A CMOS 3D Camera with Millimetric Depth Resolution

Cristiano Niclass, Alexis Rochas, Pierre-André Besse, and Edoardo Charbon
Swiss Federal Institute of Technology, Lausanne
Switzerland

Abstract- *A 3D imager is presented capable of capturing the depth map of an arbitrary scene. Depth is measured by computing the time-of-flight of a ray of light as it leaves the source and is reflected by the objects in the scene. The round-trip time is converted to digital code independently for each pixel using a CMOS time-to-digital converter. To reach millimetric accuracies an array of 32x32 highly sensitive, ultra-low jitter CMOS detectors capable of detecting a single photon is used. The scene is illuminated using a cone of low power pulsed laser light, thus no mechanical scanning devices or expensive optical equipment are required.*

1. Introduction

A number of novel and potentially high volume applications requiring fast and precise depth map evaluation have recently appeared. These applications include face recognition systems, virtual keyboards, object and person monitoring, land and sea surveyors, virtual reality games, non-ionising medical tomographic imagers, stage and choreography analysis tools, etc. Speed, compactness and especially cost concerns have prompted the emergence of a new generation of solid-state imagers.

Three main techniques have been proposed: triangulation, interferometry, and time-of-flight (TOF). In triangulation systems, distance to a precise point in the scene is derived from the angle of incidence of a known point source and the angle of reflected light [1]. The main disadvantages of such systems are the speed requirement on the sensor (usually a conventional CMOS or CCD camera), power dissipation, and a somewhat limited precision. Interferometry is being used for the high levels of accuracy it ensures. However, interferometers are usually bulky and very expensive [2].

Two variants of TOF methods are available today: those based on modulated and those based on pulsed laser sources. Modulation type TOF range finders measure the phase difference between a modulated laser source and the reflected wave. High modulation frequencies in conjunction with homodyne phase discrimination and averaging at the pixel level can be used to relax circuit specifications [3],[4]. However, relatively powerful laser or LED sources are still required and accuracy is limited by the speed at which the sensor can be clocked. In pulsed type TOF the round-trip time of a single burst of light is measured. The main advantage of this method over modulated type TOF is that a range of operation of a few meters to several kilometres can be achieved avoiding the use of different modulation frequencies.

In pulsed type TOF sub-millimetric resolutions require (a) sub-picosecond time discrimination capability, (b) high detection speed and/or (c) low ray divergence. If conventional detectors ought to be used, time discrimination can only be achieved through highly sophisticated detectors and/or very powerful laser sources [5]. Complex non-CMOS technologies are a good alternative but they generally prevent integration of accurate time discriminators on chip, unless multiple technologies are combined [6]. However this is usually an expensive proposition. Furthermore, the specification on the optical components generally leads to laboratory-stile optical tables.

The system proposed in this paper is a solid-state 3D imager based on the pulsed type TOF method. The sensor, implemented in standard CMOS technology, consists of an array of single photon avalanche diodes, capable of performing 1024 independent distance measurements, and an external low-cost CMOS time discriminator. The scene is illuminated using a single intentionally uncollimated laser beam to create a cone of light that synchronously reaches every point of the surface. This technique eliminates the need for a mechanical device to scan the entire scene [7]. Unfortunately, the optical power reaching each pixel becomes inversely proportional to the square of depth. This drawback is mitigated by very high detector sensitivity. Therefore, inexpensive laser sources with milliwatt peak power can be used for ranges up to several meters.

A time precision of a few tens of picoseconds is made possible by the jitter properties of single photon avalanche diodes when operating in Geiger mode. The 3D camera system is capable of characterizing the optical field by mapping both the depth of the scene as well as the intensity of the light reflected by it. Extremely tight speed and accuracy specifications are met, while keeping power dissipation to a minimum thanks to the inherent simplicity of the sensor design. Moreover, due to the reduced power of the laser source, it is possible to guarantee strict eye safety operation of the system even with a duty cycle of 100%.

