mode indicating purely shot noise limited single photon counting in this image sensor. The red line is a model of the shot-noise limit.

To demonstrate the sensor's quanta response the current through an LED source has been swept while data has been captured at a variety of exposure settings. For each light point, a total of 50 bit planes or fields of 96×40 pixels were spatially and temporally combined to result in a total of 192000 ensembles 'M'. For the purpose of speeding up the measurement all of the 96×40 pixels where spatially summed to contribute towards the total number of ensembles, while in a practical QIS use case a smaller subset of pixels or jots (8×8 for example [15]) would be spatially summed to represent one image element. The bit density 'D' vs the input signal 'H' curves were produced by dividing the total number of counts at each light point by M.

Fig. 9 shows the measured QIS response for a photon threshold 'K' of 1 where a pixel is assigned a binary value of '0' for no photons detected and a binary value of '1' for one or more photons detected. This binary assignment is performed by the in-pixel gating and counting logic depicted in Fig. 4. Two scenarios have been explored where 3 different exposures of ratios of 10 (0.1μs, 1μs and 10μs) and ratios of 2 (0.1μs, 0.2μs and 0.4μs) were used. The x-axis has been normalised such that an input signal of H=1 yields a bit density D=0.63 for the shortest exposure setting of 0.1μs. This is known as the 'full exposure' point as defined by [4]. The 0.1μs exposure setting has been chosen as the reference as it is the common setting across all measurements to follow. The modelled QIS response for this exposure is shown as the dashed red line where D is defined as:

$$D = 1 - e^{-H} \tag{1}$$

The measured data exhibits some deviation from the ideal model which could be attributed to non-linearity in the light source output power, illumination non-uniformity, photo-response non-uniformity and temporal variations as measurements were acquired over hours which would all contribute to the error in the spatio-temporally oversampled data. Nevertheless, the measured data offers a qualitative insight into QIS behavior. As can be seen from the results of the longest exposure setting of 10μs, it was not possible to reach low bit density values due to the limitations in the illumination source used. The authors opted for not combining data acquired by using different neutral density filters to avoid adding in more error.

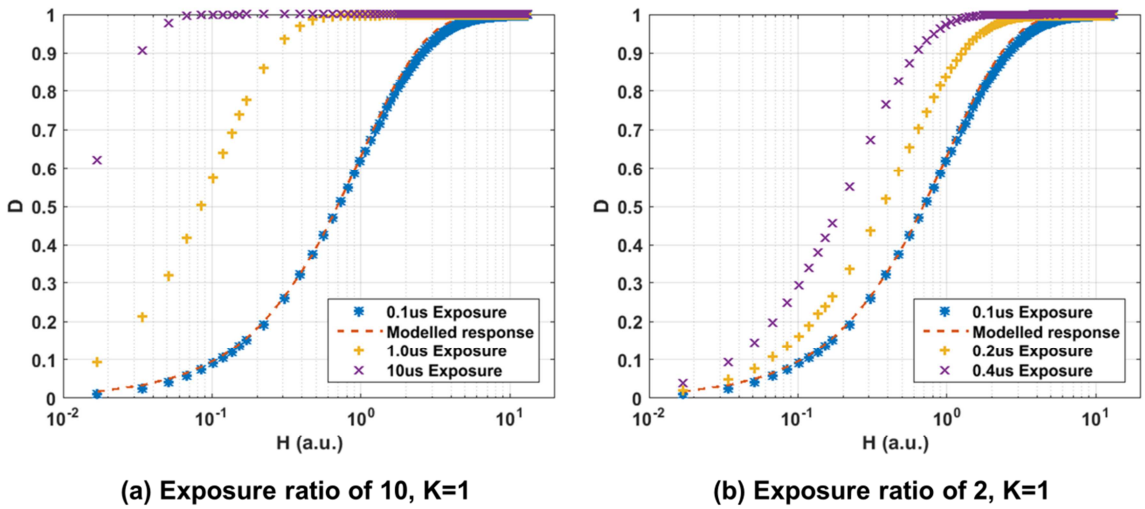

Figure 9. Measured normalized intensity (D or 'Bit plane density') to normalized input signal (H) for two sets of integration times for 1 photon threshold (K=1). (a) Exposure ratio of 10 with Short = 100ns, Mid = 1µs, Long = 10µs. (b) Exposure ratio of 2 with Short = 100ns, Mid = 200ns, Long = 400ns.

The measurement was repeated for an emulated photon threshold of K=2 (pixel assigned a binary value of '0' for no photons or one photon detected and binary value of '1' for two or more photons detected) by using linear counting mode (12 bit) and three sequential exposures. This emulation is necessary due to the latching single bit (K=1) front end in HDR mode. 50 single frames (no on-chip summation) were captured for each exposure setting where each pixel exhibits photon counts between 0 and 4095. By post processing the captured intensity frames the pixel values were re-assigned to transform the frame into a binary bit-plane or field. In the future an improved pixel design with multi-photon triggering could achieve the variable K threshold in-pixel. This variable threshold adjusts the non-linear intensity to exposure characteristic which is an interesting property of the QIS. The same exposure ratio settings were used and DlogH curves are shown in Fig. 10.

