A 50x30-pixel CMOS Sensor for TOF-based Real Time 3D Imaging

David Stoppa, Luigi Viarani, Andrea Simoni,
Lorenzo Gonzo, Mattia Malfatti, Gianmaria Pedretti

Istituto per la Ricerca Scientifica e Tecnologica (ITC-irst), Trento, Italy,
Phone: +39 0461-314531, Fax: +39 0461-314591, E-mail: stoppa@itc.it, viarani@itc.it

Abstract

A 50x30-pixel CMOS sensor aimed at real time three-dimensional vision applications is presented along with characterization results. The sensor is realized on a single mixed-signal chip fabricated in a 0.35-um, 3.3-V, 4-metal/2-poly CMOS technology. The distance measurement relies on the pulsed indirect time of flight (ITOF) technique, and a precision of 4 % is obtained in the 2-8 m range. The real time mode of operation is proved at 20 frames per second with non-cooperative targets.

Keywords

Pulsed indirect time-of-flight, laser radar, real time 3D imaging.

1. INTRODUCTION

Fast 3D acquisition is the key to success for a number of modern application fields, such as obstacle recognition, automotive industry, video surveillance and virtual reality. Although a number of systems capable of precise 3D sensing have already been built, most of them cannot produce real time imaging because of overhead due to either large data computational costs or moving parts such as scanners. Among active optical ranging techniques, state-of-art systems employ the Time-of-Flight (TOF) principle rather than interferometry or triangulation when acquisition speed and maximum detectable distance are the main issues and a distance measurement resolution in the order of the centimetre is acceptable [1]. The choice falls then between Direct and Indirect TOF (ITOF) distance extraction, the latter including pulsed and amplitude modulation variants.

The proposed sensor working principle relies on the pulsed ITOF technique, which gathers the distance information by exploiting the availability of generating delayed timings with a precision in the order of the nanosecond and does not require high-efficiency photonic mixers (as with amplitude modulated ITOF), nor avalanche photon counting devices (as with Direct TOF-based systems) [2-4]. The absence of moving parts and the possibility of adjusting the few system parameters to accommodate the desired detectable distance range make it possible to obtain a sensor throughput which is mainly limited by the light pulses integration time and the sensor data readout, with negligible off-chip computational effort.

2. DISTANCE MEASURING TECHNIQUE

The pulsed ITOF principle is sketched in Fig. 1, where the laser source is pulsed by the sensor, which is a bi-dimensional array of pixels [5]. A lens is placed in front of the emitter to obtain the flood illumination of the scene, while a camera objective is used to focus the scene onto the detector matrix.

The emitted pulses are arranged in trains for pixel-level signal accumulation and white noise averaging, and a first train is integrated according to the timing W_1, then a second one is integrated by means of the timing W_2.

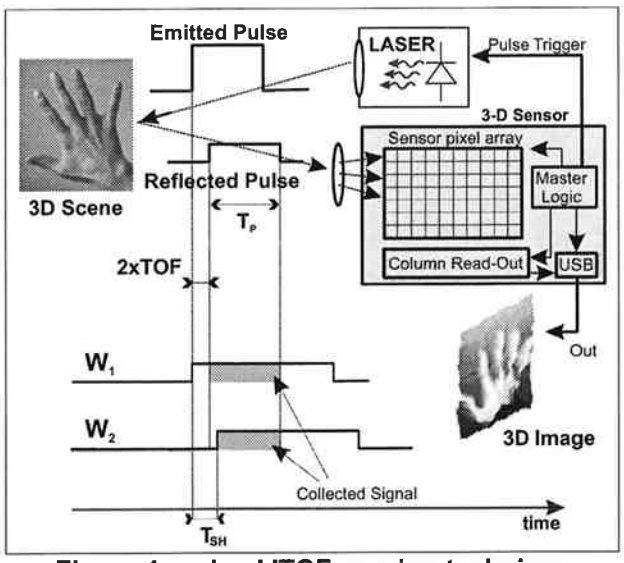


Figure 1: pulsed ITOF ranging technique

The timing W_1 is arranged so that the whole duration of the pulses (T_P) is collected for objects within the minimum and maximum distance, while the timing W_2 includes a delay (T_{SH}) which makes a portion of the pulses fall outside the pixel integration window. Depending on the time of flight of the photons reflected by the points of the scene, pixels looking at close objects get their photons earlier than pixels focused on further details, so they lose a bigger part of the reflected pulse.