The paper is organized as follows. The digital pixel is described in Section 2, while the system architecture is outlined in Section 3. The optical setup, capture methods and measurements are discussed in Section 4.

2. The 5T Digital Pixel

Operated in the so-called Geiger mode, avalanche photodiodes can count single photons [8]. A single photon avalanche diode (SPAD) is a p-n junction biased above breakdown voltage V_{bd} by an excess voltage V_e of a few volts. A primary carrier resulting from the absorption of a photon may generate an infinite number of secondary electron-hole pairs by impact ionization. In [9], an 8x4 array of SPADs integrated in a conventional CMOS technology was presented. Its potential for 3D imaging was demonstrated in [10]. However, this solid-state sensor still suffered from a limited lateral resolution due to the reduced number of pixels.

In this paper, a 32x32 array was fabricated using a high voltage 0.8µm CMOS process. The fabrication process is a 2M/2P twin-tub technology on a p-substrate allowing operating voltage from 2.5 to 50V. The SPAD sketched in Figure 1 is a dual p+/deep n-tub/p-substrate junction. The upper p+/deep n-tub junction provides the multiplication region where the Geiger breakdown occurs.

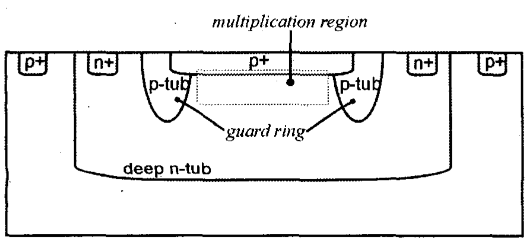

Figure 1. Single photon avalanche diode cross-section

For a SPAD to operate in avalanche mode, the central section of the junction must be protected against premature breakdown. This is achieved with a guard ring. A useful feature of this technology is the availability of a p-tub implantation to create a ring surrounding the p+ region anode [11]. The breakdown voltage V_{bd} of the p+/deep n-tub junction is typically 25.5V. A larger bias voltage $V_{bd}+V_e$ must be applied on the diode to operate in single photon detection mode.

The pixel consists of a circular SPAD and a 5T-electronic circuit as shown in Figure 2. The SPAD operates in passive quenching. The p+ anode is biased to a high negative voltage V_{p+} equal to –25.5V. This voltage is common to all the pixels in the array. The deep n-tub cathode is connected to the power supply VDD =5V through a long channel p-mos transistor T_q. The excess bias voltage is thus equal to $|V_{p+}|$ + VDD - V_{bd} = 5V.

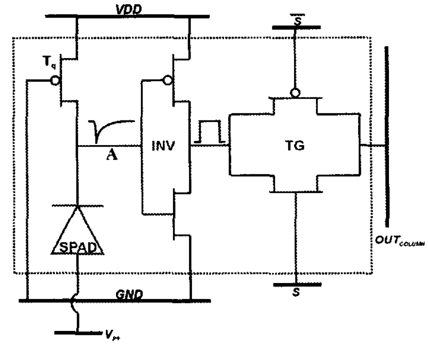

Figure 2. Schematic of 5T digital pixel

Upon photon arrival, the breakdown current discharges the depletion region capacitance, reducing the voltage across the SPAD to its breakdown voltage. The aspect ratio of T_q is set in order to provide a sufficiently resistive path to quench the avalanche breakdown. In our design, the channel resistance is larger than 200kΩ for drain-source voltage V_{ds} between GND and VDD. After avalanche quenching, the SPAD recharges through T_q and progressively recovers its photon detection capability.

The time required to quench the avalanche and restore the operating voltage is known as the dead time. It is less than 40ns for the 5T digital pixel. At node A, an analog Geiger pulse of amplitude V_e reflects the detection of a single photon. The inverter stage *INV* converts this Geiger pulse into a digital pulse. The two transistor aspect ratios of *INV* are designed to set the input threshold voltage at 3.5V. The transmission gate *TG* feeds the detection signal to the column output line when s=VDD and $\bar{s}$=GND (read phase). The pixel outputs a digital signal reflecting a photon arrival with picosecond precision. The near-infinite internal gain inherent to Geiger mode operation leads to no further amplification and the pixel output can be routed directly outside the chip. Figure 3 shows the layout of a pixel consisting of the photodiode, the quenching circuitry and the column access.

