

Microelectronic 3D Imaging and Neuromorphic Recognition for Autonomous UAVs

**Franco Zappa, Federica Villa, Rudi Lussana, Dennis Delic, Man Ching
Joyce Mau, Jean-Michel Redouté, Simon Kennedy, Daniel Morrison,
Mehmet Yuce, Tuncay Alan, Tara Hamilton, and Saeed Afshar**

Abstract The article addresses the development of highly sensitive, low-light and efficient, miniature single-photon sensor technology based on Single Photon Avalanche Diode (SPAD) arrays, its integration on a Flash Light Detection and Ranging (LiDAR) system mounted on a custom built multi-rotor Unmanned Aerial System (UAS) platform, for the collection of real time imagery and performance of neuromorphic processing for accurate target detection and classification.

17.1 Introduction

An international team of researchers have been developing highly sensitive, low-light and efficient, miniature single-photon sensor technology based on Single Photon Avalanche Diode (SPAD) arrays. A key motivation was to use Silicon CMOS-based processes and advanced 3D-IC manufacturing technologies to miniaturise arrays and digital circuits to realize affordable high definition imaging microchip sensors. Imaging cameras using smart photon sensor SPAD microchips are integral to active electro-optic systems such as 3-D Flash LiDAR (Light Detection and Ranging) for target detection and identification as well as tactical applications requiring imaging in very low light conditions. When coupled with sophisticated machine learning algorithms the work demonstrated accurate detection and classification of land-based targets from a low cost Unmanned Aerial System (UAS).

3-D Flash LiDAR systems, also known as 3-D Time of Flight (ToF) cameras which use ‘SPAD array’ sensor technology have some advantages over existing LiDAR scanning methods. They have no moving mechanical parts and scanning optics; hence they acquire a 3D depth-resolved image of a scene instantaneously allowing faster image reconstruction, which is especially useful when targets are moving or when large areas need to be surveyed quickly. They also offer an improved SWaP (Size Weight & Power) footprint over scanning systems, which means they can fit easily on power starved and mobile platforms such as UASs.

The project was able to develop state of the art SPAD sensors, successfully integrate and fly a low SWaP Flash LiDAR system on a custom built multi-rotor UAS platform, collect real time imagery and perform neuromorphic processing for accurate target detection and classification. The technology holds the potential to be developed further, both in the ability to image at greater ranges and its application in more challenging environments where targets are camouflaged and/or hidden by obscurants and clutter. The work has effectively showcased how such technology can be carried by low cost UAS’s for rapid environmental assessment of threats in low light conditions, thus providing enhanced situational awareness and decision superiority for surveillance and reconnaissance applications.

The four main project Streams are described in the subsequent sections.

17.2 Stream 1: High Density SPAD Array Design and Implementation

The main goal in this Stream was the design of high-end arrays of SPAD detectors, optimized in terms of number of pixels, detection performance, such as noise and sensitivity, and 3D ranging precision. As a first prototyping activity within this

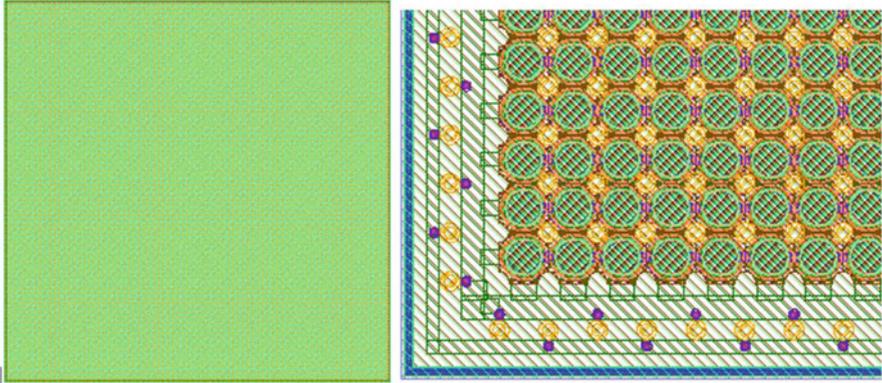


Fig. 17.1 Entire layout (left) of a 256×256 SPAD sensor array chip, with overall dimensions of 11×11 mm, and (right) zoom in of a corner

