SPADAS: a high-speed 3D Single-Photon camera for Advanced Driver Assistance Systems

D. Bronzi*[a], Y. Zou[a], S. Bellisai[a], F. Villa[a], S. Tisa[b], A. Tosi[a], and F. Zappa[a,b]

[a] Dip. Elettronica, Informazione e Bioingegneria - Politecnico di Milano - p. Leonardo Da Vinci 32, 20133 Milano, Italy
[b] Micro Photon Device S.r.l., Via Stradivari 4, 39100, Bolzano, Italy

ABSTRACT

Advanced Driver Assistance Systems (ADAS) are the most advanced technologies to fight road accidents. Within ADAS, an important role is played by radar- and lidar-based sensors, which are mostly employed for collision avoidance and adaptive cruise control. Nonetheless, they have a narrow field-of-view and a limited ability to detect and differentiate objects. Standard camera-based technologies (e.g. stereovision) could balance these weaknesses, but they are currently not able to fulfill all automotive requirements (distance range, accuracy, acquisition speed, and frame-rate). To this purpose, we developed an automotive-oriented CMOS single-photon camera for optical 3D ranging based on indirect time-of-flight (iTOF) measurements.

Imagers based on Single-photon avalanche diode (SPAD) arrays offer higher sensitivity with respect to CCD/CMOS rangefinders, have inherent better time resolution, higher accuracy and better linearity. Moreover, iTOF requires neither high bandwidth electronics nor short-pulsed lasers, hence allowing the development of cost-effective systems. The CMOS SPAD sensor is based on 64 × 32 pixels, each able to process both 2D intensity-data and 3D depth-ranging information, with background suppression. Pixel-level memories allow fully parallel imaging and prevents motion artefacts (skew, wobble, motion blur) and partial exposure effects, which otherwise would hinder the detection of fast moving objects. The camera is housed in an aluminum case supporting a 12 mm F/1.4 C-mount imaging lens, with a 40°×20° field-of-view. The whole system is very rugged and compact and a perfect solution for vehicle's cockpit, with dimensions of 80 mm × 45 mm × 70 mm, and less that 1 W consumption. To provide the required optical power (1.5 W, eye safe) and to allow fast (up to 25 MHz) modulation of the active illumination, we developed a modular laser source, based on five laser driver cards, with three 808 nm lasers each.

We present the full characterization of the 3D automotive system, operated both at night and during daytime, in both indoor and outdoor, in real traffic, scenario. The achieved long-range (up to 45m), high dynamic-range (118 dB), high-speed (over 200 fps) 3D depth measurement, and high precision (better than 90 cm at 45 m), highlight the excellent performance of this CMOS SPAD camera for automotive applications.

Keywords: 3D ranging, Single-Photon Avalanche Diode, CMOS SPAD, Time-of-Flight, automotive vision.

1. INTRODUCTION

In the last decade, global automotive industry has seen the proliferation of new control strategies – such as electronic stability control, vision-based pedestrian detection systems, lane departure warning systems, night vision systems, electronic parking assistance, adaptive cruise control and the advanced front-lighting system – thanks to improvements in microcontroller units and the low cost and wide availability of different sensor technologies. All these sensors are part of automotive systems – designed to assist in all aspects of driving, including safety, drivability, and fuel economy – that are called advanced driver assistance systems (ADAS).

The most advanced ADAS technologies to fight road accidents are the collision avoidance systems (CAS) [1], which typically employ either radar-, lidar-, ultrasonic- or camera-based depth. Fig.1 shows an overview of the functions and detection ranges typically required in CAS, along with the sensor technology likely to be used: each sensor is equally employed in a variety of applications and according to the operation principle, different sensor technologies tend to show complementary strengths in measuring certain object parameters.