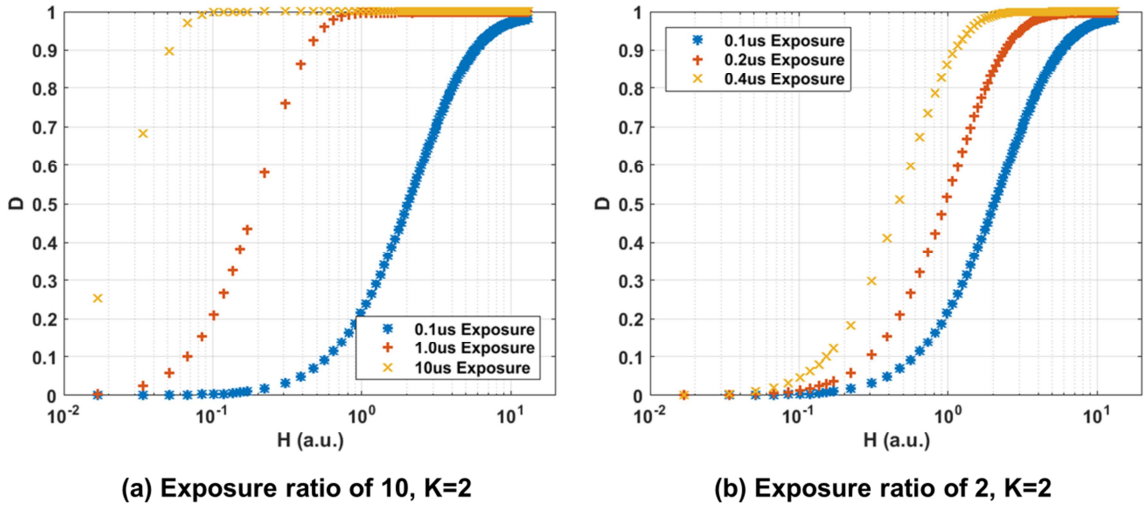

Figure 10. Measured normalized intensity (D or 'Bit plane density') to normalized input signal (H) for two sets of integration times for 2 photon threshold (K=2). (a) Exposure ratio of 10 with Short = 100ns, Mid = 1µs, Long = 10µs. (b) Exposure ratio of 2 with Short = 100ns, Mid = 200ns, Long = 400ns.

To evaluate the dynamic range (DR) and signal-to-noise ratio (SNR) of the quanta image sensor, and following from the theory presented in [4], DR hereby defined as:

$$DR = 20 \times \log(\frac{H_m}{H_n}) \quad (2)$$

Where H_m is the H value at which the measured signal reaches 99% of its saturation limit and H_n is the H value equivalent to the noise level (read + dark). Since the used digital sensor has no read noise as shown in Fig. 8, the only contribution to H_n is from the dark count rate (DCR) of the SPADs. For all carried measurements the SPADs were biased at 2V excess voltage for which the median DCR is ~150cps at room temperature [2]. Using equation 1, and taking D to be 150cps × 0.1μs, the equivalent H_n is calculated to be 1.5e-5. This value was used for all DR calculations in this work while H_m was estimated from the wanted measured signal.

It is worth noting that the number of ensembles 'M' has an effect on DR as the minimum observable signal is one photon per M ensembles (or 1/M), so for the maximum DR (DR_{max}) to be achieved it is necessary that the used number of ensembles is greater than the noise floor equivalent (i.e. $M>1/D(H_n)$), else the DR will be limited by the ability to observe a signal. Since M of 192000 used in the presented measurements satisfies this condition, all DR figures reported herein represent DR_{max} which might not be achievable in a practical QIS scenario.

For SNR calculations an alternative 'exposure referred SNR' or SNR_H definition was proposed by [4]. The objective of this definition is to project the SNR as measured in the y-axis (bit density D or 'voltage referred') onto the input x-axis (H). The reason behind this is that the voltage referred SNR will result in an artificial increase due to the compression of data by the QIS response and so SNR_H is a more meaningful measure. SNR_H is defined as:

$$SNR_H = \frac{H}{\sigma_H} \quad (3)$$

Where σ_H is defined as:

$$\sigma_H = \sigma_{Tot} \frac{dH}{dM_{Tot}} \quad (4)$$

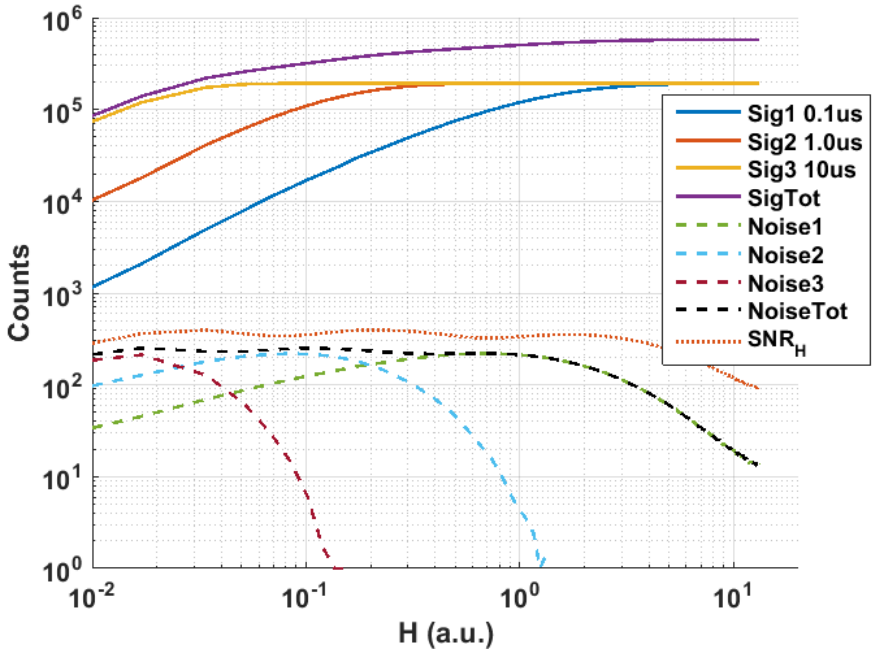

Figure 11. Measured signal, noise and SNR_H responses for 3 exposure settings with exposure ratio of 10 and photon threshold K=1.