project, a set of individual SPAD test structures and miniarrays were designed in the high-voltage CMOS $0.35 \mu\text{m}$ Fraunhofer IMS technology, to characterise the fabrication process: these designs were based on previously reported prototypes [7]. Next, four different versions of 256×256 sensor arrays were fabricated by Fraunhofer IMS. The entire layout is shown in Fig. 17.1: the overall dimensions of the IC are 11×11 mm, where each pixel has a size of $40 \times 40 \mu\text{m}$. All array pixels have $40 \mu\text{m}$ pitch and comprise a SPAD and space for a TSV (through-silicon-via) connected with the anode, whereas the cathode voltage is common among all pixels. Four 8-inches wafers were fabricated by Fraunhofer IMS. These wafers with the different “SPAD chips” and the wafers of the “SPAD front-end chip” described in next section, were sent to Fraunhofer IZM for 3D-IC assembly using TSVs and bonding.

As a monolithic alternative to the 256×256 pixel 3D-IC ensemble imager pioneered in this project, Prof. Zappa’s research team also designed a 32×32 SPAD ToF imager, which is an improved version of the SPAD array published in [8]. Each of the 1024 pixels contains a SPAD detector, an analog front end (for avalanche sensing, detector quenching and digital pulse shaping), an 8 bit digital counter (for photon-counting), and a 12 bit Time-to-Digital Converter (TDC) covering a full-scale range of $1.2 \mu\text{s}$ and a least significant bit corresponding to 310 ps. Figure 17.2 shows the chip, with overall dimensions equal to 8×8 mm. Figure 17.3 shows the resulting ToF camera, based on the 32×32 SPAD imager shown in Fig. 17.2. The weight of the body is about 300 grams, and the dimensions are 6 cm width, 8.5 cm length, and 6 cm height: the camera has been used for LiDAR measurements with a full-scale range up to 200 m and a single shoot resolution of about 5 cm.

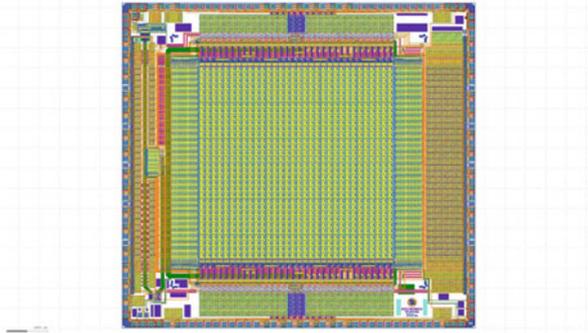
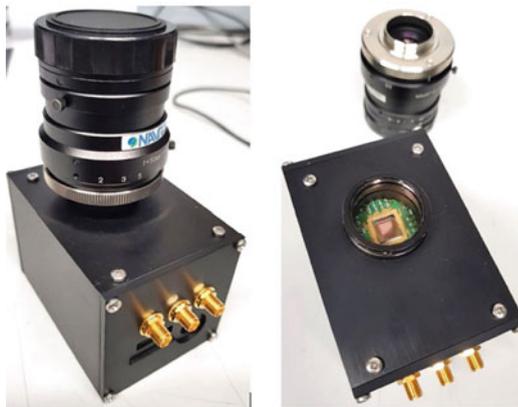


Fig. 17.2 CAD layout of the 32×32 POLIMI SPAD and TDC imager

Fig. 17.3 ToF camera designed by POLIMI using the 32×32 SPAD and TDC imager



17.3 Stream 2: High Speed Precision SPAD Front-End Design with On-Chip Photon Correlation

Stream 2 dealt with the development of integrated SPAD read-outs and front-ends with a high accuracy and resolution, and reduced chip size. As a first challenge, the “SPAD front-end chip” was designed, forming the read-out connected to the “SPAD chip” in a 3D-IC assembly process (described in the previous section). This chip contains all the necessary electronic circuits for the SPAD quenching, time measurement, photon counting and read-out of the 256×256 pixel array: each pixel contains a SPAD frontend, a counter and an interpolator, as well as input signal and clock buffering. Each SPAD front-end pixel fits on a $40 \times 40 \mu\text{m}$ pitch (matched to the “SPAD chip”). In ToF mode and at the beginning of a frame, the chip is triggered so that all pixel counters begin counting from the zero at the global clock rate. Once the SPAD detects a photon event the front-end latches on the first event. The output of the SPAD front-end is synchronized to the global clock which stops the counters, giving a final count value equal to the number of clock cycles between the start