Concerning camera-based 3D vision systems, we can consider two main categories based on different principles [2]: stereovision and time-of-flight. When evaluating this techniques for automotive safety, we should take into account not

* danilo.bronzi@polimi.it; phone +39 02 2399 4003; fax +39 02 2399 3699; www.deib.polimi.it
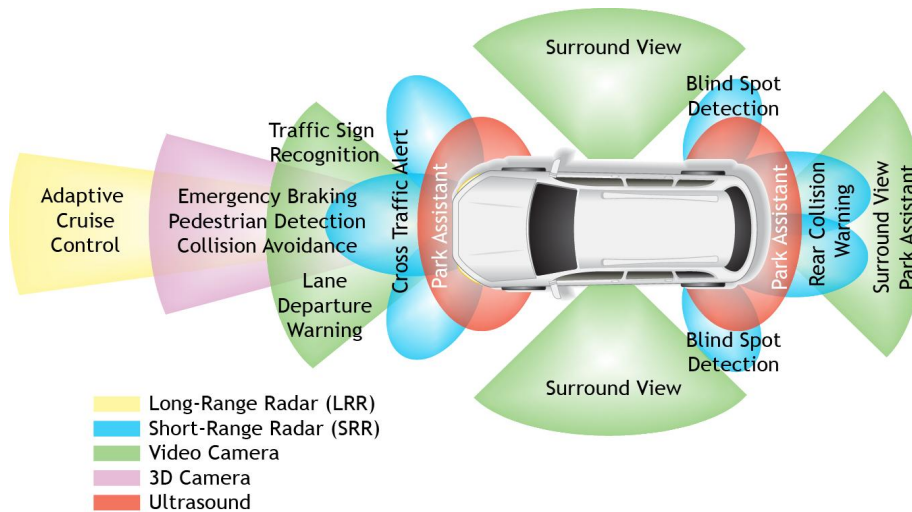
Figure 1. Vehicle functions and technologies likely to be used.

only their performance but we should also verify if they respect some constraints arising from the application. Typical performance include accuracy, precision, lateral resolution, frame-rate, field-of-view and maximum measurable distance range, while constraints mainly deal with cost, compactness and reliability [3].

Stereovision (SV) systems have the advantage of requiring neither moving parts nor illumination source, thus they can be cost-effective and easy to design. Nonetheless, the projection of a 3D scene onto the two–dimensional sensors causes a compression of spatial information and the need for high contrast scenes to identify point pairs, which are projections of the same point (correspondence problem). Additionally, SV suffers from problems resulting from shadows and moving objects, and the depth resolution is determined by the optical arrangement of the two cameras and cannot be changed [4]. Moreover, some processing is required to extract the correct distance information, thus implying a very low frame rate. Therefore, SV-based systems delivers ambiguous distance measurement in difficult environmental lighting conditions and do not fit the best in automotive applications, especially when fast 3D ranging is demanded.

Active triangulation methods are not affected by problems related to shadows and moving object, but have the major drawback of using costly and cumbersome source to generate structured light; in addition, to project light patterns on the scene, a mechanical mechanism is required, thus making the system very sensitive to vibrations and difficult to use in automotive applications [3].

Time-of-flight (TOF) cameras that, in the last decade, have gained attention thanks to their attractive characteristics overcome such limitations. TOF-vision systems are made of a single camera, which includes the imager and a light source, and can reach higher frame rates relative to SV-systems, because the depth measurement is straightforward as no matching problem has to be solved and pre-processing can therefore be performed in-pixel. The simplest TOF technique, called direct time-of-flight (dTOF), relies on the measurement of the round-trip flight time taken by a light pulse to travel from the source to objects in the scene and then back to a photodetector [5]. The measured time is then converted to distance by knowing the speed of light. This technique requires high bandwidth instrumentation and short laser pulses.

An alternative technique is represented by indirect time-of-flight (iTOF) estimation, where distance information is extracted from the phase-delay between a modulated light shone toward the target and its back-reflected echo (like in heterodyne or homodyne demodulation). Compared to dTOF technique, iTOF estimation provides higher acquisition speed with a reduced output data throughput and less external computational effort. Additionally, neither high bandwidth electronics nor short pulse lasers are required, hence allowing the development of cost-effective systems.

In this paper, we present a complete automotive-oriented camera for optical 3D ranging using indirect time-of-flight technique. The camera is based on the SPAD imager presented in the previous chapter, which is able not only to deliver two-dimensional (2D) intensity information through photon counting in free running, but also to perform smart light demodulation with in-pixel background suppression, thus enabling three-dimensional (3D) depth-resolved mapping of the objects in the scene. As a consequence, simultaneous 2D and 3D videos of rapidly changing (e.g. in shape, intensity,

distance, etc.) scenes in light-starved environments are advantageously provided. The system was in both internal and real traffic scenario, yielding high dynamic-range (118 dB) high-speed (200 fps) depth measurement with high precision (< 50 cm) up to 45 m, thus highlighting the excellent performance of CMOS SPAD cameras for automotive applications.