Fig. 11 shows the cumulative QIS signal response and SNR_H for photon threshold K=1 and three different exposures with a ratio of 10 (0.1µs, 1µs and 10µs). Sig1, Sig2 and Sig3 are the counts M_1, M_2 and M_3 of the three corresponding exposures. SigTot (or M_{Tot}) is the linear summation of the counts of the three responses:

$$SigTot = Sig1 + Sig2 + Sig3 = M_{Tot} = M_1 + M_2 + M_3 \quad (5)$$

Noise1 is the standard deviation of Sig1 and under the assumption of Poisson statistics is given by:

$$Noise1 = \sigma_1 = \sqrt{\frac{(M - M_1) \times M_1}{M}} \quad (6)$$

Where M is 192000 (50 fields × 96×40 pixels). Noise2 and Noise3 are defined similarly and NoiseTot is the total noise of the cumulative response and is defined as:

$$NoiseTot = \sigma_{Tot} = \sqrt{(\sigma_1)^2 + (\sigma_2)^2 + (\sigma_3)^2} \quad (7)$$

Hence it is possible to calculate SNR_H for the measured data from the above equations. While it is not possible to observe the rise of SNR_H at low H values due to the measurement setup limitations and the fact that the long 10µs exposure response masks the response from the shorter exposures at these low H values, it is interesting to see how SNR_H peaks forming a 'plateau' region with very smooth transitions or 'ripples' when data from different exposures are summed as opposed to the dips in SNR observed in conventional image sensors. Using the equations above, SNR_H and DR have been calculated for cases of single, double and triple exposures with a ratio of 10 showing how DR increases from ~70dB to more than a 100dB in this example (Table 1).

Table 1. Calculated SNR_H and DR_{max} from measured data for the cases of single, double and triple exposures with a ratio of 10 and K=1.

Exposure	SNR_H (dB)	DR_{max} (dB)
10µs	50.5	72
10µs + 1µs	51.8	90
10µs +1µs + 0.1µs	52	109

The same analysis was repeated for the measurements of the same exposure settings with a photon threshold of K=2 to see the effect of photon threshold on SNR_H and DR in the case of multi-photon single-bit pixels. The signal and noise plots are shown in Fig. 12 and SNR_H and DR are summarised in Table 2. It is observed that while the DR increases slightly above that of K=1 this comes at the expense of more pronounced ripples or variation in SNR_H at the plateau region when combining the three exposures. The measured SNR_H variation in this example was ~2dB. The increase in DR is attributed to the fact that the QIS response for K=2 (Fig. 10(a)) is shifted to the right with respect to the response for K=1 (Fig. 9(a)) moving the 99% saturation point further while the lower end of the response is still dominated by the noise floor. Moreover, the K=2 response exhibits a steeper slope compared to that of K=1 which reflects on the transition between the three exposure settings and hence higher variation in SNR_H.

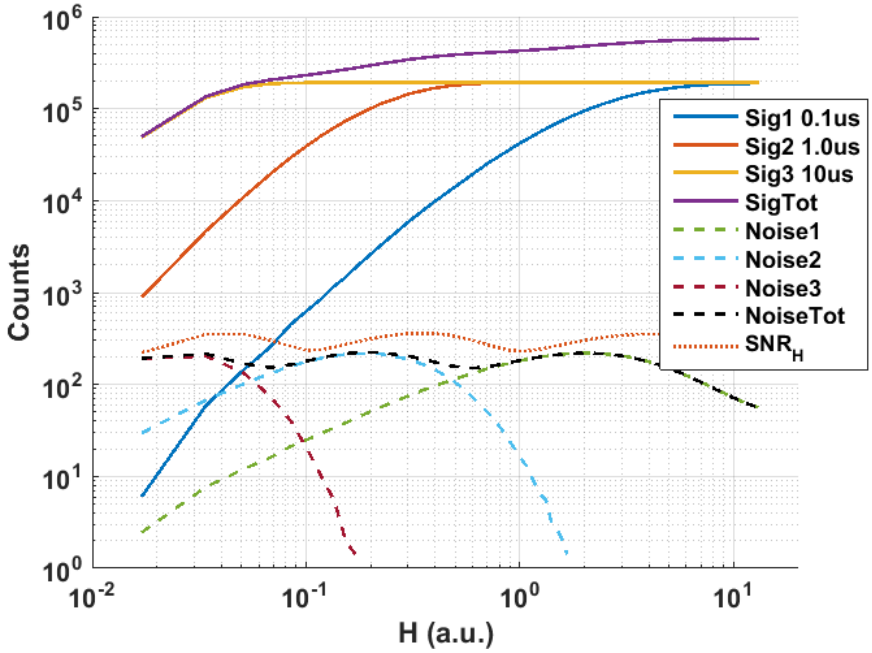

Figure 12. Measured signal, noise and SNR_H responses for 3 exposure settings with exposure ratio of 10 and photon threshold K=2.