of the frame and the first SPAD event. To improve the timing resolution beyond one clock cycle, the interpolator measures the synchronization time from the SPAD event to the counter stopping, giving a negative modifier to the measured time [5]. In photon counting mode, the pixel counter is reset at the start of the frame, but the SPAD front-end is set to self-rearm and outputs a pulse for every detected photon. The dedicated counter, requiring a minimal chip footprint then counts up one value after every SPAD event [4].

A separate 64×64 SPAD imager, containing the SPADs and the electronic front-end, was designed using the Silterra 130nm HV process, measured and implemented into a standalone camera (Fig. 17.4). The camera was measured on the DSTG laser range using an external laser, in collaboration with the DSTG team. The laser range is a 1.5 km field at DSTG Edinburgh (Adelaide) allowing testing of the camera over long distances with a high level of background light noise. The target setup and resulting image from the 250 m target are shown in Figs. 17.5 and 17.6.

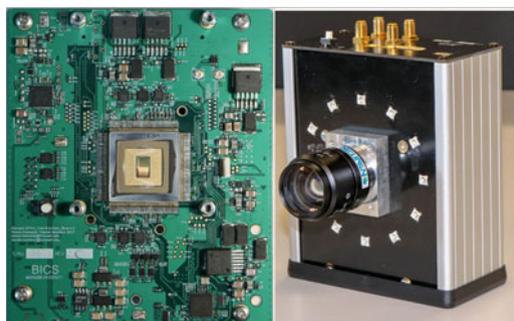


Fig. 17.4 (left) PCB interfacing with the Monash02 chips (64 by 64 SPAD array). The packaged 64×64 SPAD imager prototype is in the middle of the picture and covered with a glass lid. – (right) Photo of the camera using the 64×64 imager chip. A ring of illuminating LEDs placed on the front-panel can be used for short-range imaging

Fig. 17.5 Target setup at the 250 m target

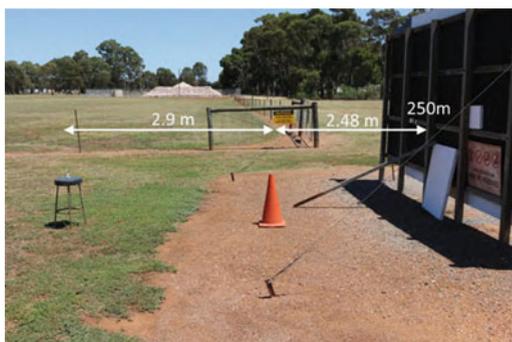
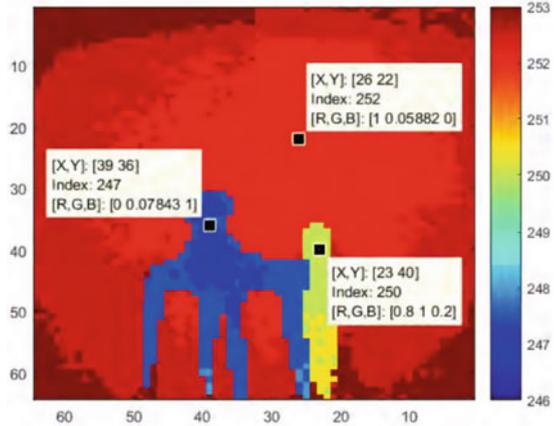


Fig. 17.6 Resulting image of the 250 m target



17.4 Stream 3: Neuromorphic Architectures of Event-Based SPAD Arrays and Neuromorphic, Event-Based Processing Algorithms

This Stream focused on developing algorithms and architectures that enable the conversion of SPAD array data into event-based signals which could subsequently be processed by event-based algorithms. These event-based algorithms are inherently neuromorphic, and are thus, low-power, real-time, and, most importantly for this medium, efficient, particularly in data generation, storage, and communications. The Recognition and Tracking System (RTS) algorithms that were developed in this project were based on previous work of the authors [1, 6]. Figure 17.7 illustrates some of the investigated scenarios.

