Arrays of Single Photon Avalanche Diodes in CMOS Technology: Picosecond Timing Resolution for Range Imaging (INVITED)

Cristiano Niclass, Pierre-André Besse and Edoardo Charbon*
Ecole Polytechnique Fédérale de Lausanne
CH-1015 Lausanne, Switzerland

Abstract
A solid-state imager fabricated in CMOS technology is presented for depth information capture of arbitrary 3D objects with millimeter resolution. The system is based on an array of 32x32 pixels that independently measure the time-of-flight of a ray of light as it is reflected back from the objects in a scene. A single cone of pulsed laser light illuminates the scene, thus no complex mechanical scanning is required. Millimetric depth accuracies can be reached thanks to the rangefinder's optical detectors that enable picosecond time discrimination. The detectors, based on a single photon avalanche diode operating in Geiger mode, utilize avalanche multiplication to enhance light detection. Optical power requirements on the light source can therefore be significantly relaxed. A number of standard performance measurements, conducted on the imager, are discussed in this paper. The 3D imaging system was also tested on real 3D subjects demonstrating the suitability of the approach.

Keywords
CMOS SPAD arrays, avalanche photodiodes, single photon detectors, photon timing, range image, time-of-flight, range-imaging camera, rangefinder, 3D vision.

1 INTRODUCTION

Until recently, cost, size, and eye safety concerns have relegated imaging systems capable of capturing 3D images to very specific applications, mostly in Computer Vision. The emergence of low-cost digital imagers has however reversed this trend, thus accelerating the push to create low-cost and compact 3D native vision systems. Consequently, a number of novel and potentially high volume applications requiring fast and precise depth map evaluation are currently under investigation. Examples of such applications include face recognition, virtual human-computer interfaces, object and person monitoring, land and sea surveying, virtual reality games, non-ionizing medical tomographic imagers, etc. Cost concerns have prompted the emergence of new generations of solid-state imagers that can achieve the desired accuracy and speed in compact and user-safe implementations.

3D vision techniques not based on stereoscopy and interferometer can be divided into two main classes: triangulation and time-of-flight (TOF). Systems based on triangulation evaluate depth by illuminating a tiny spot of the scene with a laser ray. The distance of that spot to the sensor is inferred from the known angle of incidence of the reflected light [1]. The main disadvantage of such systems is the speed requirement on the sensor (usually a conventional CMOS or CCD camera), power dissipation, and the need for highly tuned precision moving parts.

Alternatively, TOF methods are currently available in two main variants: modulation or pulsed based methods. Modulation based TOF rangefinders measure the phase difference between a modulated laser source and the reflected wave. In these imagers, megahertz modulation frequencies, in conjunction with homodyne phase discrimination and averaging at the pixel level, have been successfully used to relax circuit specifications [2], [3]. The main problems related to this method are the need for a relatively powerful laser or LED source and a limited range. In addition, accuracy is limited by the speed at which the sensor can be clocked. In pulsed type TOF methods, the round-trip time of a single burst of light is measured. The main advantage of these methods over modulated type TOF is that a range of operation of a few meters to several kilometers can be achieved avoiding the use of complex modulation schemes. In order to achieve millimetric accuracies, pulsed type TOF rangefinders must have picosecond time discrimination capability. To date, conventional CMOS detectors can approach this requirement only at a cost of highly sophisticated detection schemes and very powerful laser sources [4]. Complex non-CMOS technologies are a good alternative but they generally prevent integration of accurate time discriminators on chip, unless hybrid, potentially expensive solutions are adopted [5].

In reference [6], the potential of an 8x4 array of single photon avalanche diodes (SPAD) integrated in a conventional CMOS technology for 3D imaging was demonstrated. However, this solid-state sensor still suffered from a limited lateral resolution due to the reduced number of pixels. In this paper we report on a solid-state 3D imager based on the pulsed type TOF. The sensor is implemented in CMOS technology and it consists of an array of SPADs, capable of performing 1024 independent distance measurements. An external low-cost CMOS time discriminator based on a time-to-digital converter is employed to achieve the proper accuracies. A uniform illumination of the scene is accomplished using a single intentionally uncollimated laser beam. This method enabled us to achieve a very compact imaging solution that requires no mechanical micro-optical scanning device. The SPAD's high sensitivity enables the sensor to detect low levels of reflected light. Therefore, inexpensive laser sources with milliwatt power can be used for ranges up to several meters.

The jitter properties of SPADs allow a time discriminator to meaningfully measure picosecond time intervals. Depth and intensity can be evaluated by the same pixel, thus eliminating the potential alignment errors between the two measurements. Tight speed and accuracy specifications do not impact system complexity as in other approaches, while keeping power dissipation to a minimum thanks to the inherent simplicity of the sensor design. Not less importantly, due to the reduced power of the laser source, it is possible to guarantee strict eye safety operation of the system.