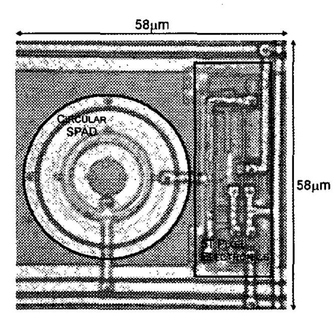

Figure 3. Photomicrograph of SPAD pixel

3. Sensor Architecture

The functional diagram of the sensor is shown in Figure 4. The sensor array consists of 32x32 pixels and requires two power supply buses *VDD*=5V and V_{p+}= -25.5V.

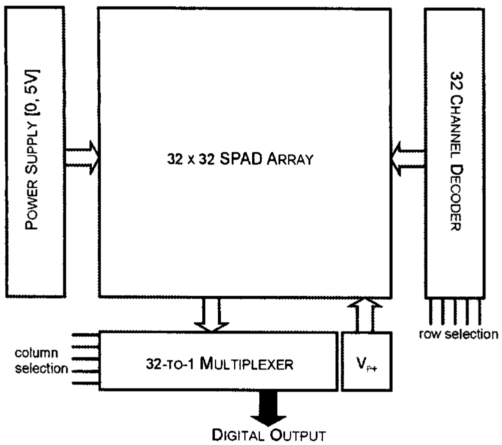

Figure 4. Functional diagram of the sensor

The readout circuitry consists of a 32-channel decoder for row selection and a 32-to-1 multiplexer for column selection. A photomicrograph of the chip is shown in Figure 5. The chip size is 7mm^2. The 5T digital pixel occupies a square area of 58µmx58µm.

At V_e=5V and room temperature, the photon detection probability (PDP), i.e. the probability for a photon impinging the photodiode to be detected, is larger than 20% between 420nm and 600nm with a peak at 28% at 470nm [11]. At 700nm the PDP is still 13% without any post process treatment of the silicon chip interface.

In 3D imaging picosecond timing precision is crucial. Timing jitter reflects the statistical fluctuations of the time interval between the arrival of the photon at the sensor and the detection output pulse leading edge. In a SPAD, the timing jitter mainly depends on the time it takes for a photogenerated carrier to be swept out of the absorption zone into the multiplication region. The micrometric thickness of the depletion region inherent to the CMOS fabrication process leads to a remarkable timing jitter at less than 50ps FWHM [11].

In a SPAD, thermal and tunneling generation effects produce pulses even in the absence of illumination. These pulses define the dark count rate (DCR) of the detector, which includes primary and secondary counts. Primary dark counts are due to carriers thermally generated or tunneling generated in the depletion region. Secondary dark counts are due to afterpulsing effects. During an avalanche breakdown, some carriers may be captured by deep levels in the forbidden band and subsequently released with a probability to re-trigger an avalanche. At room temperature, the limited active area of the SPAD and the outstanding cleanliness of the CMOS process lead to a mean value of the DCR of only 350Hz on the whole sensor array and negligible afterpulsing effects.

