

A 200kFPS, 256×128 SPAD dToF sensor with peak tracking and smart readout

Istvan Gyongy^[1], Ahmet T. Erdogan^[1], Neale A.W. Dutton^[2], Hanning Mai^[1],
 Francesco Mattioli Della Rocca^[1], Robert K. Henderson^[1]

^[1]The University of Edinburgh, Institute for Integrated Micro and Nano Systems, Edinburgh, U.K.

^[2]Imaging Division, STMicroelectronics, Edinburgh, U.K.

Istvan.Gyongy@ed.ac.uk Tel: +44 131 651 7054

Abstract—The storage, transfer and readout of large volumes of photon timing data are key challenges in the design of SPAD dToF imagers, especially if high frame rates, and long ranging distances are to be achieved. This paper describes an imager with partial histogramming units in each pixel that automatically scan, detect and track peaks. Output data rates are reduced further by sub-bin resolution peak extraction and smart readout modes which only show pixels presenting salient information.

I. INTRODUCTION

Direct time-of-flight (dToF) SPAD image sensors offer a promising all-digital solution for LIDAR, with the capability of producing accurate depth estimates even from weak signal returns, thanks to precise integrated timing electronics. In typical outdoor operating conditions, a large number of events, originating from both signal and background photons, may be recorded. Thus, capturing and reading out the relevant timing information whilst avoiding pile-up and excessive power consumption becomes a key architectural challenge. Existing approaches include time gating and photon

coincidence detection for filtering out background photon events, and data compression via on-chip histogramming and peak extraction [1-4].

II. SENSOR ARCHITECTURE

This paper presents a SPAD image sensor with integrated photon processing to enable high-speed flash LIDAR at up to 400M points per second. The chip, implemented in a standard 40nm CMOS FSI technology, features a 256×128 array of SPADs, divided into 64×32 macropixels, each with its own processing unit (Fig. 1).

A key feature of the sensor is an ability to provide depth estimates over long distance ranges, whilst reducing the effect of background photons, and keeping on-chip memory requirements and output data rates at modest levels. This is achieved by generating a partial, 8 bin photon timing histogram within each macropixel, which is scanned over the full time range, until a peak is detected by the pixel logic. Any subsequent movement in the peak (corresponding to movement in z by the detected surface) is then automatically tracked. The

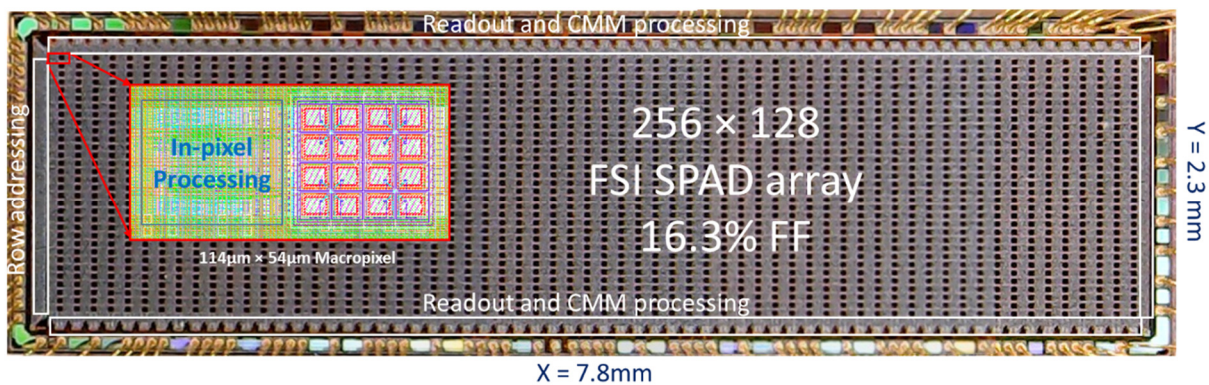


Figure 1. Chip micrograph and layout of a macropixel. Readout is over 64, 100MHz output lines (32 on the top and 32 on the bottom).