2 SINGLE PHOTON AVALANCHE DIODES

An avalanche photodiode reverse-biased above its breakdown voltage (V_{bd}) allows single photon detection [7]. When such a diode is biased above V_{bd}, it remains in a zero current state for a relatively long period of time, usually in the millisecond range. During this time, a very high electric field exists within the p-n junction forming the avalanche multiplication region. Under these conditions, if a primary carrier enters the multiplication region and triggers an avalanche process, several hundreds of thousands of secondary electron-hole pairs are generated by impact ionization, thus causing the diode's depletion capacitance to be rapidly discharged. As a result, a sharp current pulse is generated and can be easily measured. This mode of operation is commonly known as Geiger mode.

Unfortunately, typical photodiodes, as those used in conventional imagers, are not compatible with this mode of operation since they suffer from a premature breakdown when the bias voltage approaches V_{bd}. Premature breakdown occurs since the peak electric field is located only in the diode's periphery rather than in the planar region [7]. A single photon avalanche diode (SPAD), on the other hand, is a specifically designed photodiode in which premature breakdown is avoided and a planar multiplication region is formed within the whole junction area.

Linear mode avalanche photodiodes, which are biased just below V_{bd}, have a finite multiplication gain. Statistical variations of this finite gain produce an additional noise contribution known as excess noise. SPADs, on the other hand, are not concerned with these gain fluctuations since the optical gain is virtually infinite. Nevertheless, the statistical nature of the avalanche buildup is translated onto a detection probability. Indeed, the probability of detecting a photon hitting the SPAD's surface, known as the *photon detection probability (PDP)*, depends on the diode's quantum efficiency and the probability for an electron or for a hole to trigger an avalanche. Additionally, in Geiger mode, the signal amplitude does not provide intensity information since all the current pulses have the same amplitude. Intensity information is however obtained by counting the pulses during a certain period of time or by measuring the mean time interval between successive pulses. The same mechanism may be used to evaluate noise. Thermally or tunneling generated carriers within the p-n junction, which produce dark current in linear mode photodiodes, can trigger avalanche pulses. In Geiger mode, they are indistinguishable from regular photon-triggered pulses and they produce spurious pulses at a frequency known as *dark count rate* (DCR). DCR strongly depends on temperature and it is an important parameter for a rangefinder since it generates false TOF measurements.

Another source of spurious counts is represented by *after-pulses*. They are due to carriers temporarily trapped after a Geiger pulse in the multiplication region that are released after a short time interval, thus re-triggering a Geiger event. After-pulses depend on the trap concentration as well as on the number of carriers generated during a Geiger pulse. The number of carriers depends in turn on the diode's parasitic capacitance and on the external circuit, which is usually the circuit used to quench the avalanche. Typically, the quenching process is achieved by temporarily lowering the bias voltage below V_{bd}. Once the avalanche has been quenched, the SPAD needs to be recharged again above V_{bd} so that it can detect subsequent photons. The time required to quench the avalanche and recharge the diode up to 90% of its nominal excess bias is defined as the *dead time*. This parameter limits the maximal rate of detected photons, thus producing a saturation effect.

The most important property of SPADs used on a TOF rangefinder is the ability of accurately detecting the arrival time of photons. The statistical fluctuation of the time interval between the arrival of a photon at the sensor and the output pulse leading edge is defined as the *timing jitter* or timing resolution. Timing jitter mainly depends on the time a photo-generated carrier requires to be swept out of the absorption point into the multiplication region.

During an avalanche, some photons can be emitted due to the electroluminescence effect [8]. These photons may be detected by neighboring pixels in an array of SPADs and generate incorrect measurements. The probability of this effect is called optical crosstalk probability. It is worthy noticing that optical crosstalk probability is much smaller in fully integrated arrays of SPADs such the one reported in this implementation in comparison to hybrid versions. This is due to the fact that the diode's parasitic capacitance in the integrated version is orders of magnitude smaller than the hybrid solutions, thus reducing the energy dissipated during a

Geiger event. Electrical crosstalk on the other hand is produced by the fact that photons absorbed beyond the p-n junction, deep in the substrate, generate carriers which can diffuse to neighboring pixels. The probability of occurrence of this effect, whether it is optical or electrical, defines the *crosstalk probability*.

3 IMAGE SENSOR

3.1 Digital Pixel

The imager pixel consists of a circular SPAD and a 5-transistor configuration. The cross-section of the SPAD is shown in Fig. 1. The device was fabricated in a high voltage 0.8μm CMOS process. It is a dual p+/deep n-tub/p-substrate junction. The planar p+/deep n-tub junction provides the multiplication region where the Geiger breakdown occurs. The fabrication process was a 2M/2P twin-tub technology on a p-substrate allowing an operating voltage of 2.5 to 50V. Premature breakdown is avoided by surrounding the photodiode with a guard ring of relatively lightly doped diffusion. The guard ring eliminates abrupt doping profiles and corner effects by reducing electric field gradients at the diode's periphery. A useful feature of this technology is the availability of a p-tub implantation to create such a ring surrounding the p+ region anode [7]. The breakdown voltage V_{bd} of the p+/deep n-tub junction is 25.5V. A larger bias voltage $V_{bd}+V_e$ is applied on the diode to operate in single photon detection mode. V_e is known as excess bias voltage.