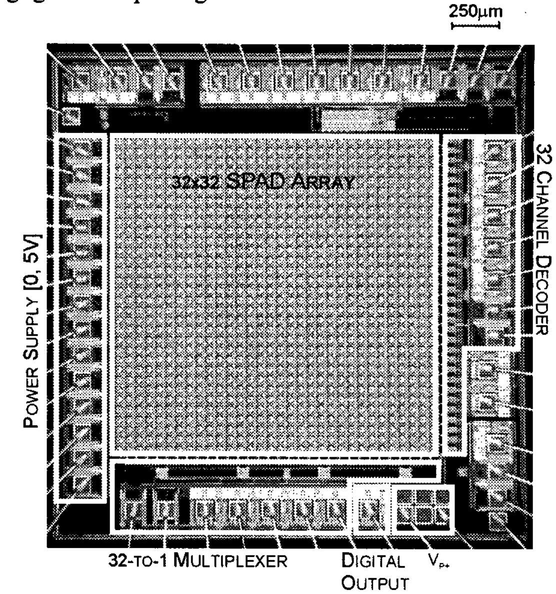

Figure 5. Photomicrograph of the complete sensor

4. Capture Setup and Measurements

The electro-optical setup is described in Figure 6. The CMOS sensor was mounted on a printed circuit board and equipped with a standard camera objective.

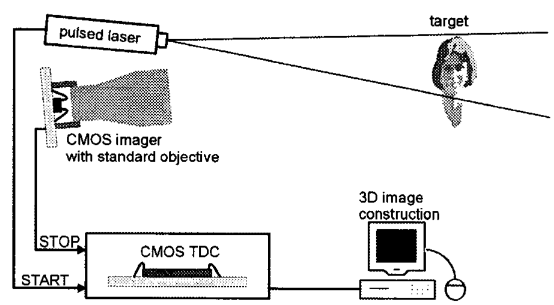

Figure 6. 3D camera set-up (not to scale)

The target was hit by uncollimated 635nm light generated by a laser source. The source was pulsed at a repetition rate f_R=50MHz, width T_P=150ps, and peak power P_S=100mW. A time to digital converter (TDC), TDC-F1 model from ACAM, Germany, was used for TOF measurement. The START signal was given by the synchronization output of the laser, the STOP signal by the digital output of the 32x32 SPAD array. The TDC, implemented in a standard CMOS technology, exhibited a timing resolution of 120ps and a 1-σ timing uncertainty of 110ps. Pixel addressing was performed by an external acquisition card.

A total timing uncertainty $\sigma(\tau)$ of 300ps was estimated for the sensor. $\sigma(\tau)$ is dominated by the TDC precision and the skews of non-uniform pixel-to-TDC propagation paths. To reach millimetric precision a statistical metrological approach was adopted, whereby a number M of measurements were averaged over the integration period. Figure 7 shows multiple depth measurements with M=100, 10^3, and 10^4 at and around 3.5m distance. The 1-σ uncertainty is also reported.

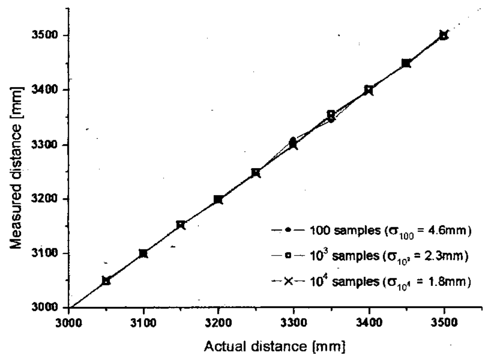

Figure 7. Distance measurements

Next, the depth map of a live-sized human mannequin was captured. Figure 8 shows a high-resolution picture of the model. To demonstrate the actual 32x32 pixel resolution, an intensity map of the model, obtained with our camera operating in 2D mode, is also shown in the figure.

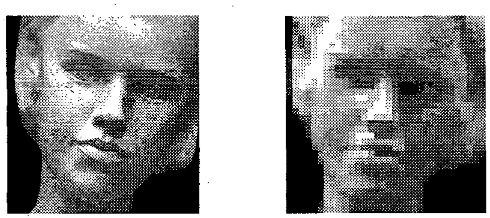

Figure 8. Model photographed with standard camera (left) and with our camera operating in 2D mode (right)

Figure 9 shows the depth map of the model's face and profile. The model was placed in front of a plane reference at 3m from the sensor. The image was obtained using the same light source parameters of the above experiments with $M=10^4$. Note that the mean optical power was only 750μW.