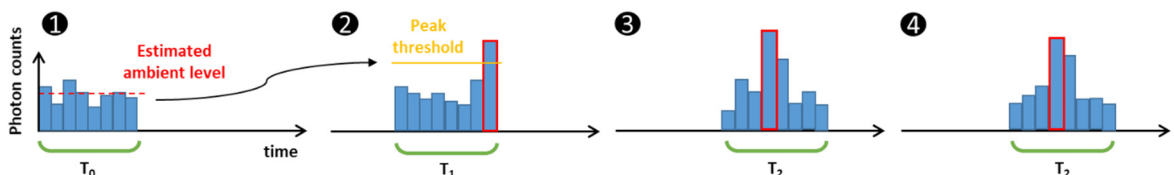


Figure 2. Peak tracking in a sequence of exposures: (1) the ambient level is estimated; there is no peak, so the time gate is shifted, (2) peak is identified at bin 8; the time gate is shifted to move the peak to the middle of the time range, (3) peak is in the middle, so no shift is required.

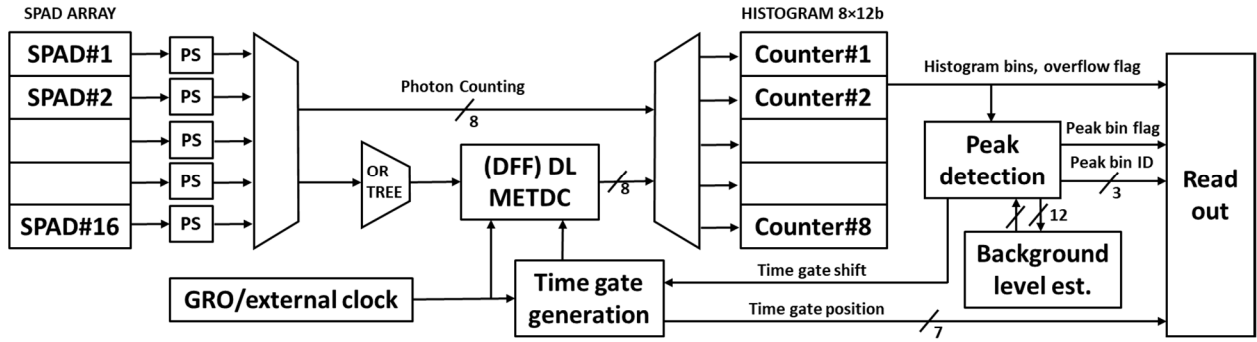


Figure 3. Macropixel block diagram.

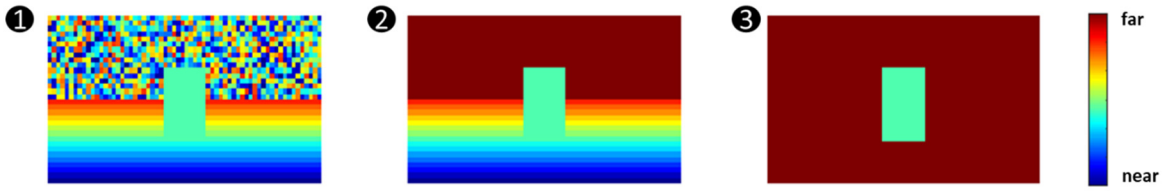


Figure 4. Illustration of the “smart” readout options, under the assumption of an approaching rectangular target on a flat horizontal surface: (1) all macropixels are read out (default), (2) only macropixels with peaks are read out (peak bin flag high), (3) only macropixels with peaks and change in time gate position are read out. Data from macropixels that are not read out is replaced by all zeros.

scheme, illustrated in Fig. 2, is implemented by assigning an independent time gate to each macropixel, whose position is then shifted (or maintained) in time after each exposure depending on the results of in-pixel processing. The leading edge of the time gate starts the fine timing 8 bin, multi-event time-to-digital converter (METDC), whose output time codes are then accumulated in in-pixel histogramming memory. Photon events outside the time gate are ignored. Compared to the alternative approach of peak finding, also using partial histogramming, which progressively reduces the bin size [5,6], this approach is expected to be more robust to low signal-to-background ratios.