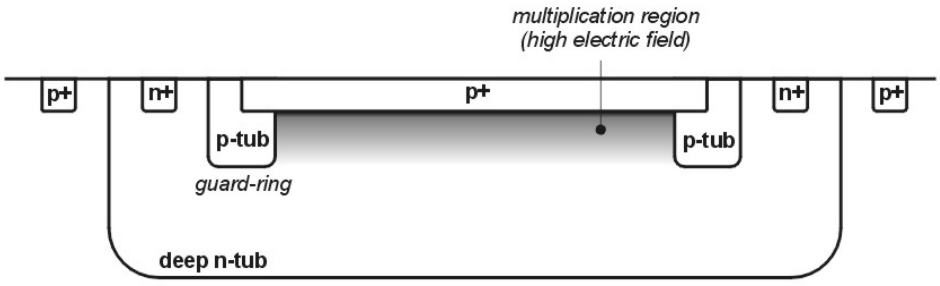

Fig. 1. SPAD cross-section

Fig. 2 shows the pixel circuit and a photomicrograph of the structure. The SPAD p+ anode is nominally biased at a negative voltage V_{op} 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 V_e is thus equal to $|V_{op}| + VDD - V_{bd} = 5$V.

The SPAD operates in passive quenching. Upon photon arrival, the breakdown current discharges its depletion region capacitance, causing the bias voltage to drop below breakdown voltage. The W/L ratio of T_q is set to provide a sufficiently resistive path to quench the avalanche process. After avalanche quenching, the SPAD recharges through T_q and progressively recovers its photon detection capability. The time required for quenching the avalanche and restoring the operating bias, i.e. the dead time, is typically less than 40ns for the digital pixel.

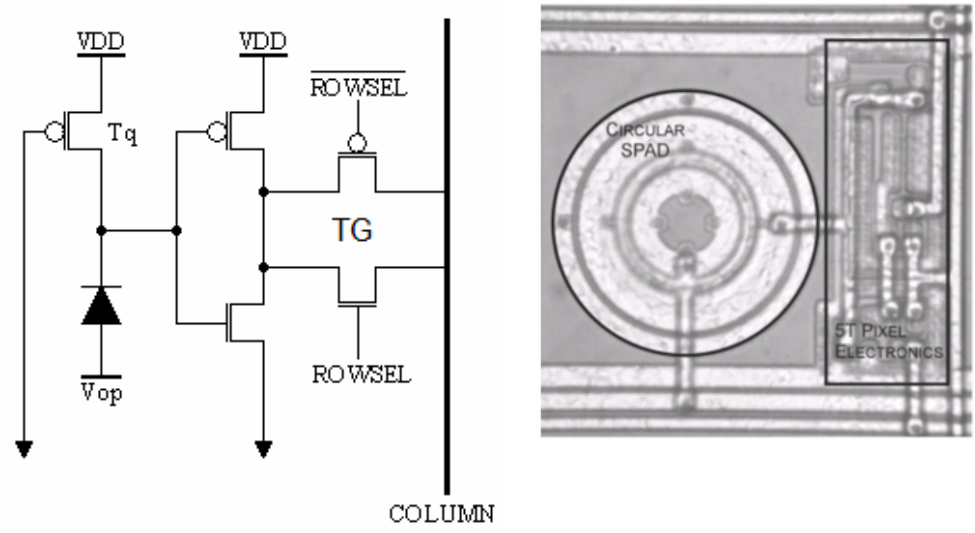

Fig. 2. 5-transistor pixel circuit and pixel photomicrograph

At the cathode, an inverted analog voltage pulse of amplitude V_e reflects the detection of a single photon. The CMOS inverter stage converts this analog voltage pulse into a digital pulse. A transmission gate (TG) is used to feed the detection signal to the nearest column output line when s=*VDD* and $\bar{s}$=*GND* (read phase). Every pixel generates a digital pulse which captures the arrival of a photon 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. The digital pixel occupies an area of 58μmx58μm. The active area of the SPAD is 38μm^2.

3.2 Image Sensor Architecture

The functional diagram of the solid-state image sensor is shown in Fig. 3. The image sensor consists of an array of 32x32 pixels and requires two power supply buses *VDD* = 5V and V_{op} = -25.5V. Digital pixels allow the readout to be designed with less stringent constraints on noise performance. Moreover, no amplification, no sample & hold, and no A/D converter are necessary. Consequently, in SPAD based image sensors, performance is not limited by analog noise and device mismatch in the readout circuitry. Therefore no particular care has to be given to the design of those components, except for minimizing digital noise and optimizing speed. In this implementation, the readout circuitry consists of a 5-bit decoder for row selection and a 32-to-1 multiplexer (MUX) for column selection. The photon timing information is obtained at the leading edge of MUX output.