5. Conclusions

A 3D camera has been proposed based on an array of 1024 CMOS single photon avalanche diodes for a number of emerging applications. The depth map is constructed by illuminating the scene with a cone of pulsed laser light and detecting the time-of-flight required by photons to complete a round-trip. Extremely reduced dark count rate and time jitter enabled high precision while at the same time allowing the use of low power laser sources and no mechanical scanning. For the first time a CMOS time-to-digital converter was used in a camera of this kind, thus allowing one to envisage high-precision high-range 3D cameras integrated on a single chip.

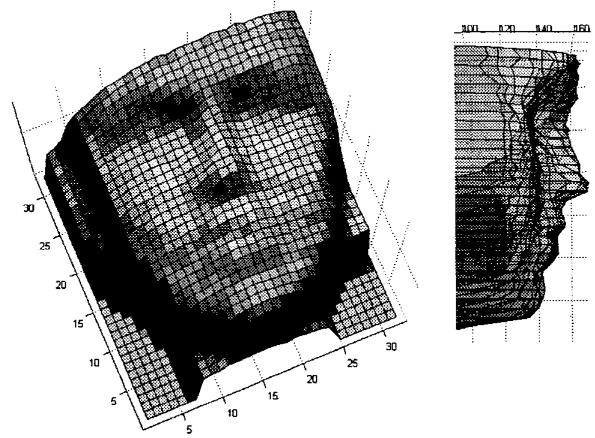

Figure 9. Human face depth map and profile (in mm)

Acknowledgements

This research was supported by a grant of the Swiss National Science Foundation – Grant Nr.: 620-066110.

References

[1] S. Yoshimura, T. Sugiyama, K. Yonemoto, K. Ueda, "A 48kfram/s CMOS Image Sensor for Real-Time 3-D Sensing and Motion Estimation", Proc. ISSCC, pp. 94-95, February 2001.

[2] B. Saleh, M. Teich, "Fundamentals of Photonics", John Wiley & Sons, 1991.

[3] R. Lange, "3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD-Technology", Ph.D. Thesis, ETH-Zürich, 2000.

[4] E. Charbon and C. Bamji, "Methods for CMOS-Compatible Three-Dimensional Image Sensing Using Quantum Efficiency Modulation", U.S. Patent 6,515,740, February 2003.

[5] R. Jeremias, W. Brockherde, G. Doemens, B. Hosticka, L. Listl, P. Mengel, "A CMOS Photosensor Array for 3D Imaging Using Pulsed LASER", Proc. IEEE ISSCC, pp. 252-253, February 2001.

[6] B.F. Aull et al., "Geiger-Mode Avalanche Photodiodes for Three Dimensional Imaging", Lincoln Laboratory Journal, Vol. 12, No. 2, 2002, pp. 335-350.

[7] J. Massa, G. Buller, A. Walker, G. Smith, S. Cova, M. Umasuthan, A.M. Wallace, "Optical Design and Evaluation of a Three-Dimensional Imaging and Ranging System based on Time-Correlated Single-Photon", Applied Optics, Vol. 41, No. 6, pp. 1063-1070, February 2002.

[8] A. Rochas, "Single Photon Avalanche Diodes in CMOS technology", Ph.D. Thesis, EPF-Lausanne, 2003.

[9] A. Rochas, M. Gösch, A. Serov, P.A. Besse, R.S. Popovic, T. Lasser, R. Rigler,"First Fully Integrated 2-D Array of Single-Photon Detectors in Standard CMOS Technology", IEEE Photonics Technology Letters, Vol. 15, No. 7, pp. 963-965, July 2003.

[10] C. Niclass, A. Rochas, P.A. Besse, and E. Charbon, "A CMOS Single Photon Avalanche Diode Array for 3D Imaging", Proc. IEEE ISSCC, pp. 120-121, February 2004.

[11] A. Rochas, M. Gani, B. Furrer, G. Ribordy, P.A. Besse, N. Gisin, and R.S. Popovic, "Single photon detector fabricated in a CMOS high voltage technology", Review of Scientific Instruments, Vol. 74, No. 7, pp. 3263-3270, 2003.