These algorithms were then implemented in the design of a neuromorphic, event-based SPAD – neuroSPAD. In this new implementation, instead of encoding, storing and transferring the ToF data off chip for processing, the calculation of the ToF from the laser pulse is abandoned entirely in favor of a neuromorphic processor that operates directly in the time domain and on the inter-spike intervals within local regions of the SPAD array (Fig. 17.8). The proposed approach motivates the development and hardware implementation of event-based feature extraction algorithms and circuits that generate local sparse event-based representations from the non-sparse event based data and in this way drastically reduces the I/O requirements of the overall system [2].

The output of the neuroSPAD is, as stated above, significantly different to the interfaces used on other SPAD designs. An event-based processor has been developed to take the sparse, event-based output of the neuroSPAD and perform hardware efficient and low-power recognition on the resulting 3D image.

Figure 17.9 shows the functionality of the event-based processor. In Fig. 17.9a we see the raw SPAD output: using the neuroSPAD architecture a small (local) section of the SPAD image is represented with local binary events (Fig. 17.9b).

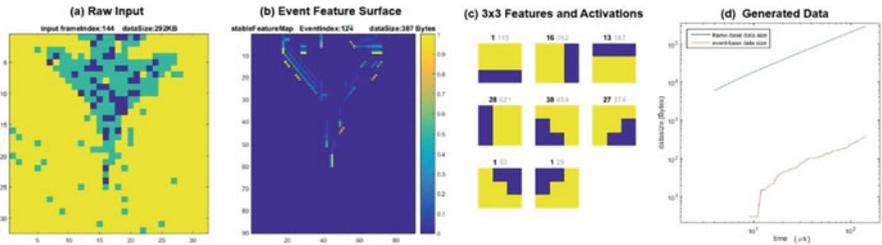


Fig. 17.7 Example of data generated from planes in motion in front of the SPAD camera: (a) raw input – (b) “surface” of the events (returning photons) – (c) subsequent 3×3 features extracted by the RTS algorithm – (d) amount of data generated by the developed event-based methodology (red curve) vs. traditional frame-based approaches (blue curve). It can be seen that the developed event-based method generates several orders of magnitude less data than traditional methods making it much more attractive for full integration as well as implementation in an autonomous system

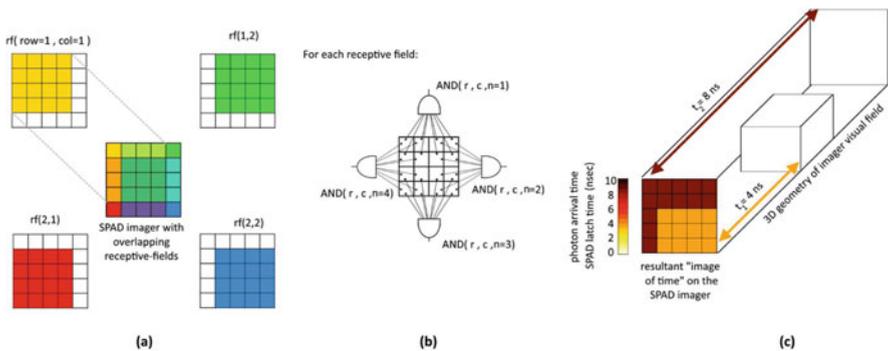


Fig. 17.8 Illustrative example of a 5×5 imager with four 4×4 receptive fields, (a). Each receptive field is connected to four AND gates – (b). An example 3D visual scene and the resultant SPAD timing pattern is shown in (c). For illustrative purposes we show a small imager with $5 \times 5 = 25$ pixels (instead of the chip’s 128×128). This imager uses four 4×4 overlapping receptive fields (same as chip). Each receptive field has four ANDs (same as chip). One of the four ANDs take as input the left/west 8 pixels of the 16 pixels of the 4×4 patch. Another AND takes as input the right/east 8 pixels, another the lower/south and another the top ones/north. If we consider the case shown in (c) with the 5×5 imager viewing a scene where the background is far away and there is a box in the foreground that is seen by pixels $(r=3:5$ and $y=2:5)$, then the lower/south AND of $rf(1, 2)$ (green) and $rf(2, 2)$ (blue) will latch which can be expressed as $AND(1, 2, 3)$ and $AND(2, 2, 3)$ latching at $t=4$ ns

The event output is temporally decayed, either linearly, exponentially, or in discrete time and then the dot-product is taken with a feature map – “layer 1” (Fig. 17.9c). The feature map is learned over time based on the 3D images that are presented to the neuroSPAD. The “learning” of the feature map can be disabled so that the feature map is set and no longer changing and thus representative of the images/objects that the system needs to classify. The resulting layer 1 surface (Fig. 17.9d) which comes from the dot product of the decayed input and the feature map can then be classified.