A compact 15-bit linear feedback shift register counter has been implemented on-chip and it is used to compute intensity images. Intensity information is obtained by counting detection pulses within the integration time.

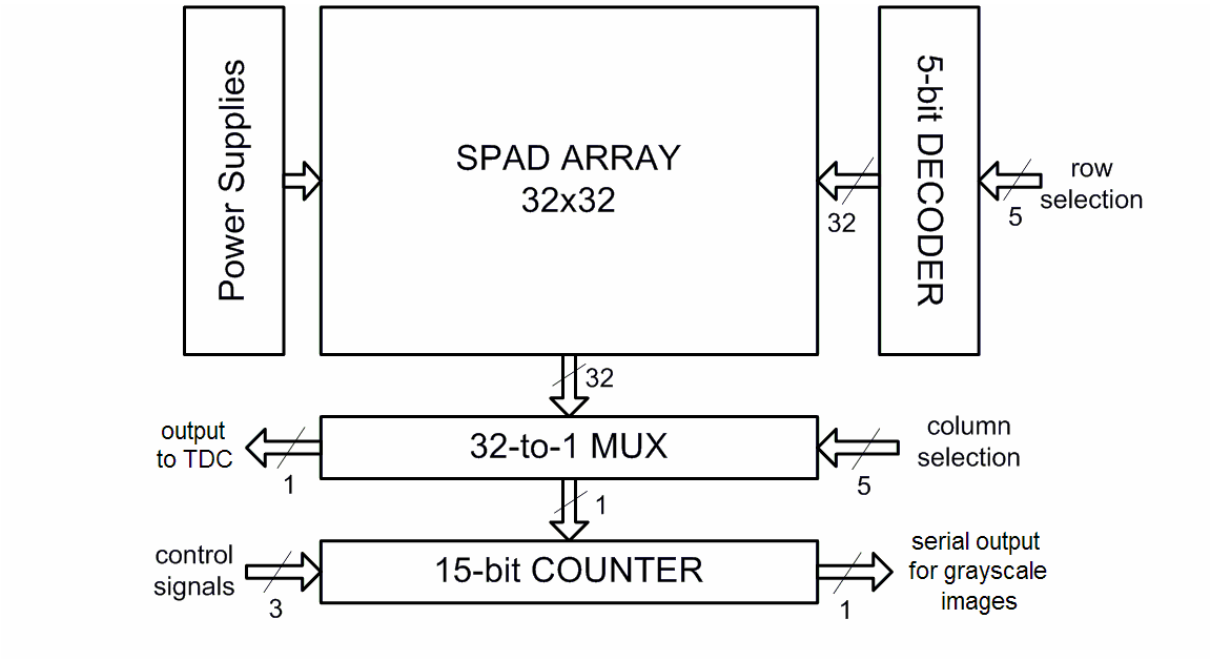

Fig. 3. Functional diagram of the image sensor

4 RANGEFINDING METHOD

4.1 Measurement Setup

The rangefinding method used in this sensor is based on pulsed type TOF. Depth is computed as d = TOF (c/2), where c is the speed of light. The system has been designed so that distance measurements can be obtained keeping the overall sensor complexity to a minimum.

The electro-optical setup is described in Fig. 4. The CMOS imager sensor was mounted on a printed circuit board and coupled with a standard camera objective. The scene was illuminated by a uniform cone of light created using a single intentionally uncollimated laser beam reaching every point of the surface. The laser source was a 635nm pulsed diode laser whose repetition rate f_R was 40MHz, pulse width T_P of 150ps, and peak power P_S of 100mW. A time-to-digital converter (TDC), TDC-F1 model from ACAM, was used for time interval measurement. The START signal was given by the synchronization output of the laser whereas the STOP signal was connected to the digital output of the 32x32 image sensor. The TDC, also implemented in CMOS technology, exhibited a timing resolution of 120ps and a 1-σ timing uncertainty of 110ps. Addressing and control signals were provided using a National Instruments PXI acquisition rack.

In this implementation, depth measurements are independently performed for each point in the scene by means of the SPAD array, though serially, due to the implemented readout circuitry. In order to avoid skews in the distance measurement for distinct pixels due to different electrical paths within the image sensor, TOF measurements are computed differentially with respect to a background plane. A reference image of a plane is captured using the same setup and is electronically stored. It is subsequently used in the computation of additional distance images, thus eliminating the pixel-to-pixel skews.