Table 2. Calculated SNR_H and DR_{max} from measured data for the cases of single, double and triple exposures with a ratio of 10 and K=2.

Exposure	SNR_H (dB)	DR_{max} (dB)
10µs	50.7	75
10µs + 1µs	50.9	92.7
10µs +1µs + 0.1µs	51.1	115.8

Another factor that has been investigated is the effect of the exposure ratio on SNR_H and DR. For that, the same measurements as above were repeated for K=1 and exposure ratios of 2 (0.1µs, 0.2µs and 0.4µs), 4 (0.1µs, 0.4µs and 1.6µs), 6 (0.1µs, 0.6µs and 3.6µs) and 8 (0.1µs, 0.8µs and 6.4µs). The 0.1µs exposure setting is the common factor across all experiments. The measured SNR_H and DR for all cases are summarised in Table 3. It is observed that while SNR_H slightly decreases as the exposure ratio increases, DR is unaffected. This suggests that the DR extension is dominated by the shortest exposure setting which in this example was the common 0.1µs. Of course this holds true due to the fact that the minimum observable signal is dominated by the noise floor as a very large number of ensembles has been used as explained previously. For a smaller number of ensembles the minimum detectable signal will then be determined by the longest exposure setting and hence influence the achievable DR. In a rolling shutter sensor the shortest exposure would be dominated by line time and in a global shutter sensor it is down to signal drivers and acceptable temporal aperture ratio. The SNR_H peak is higher for smaller exposure ratios because as can be seen from equations (3) and (4), SNR_H is dependent on the rate of change in the total signal which is higher for short exposure ratios as the individual responses are close to each other and add up together more rapidly (i.e. dM_{Tot}/dH is higher for shorter exposure ratios). On the other hand, for longer exposure ratios the individual responses are spaced apart resulting in a slower rate of change in the total signal as they are summed together.

Table 3. Measured SNR_H and DR_{max} for a three exposures scenario and K=1 with different exposure ratios of 2 (0.1µs, 0.2µs and 0.4µs), 4 (0.1µs, 0.4µs and 1.6µs), 6 (0.1µs, 0.6µs and 3.6µs) and 8 (0.1µs, 0.8µs and 6.4µs).

Ratio	SNR_H (dB)	DR_{max} (dB)
2	55	108
4	53.5	108
6	52.7	108
8	52.2	108
10	52	109

While it is possible to obtain a high DR response with a single short exposure the advantage of having longer exposures is apparent when comparing the SNR_H response for exposure ratios of 2 and 8 (Fig. 13). Both cases result in a DR of ~108dB but as the exposure ratio increases (2 to 8) the SNR_H response results in a wider plateau region spanning a larger portion of the input signal H.

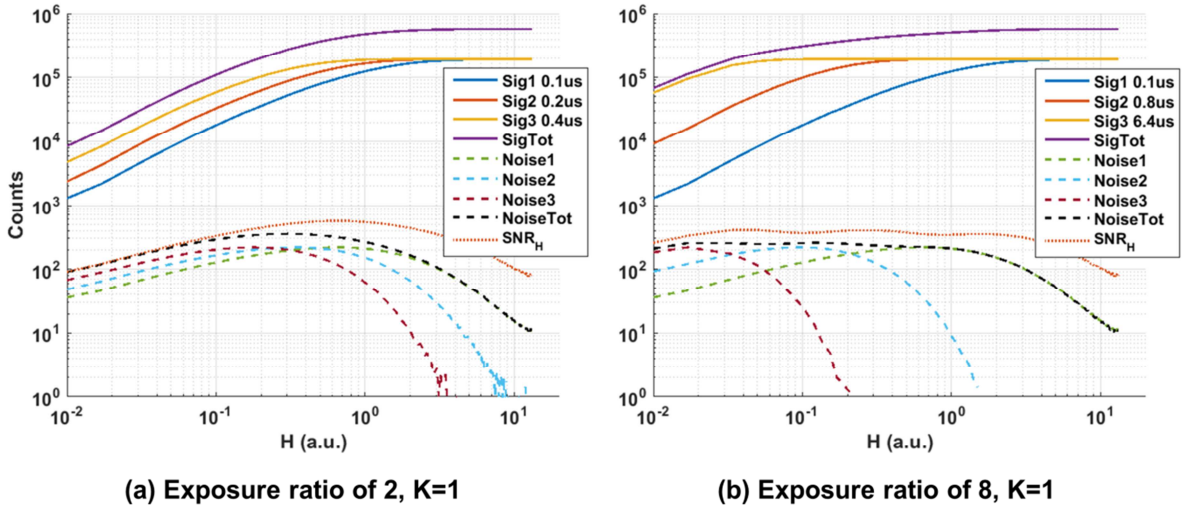

Figure 13. Measured signal, noise and SNR_H response for 3 exposure settings and K=1. (a) Exposure ratio of 2. (b) Exposure ratio of 8.

The presented results show how the dynamic range of a single frame triple-exposure sensor can be increased which is also an improvement over our previous paper [1] which required two frames to capture the three sub-exposures for the dynamic range extension. While other QIS sensors can attain similar DR performance, the partial in-pixel field summation providing 3.75x data compression and the ability to capture multiple exposure settings simultaneously significantly reduces readout requirements and offers better immunity against motion artifacts as compared to other works.

