1kFPS Time-of-Flight Imaging with a 3D-stacked CMOS SPAD Sensor

Istvan Gyongy[1], Sam W. Hutchings[1], Max Tyler[2], Susan Chan[2], Feng Zhu[2], Robert K. Henderson[1], Jonathan Leach[2]
[1]The University of Edinburgh, Institute for Integrated Micro and Nano Systems, Edinburgh, U.K.
[2]Heriot-Watt University, Institute of Photonics and Quantum Sciences, Edinburgh, U.K.
Istvan.Gyongy@ed.ac.uk Tel: +44 131 650 5620

Abstract—Current 3D Time-of-Flight (ToF) systems typically offer video frame rates. However in safety critical applications, such as automotive LIDAR, faster frame rates, and a reduced latency in 3D perception, would be of considerable benefit. This paper presents initial 3D imaging results from a direct ToF SPAD image sensor, with in-pixel photon processing, that can operate at frame rates exceeding 1kFPS.

I. Introduction

3D imagers based on Time-of-Flight (ToF) are used in a range of applications including robotics, assisted or autonomous driving systems, computer interfaces and Virtual and Augmented Reality (VR/AR) devices. There is ongoing effort to improve the accuracy, increase the pixel count, and reduce the power consumption of ToF cameras, with shorter imaging ranges (<30m) typically being served by indirect ToF devices with demodulation pixels (combined with continuous wave illumination), whilst longer ranges are attained by direct ToF using pulsed illumination and Avalanched Photodiodes (APD) or Single Photon Avalanche Diodes (SPAD).

This paper is concerned with a SPAD-based direct ToF device, implemented in 40/90nm 3D stacked technology, and featuring a re-configurable 256×256 SPAD array, with 9.2µm pitch [1]. For photon timing, SPADs are combined, in groups of 4x4, to form 64×64 macropixels. Each macropixel outputs either time codes (provided by a ring oscillator TDC), or 16-bin, 14-bit/bin histograms generated by a multi-event TDC [2]. In either mode, a multi-photon triggering functionality is available for background suppression, whereby multiple photon detections are required (within a programmable time window) for the macropixel to register an event. An additional mode returns 14-bit photon counts at the full 256×256 resolution of the array.

The focus here is on the on-chip histogramming mode, and its potential for fast 3D capture. Due to significant data reduction in this mode, histograms are read out for the 64x64 macropixels at a maximum rate of 760FPS over 8, 100MHz data lines, increasing to >1kFPS for half of the array (64×32). It is envisaged that adding further output lines would enable even higher frame rates (or larger array sizes). As a comparison, the high-speed 3D ToF SPAD sensor of [3] achieves 200FPS for 64×32 pixels, but with an indirect ToF architecture. The TI OPT8320 iToF sensor, featuring current assisted photonic demodulator pixels, is specified for a maximum 1kFPS at 80×30 [4].

The histograms from the present sensor are processed externally to obtain point clouds. Currently a centre of mass approach is applied in Matlab, enabling on the fly video-rate visualisation. However, it is anticipated that the approach could be readily implemented in hardware - such as FPGA - in a computationally modest form, with negligible processing latency (compared to the frame time).

A limitation of histograms with a low number of bins is an inherent trade-off between bin resolution and range. In characterisation experiments [1], centimetre accuracy could be obtained with a bin size of 500ps (using laser pulses spread over multiple bins). However, this then limits the range to 8ns (or 1.2 metres). Photons returns from outside this range appear in the histogram in a wrapped location. This problem can be addressed in different ways, including by resorting to the ROTDC mode of operation for range disambiguation, or by progressively reducing the bin width of the on-chip histogram (akin to the approach of [5]).

II. Experimental Data

As with any ToF system, the precision of depth estimates, and range attained, will depend strongly on the illumination parameters, as well as the receiver optics. In tests with a pulsed laser source (Picoquant LDH-Series 670nm laser diode emitting ≈120ps pulses at 60MHz with 40mW measured average optical power), and a 3.5mm/f1.4 objective lens providing a 25 degree diagonal field of view (Fig. 1), the system was able to achieve sub-cm precision over a short distance range, whilst maintaining high frame rates. For example, with a target at a distance of 2m (a flat surface), and 500µs exposure time (860FPS), the precision (pooled standard of deviation of depth estimates across SPAD array) was 0.7cm for a <10% reflectivity target, and 0.2cm for ≈80% reflectivity. The standard deviation increased to 1.5cm, and 0.7cm, respectively, when ambient illumination with irradiance 0.7W/m^2 at the target was introduced to the scene (with no optical filter on the sensor).