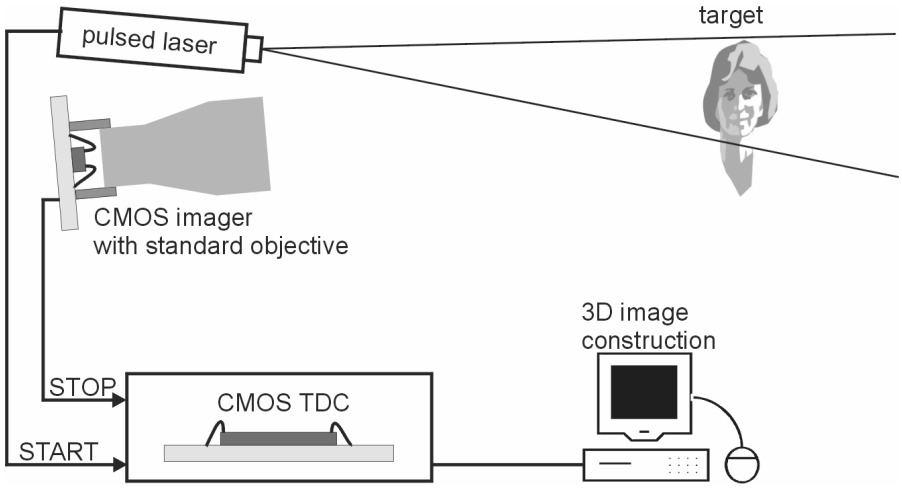

Fig. 4. 3D sensor set-up

4.2 Measurement Data Processing

Although CMOS SPADs exhibit excellent timing jitter properties, typically a few tens of picoseconds, this accuracy is insufficient to reach millimetric precision without averaging. Depth accuracy is computed as $\sigma(d) = c\sigma(\tau)/2$, where $\sigma(\tau)$ is the uncertainty of a single time measurement τ. Assuming that M TOF measurements $\{\tau_j\}$, j=1, …, M can be approximated as a collection of independent and identically distributed random variables, then, using simple averaging, the system time uncertainty is $M^{-1/2}\sigma(\tau)$, thus the distance accuracy can be theoretically increased by the square root of the number of measurements.

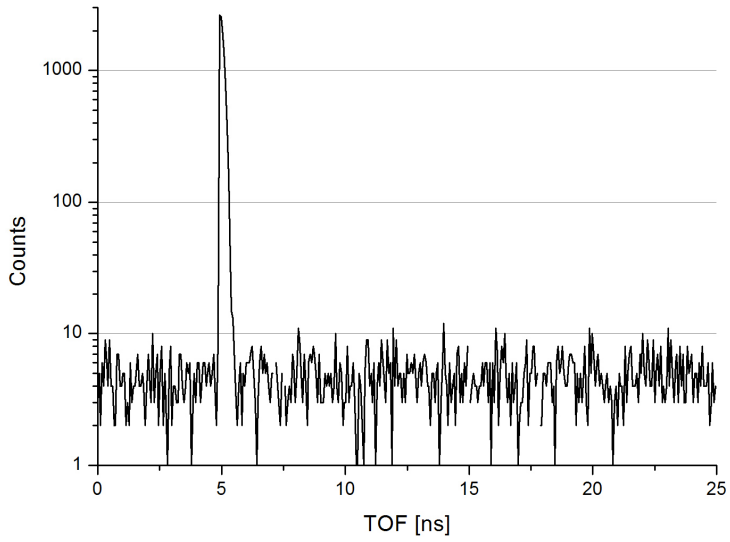

Fig. 5. Example of histogram for a typical measurement with M=10^4 an SNR=7dB

In addition, in order to reduce the influence of DCR and background illumination on the depth measurement, a simple processing algorithm was implemented. Each depth measurement was

computed building a histogram using a set of M TOF measurements. Since DCR and background illumination pulses are not correlated to the laser source, these pulses produce a noise level which is spread out over the entire histogram range. The signal on the other hand generates a peak in the histogram whose width is $\sigma(\tau)$ and can be easily isolated from noise by applying a threshold operation. Fig. 5 shows a logarithmic plot of such histogram generated for an arbitrary pixel with M=10^4 and SNR=7dB. Finally, depth is obtained by determining the average TOF value of the signal measurements. Note that histogram generation and computation are easily implemented either in software or hardware and they are well-suited for real time processing.

5 MEASUREMENT RESULTS

5.1 Image Sensor

The performance parameters of a SPAD based image sensor differ from the conventional approach which is typically limited by the analog readout circuitry noise and the mismatch parameters in the amplification and conversion stages.

Fig. 6 shows the measured PDP as a function of the photon wavelength for a typical pixel with a nominal V_e of 5V at room temperature. It is larger than 20% between 430nm and 570nm with a peak at 26% at 460nm. At 700nm the PDP is still 8% without any post process treatment of the silicon chip surface.