The 96×40 sensor is used to capture a high dynamic range scene as a demonstration of HDR QIS in operation in Fig. 14. To demonstrate this proof of principle further, Fig. 15 shows images captured by the 320×240 SPC imager from [23] which has higher resolution, wider field of view and lower DCR. Both sensors were operated with a photon threshold of K=1 and different exposures were acquired sequentially as only static scenes were imaged.

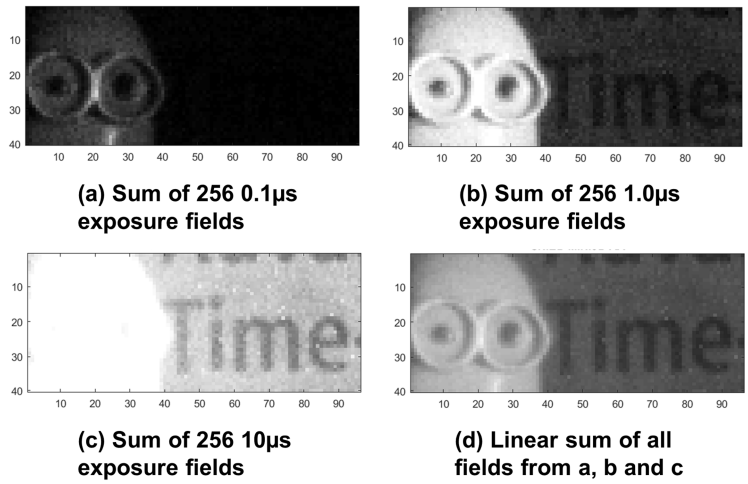

Figure 14. Images captured by 96×40 FSI sensor [2]. (a) Sum of 256 fields at 0.1µs exposure, a Minion figure is visible. (b) Sum of 256 fields at 1.0µs exposure, the Minion is visible but slightly overexposed while faint letters appear in the background. (c) Sum of 256 fields at 10µs exposure, the Minion is totally overexposed but the letters appear clearer. (d) Linear sum of all the 768 fields from a, b and c to form an HDR image preserving all details.

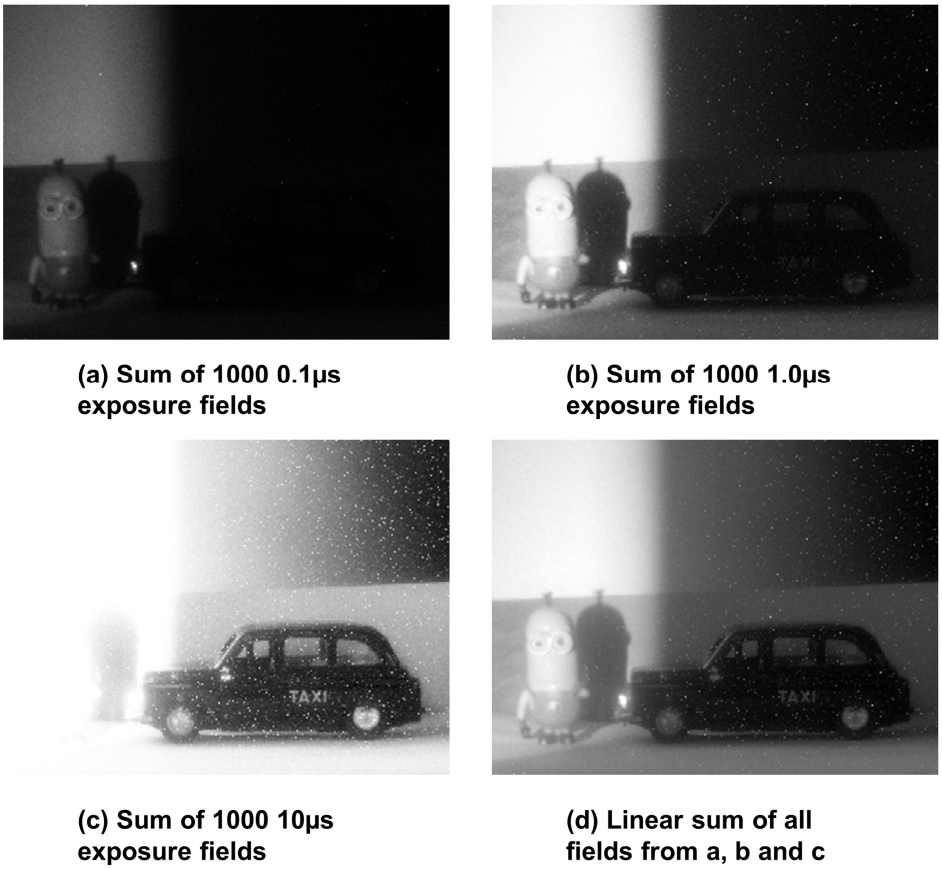

Figure 15. Images captured by 320×240 SPC sensor from [23]. (a) Sum of 1000 fields at 0.1µs exposure, a Minion appears in the lit portion of the scene. (b) Sum of 1000 fields at 1.0µs exposure, the Minion is slightly overexposed but a car figure appears in the dark region of the scene. (c) Sum of 1000 fields at 10µs exposure, Minion is completely overexposed but more detail of the car is apparent. Notice that high DCR pixels appear as white dots. (d) Linear sum of all the 3000 fields from a, b and c to form an HDR image preserving all details.