The remainder of this paper is structured as follows: in Section 2, we describe the SPAD camera and the implemented techniques; in Section 3, we briefly describe the laser source; in Section 4 we show and comment the results of experimental characterization. Finally, we comment the result from outdoor measurements in real traffic scenarios.

2. SINGLE-PHOTON CAMERA

The imager chip is based on 64 × 32 SPAD array fabricated in an automotive-grade 0.35 μm high-voltage 2P-4M CMOS technology on 8-inch wafers with local-oxide silicon based (LOCOS) insulations [6]. Each pixel integrates a 30 μm single-photon avalanche diode (SPAD), the related sensing and quenching circuitry (to manage the avalanche current ignition and to generate a synchronous voltage pulse), the shaping electronics (to control the voltage pulses duration), the processing circuitry (for iTOF, respectively) and data read-out (to store and buffer data). The processing circuit is design to implement on-chip lock-in demodulation. Indeed, the depth measurement is performed through the continuous-wave (cw) iTOF. In the cw-iTOF camera, a sinusoidal-modulated light source illuminates the scene and the reflected light is received back by the detector, phase-shifted by an amount $\Delta\varphi$, proportional to the modulation frequency f and the distance d of the object. From the phase delay, the distance is given by:

$$d = \frac{c}{4\pi f}\Delta\varphi \tag{1}$$

To retrieve phase shift information, the reflected wave is synchronously sampled with four integration windows, of same duration $T_{INT} = \frac{1}{4}f^{-1}$, thus providing C_0, C_1, C_2, C_3 as shown in Fig. 2. N modulation periods are summed together to increase SNR [5]. Then, through Discrete Fourier Transform, phase delay $\Delta\varphi$, amplitude A and offset B can be computed:

$$\Delta\varphi = \arctan\frac{C_3 - C_1}{C_0 - C_2}$$
$$A = \frac{[(C_3 - C_1)^2 + (C_0 - C_2)^2]^{1/2}}{2NT_{INT}\mathrm{sinc}(\pi f T_{INT})} \tag{2}$$
$$B = \frac{C_0 + C_1 + C_2 + C_3}{4NT_{INT}}$$

In order to implement the aforementioned technique, the present pixel consists of three counters, three internal storage latches and selection circuitry for data read-out. Two counters are 9-bit up/down and acquire differential counts ($C_{NUM} = C_3 - C_1$, $C_{DEN} = C_0 - C_2$) for phase and amplitude computation, through a common $Enable_{U/D}$ signal, which enables them selectively during the period and a $Direction$ signal which select up or down counting direction. The third 9-bit counter integrates photon counts over the whole period, hence accumulates background $C_B = C_0 + C_1 + C_2 + C_3$. An FPGA generates all control signals shown in Fig. 3.
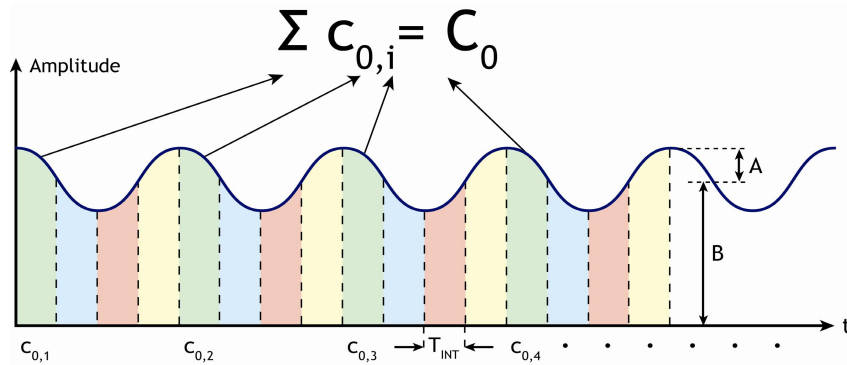

Figure 2. Sinusoidal-modulated cw light, sampled four times in a period.