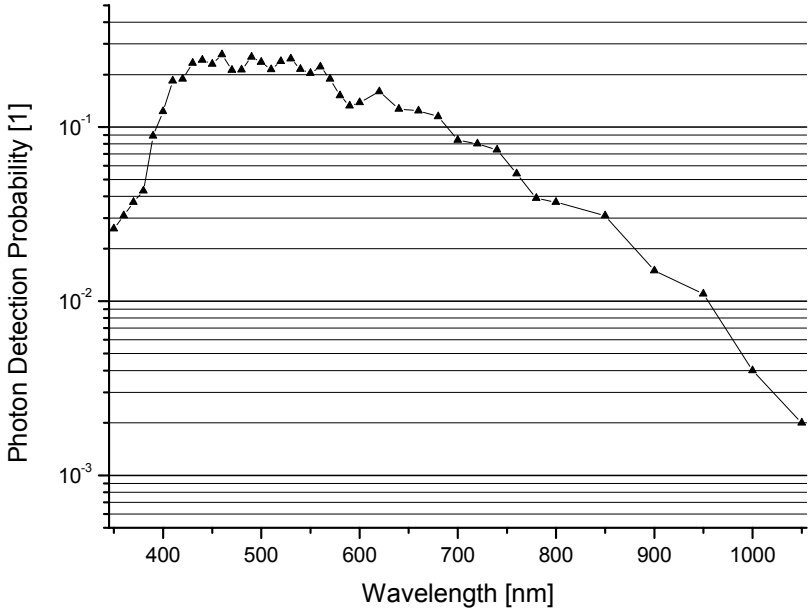

Fig. 6. Photon detection probability of a typical pixel for nominal excess voltage

Fig. 7 plots the distribution of the DCR across the sensor array for the nominal V_e at room temperature and at T=0°C. 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 350Hz on the whole sensor array and negligible after-pulsing effects. For T=0°C, the mean value of the DCR dropped below 75Hz.

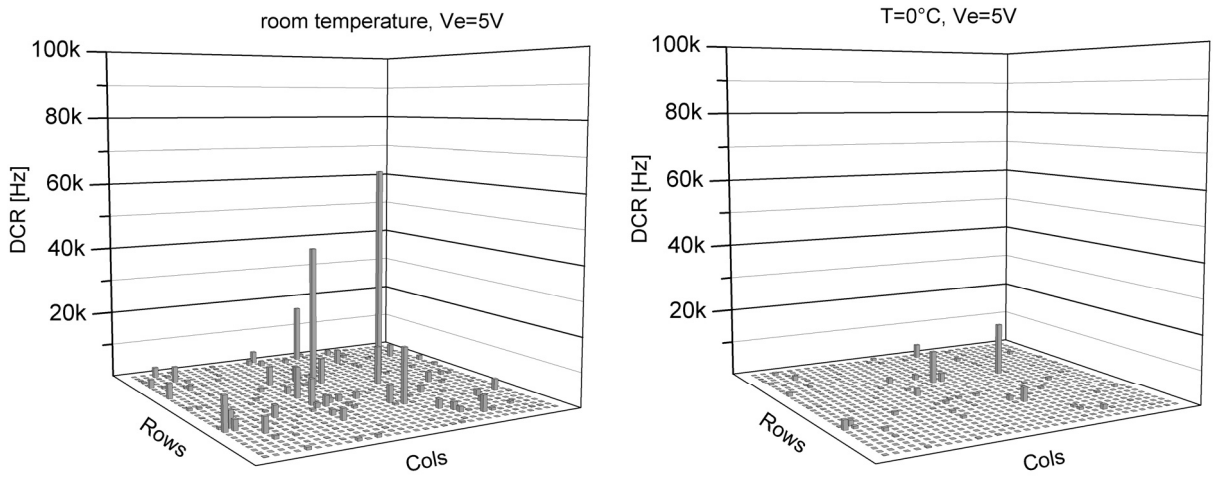

Fig. 7. Dark count rate for nominal excess voltage at room temperature and T=0°C

Fig. 8 presents the timing jitter measurement of the pixel. The measurement setup consisted of a pulsed laser source with pulse width of 30ps used to excite a single pixel through an optical fiber. A LeCroy Wavepro 7300 20GS/s oscilloscope was used to measure the time interval between the laser output trigger and the sensor output signal. The resulting overall timing jitter was 115ps FWHM even though the SPAD jitter is likely to be lower than 50ps FWHM [6]. The mismatch is likely caused by a non optimized leading edge rise time in the readout circuitry, thus a more careful design and layout techniques ought to be employed in the readout circuitry. The SPAD jitter is less than half that of the actual image sensor. Thus, a considerable margin for improvement still exists.

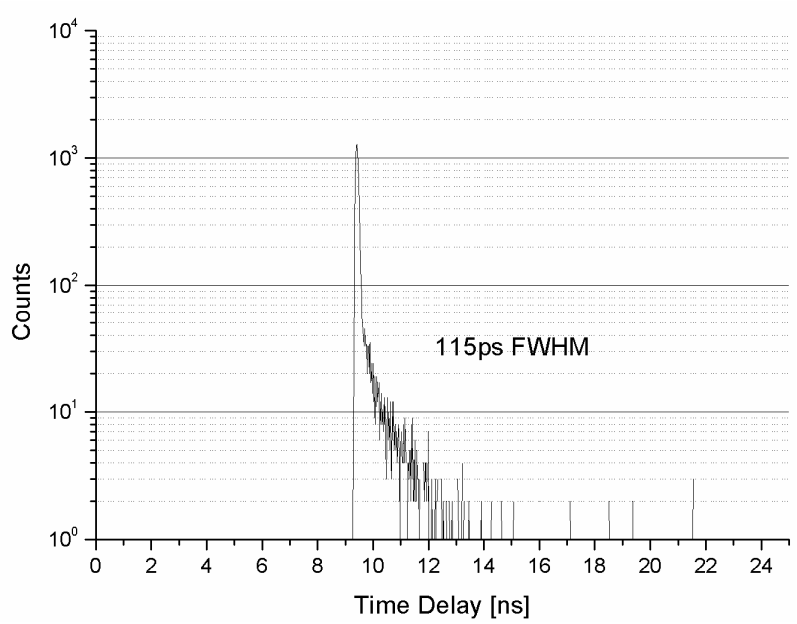

Fig. 8. Histogram of the measured timing jitter (from the slowest row)