The example given in figure 14 allows for a brief benchmarking of HDR QIS performance. The presented analysis in this work shows that for the given 96×40sensor it is possible to achieve a maximum dynamic range (DR_{max}) of 108dB. Yet the DR of the example in figure 14 is limited by the number of ensembles (M=256) rather than the noise floor, so the effective DR ($DR_{effective}$) is limited by the minimum observable signal (bit density D=1/256 for each exposure). To calculate $DR_{effective}$ the equivalent H value for this minimum signal can be calculated from equation (1), and using that as the denominator in equation (2) results in an effective DR of 99.6dB for a three exposure (0.1µs, 1µs and 10µs) scenario showing the effect of M on achievable DR.

5. Discussion and Future Projections

The multi-megapixel QIS is an attainable goal for both CIS and avalanche based detectors. Each has their own advantages: the CIS-based and APD-based have low pixel pitches but rely upon conventional active pixel sensor (APS) analogue readout so can only oversample block-by-block [15] not per-pixel. SPAD pixels have the challenge of shrinking the pixel pitch but have the benefit of all-digital pixels allowing a wide range of in-pixel functionality such as HDR oversampling. Table 4 evaluates recently published APD and SPAD-based image sensor architectures projected up to a 1M pixel HDR oversampled SPC image sensor with at least 3 HDR field exposures. For this study, the output frame rate is set at 240FPS with 256 field images summed per frame. The single bit architectures require significant data rate and power to achieve this specification. Furthermore, as they can only capture one field before readout, readout time is presumed to be a significant limiting factor in their sensitivity. In contrast, this work permits up to 15 bit planes summed in-pixel (so needing 17 readouts to obtain the oversampling ratio (OSR) of 256) which is a tradeoff between temporal resolution (for imaging fast moving objects or fast phenomena) and data rate. The analogue counter in [29] allows up to 80 bit planes to be summed but would suffer from very high motion artifacts if implementing sequential HDR timing and requiring 12 readouts (4 per exposure). The downside to all the sensors based on SPAD, is the pixel pitch remains high so the array dimensions are large.

This is addressed in the second future projection for SPAD-based pixel circuits. Table 5 presents the maximum count (or equivalent full well) of this work for a single linear exposure and for HDR dual and triple exposures. 3D-stacked SPAD image sensors have been demonstrated [11][30], and this study indicates the future path to decrease the pitch of these image sensors' bottom tier pixel circuit can be based on all digital ripple counter and leverage the shrink gained through use of future technology nodes. Assuming the challenge of shrinking the SPAD diode and the stacking interconnect can also be met, then below 20nm CMOS should realise SPAD pixels with HDR capability and pitch in the 1-3µm range. In all these cases, spatio-temporal oversampling is required as the in-pixel bit depth is low. While this work relies only on temporal (summing fields) oversampling, spatio-temporal (summing fields and groupings of pixels) oversampling was adopted by other works to create an image from a quanta image sensor for example by summing an 8×8×8 kernel (pixels×pixels×fields) [15].

6. Conclusion

This paper examines methods to best exploit the HDR of the SPAD in a high fill-factor photon counting image sensor that is scalable to megapixel arrays. The digital ripple counter split into 3 individual exposures allows HDR photon counting to be realized in a compact pixel pitch. In-pixel summing provides data compression to increase the on-chip frame rate while maintaining low data rates off-chip.

The future expansion of array sizes of SPAD image sensors relies upon both compact pixel architectures and the shrink of pixel circuit areas. With 3D stacking of top-tier SPADs on advanced digital DSM CMOS, pixel pitches below 8μm are possible. The recent advent of three tier stacking [31] allowing the interconnection of back-illuminated sensing layer, processing and memory layers will further enhance the dynamic range, frame rate and power of both SPAD-based and CIS-based oversampled image sensors while facilitating novel ISP approaches such as motion tracking and solid-state optical image stabilization [28].

Acknowledgments: The authors would like to thank Sara Pellegrini, Bruce Rae and all in ST Crolles and ST Edinburgh for silicon design support. The authors also like to thank Hanning Mai from The University of Edinburgh for the helpful discussions. The authors are grateful to The University of Edinburgh and PROTEUS project (http://proteus.ac.uk) for funding this work (EPSRC grant number EP/K03197X/1) and POLIS project (http://polis.minalogic.net) and ST Crolles for providing silicon. This research received funding from the ERC TOTALPHOTON grant (grant no. 339747).

Author Contributions: N.D., T.A. wrote the paper; all authors contributed edits to the paper; N.D., T.A., I.G., and F.M.D.R., R.H. conceived and designed and performed the experiments; N.D., T.A. analyzed the data; T.A., N.D., R.H. contributed to the silicon design.

Conflicts of Interest: The authors declare no conflict of interest.

460 **Table 4.** Projection of different APD and SPAD pixel architectures to 1MPixel HDR QIS with ≥ 3
461 HDR field exposures.

	This Work	[23]	[1] [11]	[29]	[24]	[14]
Pixel Pitch	8.25	8	7.83	15	24	3.8
Circuit Type	Digital Ripple Counter	NMOS Dynamic Memory	Digital Ripple Counter	Analogue Counter	NMOS Static Memory	APD + CIS APS
Oversampling in-pixel	✓	×	✓	✓	×	×
Exposures In-Pixel	3	1	2	1	1	1
Counter Depth	4b	1b	6b	7b	1b	1b
Summing in-pixel per Exposure	15	1	63	80	1	1
Pixel data output width	12b	1b	12b	7b	1b	1b
Projection to 1MPix (1024 × 1024) 3D Stacked QIS with >100dB HDR						
Array Dimension (μm)	8448	8192	8017	15360	24576	3891
QIS Output Frame Rate (FPS)	240	240	240	240	240	240
Total OSR per Frame	256	256	256	256	256	256
Sensor Field Rate (FiPS) *	4,096	184,320	2,926**	2,304	184,320	184,320
Interface Data Rate (Gpbs)	48	180	34.3	15.8	180	180
In-pixel Data Compression Ratio ***	3.75	None	10.5	11.4	None	None
Motion Artifact	Best – V.Low	High	Low	V.High	High	High
Multiple HDR exposures in-pixel	✓	×	✓	×	×	×