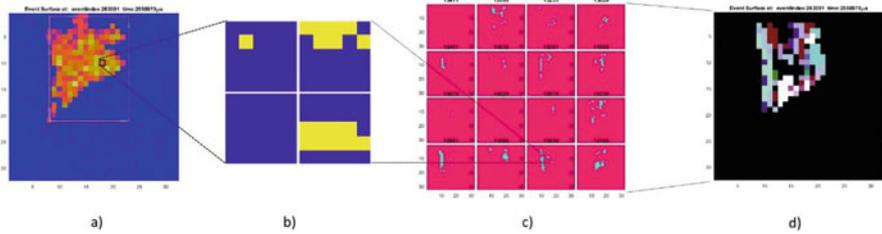


Fig. 17.9 Functionality of the neuromorphic, event-based processor. (a) Raw SPAD output. (b) Local SPAD event output. (c) Layer 1 activation. (d) Layer 1 surface

Measurements of the combined NeuroSPAD (or a conventional SPAD can be used in its place), RTS algorithm and neuromorphic, event-based processor, are presented in the next section.

17.5 Stream 4: SPAD Microlens Bonding and UAS Multi-copter Platform Development

A first challenge was to provide better sensitivity and selectivity of the SPAD arrays to the laser wavelength used for the LiDAR system. Various commercially available microlens arrays were evaluated and those found suitable were aligned and bonded on top of a previously designed SPAD imager (32×32 pixels), refer to Fig. 17.10a. This was done to improve the fill factor and hence sensitivity of the sensor. In addition, custom manufactured micro-lens array technology and techniques were developed which could be directly patterned to the SPAD array imaging sensors, see Fig. 17.10b, c.

A second challenge tackled in this Stream was to interface SPAD-based sensor technology (operating in ToF mode LiDAR mode) with neuromorphic-like image processing algorithmic techniques, to perform classification of imaged targets in real time, with improved target detection and recognition performance [9]. Indoor laboratory experiments of a low powered LiDAR system (a 100 mW 660 nm Coherent CUBE diode laser in conjunction with a 32×32 SPAD camera) and NVIDIA TX2 embedded board was used for data capture and image processing to verify operation of the real time classifier. Classification of 4 different model airplanes imaged by 32×32 SPAD camera with accuracy up to 98.7% was achieved. The power consumption of the embedded board was measured to be 5.1 W, which is ideal for low SWaP UAS applications [3].

Finally, DSTG was able to successfully integrate and fly a low SWaP, high powered Flash LiDAR system on a custom-built multi-rotor UAS platform (Fig. 17.11). In collaboration with partners a custom octo-copter multirotor aircraft was designed and developed. It could achieve a maximum take-off weight of 40 kg with full propulsion redundancy. All the components of a high-powered LiDAR system were