By considering a specific pixel, the punctual distance information can be easily extracted by dividing the signals collected after the integration of the two trains (S_{W1} and S_{W2}), thus obtaining a measurement which is independent from the reflectivity conditions of the focused point, as follows:

$$D = \frac{c}{2} \cdot \left[T_P \cdot \left(\frac{S_{W2}}{S_{W1}} - 1 \right) + T_{SH} - \tau_L \right] \quad (1)$$

where c is the speed of light and τ_L is the delay associated to the laser module time constant.
Additional trains of pulses (W_3, W_4, ...) can be generated and integrated to detect further distance values.
Basically, objects within a specified range can be revealed by means of only two trains by properly setting the values of T_{SH} and τ_L. This allows for fast 3D imaging, since the time needed to get a 3D image is mainly due to the light integration phase and the sensor readout, this being accomplished just two times per frame.

3. PIXEL ARCHITECTURE

The pixel schematic diagram shown in Fig. 2 is based on a fully differential architecture with two N-well/ P-substrate photodiodes (D_2 is a dummy diode covered with a metal light shield) and a switched-capacitor (SC) charge amplifier based on an Operational Transimpedance Amplifier (OTA), whose output common mode voltage is set by a SC network (not shown).

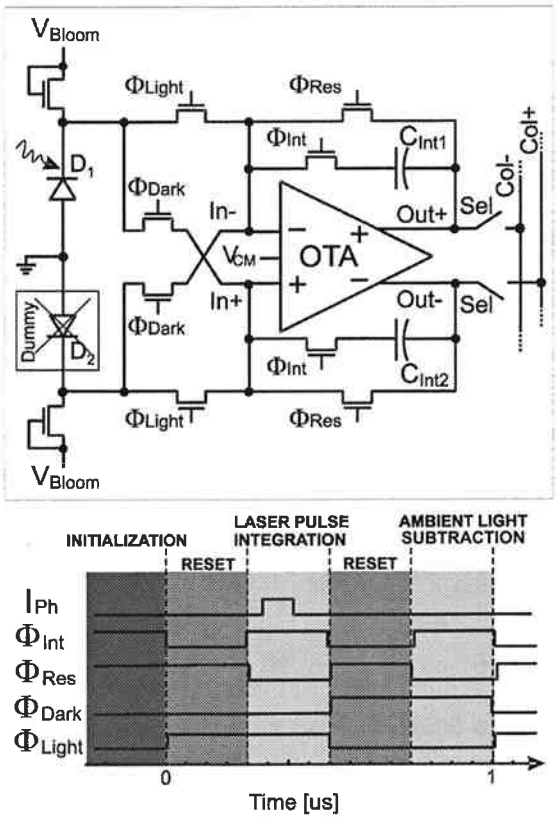


Figure 2: pixel schematic diagram

The architecture derives from the pixel implementation proposed in [6], and converts the received light pulses into voltage packets while removing the ambient light contribution and summing each packet to those previously accumulated within the same train. This cumulative integration allows to reach a reasonable signal level while increasing the S/N ratio due to the implicit white noise averaging of the accumulation operation. Nevertheless continuous-wave components must be rejected in order to avoid the pixel output saturation due to the background light, which does not provide useful signal for the punctual distance estimation.
The pixel working principle is depicted in the timing diagram of Fig. 2 and is as follows:

1. an initialization phase is performed with the OTA in the source follower configuration resetting the integration capacitances C_{Int1}, C_{Int2};
2. a first reset period occurs with (Φ_{Light}=high, Φ_{Dark}=low) for symmetry of operation with respect to the successive repetitions of steps 2-5;
3. the pixel integration window is opened and the laser pulse is collected together with the ambient light (the timing of Fig. 2 refers to timing W_1 and the whole pulse is collected);
4. a second reset period occurs and the diodes terminals are exchanged between the OTA input nodes (Φ_{Light}=low, Φ_{Dark}=high);
5. a second light integration phase is performed with the switched diodes, so that the continuous background contribution is cancelled from the output capacitances C_{Int1}, C_{Int2}.