The block diagram of a macropixel unit is given in Fig. 3. The outputs from the 16 SPADs are first fed through pulse shorteners (PS), before being combined by an OR tree and sent to the METDC, which receives fine and

Option	Coarse timing (for time gate positioning)	Fine timing (for TDC bins)
1	GRO clock	GRO clock
2	GRO clock	Delay cell DL
3	External clock	GRO clock
4	External clock	Delay cell DL
5	External clock	External clock

Table 1. Clock options for timing. There is a choice of two METDCs: a D-type flip-flop delay line using a reference clock from a gated ring oscillator (GRO) or an external reference, and a tuneable delay line composed of delay cells.

coarse timing reference signals. A number of options are available for generating these signals (see Table 1.), enabling bin sizes ≥ 0.5 ns. There are 128 time gate positions, with 50% overlap, leading to an unambiguous timing range of 512ns or ~ 77 m at the default bin size of 1ns (the timing range can be increased arbitrarily by widening the bin size). The output time codes populate a 12-bit/bin histogram based on a counter array with overflow protection. At the end of each exposure, the peak bin is identified and its value compared to a background level estimate held in memory for statistical significance [7]. If a significant peak is identified, then the time gate position is adjusted, if necessary, so as to keep the peak within the four middle bins of the histogram, the nominal 50% overlap between time gate positions guaranteeing a position which captures the peak in its entirety. Every time the time gate is moved,

Mode	Output format	Total no. bits/macro pixel	Frame rate (FPS)
Histogram mode (default)	Histogram bins and histogram peak data	108	29k
Bin-resolution depth	Histogram peak data (peak bin flag, peak bin ID and overflow flag)	12	260k
Sub-bin resolution depth	Centre-of-mass of background-corrected histogram bins	15	208k

Table 2. Outline of data compression options in time-resolved mode (in terms of data read out in addition to 7-bit time gate position). In the histogram mode, row-skipping (partial readout) is possible, if a row is found not to have any macropixels with peaks.

the background estimate is updated. There is a dual exposure mode whereby macropixels that are peak searching operate with shorter exposure times to allow for fast convergence to the peaks. The rationale is that much fewer photons are typically required for accurate peak detection compared with estimating the position of the peak with suitable precision. The chip can also be configured for continuous scanning (the time gate positions being incremented after every exposure), or programmed with fixed gate positions.

To provide further data compression, and direct depth readings with sub-bin precision, centre-of-mass (CMM) processing with background compensation is implemented in column parallel logic (Table 2). Furthermore, taking advantage of the results of in-pixel processing, smart readout options are available which only show detected surfaces, or surfaces that are moving in z (Fig. 4), reducing I/O toggling and downstream processing requirements. This has parallels with Dynamic Vision Sensors [8] and may suit robotic vision-type applications.

A 128×128 , 12-bit photon counting (SPC) mode is also available, enabling off-chip intensity-guided depth upscaling [9]. In this mode, SPAD detectors are binned in pairs horizontally. The sensor can alternate between time-resolved and photon counting modes, with gate positions being preserved. The 4×4 detector array within macropixels matches the size of the adjacent processing unit, making the design 3D stacking ready, for higher fill factor.

III. PRELIMINARY RESULTS

Preliminary test results in the SPC and histogram modes suggest correct operation in these modes (Fig. 5 and 6). For the latter trial, the sensor was coupled with an 850nm VCSEL source operated with ~ 6 ns pulse width, 800kHz repetition rate and ~ 50 mW average optical power. The field of view was approximately $80^\circ \times 20^\circ$ (H \times V). As the sensor and illuminator optics were unoptimised, only the central portion of the field of view (featuring a box on top of a table) was illuminated. Fig. 7 shows the evolution, over four successive exposures, of the in-pixel histogram for three pixel locations: (1) outside the illuminated area, (2) on the box, and (3) on the wall behind the box. For (1), as no peak is detected, the pixel continues scanning by incrementing the time gate position after each exposure. In the case of (2), a peak is detected in the second exposure, and as it is in one of the central bins (bin no. 3-6), the time gate position is maintained subsequently. The pixel at (3) also detects a peak in exposure 2, but because it is in an outlying bin (the wall being at a further distance than the box), the time gate position is incremented again to move the peak

to one of the central bins for exposure 3. The time gate is then maintained for exposure 4.