The total chip power consumption P_{chip} was measured at <100mW. Assuming a laser wall-plug efficiency of 50%, P_{laser}≈80mW, which equates to the following system energy E_p requirement per macropixel for a single depth frame:

$E_p=(P_{laser}\times T_{exp}+P_{chip}\times T_{frame})/(N_{pixels})$=78nJ,

where T_{exp} is the exposure time, and T_{frame} is the total frame time. As a comparison, the megapixel indirect ToF of [6], running at 30fps, has E_p=30nJ. Ref. [7] proposes figures of merit for iToF in terms of responsivity, precision and background rejection, which, if generalised to dToF, could aid in making fuller comparisons.

Example 3D data are given of fast evolving scenes (Figures 2-6). The data were captured either in the 16-bin TCSPC mode (at 64×32 resolution), or a hybrid configuration alternating between TCSPC, and photon counting (256×128 resolution) frames. Global shutter was used throughout. The first sequence (Fig. 2) is of bursting a balloon, the camera successfully capturing the moment the balloon ruptures. The second sequence (Fig. 3-4) shows a table tennis ball hitting a plate of milk, the milk surface (prior to being perturbed) showing a strong reflection from the background wall (resulting in a second peak in the TCSPC histogram). A third sequence (Fig. 5) captures a 1000RPM fan; the double peak content of macropixels at the blade edge is highlighted. Fig 6. has depth and intensity images of a person juggling, captured indoors and outdoors (with an optical filter). The results indicate that high frame rates can be attained even under strong sunlight.

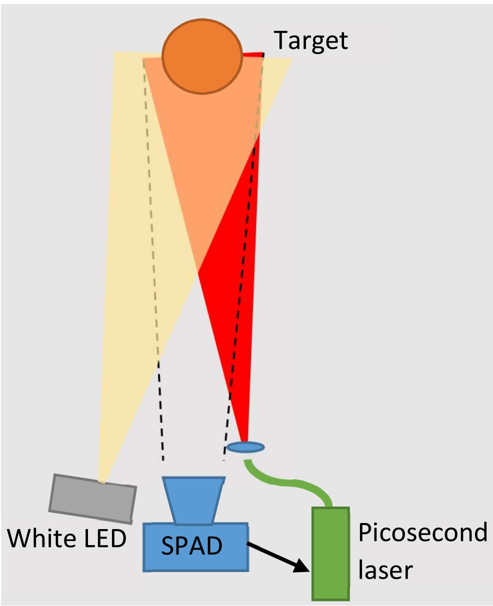


Figure 1: Setup for 3D data capture. In addition to the laser, the scene was illuminated by a high intensity (continuous wave) white LED source to provide ambient illumination

III. Conclusions

High speed imaging results have been presented from a 3D-ToF SPAD sensor, with frame rates exceeding 1kFPS. The results were obtained over a distance range of a few meters, and further work is planned to extend operability to longer ranges (whilst maintaining high frame rates) by optimising the illumination source.

Acknowledgments

This work was supported by EPSRC through grants EP/M01326X/1, EP/S001638/1 and EP/L016753/1. The authors are grateful to STMicroelectronics and the ENIAC-POLIS project for chip fabrication.

References

[1] Henderson, R.K. et al. "A 256×256 40nm/90nm CMOS 3D-Stacked 120dB-Dynamic Range Reconfigurable Time-Resolved SPAD Imager" ISSCC 2019

[2] Dutton, N.A.W. et al. "A Time-Correlated Single-Photon-Counting Sensor with 14GS/S Histogramming Time-to-Digital Converter" ISSCC 2015

[3] Bronzi, D. et al. "SPADAS: A High Speed 3D Single-Photon Camera for Advanced Driver Assistance Systems" Proc. SPIE 9366, 2015

[4] http://www.ti.com/product/OPT8320

[5] Linder, S. et al. "A 252 × 144 SPAD Pixel FLASH LiDAR with 1728 Dual-clock 48.8 ps TDCs, Integrated Histogramming and 14.9-to-1 Compression in 180nm CMOS Technology" VLSI 2018

[6] Bamji, C.S. et al. "1Mpixel 65nm BSI 320MHz Demodulated TOF Image Sensor with 3.5µm Global Shutter Pixels and Analog Binning" ISSCC 2018

[7] Perenzoni, M. et al. "Figures of Merit for Indirect Time-of-Flight 3D Camera: Definition and Experimental Evaluation" Remote Sense. 3 (2011)