Apart from the first step (performed only at the beginning of a train integration) the integration of a pulse train relies on the repetition of steps '2-5', where each step '5' ends with the addition of a voltage packet to the previous ones. All the pixels are read at the end of the train, then a new initialization step occurs and the four digital phases are delayed by T_{SH} with respect to the timing applied to the previous train.
The reset periods of steps '2' and '4' serve to restore the OTA input common mode voltage, which drifts because of the intrinsic unbalanced nature of the input photogenerated current. A value of 56 fF is chosen for C_{Int1}, C_{Int2} to guarantee a good device matching, a low KT/C noise level and a capacitive ratio with respect to the photodiode junction capacitance that keeps the input common mode drift negligible even in case of strong light intensity (tens of nA).
The main guidelines adopted for the OTA design are summarized in [6] and are dictated by the search for a low pixel area occupancy and power consumption, a reasonable bandwidth in response to the short laser pulses (50-200 ns), and a sufficiently high gain for proper operation of the charge amplifier. Typical conditions simulations run with the Cadence design tools report a GBP of 40 MHz, a 60-deg phase margin with 200-fF loads, and 90 dB for the DC open loop voltage gain. The pixel pitch is of 81.9x81.7 um^2 with a fill factor of 20 %, and the static power consumption is of 129 uW, which is acceptable for a bidimensional array implementation.

4. FABRICATED CHIP

A 50x30-pixel CMOS sensor has been fabricated in the standard AMS 0.35-um, 3.3-V, 4M-2P technology. Fig. 3 shows the die microphotograph and the PCB board realized to interface the chip to a personal computer through a Universal Serial Bus (USB) port.

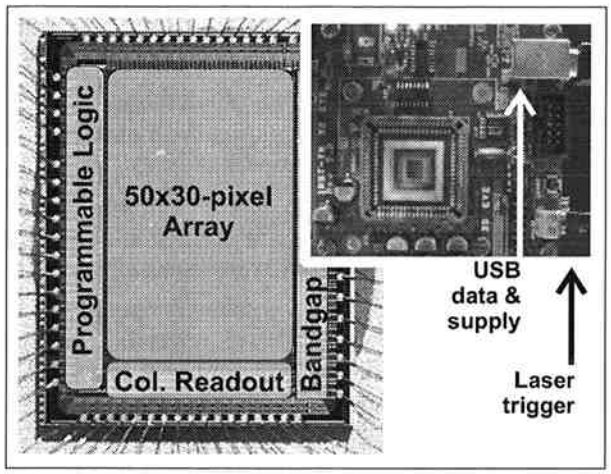
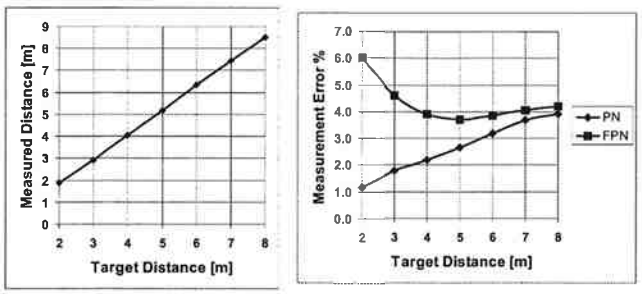


Figure 3: chip microphotograph and PCB

The chip measures 5.4x4.1 mm^2 and integrates on the same die the pixel array, a dedicated column readout stage for fixed pattern noise reduction, an internal bandgap voltage and current reference block and a programmable master logic. The main system parameters like the timings delay (T_{SH}), the laser pulse duration (T_P), the number of pulses per train and the number of trains per frame can be set by programming the chip master logic through the USB port, which provides also the data bus for chip readout and the power supply.