Under low photon activity, the total chip power consumption is ~ 7 mW in the histogram mode (with a 40MHz external clock), increasing to ~ 9 mW for clock options 1 and 2.

IV. CONCLUSIONS

We have presented a SPAD dToF imager that generates an 8-bin partial histogram in each pixel, which scans the time range of interest and tracks peaks. The characterisation of the chip, as well as the optimisation of the sensor optics is ongoing, and should yield improved ToF results in the future. Of particular interest is a potential demonstration of the device in imaging high-speed scenes.

ACKNOWLEDGMENTS

This research was supported by EPSRC via grants EP/M01326X/1, EP/S001638/1. The authors are grateful to STMicroelectronics for chip fabrication and in particular to Sara Pellegrini and Thierry Lachaud for their support.

REFERENCES

- [1] Padmanabhan et al., "A 256×128 3D-Stacked (45nm) SPAD FLASH LiDAR with 7-Level Coincidence Detection and Progressive Gating for 100m Range and 10klux Background Light," ISSCC 2021
- [2] Hutchings et al., "A reconfigurable 3-D-stacked SPAD imager with in-pixel histogramming for flash LiDAR or high-speed time-of-flight imaging," JSSC 2019, 54(11)
- [3] Niclass et al., "A 0.18- μ m CMOS SoC for a 100-m-Range 10-Frame/s 200×96 -Pixel Time-of-Flight Depth Sensor," JSSCC 2013, 49(1)
- [4] Kumagai et al., "A 189×600 Back-Illuminated Stacked SPAD Direct Time-of-Flight Depth Sensor for Automotive LiDAR Systems," ISSCC 2021
- [5] Zhang et al., "A 30-frames/s, SPAD Flash LiDAR with 1728 Dual-Clock 48.8-ps TDCs, and Pixel-Wise Integrated Histogramming," JSSC 2018, 54(4)
- [6] Kim, B., et al., "A 48×40 13.5 mm Depth Resolution Flash LiDAR Sensor with In-Pixel Zoom Histogramming Time-to-Digital Converter," ISSCC 2021
- [7] Gneccchi et al., "Long Distance Ranging Performance of Gen3 LiDAR Imaging System based on 1×16 SiPM Array," IISW 2019
- [8] Yang et al., "A dynamic vision sensor with 1% temporal contrast sensitivity and in-pixel asynchronous delta modulator for event encoding," JSSC 2015, 50(9)
- [9] de Lutio et al., "Guided super-resolution as pixel-to-pixel transformation," ICCV 2019

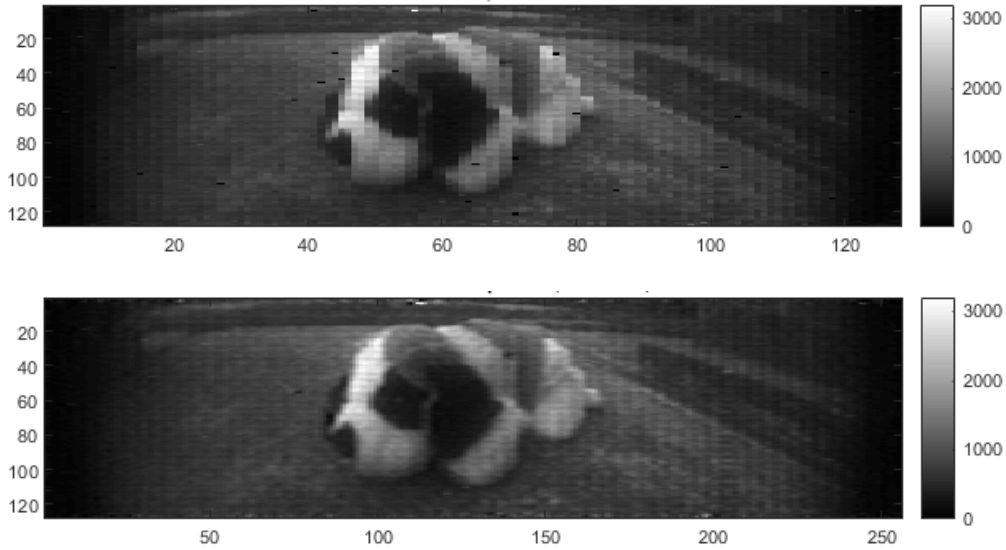


Figure 5. Photon counting image captured outdoors with 2ms exposure time showing a plush dog. The top image shows the raw pixel data after correcting for the dark count rate of pixels; in the bottom image interpolation is applied to correct for the uneven pixel spacing in the horizontal direction.