462 * Field Rate = [OSR] x [Required Frame Rate] x [3 HDR Exposures] / ([Exposures in Pixel] x [Bit Planes in Pixel]). ** N.B. For 2
463 exposures in pixel, 2 readouts are required for 3 HDR exposures. *** Compression Ratio = [Bit Planes in Pixel] x [Exposures In
464 Pixel] / [Output Data Width]

465 **Table 5.** Projected pixel shrink trends, of 2D monolithic and 3D-stacked SPAD photon counting
466 pixels based on an all-digital ripple counter architecture, towards 3D-stacked SPAD or Avalanche
467 based Quanta Image Sensors with multi-megapixel arrays.

CMOS Tech' Node (nm) for Digital Logic	**Circuit Pixel Pitch (μm)**	**2D Monolithic**	**3D Stacked**	**Single Exposure Linear Counter**		**Dual Exposure HDR**		**Triple Exposure HDR**	
				Max Count	**Counter Bit Depth**	**Max Count**	**Counter Bit Depth**	**Max Count**	**Counter Bit Depth**
This Work, [1] and [11]									
40	7.83 - 8.25	✓	✓	4k[1]	12	63	6	15	4
Future Projected Trends									
32	6	✓	✓	16k[1]	14	127	7	15	4
22	3		✓	127	7	7	3	3	2
16	1.5		✓	7	3	1	1	1	1
11	1.0		✓	7	3	1	1	1	1
11	0.75		✓	1	1	-	-	-	-

468 [1] With the exception of these two high maximum counts, spatio-temporal oversampling is required as the in-pixel max count
469 (or equivalent full well) is low.

470

Bibliography

1. Dutton, N.A.W; Al Abbas, T.; Gyongy, I; Henderson, R.K. "Extending the Dynamic Range of Oversampled Binary SPAD Image Sensors" Proceedings of the International Image Sensors Workshop, Hiroshima, Japan, 2017; P07.
2. Al Abbas, T.; Dutton, N.A.W; Almer, O.; Della Rocca, F.M; Pellegrini, S., Rae, B.; Golanski, D.; Henderson, R.K. "8.25μm Pitch 66% Fill Factor Global Shared Well SPAD Image Sensor in 40nm CMOS FSI Technology" Proceedings of the International Image Sensors Workshop, Hiroshima, Japan, 2017; P10.
3. Seitz, P.; Theuwissen, A. "Single Photon Imaging", 1st ed.; Springer: Heidelberg, Germany, 2011.
4. Fossum, E.R. Modeling the Performance of Single-Bit and Multi-Bit Quanta Image Sensors. *IEEE J. Electron Devices Soc.* 2013, *1*, 166–174.
5. Teranishi, N. "Required Conditions for Photon-Counting Image Sensors. *IEEE Trans. Electron Devices* 2012, *59*, 2199–2205.
6. Dutton, N.A.W; Gyongy, I; Parmesan, L.; Henderson, R.K. "Single Photon Counting Performance and Noise Analysis of CMOS SPAD-Based Image Sensors" Sensors, 2016, Jul; 16(7): 1122.
7. Ma, J.; Fossum, E. "Quanta Image Sensor Jot with Sub 0.3 e-rms. Read Noise and Photon Counting Capability". IEEE Electron Device Lett. 2015, 36, 926–928.
8. Seo, M.-W.; Kawahito, S.; Kagawa, K.; Yasutomi, K. "A 0.27 e-– Read Noise 220-uV Conversion Gain Reset-Gate-Less CMOS Image Sensor". IEEE Electron Device Lett. 2015, 36, 1344–1347.
9. Gyongy, I.; Dutton, N.A.W.; Parmesan, L.; Davies, A.; Saleeb, R.; Duncan, R.; Rickman, C.; Dalgarno, P.; Henderson, R.K. "Bit-plane Processing Techniques for Low-Light, High Speed Imaging with a SPAD-based QIS". In Proceedings of the International Image Sensor Workshop, Vaals, The Netherlands, 8–11 June 2015.
10. Pellegrini, S.; Rae, B.; Pingault, A.; Golanski, D.; Jouan, S.; Lapeyre, C.; .Mamdy, B. "Industrialised SPAD in 40 nm Technology" Proceedings of International Electron Devices Meeting, San Francisco, CA, 2017; 16.5
11. Al Abbas, T.; Dutton, N.A.W; Almer, O.; Pellegrini, S.; Henrion, Y.; Henderson, R.K. "Backside illuminated SPAD image sensor with 7.83μm pitch in 3D-stacked CMOS technology" Proceedings of International Electron Devices Meeting, San Francisco, CA, 2016; 8.1
12. Henderson, R.K.; Webster, E. A. G.; Walker, R.; Richardson, J. A.; Grant, L. A. "A 3×3, 5μm pitch, 3-transistor single photon avalanche diode array with integrated 11V bias generation in 90nm CMOS technology," 2010 International Electron Devices Meeting, San Francisco, CA, 2010, pp. 14.2.1-14.2.4.
13. You, Z.; Parmesan, L.; Pellegrini, S.; Henderson, R.K. "3 μm Pitch, 1 μm Active Diameter SPAD Arrays in 130nm CMOS Imaging Technology" Proceedings of the International Image Sensors Workshop, Hiroshima, Japan, 2017; R20.
14. Mori, M.; et al. "6.6 A 1280 x 720 single-photon-detecting image sensor with 100dB dynamic range using a sensitivity-boosting technique," 2016 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, 2016, pp. 120-121.
15. Ma., J; Masoodian, S.; Wang, T.J; Fossum, E.R. "A 1Mjot 1040fps 0.22e-rms Stacked BSI Quanta Image Sensor with Cluster-Parallel Readout" Proceedings of the International Image Sensors Workshop, Hiroshima, Japan, 2017; R18
16. Della Rocca, F.M; Al Abbas, T.; Dutton, N.A.W; Henderson, R.K. "A High Dynamic Range SPAD Pixel for Time of Flight Imaging" Proceedings of IEEE Sensors, Glasgow, UK, 2017, Paper 5.4.
17. Pancheri, L.; Stoppa, D. "A SPAD-based pixel linear array for high-speed time-gated fluorescence lifetime imaging," 2009 Proceedings of ESSCIRC, Athens, 2009, pp. 428-431.
18. Niclass, C.; Sergio, M.; Charbon, E. "A CMOS 64x48 Single Photon Avalanche Diode Array with Event-Driven Readout," Proc. Eur. Solid-State Circuits Conf., pp. 556–559, 2006.
19. Pancheri, L.; Massari, N.; Borghetti, F.; Stoppa, D. "A 32x32 SPAD Pixel Array with Nanosecond Gating and Analog Readout," in Proceedings of International Image Sensor Workshop, Japan, 2011, R40
20. Stoppa, D.; et. al. "A 32x32-pixel array with in-pixel photon counting and arrival time measurement in the analog domain," in Proceedings of the European Solid-State Circuits Conference (ESSCIRC), 2009, pp. 204–207.