5. DISTANCE MEASUREMENT RESULTS

A first test has been carried out to investigate the pulsed ITOF-based distance measurement (Fig. 4). A cardboard slab with a reflectivity of 40 % has been positioned with steps of 1 m in the 2-8-m range in front of the sensor and a 905-nm laser source has been employed to emit trains of 200-ns pulses to illuminate the target with a total power on the scene of 1 W. The distance values are extracted by averaging 100 frames, each frame being made of the integration of the two trains associated to timings W_1 and W_2.



(a) (b)

Figure 4: distance measurement performances

Inset 'a' shows the measured distance points, whose agreement with the actual target distance worsens for high distance values, as it can be expected from Eq. (1) in the case where the signal collected with timing W_2 becomes as large as the one collected with W_1. Inset 'b' reports the percentual Pixel Noise (PN) and Fixed Pattern Noise (FPN) of the pixel array. The PN is representative of the ranging precision and increases with the distance, remaining less than 4 % for the 8-m step. The FPN quantifies the overall sensor non-uniformity and is higher for lower distance values due to the strong non-proportional increase of the light intensity received from close objects.

6. 3D IMAGING RESULTS

Fig. 5 shows 3D pictures taken with the proposed sensor of a lamp switched 'on' and 'off'. Insets 'a' and 'd' have been taken with a short-exposure commercial camera, while 'b', 'c', 'e' and 'f' are different representations of the 20-time averaged pictures obtained with the presented imager. In particular, 'b' and 'e' are obtained superimposing the distance information acquired (reported colour-coded in 'c' and 'f') with the grey-level 2-D images of the scene taken operating the sensor with an additional 100-us continuous-light integration cycle.

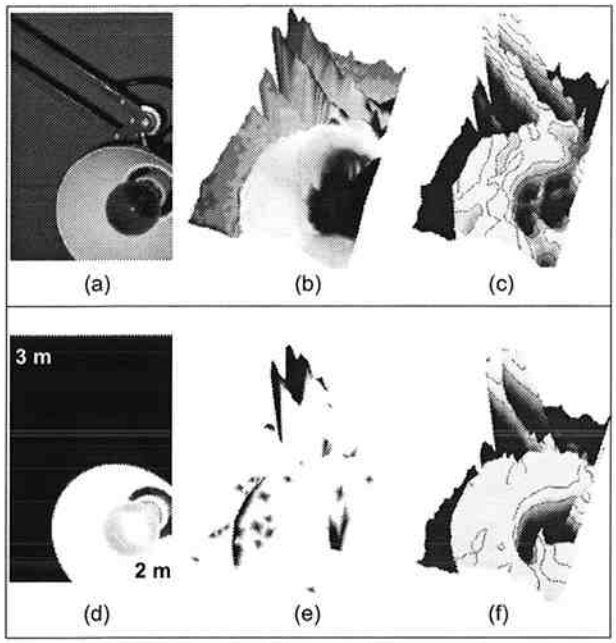


Figure 5: 3D pictures of a lamp 'off' and 'on'

It can be noticed that when the lamp is in the 'on' state the 2-D images taken by the commercial camera and the sensor are saturated (insets 'd' and 'e' exhibit detail loss) while the sensor ranging functionality is still properly working (details of inset 'f' are clearly visible) due to the implemented pixel ambient light rejection. The maximum intensity of continuous light that is cancelled with negligible effects on the output signal has been measured with a photographic lux-meter and has been found to be about 40 klux.