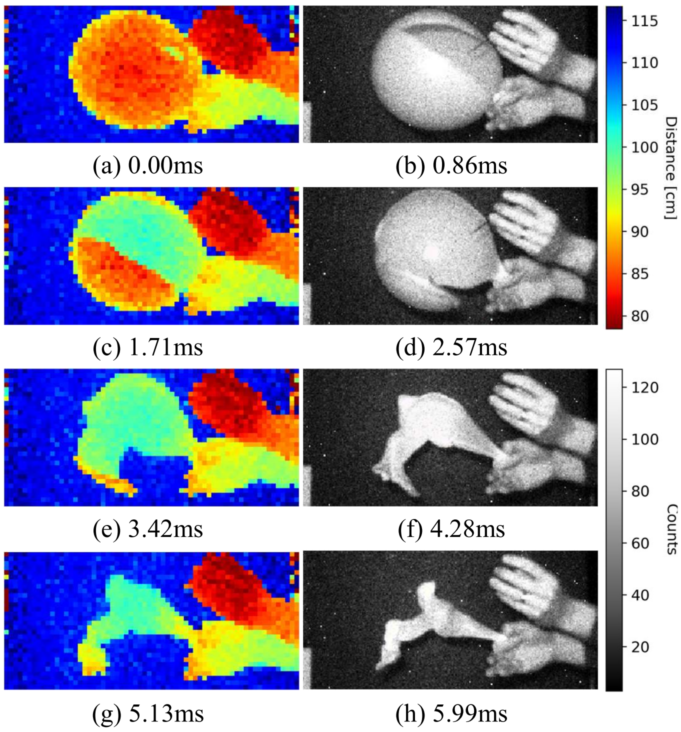
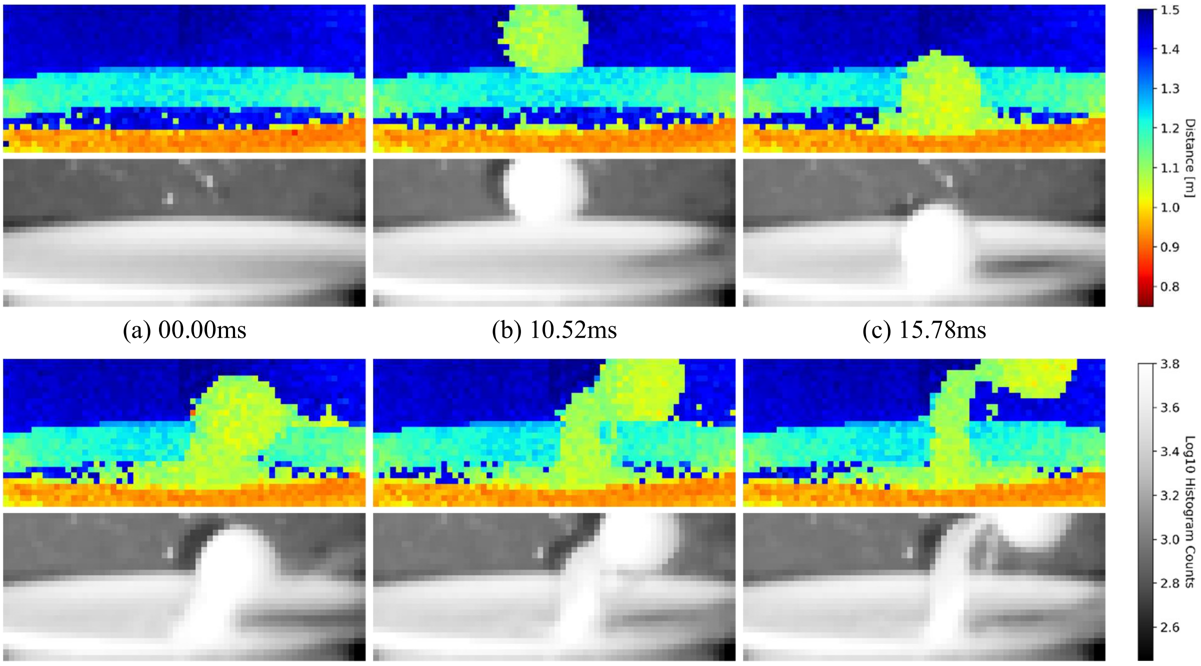