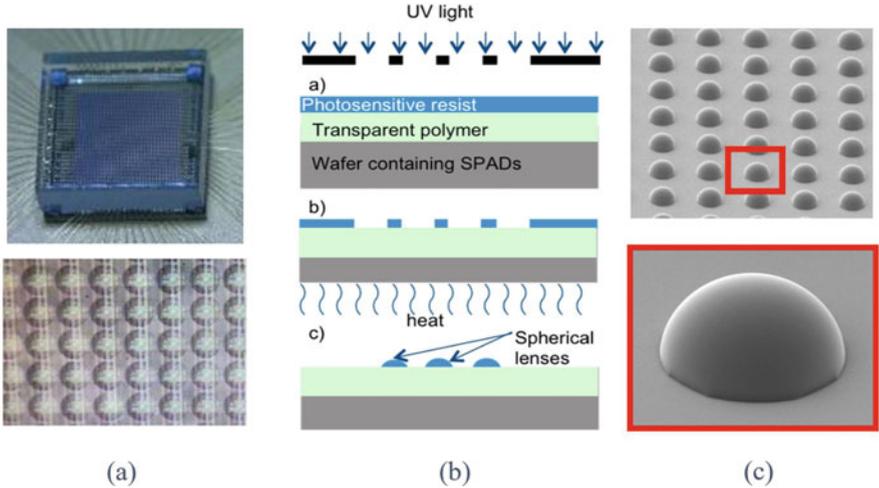


Fig. 17.10 (a) (Top) Commercial microlens array aligned and bonded to a 32×32 SPAD array sensor. (Bottom) Magnified top view showing each lens aligned directly above each SPAD detector. – (b) Customised polymeric microlenses were developed using lithographical techniques and applied directly to SPAD array sensor. Alignment $\pm 2 \mu\text{m}$ possible. – (c) SEM image of an individual microlens directly patterned on the SPAD array sensor



Fig. 17.11 (Left) – Photograph of the LiDAR system integrated and balanced onto a 3 axis Ronin Gimbal. (Right) – Complete system successfully operating in-flight

then integrated so it would fit on a 3-axis Ronin gimbal and be balanced (i.e. not prone to oscillations or drift). This included: a 40 mJ 532 nm Arete laser, optics, 32×32 SPAD camera, optical camera and power electronics. The total payload weight (including gimbal) was measured to be 7.5 kg. Various successful tests were conducted to ensure the gimbal was controllable, balanced and immune to vibration. The power conversion, wireless communication and kill switch (for eye safety precautions and if the UAS was no longer obeying commands) hardware and corresponding software interfaces to enable integration of the LiDAR fitted gimbal to the UAS platform were developed at DSTG.

References

1. Afshar S, George L, Thakur CS, Tapson J, van Schaik A, de Chazal P, Hamilton TJ (2015) Turn down that noise: synaptic encoding of afferent SNR in a single spiking neuron. *IEEE Trans Biomed Circuits Syst* 9(2):188–196
2. Afshar S, Hamilton TJ, Tapson J, van Schaik A, Cohen G (2019) Investigation of event-based surfaces for high-speed detection, unsupervised feature extraction, and object recognition. *Front Neurosci* 12:1047
3. Mau J, Afshar S, Hamilton TJ, Schaik A, Lussana R, Panella A, Trunpf J, Delic D (2019) Embedded implementation of a random feature detecting network for real time classification of time-of-flight SPAD array recordings. In: *SPIE defence + commercial sensing conference*
4. Morrison D, Delic D, Yuce MR, Redouté J-M (2019) Multistage linear feedback shift register counters with reduced decoding logic in 130 nm CMOS for large scale array applications. *IEEE Trans Very Large Scale Integr Syst* 27(1):103–115
5. Morrison D, Kennedy S, Delic D, Yuce MR, Redouté J-M (2018) A triple integration timing scheme for SPAD time of flight imaging sensors in 130 nm CMOS. In: *Proceedings of IEEE international conference on electronics, circuits and systems*, pp 13–16
6. Sofatzis RJ, Afshar S, Hamilton TJ (2014) The synaptic kernel adaptation network. In: *IEEE international symposium on circuits and systems (ISCAS)*, pp 2077–2080
7. Villa F, Bronzi D, Zou Y, Scarcella C, Boso G, Tisa S, Tosi A, Zappa F, Durini D, Weyers S, Paschen U, Brockherde W (2014) CMOS SPADs with up to 500 μm diameter and 55% detection efficiency at 420 nm. *J Mod Opt* 61(2):102–115
8. Villa F, Lussana R, Bronzi D, Tisa S, Tosi A, Zappa F, Mora AD, Contini D, Durini D, Weyers S, Brockherde W (2014) CMOS imager with 1024 SPADs and TDCs for single-photon timing and 3-D time-of-flight. *IEEE J Sel Top Quantum Electron* 20(6):364–373
9. Woods WF, Delic DV, Smith BW, Swierkowski L, Day GS, Devrelis V, Joyce RA (2019) Object detection and recognition using laser radar incorporating novel SPAD technology. In: *Proceedings of SPIE 11005, laser radar technology and applications XXIV*, vol 1100504