Figure 3. Simplified block diagram of the cw-TOF pixel (top) and timings of the control signals (bottom).

In order to easily operate the 64 × 32 SPAD imager, we developed a compact camera, which includes a first board with the imager (bonded onto a 72-pin custom-made Pin Grid Array), I/O signal conditioning electronics, and power supply. A second board contains the FPGA (Xilinx Spartan-3) that manages timings of I/O signals, processes data read out from the chip and uploads processed data to a remote PC through a USB 2.0 link, which also provides +5 V bias (USB-powered system). A direct digital synthesizer (DDS), is also implemented to providing arbitrary analog waveform modulation for the light source. The camera is housed in a case supporting a 12 mm F/1.4 C-mount imaging lens, with a 40° × 20° (H × V) field-of-view. The whole system is very rugged and compact (80 mm × 45 mm × 70 mm). A MATLAB[®] interface is used for setting the measurement parameters (frame duration, integration time, etc…), the system clock frequency, and for data acquisition and post-processing. The power consumption is about 1 W, mostly dissipated by the FPGA board (240 mA), with negligible contribution from the SPAD imager (10 mA).

3. LASER DIODE SOURCE

The illumination source is a critical component, because the distance error is strictly dependent on it, as well as the measurement linearity. The illuminator has a modular design based on a power supply board and five laser driver cards, each one mounting 3 laser diodes (LDs). The laser driver enables the switching of the laser with well-defined current pulses, with frequency up to 200 MHz. The current into each channel is controlled with two input signals fed by the camera: the enable signal (*EN*) switches on and off the current in the laser, while the Current control voltage Input (*CI*) is an analog signal that controls the LD current by varying the cathode voltage. A temperature protection circuit, heat sinkers and a fan are also used to lower and stabilize the temperature. The total output of the LD source is 1.5 W, which ensures high precision up to a distance of 40 m, as it will be shown in the reminding of the paper.

4. SYSTEM CHARACTERIZATION

We performed static tests by varying the distance (2-40 m) between the camera and a wall in a controlled-light environment. Measurements were performed at 100 fps for two different modulation frequencies: 5 MHz and 8.33 MHz.
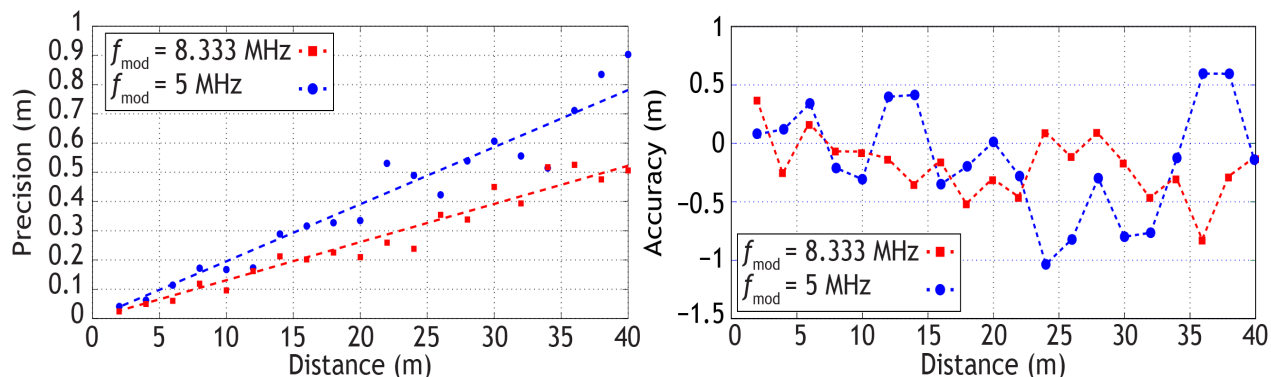
Figure 4. Precision (left) and accuracy (right) as a function of distance, for two different modulation frequencies.

The 5 MHz modulation ensures a non-ambiguous range of 30 m and a minimum precision of 0.8 m with an accuracy of less than 1 m, while the 8.33 MHz modulation ensures a non-ambiguous range of 18 m and a minimum precision of 0.5 m with an accuracy better than 80 cm (Fig. 4). The two modulations can be used in a multiple frequency approach to ensure a non-ambiguous range of 45 m but with a precision better than 0.6 m, according to [7] . The data from precision measurements are interpolated to fit the theoretical linear waveform.