21. Panina, E.; Pancheri, L.; Dalla Betta, G.; Massari, N.; Stoppa, D. “Compact CMOS Analog Counter for SPAD Pixel Arrays,” IEEE Trans. Circuits Syst. II Express Briefs, vol. 61, no. 4, pp. 214–218, Apr. 2014.
22. Dutton, N.A.W.; Grant, L.A.; Henderson, R.K. 9.8 μm SPAD-based Analogue Single Photon Counting Pixel with Bias Controlled Sensitivity. In Proceedings of the International Image Sensors Workshop, Snowbird, UT, USA, 12–16 June 2013.
23. Dutton, N.A.W.; Gyongy, I.; Parmesan, L.; Gnecchi, S.; Calder, N.; Rae, B.R.; Pellegrini, S.; Grant, L.A.; Henderson, R.K. A SPAD-Based QVGA Image Sensor for Single-Photon Counting and Quanta Imaging. IEEE Trans. Electron Devices 2016, 63, 189–196.
24. Burri, S.; Maruyama, Y.; Michalet, X.; Regazzoni, F.; Bruschini, C.; Charbon, E. Architecture and applications of a high resolution gated SPAD image sensor. *Opt. Express* **2014**, *22*, 17573–17589.
25. Lule, T.; Keller, H.; Wagner, M.; Bohm, M. "100.000 pixel 120 dB imager in TFA-technology," 1999 Symposium on VLSI Circuits. Digest of Papers (IEEE Cat. No.99CH36326), Kyoto, Japan, 1999, pp. 133-136.
26. Decker, S.; McGrath, R.; Brehmer, K.; Sodini, C. "A 256/spl times/256 CMOS imaging array with wide dynamic range pixels and column-parallel digital output," 1998 IEEE International Solid-State Circuits Conference. Digest of Technical Papers, San Francisco, CA, USA, 1998, pp. 176-177.
27. Lulé, T.; Mandier, C.; Glais, A.; Roffet, G.; Monteith, R.; Deschamps, B. “High Performance 1.3MPix HDR Automotive Image Sensor” in Proceedings of the International Image Sensor Workshop, Vaals, The Netherlands, 8–11 June 2015.
28. Gyongy, I.; Al Abbas, T.; Dutton, N.A.W.; Henderson, R.K. “Object Tracking and Reconstruction with a Quanta Image Sensor” Proceedings of the International Image Sensors Workshop, Hiroshima, Japan, 2017; R22.
29. Perenzoni, M.; Massari, N.; Perenzoni, D.; Gasparini, L.; Stoppa, D. "A 160 x 120 Pixel Analog-Counting Single-Photon Imager With Time-Gating and Self-Referenced Column-Parallel A/D Conversion for Fluorescence Lifetime Imaging," in IEEE Journal of Solid-State Circuits, vol. 51, no. 1, pp. 155-167, Jan. 2016.
30. Aull, B.; Schuette, D.; Young, D.; Craig, D.; Felton, B.; Warner, K. “A Study of Crosstalk in a 256x256 Photon Counting Imager Based on Silicon Geiger-Mode Avalanche Photodiodes,” IEEE Sens. J., vol. 15, no. 4, pp. 2123–2132, Apr. 2015.
31. Haruta, T.; et al. "4.6 A 1/2.3inch 20Mpixel 3-layer stacked CMOS Image Sensor with DRAM," 2017 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, 2017, pp. 76-77.