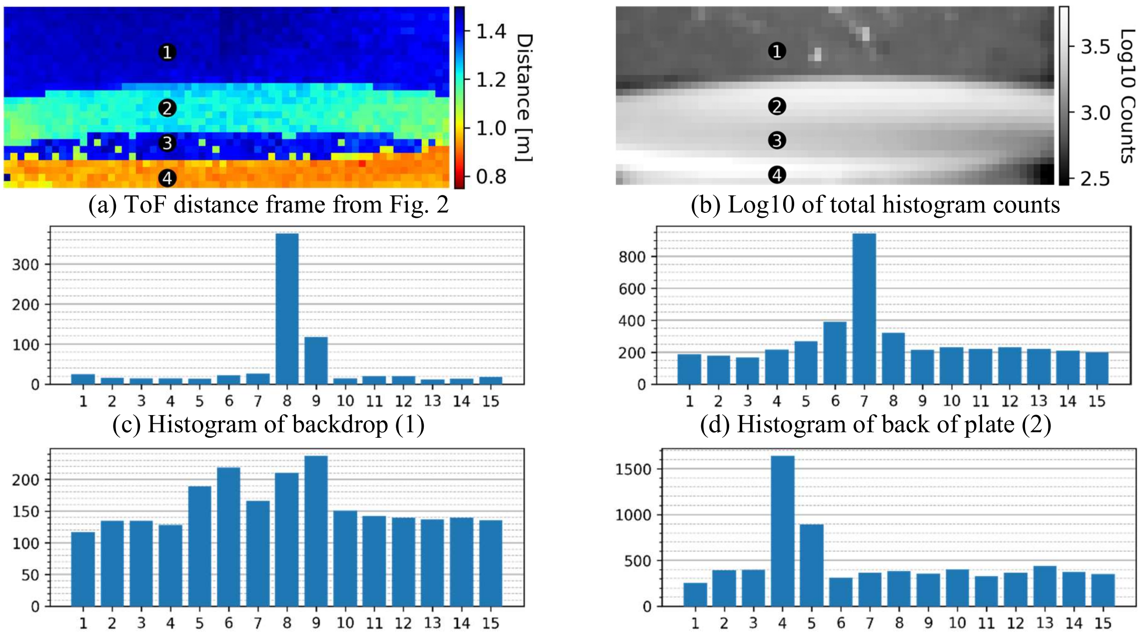
Figure 2: Alternating ToF/photon counting frames of a balloon popping with a 200µs exposure time (≈1170FPS).



(a) 00.00ms (b) 10.52ms (c) 15.78ms

(d) 28.93ms (e) 39.45ms (f) 47.34ms

Figure 3: Frames extracted from a 380FPS ToF video of a table tennis ball being dropped on a plate of milk (1ms exposure time was used).



(a) ToF distance frame from Fig. 2

(b) Log10 of total histogram counts

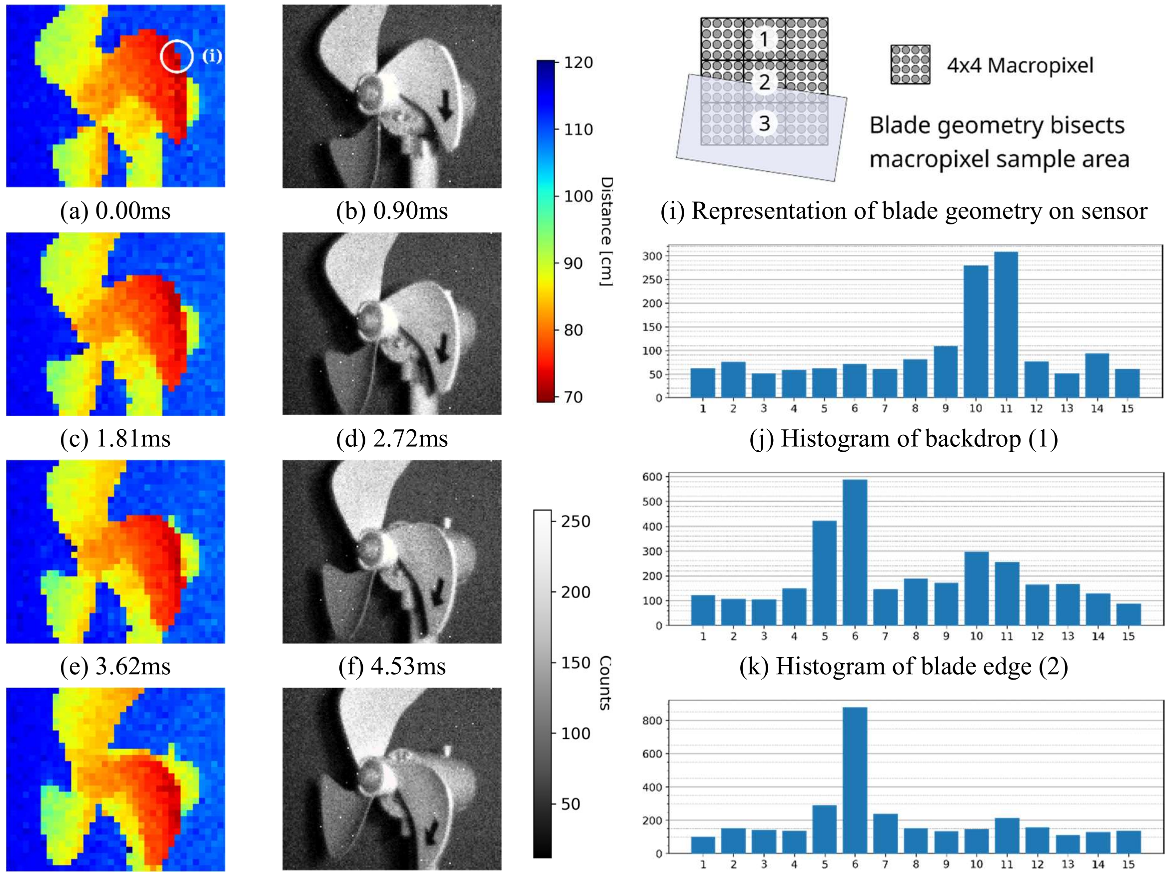
(c) Histogram of backdrop (1)