5. OUTDOOR TESTS

The 3D SPAD camera is intended to be used on a vehicle in order to detect other vehicles, objects or pedestrians too close to the vehicle itself avoiding a possible accident. In order to assess the correct operation of the camera, the entire system (SPAD camera and illumination source) was installed on a vehicle as shown in Fig. 5, together with the LD source and a webcam for co-registration of the images.

Different road conditions were identified as possible scenarios of use for the camera: large square with trees and other fixed objects, halted car and object entering within the FOV of the camera, normal traffic condition on road and pedestrian crossing in front of the vehicle. All the following measurements were acquired with a frame rate of 100 fps. The background level depends on the acquired objects but is attested about 1000-2000 photons per pixel per frame. Thanks to a narrow optical filter (10 nm), centered at the emission wavelength of the laser diodes (808 nm), the background is reduced to an optimal level. Background light sources were the projectors and the taillights of the other vehicles, the traffic lights and the road artificial illumination (mostly sodium-vapor lamps).

Fig. 6 shows a frame taken form a 100 fps acquisition, where a bus crossing the camera FOV is imaged. The distance on the right represents in false color the distance of the objects in the scene. The image is used combined both the recorded amplitude of the reflected light (A) and the measured distance. As shown in the bottom of Fig. 6, one of problem encountered was due to the not perfect alignment to the roof of the car, which causes a partial reflection of the modulated light, as shown in the bottom of the right picture. However, this problem could be easily overcome with a proper design on the mechanical housing for the camera and the illuminator.
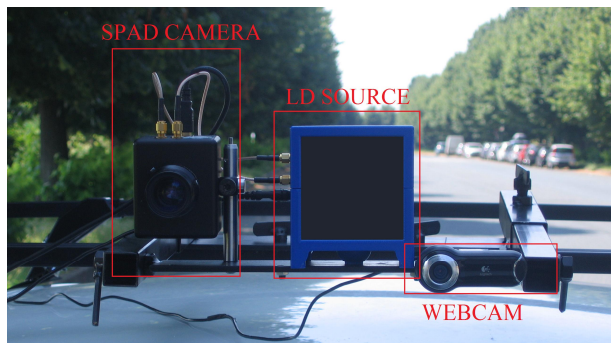

Figure 5. 3D SPAD Camera and illuminator installed together with a standard webcam for outdoor tests.
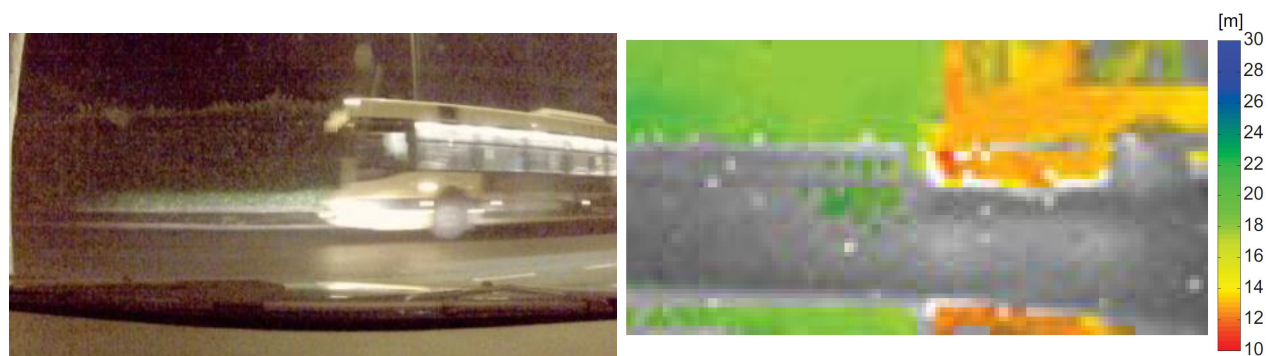
Figure 6. : Bus passing in front of the 3D vehicle. The background wall actually is a hedge.