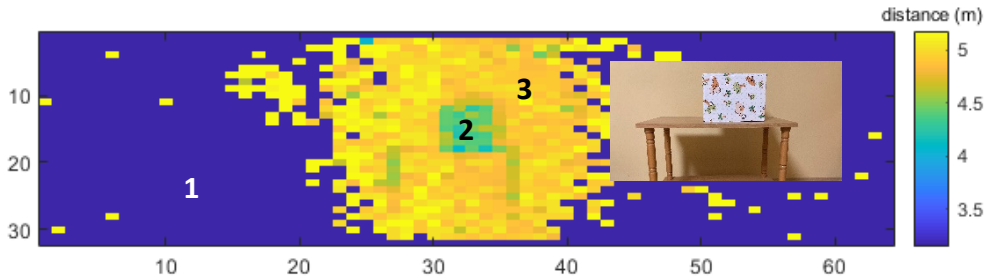


Figure 6. Uncalibrated ToF data obtained in histogram mode, with clocking option 4 (see Table 1.) and bin size configured to 3ns. The data corresponds to the 4th exposure in a sequence of 10ms exposures. Only pixels with peak bin flag=1 are shown. Depth is computed by applying a centroid calculation to the raw histogram data. Three pixel positions are highlighted: (1) outside the illuminated area, (2) on the box, and (3) on the wall.

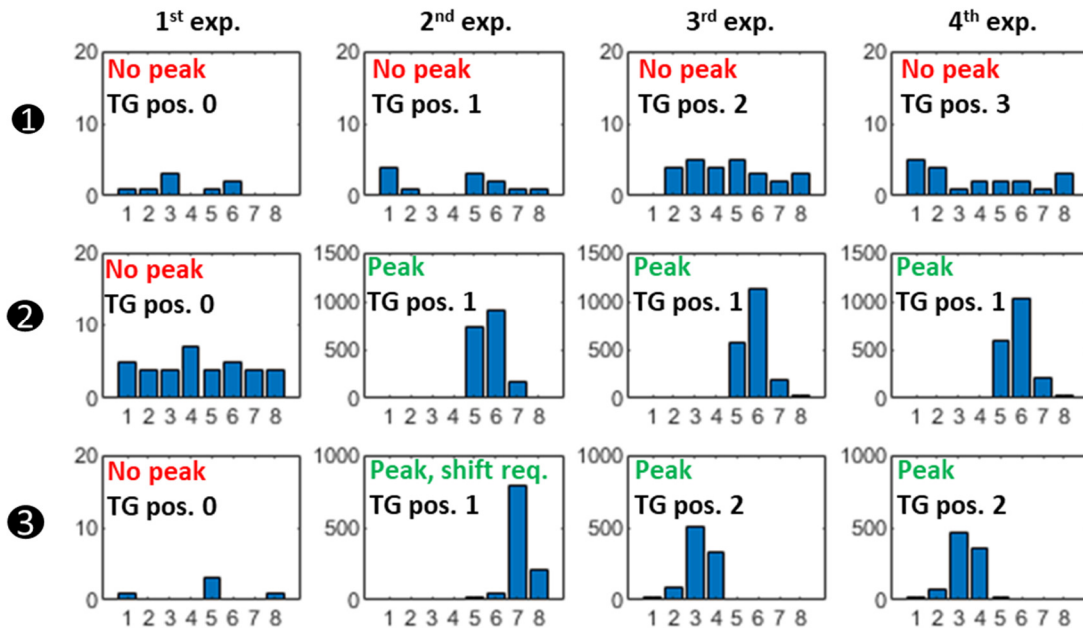


Figure 7. Histograms for the pixel positions indicated in Figure 6 across the four exposures. The time range of the histogram is $8 \times 3 = 24\text{ns}$ and there is a shift of 12ns between successive time gate positions.