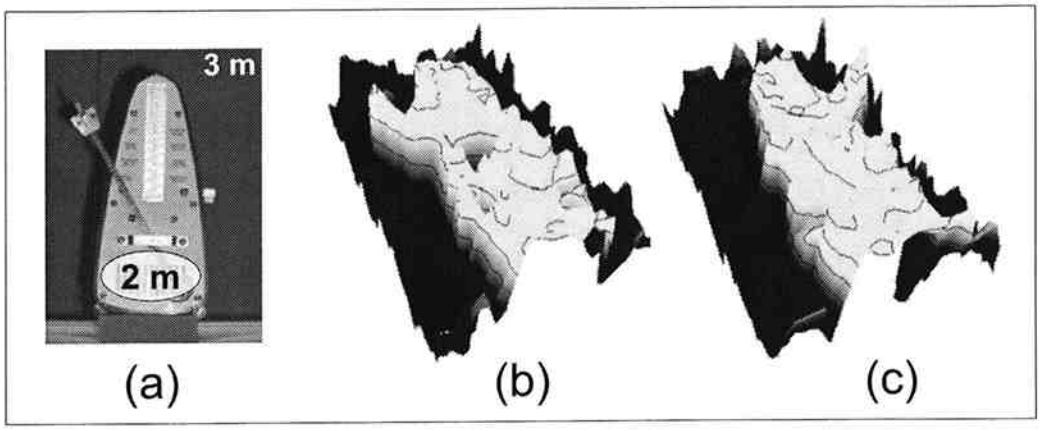


Figure 6: Two frames extracted from a real time 3D acquisition (10 fps)

The sensor has been tested also with moving objects by reducing the number of averaged frames to 1, 2 frames in order to achieve a refresh rate of 20, 10 Hz respectively. Fig. 5 reports 2-time averaged frames of a 3D stream taken at 10 fps.

7. CONCLUSIONS

A 50x30-pixel CMOS sensor aimed at real time 3D vision applications has been presented. The chip is fabricated in a standard CMOS technology and its main characteristics are summarized in Table 1 below.

TABLE 1: sensor main characteristics.

Technology	0.35 um, 3.3 V, 4M-2P
Die size	5.4x4.1 mm^2
Photodetector type	Nwell/Psubstrate
Pixel resolution	50x30
Pixel pitch, fill factor	81.9x81.7 um^2, 20 %
Laser source	905-nm, class 1
PN, FPN (worst case)	4 % (@ 8 m), 6 % (@ 2 m)
Ambient light rejection	< 40 klux
Maximum refresh rate	20 fps

The particular ranging technique implemented (pulsed ITOF) permits a distance extraction process which is independent from target reflectivity conditions, while the ambient light rejection is performed at the pixel level, thus opening for a variety of possible application fields.
The embedded programmable master logic allows for a flexible chip configuration and the use of the USB PC interface eliminates the need of a data acquisition card and of a 3.3-V supply cable for the chip, so that the complete system (sensor plus laser source) is very compact and the only connections needed are the USB and the laser emitter power supply cables.

The total absence of external 'aid' devices (such as scanners, image intensifiers, optical filters) and the minimal off-chip computational effort required make it possible to achieve 3D visualization streams with a refresh rate of up to 20 fps by disabling the average operation on the acquired frames.
As for eye-safety conditions, the target scene is illuminated with a total incident power of 1 W obtained with a 905-nm "class 1" laser source, which is operated with 200-ns pulses and a very low duty cycle (0.1 %).

8. REFERENCES

[1] A. Simoni, L. Gonzo, M. Gottardi, "Integrated Optical Sensors for 3D vision", *Proc. of IEEE Sensors*, 2002, session paper 2.1.

[2] M. R. Muguira, J. T. Sackos and B. D. Bradley, "Scannerless range imaging with a square wave", *Proc. of SPIE*, 1995, vol. 2472.

[3] R. Lange, P. Seitz, "Solid-State Time-of-Flight Range Camera", *IEEE Journal of Quantum Electronics*, March 2001, vol. 37, no. 3, pp. 390-397.

[4] C. Niclass, E. Charbon, "A Single Photon Detector Array with 64x64 Resolution and Millimetric Depth Accuracy for 3D Imaging", *ISSCC Digest of technical papers*, February 2005, pp. 300, 301.

[5] L. Viarani, D. Stoppa, L. Gonzo, M. Gottardi, A. Simoni, "A CMOS Smart Pixel for Active 3D Vision Applications", *IEEE Sensors Journal*, February 2004, vol. 4, no.1, pp. 145-152.

[6] D. Stoppa, L. Viarani, A. Simoni, L. Gonzo, M. Malfatti, G. Pedretti, "A 16×16-Pixel Range-Finding CMOS Image Sensor", *Proc. of IEEE ESSCIRC 2004*, Leuven, Belgium, September 2004, pp. 419-422.