Taking into account the reflected signal amplitude (about 2500 photons per frame) and background amplitude (about 6900 photons per frames, taking into account both average value of the illuminator and ambient light sources), the expected precision for Fig. 6 is about 25 cm. Another situation is depicted in Fig. 7 where the car preceding the vehicle is approaching. The distance information is correctly reported, with the orange color indicating a distance below 12 m and the yellow indicating distance about 14 m.

Figure 7. Preceding car approaching the 3D vehicle and measured distances in the scene.

Consider now the vehicle halted at a crossroad and a car passing on the crossing road as shown in Fig. 8. The wall on the background is about 25 m far away from the vehicle and is correctly reported in blue color, whereas the car passing in the nearer lane is shown in yellow. Another effect was noticed during the road acquisitions: the reflection of the NIR light depends on both the color and on the material used in the bodywork of the of the vehicle. Actually, a multi wavelength acquisition could be performed, by fusing 2D and 3D data coming from other sensors (e.g. radar).

Figure 8. Car passing on the crossing road in front of the 3D vehicle.

6. CONCLUSIONS

The main goal of this paper was to study, design and develop a 3D ranging camera able to acquire depth maps at a frame rate up to 100 fps with a precision of 50 cm over a wide-angle 40 m range. The design of the ultimate 3D camera is based on a 64×32 SPAD based sensor. Every pixel contains three programmable counters for many indirect Time-of-Flight measurements. The low noise (under 100 cps per pixel) allows performing a measurement without introducing further background counts. The camera enclosure is $80 \times 70 \times 50$ cm^3, provides C-type lens mount, and USB 2.0 connection. The camera USB-bus-powered and fully PC controlled, and it is able to acquire 3D movies with a frame rate up to 200 fps. The illuminator is directly controlled by the camera, in order to guarantee proper synchronization. The camera was tested in outdoor scenarios measurements in order to assess the performance and the limitations of the system, particularly concerning the mounting within the vehicle.

In preliminary tests, the camera was characterized in indoor controlled conditions providing very good performances, with a precision of less than 60 cm at 45 m and an accuracy of less than 1 m.

The overall design and development of the short-range camera and illuminator for the automotive application based on a low-cost and compact SPAD-based system has been proved to be effective in providing accurate and fast acquisition of depth maps. Of course, the development of the system can be further expanded in the exploitation phase; in fact, the camera and the entire system can be further developed in order to improve functionality. As an example, object recognition could be implemented within the FPGA in order to disengage to the need of a PC and have a real standalone system.

REFERENCES

[1] "Collision Avoidance Systems." http://www.collisionavoidancesystems.net/. Accessed: 06–28–2014.
[2] R. Lange, "3D time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology," Ph.D. Thesis, University of Siegen, 2000.
[3] C. Niclass, "Single-Photon Image Sensors in CMOS: Picosecond Resolution for Three-Dimensional Imaging," Ph.D. Thesis, École Polytechnique Fédérale De Lausanne, 2008.
[4] S. Hussmann, T. Ringbeck, and B. Hagebeuker, "A performance review of 3D TOF vision systems in comparison to stereo vision systems," Stereo Vision, pp. 103–120, 2008.
[5] F. Villa, R. Lussana, D. Bronzi, S. Tisa, A. Tosi, F. Zappa, A. Dalla Mora, D. Contini, D. Durini, S. Weyers, and W. Brockherde, "CMOS Imager With 1024 SPADs and TDCs for Single-Photon Timing and 3-D Time-of-Flight," IEEE J. Sel. Top. Quantum Electron., vol. 20, no. 6, pp. 1–10, Nov. 2014.
[6] D. Bronzi, F. Villa, S. Tisa, A. Tosi, F. Zappa, D. Durini, S. Weyers, and W. Brockherde, "100 000 Frames/s 64 × 32 Single-Photon Detector Array for 2-D Imaging and 3-D Ranging," IEEE J. Sel. Topics Quantum Electron, vol. 20, no. 6, pp. 354-363, Nov.-Dec. 2014.
[7] A.P.P. Jongenelen et al., "Maximizing precision over extended unambiguous range for TOF range imaging systems," Instrumentation and Measurement Technology Conference (I2MTC), 2010 IEEE, pp.1575-1580 (2010).