(d) Histogram of back of plate (2)

(e) Histogram of milk surface (3)

(f) Histogram of front of plate (4)

Figure 4: Example histograms for “milk and ball” sequence. The data was captured without an optical filter in place, hence the considerable background level on the histograms. It is interesting to note that the photon returns from the milk (panel e) show a second peak (due to specular re-reflection of laser photons reflected off the wall). This may have implications in LIDAR imaging in wet conditions.



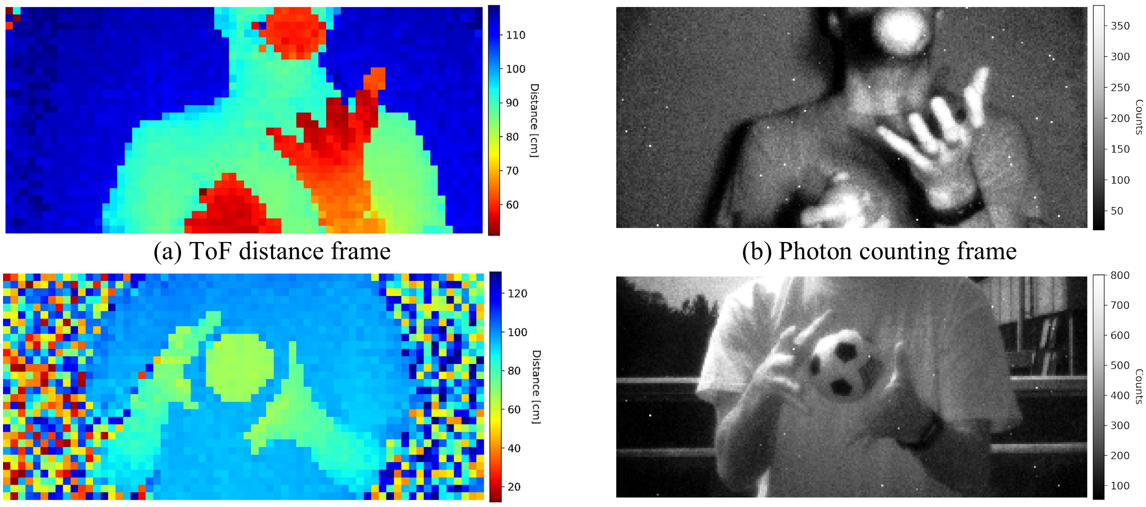
(a) 0.00ms (b) 0.90ms (i) Representation of blade geometry on sensor

(c) 1.81ms (d) 2.72ms (j) Histogram of backdrop (1)

(e) 3.62ms (f) 4.53ms (k) Histogram of blade edge (2)

(g) 5.43ms (h) 6.34ms (l) Histogram of blade body (3)

Figure 5: Eight consecutive frames of a 1000RPM fan swapping between ToF and photon counting capture modes both with 250µs exposure time (≈1100FPS). Panel k shows distinct histogram peaks for both the fan blade edge and the backdrop, with adjacent points shown in panels j and l (region indicated in panel a).



(a) ToF distance frame (b) Photon counting frame

(c) ToF distance frame (d) Photon counting frame

Figure 6: "Juggling person", captured indoors (panels a and b, 500µs exposure time, 860FPS, 1.6m stand-off distance) and outdoors (panels c and d, 300µs exposure time, 1050FPS, 3m stand-off distance). The outdoor scene was imaged under a clear, midday sun (the noise seen in the depth map around the person resulting, in part, from incomplete illumination of the scene).