Experimental measurements demonstrated that the SPAD array is completely free of crosstalk effects. Crosstalk was measured by illuminating a single pixel in the center of the array using a highly focused laser beam through the optics of a microscope. These measurements showed

a negligible crosstalk probability lower than 0.01%. This value was below DCR levels. As expected, optical crosstalk is negligible since the parasitic capacitance is very small. Electrical crosstalk was eliminated by design. Minor carriers diffusing in the substrate cannot reach the multiplication region of a pixel since they would be collected by the deep n-tub/p-sub junction.

5.2 Rangefinder

Operating in rangefinding mode, a total timing uncertainty $\sigma(\tau)$ of 250ps was expected whereas 300ps has been measured. $\sigma(\tau)$ is believed to be dominated by the TDC resolution and the extra jitter produced by digital switching on-chip. Although the measurement was carried out serially, all 1024 SPADs operated simultaneously during the measurement, thus generating noise on substrate and VDD lines. These noise sources may potentially increase the overall sensor timing jitter. To reach millimetric precision, the proposed processing algorithm was adopted, whereby the threshold value was set to 10% of the peak's maximum value. Fig. 9 shows multiple depth measurements with M=100, 10^3, and 10^4 at and around 3m distance. The 1-σ uncertainty is also reported.

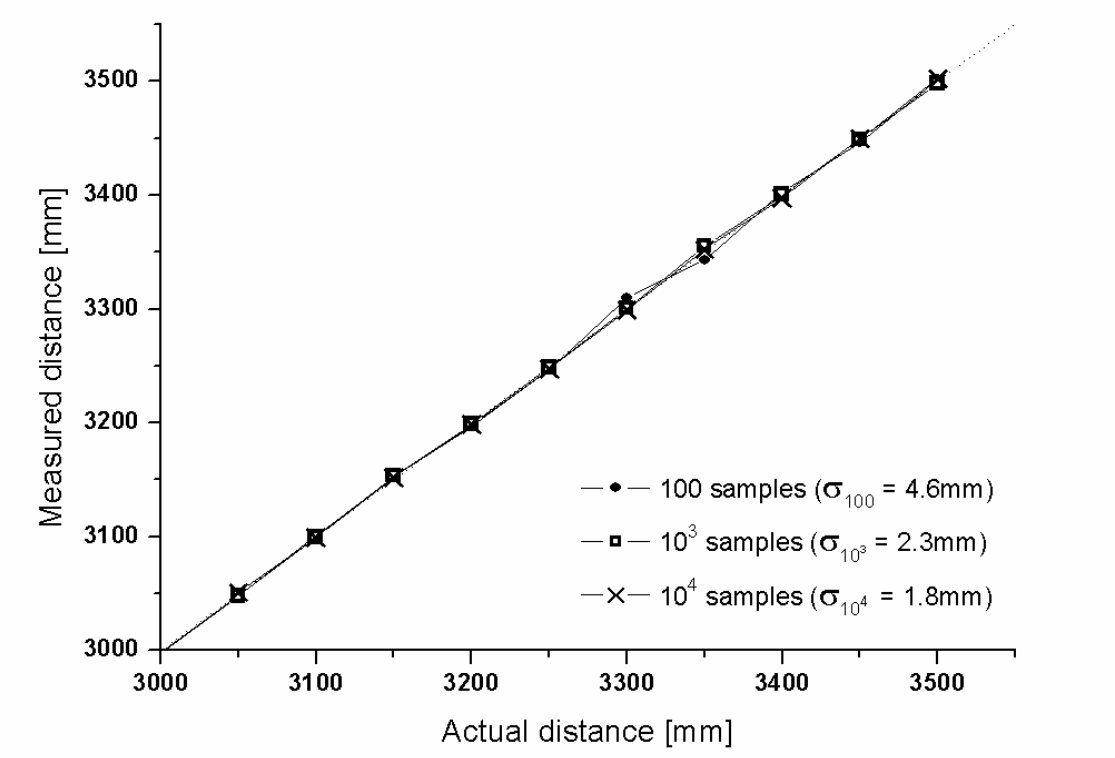

Fig. 9. Distance measurements

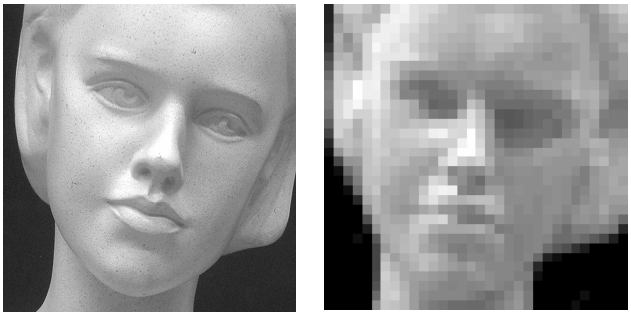

Fig. 10. Model photographed with high-resolution camera (left) and with our image sensor operating in 2D mode (right)

Subsequently, the depth map of a live-sized human mannequin was captured. Fig. 10 shows a high-resolution picture of the model taken by a digital still camera. To demonstrate the actual 32x32 pixel resolution, an intensity map of the model, obtained with our sensor operating in 2D mode, is also shown in the figure.

Fig. 11 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. In the current implementation, the average measurement rate was 50,000 measurements per second, which gives a pixel acquisition time (t_{acq}) of 200ms. Table 1 lists the salient performance measurements of the image presented in this paper.

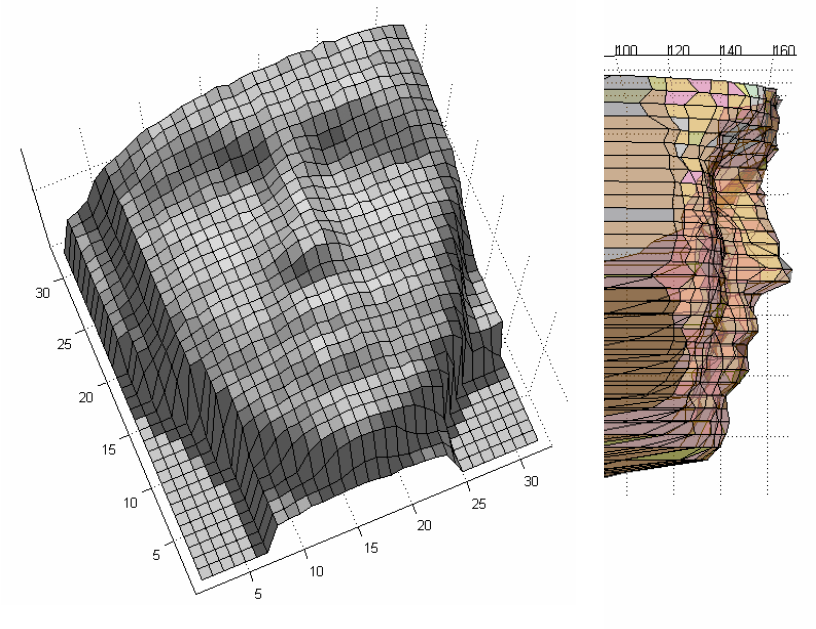

Fig. 11. Human face depth map and profile (in mm)

Table 1. Performance Summary

Level	Measure	Symbol	Value	Unit
Pixel	Photon detection probability @ 635nm	PDP	12	%
	Dark count rate (average)	DCR	350	Hz
	Fill factor	-	1.1	%
	FWHM jitter	-	115	ps
	Dead time	t_d	< 40	ns
Sensor	Distance range	D	3	m
	Time uncertainty	$\sigma(\tau)$	300	ps
	RMS Distance accuracy	$\sigma(d)$	1.8	mm
	Number of measurements per 3D point	-	10^2~10^4	-
	Power dissipation	P_{tot}	< 6	mW
Source	Repetition rate	f_R	40	MHz
	Peak power	$\hat{P}_S$	100	mW
	Average power	P_S	750	µW
	Wavelength	λ	635	nm
	RMS Pulse width	T_p	< 150	ps

6 CONCLUSION

Accurate depth mapping of arbitrary 3D scenes is performed using an imaging system based on time-of-flight. A cone of pulsed light generated by uncollimated laser illuminates the scene. The photons reflected by an object in the scene are collected by the imager that accurately detects their arrival time, thereby inferring the distance to the reflection point. The core of the system is an array of 32x32 single photon avalanche diodes that allows picosecond-accurate timing measurements at low levels of light intensity, thus achieving millimetric depth resolutions. The range of the depth map is 3m, while no mechanical micro-optical scanning device is required. Due to the high sensitivity of the detectors, very low output optical power is needed by the laser source, thus reducing the overall cost of the system and ensuring strict eye-safe operation.

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 48kframe/s CMOS Image Sensor for Real-Time 3-D Sensing and Motion Estimation", Proc. ISSCC, pp. 94-95, February 2001.

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

[3] 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.

[4] 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.

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

[6] 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.

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

[8] R. H. Haitz, "Studies on optical coupling between silicon p-n junctions", Solid State Electronics, 8, pp. 417-425, 1965.

* e-mail: cristiano.niclass@epfl.ch, phone +41 21 693 6431.